



УНИВЕРЗИТЕТ У НОВОМ САДУ  
ФАКУЛТЕТ ТЕХНИЧКИХ НАУКА У  
НОВОМ САДУ



Buda Bajić Papuga

**Methods for image restoration and  
segmentation by sparsity promoting  
energy minimization**

DOCTORAL DISSERTATION

Буда Бајић Папуга

**Методе за рестаурацију и сегментацију  
дигиталне слике засноване на  
минимизацији функције енергије која  
фаворизује ретке репрезентације  
сигнала**

ДОКТОРСКА ДИСЕРТАЦИЈА

Нови Сад, 2019



УНИВЕРЗИТЕТ У НОВОМ САДУ • ФАКУЛТЕТ ТЕХНИЧКИХ НАУКА  
21000 НОВИ САД, Трг Доситеја Обрадовића 6

## КЉУЧНА ДОКУМЕНТАЦИЈСКА ИНФОРМАЦИЈА

Редни број, <b>РБР:</b>	
Идентификациони број, <b>ИБР:</b>	
Тип документације, <b>ТД:</b>	Монографска документација
Тип записа, <b>ТЗ:</b>	Текстуални штампани запис
Врста рада, <b>ВР:</b>	Докторска дисертација
Аутор, <b>АУ:</b>	Буда Бајић Папуга
Ментор, <b>МН:</b>	Проф. др Наташа Сладоје Матић, проф. др Јоаким Линдبلاد
Наслов рада, <b>НР:</b>	Методe за рестаурацију и сегментацију дигиталне слике засноване на минимизацији функције енергије која фаворизује ретке репрезентације сигнала
Језик публикације, <b>ЈП:</b>	Енглески
Језик извода, <b>ЈИ:</b>	Српски, енглески
Земља публиковања, <b>ЗП:</b>	Република Србија
Уже географско подручје, <b>УГП:</b>	
Година, <b>ГО:</b>	2019
Издавач, <b>ИЗ:</b>	Ауторски репринт
Место и адреса, <b>МА:</b>	Нови Сад, Факултет техничких наука, Трг Доситеја Обрадовића 6
Физички опис рада, <b>ФО:</b> (поглавља/страна/ цитата/табела/слика/ графика/прилога)	8/207/189/5/34/13/6
Научна област, <b>НО:</b>	Примењена математика
Научна дисциплина, <b>НД:</b>	Рачунарска обрада слика
Предметна одредница/Кључне речи, <b>ПО:</b>	Рачунарска обрада слика, рестаурација слике, реконструкција слике у супер-резулацији, отклањање комбинованог Поасоновог и Гаусовог шума, сегментација, минимизација функције енергије
<b>УДК</b>	
Чува се, <b>ЧУ:</b>	Библиотеци Факултета техничких наука, Трг Доситеја Обрадовића 6, Нови Сад
Важна напомена, <b>ВН:</b>	



## KEY WORDS DOCUMENTATION

Извод, **ИЗ:**

Поступак минимизације функције енергије је често коришћен за решавање проблема у обради дигиталне слике. Предмет истраживања тезе су два круцијална задатка дигиталне обраде слике: рестаурација и сегментација слика деградираних шумом и замагљењем. И рестаурација и сегментација су моделовани као проблеми минимизације функције енергије која представља збир две функције: функције фитовања података и регуларизационе функције. Главни допринос тезе је развој нових функција фитовања података и нових регуларизационих функција за рестаурацију и сегментацију.

Методe за рестаурацију (оне код којих је функција замагљења позната и код којих је функцију замагљења потребно оценити на основу датих података као и методе за реконструкцију слике у супер-резољуцији) развијене у оквиру ове тезе третирају мешавину Поасоновог и Гаусовог шума која се појављује у многобројним реалистичним сценаријима. За третирање такве врсте шума користили смо нелинеарну трансформацију и предложили смо нову функцију фитовања података која узима у обзир такву трансформацију. У вези са регуларизационим функцијама смо тестирали хипотезу да се функција Тоталне Варијације која промовише ретку слику у градијентном домену може побољшати уколико се користе тзв. потенцијалне функције. Показали смо да се употребом Хуберове потенцијалне функције може значајно побољшати квалитет рестауриране слике која је деградирана замагљењем и мешавином Поасоновог и Гаусовог шума.

У оквиру тезе смо предложили нову методу сегментације која допушта делимичну покривеност пиксела објектом. Функција фитовања података ове методе укључује и модел замагљења и смањења резолуције. На тај начин је постигнута робустност сегментације у присуству замагљења и добијена могућност сегментирања слике у супер-резољуцији. Додатно, нове регуларизационе функције које промовишу ретке репрезентације слике су предложене.

Предложене методе рестаурације и сегментације која допушта делимичну покривеност пиксела објектом су примењене на слике добијене помоћу електронског микроскопа, хиперспектралне слике и медицинске ЦТ слике.

Датум прихватања теме, **ДП:**

27.12.2018.

Датум одбране, **ДО:**

Чланови комисије, **КО:**

Председник:

др Тибор Лукић, ванредни професор ФТН

Члан:

Др Татјана Давидовић, научни саветник,  
Математички институт Београд

Члан:

Др Лидија Чомић, доцент ФТН

Члан,  
ментор:

др Наташа Сладоје Матић, гостујући професор ФТН

Члан,  
ментор:

др Јоаким Линдблад, истраживач Uppsala University

Потпис ментора



UNIVERSITY OF NOVI SAD • FACULTY OF TECHNICAL SCIENCES  
21000 NOVI SAD, Trg Dositeja Obradovića 6

## KEY WORDS DOCUMENTATION

Accession number, <b>ANO</b> :	
Identification number, <b>INO</b> :	
Document type, <b>DT</b> :	Monographic publication
Type of record, <b>TR</b> :	Textual printed material
Contents code, <b>CC</b> :	PhD thesis
Author, <b>AU</b> :	Buda Bajić Papuga
Mentor, <b>MN</b> :	Prof. dr Nataša Sladoje Matić, prof. dr Joakim Lindblad
Title, <b>TI</b> :	Methods for image restoration and segmentation by sparsity promoting energy minimization
Language of text, <b>LT</b> :	English
Language of abstract, <b>LA</b> :	Serbian, english
Country of publication, <b>CP</b> :	Republic of Serbia
Locality of publication, <b>LP</b> :	
Publication year, <b>PY</b> :	2019
Publisher, <b>PB</b> :	Author's reprint
Publication place, <b>PP</b> :	Novi Sad, Faculty of Technical Sciences, Trg Dositeja Obradovića 6
Physical description, <b>PD</b> : (chapters/pages/ref./tables/pictures/graphs/appendices)	8/207/189/5/34/13/6
Scientific field, <b>SF</b> :	Applied mathematics
Scientific discipline, <b>SD</b> :	Digital image processing
Subject/Key words, <b>S/KW</b> :	Digital image processing, image restoration, mixed Poisson-Gaussian noise, coverage segmentation, energy minimization
<b>UC</b>	
Holding data, <b>HD</b> :	Library of the Faculty of Technical Sciences, Trg Dositeja Obradovića 6, Novi Sad
Note, <b>N</b> :	



## KEY WORDS DOCUMENTATION

Abstract, **AB**:

Energy minimization approach is widely used in image processing applications. Many image processing problems can be modelled in a form of a minimization problem. This thesis deals with two crucial tasks of image analysis workflows: image restoration and segmentation of images corrupted by blur and noise. Both image restoration and segmentation are modelled as energy minimization problems, where energy function is composed of two parts: data fidelity term and regularization term. The main contribution of this thesis is development of new data fidelity and regularization terms for both image restoration and segmentation tasks.

Image restoration methods (non-blind and blind deconvolution and super-resolution reconstruction) developed within this thesis are suited for mixed Poisson-Gaussian noise which is encountered in many realistic imaging conditions. We use generalized Anscombe variance stabilization transformation for removing signal-dependency of noise. We propose novel data fidelity term which incorporates variance stabilization transformation process into account. Turning our attention to the regularization term for image restoration, we investigate how sparsity promoting regularization in the gradient domain formulated as Total Variation, can be improved in the presence of blur and mixed Poisson-Gaussian noise. We found that Huber potential function leads to significant improvement of restoration performance.

In this thesis we propose new segmentation method, the so called coverage segmentation, which estimates the relative coverage of each pixel in a sensed image by each image component. Its data fidelity term takes into account blurring and down-sampling processes and in that way it provides robust segmentation in the presence of blur, allowing at the same time segmentation at increased spatial resolution. In addition, new sparsity promoting regularization terms are suggested: (i) Huberized Total Variation which provides smooth object boundaries and noise removal, and (ii) non-edge image fuzziness, which responds to an assumption that imaged objects are crisp and that fuzziness is mainly due to the imaging and digitization process.

The applicability of here proposed restoration and coverage segmentation methods is demonstrated for Transmission Electron Microscopy image enhancement and segmentation of micro-computed tomography and hyperspectral images.

Accepted by the Scientific Board on,  
**ASB**:

December 27, 2018

Defended on, **DE**:

Defended Board, **DB**:

President:	dr Tibor Lukić, associate professor at FTN
Member:	dr Tatjana Davidović, research professor at Mathematical Institute, Belgrade
Member:	dr Lidija Čomić, assistant professor at FTN
Member, Mentor:	dr Nataša Sladoje Matić, guest professor at FTN
Member, Mentor:	dr Joakim Lindblad, researcher at Uppsala University

Menthor's sign



*To my father*





# Abstract

Energy minimization approach is widely used in image processing applications. Many image processing problems can be modelled in a form of a minimization problem. This thesis deals with two crucial tasks of image analysis workflows: image restoration and segmentation of images corrupted by blur and noise. Both image restoration and segmentation are modelled as energy minimization problems, where energy function is composed of two parts: data fidelity term and regularization term. The main contribution of this thesis is development of new data fidelity and regularization terms for both image restoration and segmentation tasks.

Image restoration methods (non-blind and blind deconvolution and super-resolution reconstruction) developed within this thesis are suited for mixed Poisson-Gaussian noise which is encountered in many realistic imaging conditions. We use generalized Anscombe variance stabilization transformation for removing signal-dependency of noise. We propose novel data fidelity term which incorporates variance stabilization transformation process into account. Turning our attention to the regularization term for image restoration, we investigate how sparsity promoting regularization in the gradient domain formulated as Total Variation, can be improved in the presence of blur and mixed Poisson-Gaussian noise. We found that Huber potential function leads to significant improvement of restoration performance.

In this thesis we propose new segmentation method, the so called coverage segmentation, which estimates the relative coverage of each pixel in a sensed image by each image component. Its data fidelity term takes into account blurring and down-sampling processes and in that way it provides robust segmentation in the presence of blur, allowing at the same time segmentation at increased spatial resolution. In addition, new sparsity promoting regularization terms are suggested: (i) Huberized Total Variation which provides smooth object boundaries and noise removal, and (ii) non-edge image fuzziness, which responds to an assumption that imaged objects are crisp and that fuzziness is mainly due to the imaging and digitization process.

The applicability of here proposed restoration and coverage segmentation methods is demonstrated for Transmission Electron Microscopy image enhancement

---

and segmentation of micro-computed tomography and hyperspectral images.

# Rezime

## Formiranje digitalne slike

Rad većine digitalnih foto-aparata, mikroskopa, CT skenera, mamografa i ostalih uređaja za snimanje digitalnih slika se zasniva na reakciji foto-osetljivog senzora na svetlost. CCD (Charge Coupled Device) [1] i CMOS (Complementary Metal-Oxide Semiconductor) [2, 3] su najzastupljeniji foto-osetljivi senzori. Ovi senzori pretvaraju broj fotona koji dospe na svaku od nekoliko miliona foto osetljivih ćelija (piksela) u električni napon/analogni signal koji se dalje pomoću analogno-digitalnog konvertora pretvara u digitalnu sliku. Intenzitet svakog piksela u digitalnoj slici odgovara broju fotona koju je odgovarajuća foto-osetljiva ćelija prikupila. Digitalne slike dobijene pomoću ovakvih senzora sadrže šum čija jačina zavisi od jačine signala [4]. Takav šum nije aditivan i matematički se modeluje kao linearna kombinacija Poasonove i Gausove slučajne promenljive. Digitalne slike koje imaju široku primenu u astronomiji [5], medicini [6] i biologiji [7] su degradirane mešavinom Poasonovog i Gausovog šuma. Pored degradacije zbog prisustva šuma, digitalne slike često mogu biti i zamagljene (kao rezultat npr. nedovoljno kvalitetnih optičkih komponenti, pomeranja uređaja i/ili objekta prilikom snimanja), što dodatno otežava vidljivost relevantnih detalja. Foto-osetljivi senzori sa relativno malim brojem foto-osetljivih ćelija proizvode digitalne slike male rezolucije u kojima je takođe ograničena vidljivost objekata malih dimenzija.

Proces formiranja digitalne slike se matematički modeluje linearnim modelom:

$$v = \psi(Lu^*) + \eta, \quad (1)$$

koji je baziran na pretpostavci da je snimljena digitalna slika  $v$  nastala linearnom transformacijom originalne slike  $u^*$  koja je dodatno degradirana šumom  $\psi$  i šumom  $\eta$  čije jačine, redom, zavise, odnosno ne zavise, od jačine signala. Linearni operator  $L$  se naziva direktni operator. Primeri direktnih operatora su identičko preslikavanje (koji ne utiče na sliku), operator zamagljenja (tada  $Lu^*$  predstavlja konvoluciju slike i funkcije zamagljenja i slika  $v$  je zamagljena) i operator uzorkovanja (tada je slika  $v$  manje rezolucije u poređenju sa  $u^*$ ). Neki



(a) Originalna slika



(b) Slika degradirana mešavinom Poasonovog i Gausovog šuma.



(c) Zamagljena slika bez šuma.



(d) Zamagljena slika degradirana mešavinom Poasonovog i Gausovog šuma.



(e) Zamagljena slika smanjene rezolucije, degradirana mešavinom Poasonovog i Gausovog šuma.

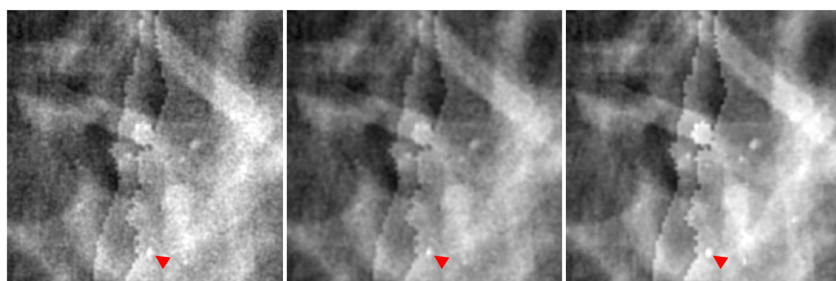
Slika 1: Primeri degradiranih slika nastalih primenom različitih operatora  $L$ .

primeri degradiranih slika nastalih primenom različitih operatora  $L$  su prikazani na Slici 1.

## Restauracija digitalne slike

Metode za restauraciju digitalne slike imaju za cilj da generišu digitalnu sliku bez šuma i zamagljenja polazeći od snimljene digitalne slike koja je zamagljena i degradirana šumom. U nekim slučajevima je, istovremeno sa uklanjanjem šuma i zamagljenja, od interesa i povećati rezoluciju slike u odnosu na onu na kojoj je snimljena; to se postiže metodama za rekonstrukciju slike u super-rezoluciji.

Razvoj metoda za restauraciju je motivisan potrebom da se boljim kvalitetom slika obezbede bolji rezultati njihove dalje, *manuelne* ili *automatske*, analize.



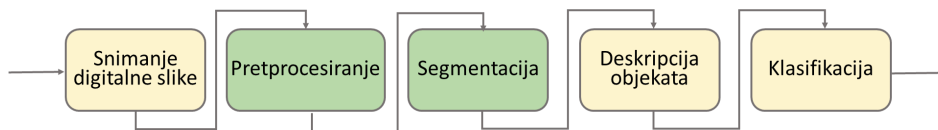
(a)



(b)

Slika 2: (a) Primer slike degradirane šumom u medicini - mamografska slika degradirana šumom (levo), restaurirana slika (sredina) i originalna slika bez šuma (desno). Strelice pokazuju mikrokalcifikacije, male naslage kalcijuma, koje su rani indikatori raka dojke. Njihova vidljivost je smanjena u slikama degradiranim šumom (izvor [8], adaptirana slika). (b) Primer zamagljene slike u forenzici - zamagljene registarske tablice (gore) i restaurirane tablice (dole) (izvor [9], adaptirana slika).

Restauracija digitalnih slika ima veoma važnu ulogu kod *manuelnih* metoda za interpretaciju sadržaja digitalne slike u medicini, biologiji, forenzici, astronomiji i mnogim drugim oblastima. Medicinske digitalne slike snimljene CT aparatima i mamografima (Slika 2 (a)) pomoću X-zraka neizostavno su degradirane šumom usled niskih doza radijacije koje se koriste kako bi se smanjile negativne posledice zračenja na zdravlje pacijenata [8, 10]. Takođe, često su slike u medicini zamagljene usled pokreta pacijenta prilikom snimanja. Metode restauracije se često primenjuju u takvim situacijama kako bi se izoštrila vidljivost detalja i radiolozima pomoglo prilikom vizuelne analize digitalnih slika i davanja dijagnoze na osnovu istih. U biologiji, prilikom snimanja digitalne slike elektronskim mikroskopom često dolazi do pomeranja biološkog materijala koji se analizira i/ili do njegovog oštećenja usled bombardovanja elektronima, pogotovo kada se koriste velike doze elektrona i ukoliko se snimanje vrši dugo



Slika 3: Šematski prikaz automatske obrade digitalne slike. Istraživanje sprovedeno u tezi pripada zelenim blokovima u dijagramu.

vremena. Kako bi se ove negativne posledice izbegle, najčešće se smanjuje vreme snimanja i doza elektrona što rezultuje velikim prisustvom šuma u slikama i povećava potrebu za restauracijom [11, 12]. Primena metoda za restauraciju povećava vidljivost detalja na slikama biološkog materijala i pomaže pouzdanijoj kliničkoj dijagnostici. U forenzici, kako bi se precizno identifikovale osobe koje su bile na licu mesta zločina koji se dogodio u okruženju smanjene vidljivosti, od ključne važnosti može biti upotreba metoda za restauraciju kako bi se digitalna slika izoštrila i detalji bolje videli. Identifikacija oznaka na zamagljenim digitalnim slikama registraskih tablica automobila je drugi tipičan primer gde je restauracija od velike pomoći u forenzici, Slika 2 (b).

Restauracija digitalnih slika takođe ima veoma važnu ulogu i u *automatskim* metodama za obradu digitalne slike i računarske vizije. Pomenute automatske metode najčešće počinju uklanjanjem šuma i zamagljenja u digitalnoj slici (pretprocesiranje), zatim se vrši segmentacija slike, prepoznavanja objekata u slici, klasifikacija objekata, i tome slično (Slika 3). Segmentacija digitalne slike ima za cilj da identifikuje celine u digitalnoj slici koje sadrže piksele zajedničkih karakteristika, npr. poput intenziteta, ili pak pripadaju istoj povezanoj komponenti. Prisustvo šuma i zamagljenja u digitalnoj slici rezultira smanjenom preciznošću kako segmentacije, tako i ostalih daljih postupaka koji se primenjuju u automatskoj obradi digitalne slike. Slika 4 (e, f) ilustruje loše rezultate dobijene metodama za detekciju ivica i segmentaciju kada se primene na sliku degradiranu šumom (ivice i segmenti su detektovani tamo gde ne postoje).

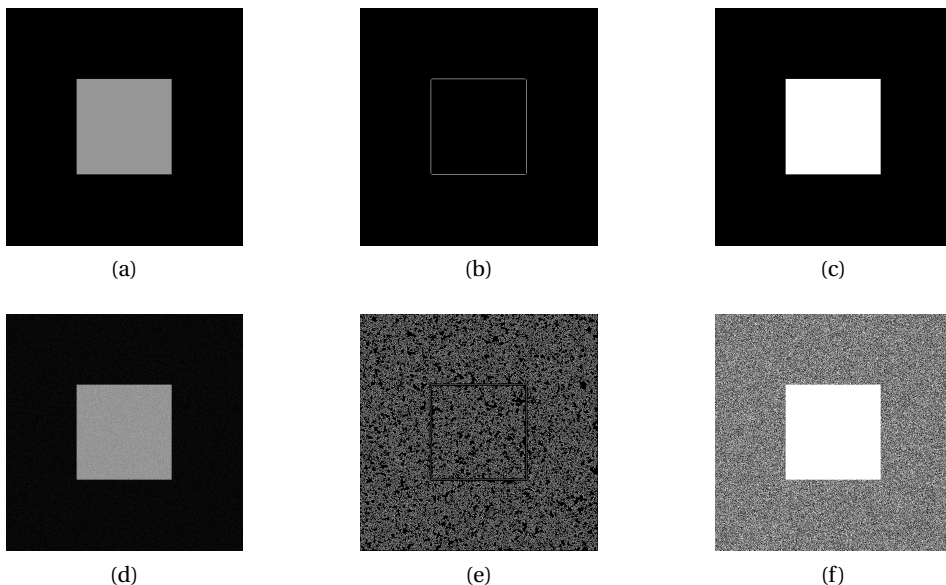
## Restauracija zasnovana na minimizaciji funkcije energije

Mnoge efikasne metode za restauraciju digitalnih slika i rekonstrukciju slika u super-rezoluciji su bazirane na minimizaciji, tzv. funkcije energije oblika

$$E(u) = D(u; v) + \lambda R(u). \quad (2)$$

Funkcija energije predstavlja zbir dve funkcije:

- funkcije fitovanja podataka  $D$  koja meri odstupanje polazne zamagljene slike, degradirane šumom, i male rezolucije, od željene slike bez šuma i



Slika 4: Rezultati metoda za detekciju ivica i segmentaciju primenjene na sliku degradiranu šumom. Slika (a) je originalna slika bez šuma, (d) je slika (a) degradirana šumom. Slike (b, c) i (e, f) su rezultati Canny metode za detekciju ivica i Otsu segmentacije primenjene na (a) i (d), respektivno.

zamagljenja, eventualno povećane rezolucije, koja je definisana na osnovu pretpostavljenog modela formiranja digitalne slike;

- regularizacione funkcije  $R$  koja ima za cilj da nametne željene osobine krajnjem rešenju.

Argument koji minimizuje funkciju energije,

$$\hat{u} = \underset{u}{\operatorname{argmin}} E(u) \quad (3)$$

predstavlja sliku poboljšanog kvaliteta (koja sadrži manje šuma i ima oštrije ivice u poređenju sa polaznom slikom), uspostavljajući ravnotežu između datih podataka i postavljenih uslova.

Različite funkcije fitovanja podataka i regularizacione funkcije postoje u literaturi. Neodgovarajući izbor bilo koje od ovih funkcija rezultuje nedovoljno kvalitetnom restauriranom slikom. Funkcija fitovanja podataka zavisi od pretpostavke o raspodeli koju šum prati, dok regularizaciona funkcija zavisi od odluke koju vrstu efekata u slici želimo da uklonimo/sačuvamo.

---

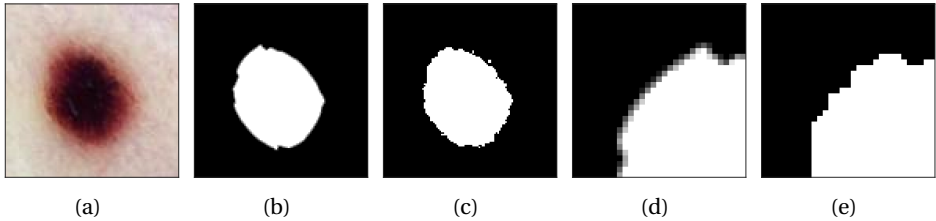
U literaturi su poznate dve vrste metoda za restauraciju: one kod kojih je poznata funkcija zamagljenja [13–15], i one kod kojih ta funkcija nije poznata, pa je dodatno potrebno i nju oceniti iz podataka [16–19]. Većina postojećih metoda, kao i metoda za rekonstrukciju digitalne slike u super-rezoluciji [20–22], zasniva se na pretpostavci da prisutni šum prati Gausovu raspodelu ili pak da se raspodela šuma može aproksimirati Gausovom raspodelom. Ovakva pretpostavka znatno olakšava optimizaciju i rešavanje problema jer je odgovarajuća funkcija fitovanja podataka kvadratna i ima jedan lokalni minimum (koji je ujedno i globalni minimum) za čije pronalaženje postoje efikasni numerički postupci. Intenzitet kombinovanog Poasonovog i Gausovog šuma u digitalnoj slici zavisi od jačine signala, a pretpostavka o prisustvu isključivo Gausovog šuma vodi do lošijih performansi metoda za restauraciju digitalne slike i rekonstrukciju digitalne slike u super-rezoluciji, pogotovo u uslovima kada foto-osetljive ćelije kolektuju mali broj fotona. U literaturi je pretpostavka o mešavini Gausovog i Poasonovog šuma retko postavljena zbog teoretskih i praktičnih teškoća koje prouzrokuje. Funkcija fitovanja podataka koja teorijski odgovara ovoj vrsti šuma se sastoji od beskonačnog broja sabiraka i kao takva se ne može računarski implementirati, već zahteva izvesnu vrstu aproksimacije. Metoda predložena u radu [23] ovakvu beskonačnu sumu aproksimira konačnim brojem sabiraka, što rezultira izuzetno neefikasnom i sporom restauracijom. Nijedna od postojećih metoda za restauraciju slike kod kojih je dodatno potrebno iz podataka oceniti i funkciju zamagljenja, kao ni metoda za rekonstrukciju digitalnih slika u super-rezolucije, nije precizno tretirala ovu vrstu šuma.

Pored najčešće korišćene kvadratne funkcije fitovanja podataka, funkcije energije u velikom broju slučajeva uključuju tzv. funkciju Totalne Varijacije [17] kao regularizacionu funkciju. Funkcija Totalne Varijacije nije diferencijabilna, što predstavlja problem pri optimizaciji. Pored funkcije Totalne Varijacije predložene su i druge regularizacione funkcije, koje se generišu korišćenjem tzv. potencijalnih funkcija. Međutim, u literaturi ne postoji studija o efikasnosti funkcija za restauraciju zamagljenih slika degradiranih različitim vrstama šuma koja bi omogućila odgovarajući izbor metode u datim uslovima.

## **Segmentacija koja dopušta delimičnu pokrivenost piksela**

Raznovrsne metode obrade digitalne slike poput segmentacije [24, 25], rekonstrukcije CT [26, 27] i MRI [28, 29] slika, metode popunjavanja nedostajućih delova u slici [30–32] i druge, mogu takođe biti zasnovane na minimizaciji funkcije energije sa odgovarajuće dizajniranim regularizacionim funkcijama i funkcijom fitovanja podataka. Fokus ove teze je na metodama za restauraciju (deo preprocesiranja) i segmentaciju (zeleni blokovi Slike 3) digitalne slike zasnovanim na minimizaciji funkcije energije koja favorizuje retke reprezentacije signala.





Slika 5: Segmentacija koja dopušta delimičnu pokrivenost piksela: (a) originalna slika, (b) segmentacija koja dopušta delimičnu pokrivenost piksela i (c) klasična binarna segmentacija od (a), (d) i (e) zumirani delovi slika (b) i (c).

Klasične metode segmentacije generišu binarne slike u kojima pikseli koji pripadaju objektu imaju vrednost 1, a pikseli koji pripadaju pozadini objekta imaju vrednost 0. U ovakvoj reprezentaciji slike svaki piksel je ili u potpunosti pokriven objektom, ili nije uopšte pokriven objektom. Daleko prirodnija reprezentacija digitalne slike je ona u kojoj pikseli mogu biti delimično pokriveni objektom i pozadinom (Slika 5). Segmentacija bazirana na ovakvoj reprezentaciji generiše slike u kojima pikseli koji pripadaju objektu imaju vrednost 1, pikseli koji pripadaju pozadini objekta imaju vrednost 0, dok pikseli koji su delimično pokriveni objektom imaju vrednost između 0 i 1, koja odgovara nivou pokrivenosti piksela objektom. U radovima [33–37] je pokazano da je reprezentacija digitalne slike u kojoj pikseli mogu biti delimično pokriveni objektom bogatija informacijom. Deskriptori segmentovanih oblika razvijeni za ovakvu reprezentaciju su precizniji u poređenju sa deskriptorima koji se koriste u tradicionalnoj binarnoj reprezentaciji slike [33–37].

## Segmentacija koja dopušta delimičnu pokrivenost piksela zasnovana na minimizaciji funkcije energije

U radu [38] je predloženo da se segmentacija  $\hat{A}$  koja dopušta delimičnu pokrivenost piksela slike  $\nu$  računa kao

$$\hat{A} = \operatorname{argmin} E(A), \quad (4)$$

gde je funkcija energije definisana

$$E(A) = D(A; \nu) + \mu R(A). \quad (5)$$

Ovde se pretpostavlja da je snimljena slika formirana kao linearna kombinacija reprezentativnih predstavnika pojedinih klasa (delova slike), tj. slika  $\nu$  se može modelovati

$$\nu = A \cdot C + \text{noise}, \quad (6)$$

---

gde je  $C$  matrica reprezentata pojedinih klasa. Ovakav model je baziran na rigidnoj pretpostavci da je rub objekta širine ne veće od jednog piksela, tj. da najviše jedan piksel na rubu između objekta i pozadine može biti delimično pokriven objektom. Takva pretpostavka nije realistična u uslovima kada je slika degradirana zamagljenjem tj. kada ivica objekta nije dovoljno oštra. Metoda predložena u radu [38] zasnovana na minimizaciji funkcije energije (5) nije dovoljno precizna kada se primeni na zamagljene slike. Ovo motiviše da se razvije nova metoda segmentacije koja će uključiti linearni direktni model formiranja slike (1) u postojeći model linearne kombinacije reprezentata klasa (6). Ovim se poboljšava metoda predložena u [38] i povećava se njena primenljivost na zamagljene slike. Dodatno, takva metoda omogućava segmentaciju zasnovanu na delimičnoj pokrivenosti piksela i u super-rezoluciji.

## Naučni doprinos i originalni rezultati istraživanja

1. U tezi je pokazana važnost uključivanja realistične pretpostavke o tome da šum prisutan u digitalnim slikama ima raspodelu koja se modeluje kao linearna kombinacija Gausove i Poasonove slučajne promenljive, pri rešavanju problema restauracije digitalne slike, kao i problema rekonstrukcije digitalnih slika u super-rezoluciji. U radovima koji su uključeni u tezu, Publikacije I-III, pokazana je važnost preciznog tretiranja mešavine Gausovog i Poasonovog šuma i predložene su nove metode za restauraciju digitalnih slika i rekonstrukciju slike u super-rezoluciji, u slučaju kad funkcija fitovanja podataka uključuje pretpostavku o takvoj mešavini šuma. Predloženi način tretiranja ovakvog šuma se sastoji u primeni nelinearne transformacije podataka koja rezultuje time da šum prisutan u transformisanim podacima približno prati Gausovu raspodelu. Uključivanje pomenute transformacije u funkciju fitovanja podataka rezultuje bržim i preciznijim metodama za restauraciju slika koje su degradirane mešavinom Poasonovog i Gausovog šuma, u poređenju sa metodama koje uopšte ne tretiraju ovu vrstu šuma ili pak onima koje, poput metode predložene u radu [23], tretiraju pomenuti šum na manje efikasan način.
2. U tezi je pokazano da se performanse metoda za restauraciju zasnovanih na minimizaciji funkcije energije mogu značajno poboljšati ukoliko se klasična regularizaciona funkcija Totalne Varijacije zameni regularizacijom koja uključuje Huberovu potencijalnu funkciju. Ovi rezultati su objavljeni u Publikacijama I i IV i uopštavaju prethodno poznate rezultate tako što uključuju degradaciju zamagljenjem i smanjenjem rezolucije, uz prethodno posmatrane slučajeve prisustva šuma.
3. Dodatno, u tezi je pokazana tačnost hipoteze da se, ukoliko se pri definisanju funkcije energije za segmentaciju slike koja dopušta delimičnu pokri-

---

venost piksela uzme u obzir pretpostavka o formiranju zamagljene slike, odnosno ukoliko se omogući uklanjanje tog zamagljenja, segmentacija može značajno poboljšati. Metoda koja istovremeno otklanja zamagljenje iz slike i daje ocenu o pokrivenosti svakog piksela objektom je predložena u Publikaciji VI. Razvijena metoda dodatno uzima u obzir i pretpostavku da je data slika formirana na maloj rezoluciji, i modeluje taj proces, te daje ocenu pokrivenosti piksela pri većoj rezoluciji od polazne (daje ocenu pokrivenosti piksela u super-rezoluciji).

4. U tezi su predložene nove regularizacione funkcije koje favorizuju retke reprezentacije slika za segmentaciju koja dopušta delimičnu pokrivenost piksela, koje su značajno poboljšale performanse predložene metode.
5. Efikasnost metoda numeričke optimizacije koje se primenjuju za pronalazanje minimuma funkcija energije značajno utiče na efikasnost restauracije i segmentacije. U svim predloženim metodama restauracije (Publikacije I-V) funkcija energije je optimizovana pomoću metode Spektralnog Projektovanog Gradijenta [39] (SPG). Ova metoda je veoma fleksibilna u pogledu promene ili uključivanja novih funkcija fitovanja podataka i regularizacionih funkcija. Za rešavanje složenijih nekonvexnih problema optimizacije u vezi sa restauracijom kod koje funkcija zamagljenja nije poznata i segmentacijom koja dopušta delimičnu pokrivenost piksela, predloženi su složeniji algoritmi optimizacije. Za metodu restauracije kod koje funkcija zamagljenja nije poznata (Publikacija II) predložen je postupak Alternativne Minimizacije (AM), gde je problem podeljen u konveksne potprobleme. Za segmentaciju koja dopušta delimičnu pokrivenost piksela (Publikacija VI) predložen je optimizacioni algoritam Graduated Non-Convexity (GNC) koji se sastoji od postepenog povećanja složenosti problema. Pomenuti algoritam koristi rešenja numerički lakših potproblema za inicijalizaciju optimizacije kompleksnijih problema. Potproblemi unutar AM i GNC algoritama su optimizovani pomoću SPG-a.

## Struktura teze

Teza je podeljena u 8 poglavlja. Prvo poglavlje objašnjava važnost restauracije i precizne segmentacije ističući istovremeno izazove vezane za oba problema. Ovo poglavlje uvodi pojam funkcije energije koja se koristi za rešavanje oba problema u tezi. Većina pomenutih izazova je rešena u okviru izrade doktorske disertacije, i rezultati su sumirani u Publikacijama I-VI koje se mogu naći u Dodatku teze.

---

Drugo poglavlje prezentuje statistički pristup baziran na Bajesovoj teoremi koji se koristi kako bi se izvele Maximum a Posteriori (MAP) ocene za inverzne probleme kod kojih su slike modelovane kao slučajne promenljive. Ovo poglavlje takođe povezuje MAP ocene i funkcije energije.

Treće poglavlje prezentuje različite funkcije fitovanja podataka za restauraciju i segmentaciju koja dopušta delimičnu pokrivenost piksela, koje se javljaju u literaturi, kako one koje su jednake funkciji verodostojnosti tako i njihove različite aproksimacije. Ovo poglavlje uvodi nove funkcije fitovanja podataka koje su predložene u okviru disertacije. Konkretno, za tretiranje kombinovanog Gausovog i Poasovanog šuma, predložena je nova funkcija fitovanja podataka. Predložena funkcija uključuje nelinearnu transformaciju koja rezultuje time da šum prisutan u transformisanim podacima prati približno Gausovu raspodelu. Za segmentaciju koja dopušta delimičnu pokrivenost piksela, u model linearne kombinacije reprezentata klasa je uključen linearni direktni model formiranja slike koji u sebi sadrži pretpostavku o zamagljenju i smanjenju rezolucije snimljene slike.

Četvrto poglavlje predstavlja različite regularizacione funkcije za restauraciju i segmentaciju. Regularizacione funkcije koje favorizuju retke reprezentacije signala su uvedene sa fokusom na funkciju Totalne Varijacije koja favorizuje retku gradijentnu sliku. Potencijalne funkcije koje očuvavaju ivice su definisane i postavljena je hipoteza da se, ukoliko se ove funkcije koriste u restauraciji umesto identičkog preslikavanja (koje odgovara Totalnoj Varijaciji), mogu postići bolji rezultati, odnosno kvalitetnija restauracija. Za segmentaciju, posmatrana je Mumford-Shah regularizacija. Objasnjena je veza između funkcije Totalne Varijacije i Mumford-Shah regularizacije. Ta veza je motivisala da umesto prethodno korišćene regularizacione funkcije za segmentaciju koja dopušta delimičnu pokrivenost piksela iz rada [38] predložimo korišćenje regularizacione funkcije bazirane na Totalnoj Varijaciji. Dodatno, predložena je i nova regularizaciona funkcija koja promoviše da većina piksela bude potpuno pokrivena objektom ili pozadinom. Time je značajno pojednostavljen postupak optimizacije posmatrane nekonveksne funkcije energije.

Peto poglavlje navodi najvažnije aspekte algoritama dizajniranih za efikasnu numeričku optimizaciju funkcija energije. Navedeni su mnogobrojni numerički postupci razvijeni u skorije vreme koji se koriste za rešavanje problema restauracije i segmentacije. U disertaciji funkcija energije je optimizovana pomoću metode Spektralnog Projektovanog Gradijenta [39] (SPG). U ovom poglavlju je prezentovan SPG i dva nova optimizaciona algoritma iz Publikacija II i VI, AM i GNC.

Šesto poglavlje prezentuje skup slika korišćenih za evaluaciju predloženih metoda

---

u okviru disertacije. Ovo poglavlje takođe prezentuje najvažnije kvantitativne i kvalitativne rezultate dobijene pomoću predloženih metoda.

Sedmo poglavlje prezentuje primenu metoda za restauraciju na slikama dobijenim pomoću elektronskog mikroskopa (Publikacije II i V) i metode za segmentaciju koja dopušta delimičnu pokrivenost piksela u analizi hiperspektralnih satelitskih slika.

Osmo, poslednje, poglavlje, sumira zaključke istraživanja i prezentuje potencijalne buduće pravce istraživanja.



# Contents

<b>Abstract (English/Srpski)</b>	<b>iii</b>
<b>Rezime</b>	<b>v</b>
<b>List of Publications</b>	<b>xxi</b>
<b>Abbreviations</b>	<b>xxiii</b>
<b>Notation</b>	<b>xxv</b>
<b>List of figures</b>	<b>xxvi</b>
<b>List of tables</b>	<b>xxviii</b>
<b>Introduction</b>	<b>1</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Image formation process . . . . .	1
1.2 Image restoration . . . . .	3
1.2.1 Image restoration by energy minimization . . . . .	6
1.3 Coverage segmentation . . . . .	8
1.3.1 Coverage segmentation by energy minimization . . . . .	8
1.4 Problem Statement and Research Objectives . . . . .	9
1.5 Contribution of This Thesis . . . . .	10
1.6 Thesis Outline . . . . .	11
<b>Inverse problems</b>	<b>13</b>
<b>2 Inverse problems</b>	<b>13</b>
2.1 Maximum a posteriori (MAP) estimator . . . . .	14
2.2 MAP estimators for inverse problems and their link to energy functions . . . . .	15
<b>Data fidelity terms</b>	<b>17</b>

<b>3</b>	<b>Data fidelity terms</b>	<b>17</b>
3.1	Preliminaries and notation . . . . .	17
3.2	MAP based data fidelity terms for super-resolution image reconstruction . . . . .	18
3.2.1	Data fidelity term suited for Gaussian noise . . . . .	18
3.2.2	Data fidelity term suited for Poisson noise . . . . .	20
3.2.3	Data fidelity term suited for Poisson-Gaussian noise . . . . .	22
3.3	Alternative data fidelity terms for non-Gaussian noise . . . . .	23
3.3.1	Data fidelity term based on Variance Stabilization Transformation . . . . .	24
3.4	Data fidelity term for coverage segmentation . . . . .	27
3.4.1	Preliminaries and notation . . . . .	27
3.4.2	Linear unmixing . . . . .	28
3.4.3	Data fidelity term for coverage segmentation based on linear unmixing . . . . .	28
3.4.4	Data fidelity term for super-resolution coverage segmentation . . . . .	29
	<b>Regularization terms</b>	<b>31</b>
<b>4</b>	<b>Regularization terms</b>	<b>31</b>
4.1	Regularization terms for image restoration . . . . .	31
4.1.1	Sparse signals . . . . .	31
4.1.2	Sparsity promoting norms . . . . .	32
4.1.3	Total Variation based regularization . . . . .	34
4.1.4	Edge preserving potential functions . . . . .	35
4.2	Regularization terms for segmentation . . . . .	38
4.2.1	Mumford-Shah penalty . . . . .	38
4.2.2	Link between Mumford-Shah penalty and Total Variation via level set functions . . . . .	40
4.2.3	Regularization terms for coverage segmentation . . . . .	42
4.2.4	Regularization terms for super-resolution coverage segmentation . . . . .	43
	<b>Numerical optimization</b>	<b>46</b>
<b>5</b>	<b>Numerical optimization</b>	<b>47</b>
5.1	Minimization of energy functions . . . . .	47
5.2	Spectral Projected Gradient . . . . .	50
5.3	Alternating minimization scheme for blind deconvolution . . . . .	52
5.4	Graduate non-convexity for super-resolution coverage segmentation . . . . .	53
	<b>Performance evaluation</b>	<b>54</b>



<b>6 Performance evaluation</b>	<b>55</b>
6.1 Performance evaluation of image restoration suited for PG noise .	55
6.1.1 Quantitative evaluation . . . . .	55
6.1.1.1 Data set . . . . .	55
6.1.1.2 Quality measures . . . . .	56
6.1.1.3 Optimal parameter selection . . . . .	58
6.1.2 Importance of an appropriate treatment of Poisson-Gaussian noise . . . . .	58
6.1.3 Comparison of MAP and VST approaches . . . . .	60
6.1.4 Evaluation of different potential functions . . . . .	63
6.1.4.1 Deconvolution and denoising in presence of Gaussian noise . . . . .	63
6.1.4.2 Deconvolution and denoising in presence of Poisson-Gaussian noise . . . . .	65
6.1.4.3 On optimal parameter selection for Huber potential	66
6.1.5 Blind deconvolution in presence of Poisson-Gaussian noise	67
6.2 Performance evaluation of super-resolution coverage segmentation	68
6.2.1 Synthetic tests . . . . .	68
6.2.2 An illustrative example on a naturally degraded image . . .	73
<b>Applications</b>	<b>75</b>
<b>7 Applications</b>	<b>75</b>
7.1 Restoration of transmission electron microscopy images . . . . .	75
7.1.1 Transmission electron microscopy . . . . .	75
7.1.2 Blind deconvolution of long exposure TEM images . . . . .	77
7.1.3 Denoising of short exposure TEM images . . . . .	80
7.2 Application of super-resolution coverage segmentation in medicine	84
7.3 Application of super-resolution coverage segmentation in remote sensing . . . . .	85
<b>Conclusion and future work</b>	<b>89</b>
<b>8 Conclusion</b>	<b>89</b>
8.1 Summary of contributions . . . . .	89
8.2 Future work . . . . .	90
<b>Acknowledgements</b>	<b>93</b>
<b>Bibliography</b>	<b>112</b>

## Contents

---

<b>Appendix</b>	<b>113</b>
Publication I . . . . .	114
Publication II . . . . .	125
Publication III . . . . .	131
Publication IV . . . . .	139
Publication V . . . . .	148
Publication VI . . . . .	154

# List of Publications

The work of this thesis is based on the following publications; copies of these publications I-VI can be found in Appendix:

- I **Bajić B**, Lindblad J and Sladoje N. Restoration of images degraded with signal-dependent noise based on energy minimization: an empirical study. *SPIE Journal of Electronic Imaging*, 2016, 25(4), 043020.
- II **Bajić B**, Lindblad J and Sladoje N. Blind restoration of images degraded with mixed Poisson-Gaussian noise with application in Transmission Electron Microscopy. In *Proceedings of the 13<sup>th</sup> IEEE International Symposium on Biomedical Imaging (ISBI)*, Prague, April 2016, pp. 123–127.
- III **Bajić B**, Lindblad J and Sladoje N. Single image super-resolution reconstruction in presence of mixed Poisson-Gaussian noise. In *Proceedings of the 6<sup>th</sup> IEEE International Conference on Image Processing Theory, Tools and Applications (IPTA)*, Oulu, Finland, December 2016, pp. 1–6.
- IV **Bajić B**, Lindblad J and Sladoje N. An Evaluation of Potential Functions for Regularized Image Deblurring. In *Proceedings of International Conference on Image Analysis and Recognition (ICIAR)*, Algarve, Portugal, October 2014, Lecture Notes on Computer Science, 8814, pp. 150–158.
- V **Bajić B\***, Suveer A\*, Gupta A\*, Pepić I, Lindblad J, Sladoje N, and Sintorn I-M. Denoising of Short Exposure Transmission Electron Microscopy Images for Ultrastructural Enhancement. In *Proceedings of the 15<sup>th</sup> IEEE International Symposium on Biomedical Imaging (ISBI)*, Washington D.C., USA, April 2018.
- VI **Bajić B**, Lindblad J and Sladoje N. Super-resolution coverage segmentation by linear unmixing in presence of blur and noise. *SPIE Journal of Electronic Imaging*, 2019, 28(1), 013046.

---

\* These authors have contributed equally.

## List of Publications

---

Work presented in Publications I-III was mainly done by the author with helpful discussions with co-authors. The author designed, implemented and evaluated the proposed methods and wrote the main part of the text. In Publication IV, the author contributed to the design, implementation and evaluation of the proposed method while writing was shared between all co-authors. Publication V resulted from the equally shared work between the co-authors. The author's contribution to the Publication VI was in developing the method and performing experiments with valuable feedback from co-authors, while writing was equally shared between all three co-authors.

# Abbreviations

PG	Poisson-Gaussian (noise)
MAP	Maximum a posteriori approach
VST	Variance stabilization transformation
MAPG	MAP based restoration method suited for Gaussian noise
MAPP	MAP based restoration method suited for Poisson noise
VSTPG	Restoration method suited for Poisson-Gaussian noise, based on VST approach
MAPPG	MAP based restoration method suited for Poisson-Gaussian noise
HR	High resolution
LR	Low resolution
SR-G	Super-resolution reconstruction suited for Gaussian noise
SR-PG	Super-resolution reconstruction suited for Poisson-Gaussian noise, based on VST approach
SR-LOR	Super-resolution reconstruction with Lorentzian data fidelity term
SR-HUB	Super-resolution reconstruction with Huber data fidelity term
TV	Total Variation
PSNR	Peak-Signal-to-Noise Ratio
SNR	Signal-to-Noise Ratio
SSIM	Structural Similarity Index Measure
SPG	Spectral Projected Gradient
GNC	Graduated Non-convexity
AM	Alternating Minimization
TEM	Transmission Electron Microscopy
SEM	Scanning Electron Microscopy
CT	Computed Tomography
MRI	Magnetic Resonance Imaging
CCD	Charge Coupled Device
CMOS	Complementary Metal-Oxide Semiconductor
PCD	Primary Ciliary Dyskinesia
CNN	Convolutional Neural Networks



# Notation

## Vectors

$$u^* = [u_1^*, u_2^*, \dots, u_{N_h}^*]^T$$

$$\hat{u} = [\hat{u}_1, \hat{u}_2, \dots, \hat{u}_{N_h}]^T$$

$$v = [v_1, v_2, \dots, v_{N_l}]^T$$

$$u = [u_1, u_2, \dots, u_{N_h}]^T$$

small letter:

original HR noise-free and blur-free image with  $N_h$  pixels

restored HR image obtained as  $\arg \min_u E(u)$

sensed single-band LR image of  $N_l$  pixels

unknown HR noise-free and blur-free image

## Random variable

$\mathcal{V}$

capital calligraphic letter :

random variable of a sensed image

$\mathcal{U}^*$

random variable of the original image

$\mathcal{U}$

random variable corresponding to unknown image

$\mathcal{N}$

random variable with Gaussian distribution

$\mathcal{P}$

random variable with Poisson distribution

## Matrices

$$V = [v_{i,k}]_{N_l \times b}$$

capital letter:

sensed multi-band LR image,  $b$  number of bands

$$A = [\alpha_{i,j}]_{N_l \times m}$$

coverage segmentation of LR image into  $m$  components

$$C = [c_{j,k}]_{m \times b}$$

end-member matrix

$$S_{N_l \times N_h}$$

down-sampling matrix

$$H_{N_h \times N_h}$$

blurring matrix

## Functions:

$$E(\cdot)$$

energy function

$$D(\cdot)$$

data fidelity term

$$R(\cdot)$$

regularization term

$$\mathbf{u} : \Omega \rightarrow \mathbb{R}^b$$

continuous representation of an image

$$\Omega \subset \mathbb{R}^2$$

image domain

$\ell_p$  norm

$$\ell_p(u) = \left( \sum_{i=0}^{N_h} |u_i|^p \right)^{\frac{1}{p}}$$





# List of Figures

1	Primeri degradiranih slika nastalih primenom različitih operatora $L$ .	vi
2	Primeri slika degradiranih šumom i zamagljenjem u medicini i forenzici. . . . .	vii
3	Automatizovana obrada digitalne slike. . . . .	viii
4	Rezultati metoda za detekciju ivica i segmentaciju primenjene na sliku degradiranu šumom. . . . .	ix
5	Segmentacija koja dopušta delimičnu pokrivenost piksela dermoskopske slike. . . . .	xi
1.1	Examples of corrupted images for different forward operators $L$ .	2
1.2	Examples of noisy and blurred medical and forensic images. . . .	4
1.3	Automated image analysis work-flow. . . . .	5
1.4	Poor edge detection and segmentation performances on noisy image. . . . .	6
1.5	Coverage segmentation of dermoscopic image. . . . .	8
3.1	Down-sampling process. . . . .	18
3.2	Examples of corrupted images with different types of noise. . . .	21
3.3	Variance stabilization transformations for Poisson random variable.	25
4.1	Plot of discrete sinus signal in time and Fourier domain. . . . .	32
4.2	Sparsity promotion achieved by minimization of $\ell_1$ norm. . . . .	33
4.3	Illustration of sparsity of the gradient magnitude image. . . . .	35
4.4	Potential functions - plot. . . . .	36
4.5	Illustration of Mumford-Shah image segmentation. . . . .	38
4.6	Piecewise smooth approximation using the Mumford-Shah functional. . . . .	40
4.7	Illustration of level set function of a circle. . . . .	41
6.1	Test images used for quantitative evaluation of restoration methods.	55
6.2	Improvement in PSRN due to appropriate treatment of PG noise in the image deconvolution task. . . . .	59

## List of Figures

---

6.3	Improvement in PSNR due to appropriate treatment of PG noise in the super-resolution task. . . . .	60
6.4	Super-resolution restoration of a test image, degraded by blur and PG noise. . . . .	61
6.5	Performance of VSTPG on confocal microscopy image. . . . .	62
6.6	Improvement in PSNR and SSIM for all potential functions in image deconvolution task. . . . .	64
6.7	Illustration of improved edge preservation by Huber potential in image deconvolution task. . . . .	65
6.8	PSNR improvement achieved by Huber potential in comparison to TV, for VSTPG deconvolution method. . . . .	65
6.9	Relation between regularization parameter $\lambda$ for TV and Huber potentials. . . . .	66
6.10	Quantitative performances of the blind deconvolution method for PG noise. . . . .	67
6.11	Qualitative performances of the blind deconvolution method for PG noise. . . . .	67
6.12	Test image used in synthetic tests for coverage segmentation. . . . .	68
6.13	Qualitative performances of 5 considered coverage segmentation methods. . . . .	70
6.14	Quantitative performances of 5 considered coverage segmentation methods. . . . .	71
6.15	Performance of SRCS-HTV $\tilde{F}$ at 3 times increased resolution. . . . .	71
6.16	Robustness of SRCS-HTV $\tilde{F}$ w.r.t.to end-member matrix and blur matrix mismatch. . . . .	72
6.17	Coverage segmentation of a blurry and noisy real color image. . . . .	74
7.1	Transmission Electron Microscope . . . . .	76
7.2	TEM image typically degraded by noise and blur. . . . .	77
7.3	TEM long and short exposure images . . . . .	78
7.4	Blind deconvolution of blurred and noisy high magnification TEM images. . . . .	79
7.5	CNN architecture for denoising of short exposure TEM images. . . . .	81
7.6	Series of short exposure TEM image (2048 $\times$ 2048 pixels). . . . .	82
7.7	Qualitative results of proposed method for three different denoising strategies for short exposure TEM images. . . . .	83
7.8	Segmentation of micro structures of a trabecular bone at 4 times increased resolution. . . . .	84
7.9	A satellite image illustrating pure and mixed pixels. . . . .	86
7.10	Coverage segmentation of a hyperspectral AVIRIS image. . . . .	88

# List of Tables

4.1	Potential functions - analytical expression. . . . .	37
6.1	Comparison of VSTPG and MAPPG. . . . .	62
7.1	Quantitative results of the three different denoising strategies for short exposure TEM images. . . . .	83
7.2	Performance on micro-CT image segmentation. . . . .	85
7.3	Performance of the super-resolution coverage segmentation method on the AVIRIS hyperspectral data set . . . . .	87



# 1 Introduction

Digital images are electronically acquired representations of a scene or scanned documents, such as photographs, manuscripts, printed texts, and artwork. In an ideal world the imaging device used to acquire an image of a scene should provide a *perfect image*. A perfect digital image from the point of view of image analysis and computer vision does not contain noise, blur, or any other artifacts (undesirable features). In reality, images are generally degraded in various ways in the acquisition process.

## 1.1 Image formation process

To give a better insight in the correspondence between steps of the acquisition process and different artifacts, a brief explanation how a digital image is typically formed is presented in the following lines.

Most digital imaging devices, such as cameras, and imaging modalities, such as different types of microscopy, computed tomography (CT), mammography and others, capture the images by CCD (Charge Coupled Device) [1] and CMOS (Complementary Metal-Oxide Semiconductor) [2, 3] sensors. These sensors convert the photons registered by millions of light-sensitive photodiodes (pixels) into an electrical voltage (analog signal). The analog signal is digitized using an analog-to-digital converter. The intensity of each pixel in a digital image is proportional to the number of photons registered by the pixel's corresponding light-sensitive photodiode. Due to random fluctuations of photons as well as thermal and electronic fluctuations of the acquisition devices, the obtained digital images are degraded by noise. Stochastic photon counting process typically follows a Poisson distribution, as observed in [4]. Therefore, the digital images acquired by photon (or electron) counting devices are degraded by



(a) The original image



(b) Image corrupted by Poisson-Gaussian noise.



(c) Noise-free image corrupted by Gaussian blur.



(d) Image corrupted by Gaussian blur and Poisson-Gaussian noise.



(e) Image corrupted by Gaussian blur and down-sampled in the presence of Poisson-Gaussian noise.

Figure 1.1: Examples of corrupted images for different linear forward operators  $L$  and different types of noise.

signal-dependent Poisson noise. On the other hand, thermal and electronic fluctuations of the acquisition devices follow a Gaussian distribution. If the noise present in an image is modeled only by Poisson random variable, the Gaussian signal-independent component is neglected. A more elaborate approach is to take into account the noise coming from signal-independent noise sources, and to model the noise in digital image with a mixed Poisson-Gaussian distribution. Mixed Poisson-Gaussian noise (PG) is often encountered in digital images used in astronomy [5], biology [7], and medicine [6]. In addition to noise, acquired images are typically degraded by blur, as a consequence e.g., of imperfect optics of imaging device and/or object/camera movements during the acquisition process. Moreover, images acquired by CCD and CMOS sensors with relatively small number of photo-sensitive cells are of limited (low) resolution (LR) and the visibility of small objects is additionally decreased.

A typical mathematical model of the imaging process assumes that the acquired image  $v$  is a corrupted version of the perfect original image  $u^*$ , where  $u^*$  is transformed by a linear operator  $L$  and additionally corrupted by random noise. This implies that the observed image  $v$  and the original image  $u^*$  satisfy the model

$$v = \psi(Lu^*) + \eta. \quad (1.1)$$

Here  $\psi$  and  $\eta$  denote signal-dependent and signal-independent (additive) random noise components, respectively (e.g., Poisson and Gaussian noise). The operator  $L$  is referred to as the forward operator of the problem. Some typical examples of the forward operator are the identity operator (when the only undesirable feature is noise), blurring operators (in that case  $Lu^*$  denotes the convolution with a blurring kernel and the undesirable feature is blur) or decimating operator (when the undesirable feature is low resolution). Some examples of corrupted images for different forward operators  $L$  are shown in Fig. 1.1.

## 1.2 Image restoration

Image restoration aims to reverse the effects of imperfect imaging and to recover noise-free and blur-free image  $u^*$  from the degraded image  $v$ , utilizing the assumption about the aforementioned imaging process.

Image restoration is a fundamental task in image processing, since both *manual* and *automated* analysis of digital images are negatively affected by the presence of the undesirable, but inevitable features.

Image restoration has an important role in *manual* interpretation of information contained in digital images, in fields such as medicine, biology, forensics, astronomy, and many others. In medicine, for instance, in imaging based on X-rays, such as CT or mammography (Fig. 1.2 (a)), the presence of noise is unavoidable because of the low doses of radiation that are preferably used in order to protect the patients [8, 10]. Images can be additionally corrupted by blur due to the patient's motion during the acquisition process. To facilitate the radiological diagnostic procedures, image restoration methods are often applied to enhance the important details in the digital imaging content. In biomedicine, in imaging based on Transmission Electron Microscopy (TEM), in order to avoid sample drift and decrease the amount of electrons interacting with, and modifying, the sample, the electron dose and acquisition time have to be decreased. This results in images with large amount of noise and increases the need for restoration [11, 12]. Restoration enhances structural details in biological samples and tissue sections and leads to more reliable clinical diagnostics. In forensics, to correctly identify someones face at the crime scene in dark and noisy imaging

conditions, it may be essential to make use of image restoration to enhance the degraded image. Reading the letters on a motion-blurred number-plate is another classic example where image restoration can help, Fig. 1.2 (b).

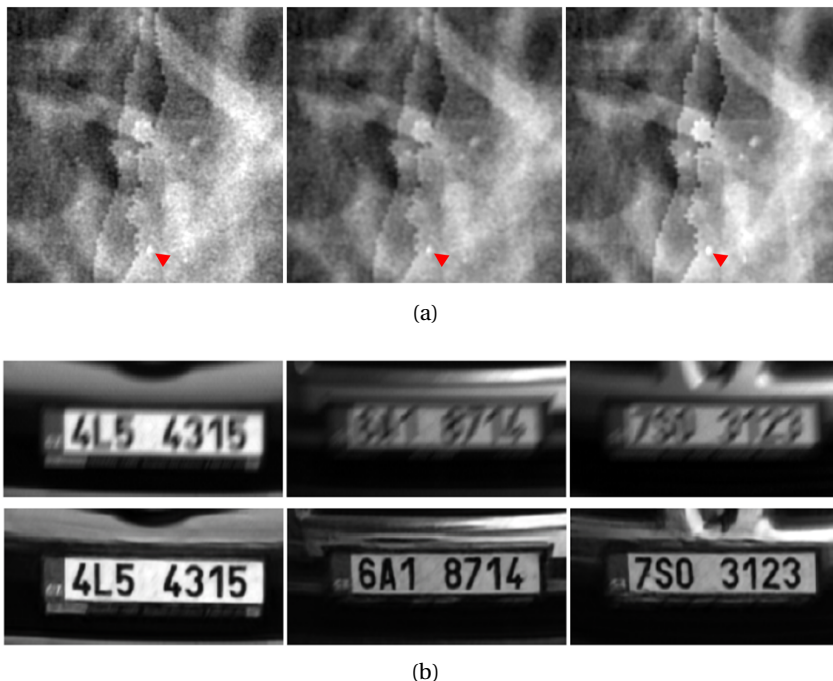


Figure 1.2: (a) Examples of noisy medical images - noisy mammography image (left), restored image (middle) and noise-free image (right). Arrows indicate microcalcification, small deposits of calcium, which can be an early sign of breast cancer. Their visibility is decreased in noisy images (source [8], adapted). (b) Examples of blurry forensic images - motion-blurred car plates (up) and restored plates (down) (source [9], adapted).

Image restoration is also a crucial part of an *automated* image analysis and computer vision work-flows. Such a work-flow typically starts with removing or reducing noise, blur and other distortions from an image (preprocessing) and it is followed by image segmentation, object description and recognition and/or classification (Fig. 1.3). Image segmentation aims to identify different segments of the image which correspond to the imaged objects of interest. The segments usually contain pixels with similar intensities and form connected components. For precise segmentation it is essential that the restoration process efficiently suppresses blur and noise while preserving edges, since they define the location of different segments and objects in an image. In addition, many classic edge detection or segmentation methods rely on derivatives. Numerical



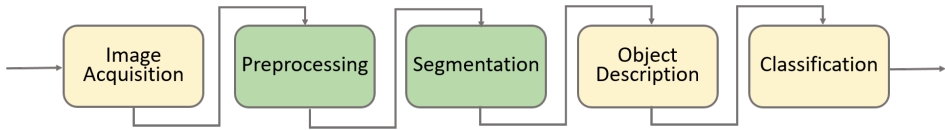


Figure 1.3: Schematic overview of the main building blocks of a typical automated image analysis work-flow. Thesis contributions belong to the green building blocks.

differentiation is not robust in the presence of noise (small perturbations in the function to be differentiated may lead to large variations in the computed derivative) and, consequently, the derivatives of noisy images do not contain reliable information and can be of limited use for edge detection or segmentation purposes. Hence, the performances of image segmentation and classification methods depend to a large extent on the quality of the input images and in the case when noise and blur are not efficiently suppressed in the acquired images, the success of these techniques is reduced. Fig. 1.4 (e, f) illustrates the poor performances of edge detection and segmentation methods performed on noisy image.

Depending on the nature of the operator  $L$  from model (1.1), image restoration is referred to as:

- (i) Image denoising:  $L$  is the identity operator, i.e., only noise is present, see Fig. 1.1(b).
- (ii) Image deconvolution:  $L$  represents a convolution (blurring) operator and noise may or may not be present, see Fig. 1.1(c) and Fig. 1.1(d). If operator  $L$  is unknown, then the deconvolution is called blind; otherwise the deconvolution is called non-blind.
- (iii) Single image super-resolution:  $L$  is a decimating operator and noise and blur may, or may not, be present, see Fig. 1.1(e).

The most general among the considered tasks is the single image super-resolution reconstruction of images degraded by blur and noise which, for the case when decimating operator is identity operator, narrows down to deconvolution (in the presence of noise). The deconvolution task further narrows down to denoising, when blurring operator equals to identity operator.

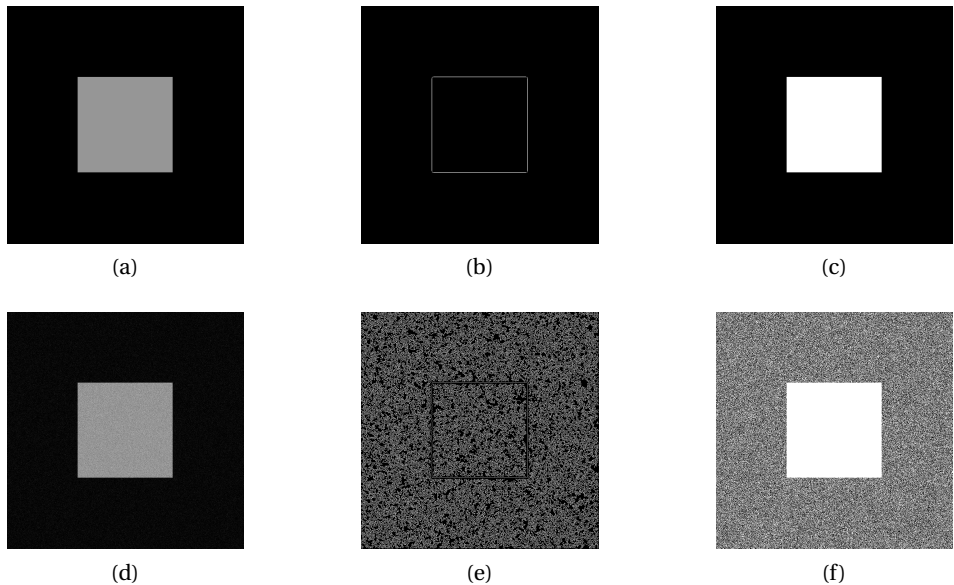


Figure 1.4: Poor edge detection and segmentation performances on noisy image. Image (a) is the original noise-free image and (d) is its noisy observation. Images (b, c) and (e, f) are results of Canny edge detection and Otsu thresholding segmentation applied on (a) and (d), respectively.

### 1.2.1 Image restoration by energy minimization

Image restoration can be performed by regularized energy minimization, due to its simplicity and generally good performance. In the most generic setting, one tries to recover a blur-free and noise-free version of  $v$  by minimizing an energy function which has the following form:

$$E(u) = D(u; v) + \lambda R(u). \tag{1.2}$$

An argument which minimizes the energy function,

$$\hat{u} = \arg \min_u E(u) \tag{1.3}$$

is considered to be an estimate of the original image  $u^*$ .

The function  $D$  is called the *data fidelity term*. It measures the distance between the sensed image  $v$  and its reconstruction  $u$  after the forward operator has acted on it. It utilizes the model of image formation process given by Eq. (1.1). Small values of this term indicate that the transformed image  $Lu$  is close to the data

$v$  in a suitable sense. The function  $R$  is called the *regularization term* and it imposes a priori knowledge on the solution  $u$ . It is expected that small values of  $R$  lead, up to a certain extent, to the elimination of the undesirable features (e.g., noise) in the corrupted data  $v$ . Regularization also provides numerical stabilization of the image restoration problem. The regularization parameter  $\lambda$  controls the trade-off between the two terms i.e., the level of regularization vs. the faithful recovery of the (possibly noisy) image detail.

Different choices can be made concerning both data fidelity and regularization terms. An unsuitable choice of any of the terms leads to poor restoration results. The choice of *data fidelity term* depends on the noise distribution while the choice of *regularization term* depends on our decision which features to eliminate/preserve in the images.

Most of existing non-blind [13–15] and blind [16–19] deconvolution and super-resolution [20–22] methods assume that noise present in the image is modeled by (only) Gaussian distribution due to the simplicity of this assumption. The corresponding data fidelity term is quadratic ( $\ell_2$ ) which is easy to numerically optimize. In the presence of mixed Poisson-Gaussian noise, the minimization of energy functions with the quadratic (Gaussian) data fidelity term, leads to significantly reduced performance. Up till now signal-dependent PG noise has not been widely considered because of theoretical and practical difficulties which arise from the fact that data fidelity term includes an infinite sum which can not be accurately computed. None of the existing blind deconvolution and super-resolution methods have treated PG noise appropriately, while only few recently proposed non-blind deconvolution methods [5,23,41–43] are developed for PG noise. These methods approximate the infinite sum with a finite number of summands. Such an approach leads to slow and complicated algorithms. This motivates our research aiming at finding alternative ways to adequately tackle PG noise in presence of blur and low spatial resolution.

The most commonly utilized regularization term is non-differentiable sparsity promoting Total Variation (TV) [44]. In [45] it is indicated that smoothing the TV with a so-called potential function can lead to improved performances of methods suited for Gaussian noise. However, it was not explored whether it would be possible to further improve TV performances in a similar way in the presence of blur in the image and for different types of noise, e.g., Poisson or mixed PG noise.

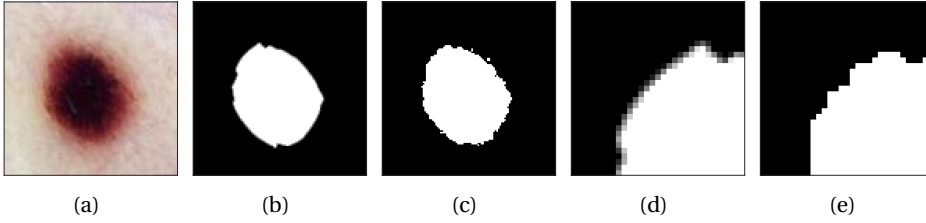


Figure 1.5: Segmentation of dermoscopic image: (a) original image, (b) coverage and (c) crisp segmentation of (a); (d) and (e) zoomed-in region from (b) and (c).

### 1.3 Coverage segmentation

Numerous image processing techniques of different nature, for example segmentation [24,25], CT [26,27] or MRI [28,29] image reconstruction, inpainting [30–32] and many others, can be performed by minimizing an appropriate energy function with suitably designed data and regularization terms. This thesis focuses on two sparsity promoting energy minimization based tasks:

1. TV-based image restoration suited for PG noise and
2. TV-based super-resolution coverage segmentation in presence of blur and noise (Fig. 1.3).

Classical crisp segmentation assigns a label to each pixel in an image, indicating the single component to which the pixel belongs. In a crisp representation each pixel is fully and exclusively covered either by one of the object classes or by background. Coverage representation allows for each pixel to be partially covered by both object and background (Fig. 1.5). This type of segmentation assigns a vector with elements in the range  $[0, 1]$  to each pixel, indicating the degree of (partial) coverage of that pixel by each of the image components (including background). Several studies show advantages of coverage segmentation and confirm that the coverage information enables increased precision and accuracy of object descriptor estimates, as well as their decreased noise sensitivity, [33–37].

#### 1.3.1 Coverage segmentation by energy minimization

In [38], it is proposed to compute the coverage segmentation  $\hat{A}$  of an image  $\nu$  as

$$\hat{A} = \arg \min E(A), \quad (1.4)$$

where

$$E(A) = D(A; \nu) + \mu R(A) \quad (1.5)$$

## 1.4. Problem Statement and Research Objectives

---

is a regularized energy function. Here an assumption about a sensed image is that it is formed such that pixel intensities are obtained as per pixel linear mixtures of pure class representatives (end-members), i.e., that the sensed image  $v$  can be modeled as

$$v = A \cdot C + \text{noise}, \quad (1.6)$$

where  $C$  is the end-member matrix. This model assumes that the mixed (partly covered) pixels form a one-pixel-thin boundary between homogeneous regions of pure pixels. However, this assumption is often violated in the presence of blur, which has a negative impact on the resulting segmentation. This motivates to develop a coverage segmentation method which incorporates the linear forward imaging model with blur and decimation (Eq. (1.1)) into the existing linear mixture model (Eq. (1.6)) in order to improve poor performance of the coverage segmentation method from [38] in presence of blur in the acquired image, as well as to allow an estimation of coverage values (values from the interval  $[0, 1]$ ) at an increased spatial resolution, higher than that of the acquired image.

### 1.4 Problem Statement and Research Objectives

The main focus of this thesis is to address some of the challenges associated with image restoration and coverage segmentation using sparsity promoting energy minimization based techniques. To do so, the overall objectives of the thesis are structured as follows:

1. To develop image restoration methods suited for Poisson-Gaussian noise, which is less explored regardless of its high significance for practical application purposes.
2. To design a data fidelity term which properly treats Poisson-Gaussian noise.
3. To explore different choices of regularization terms that promote sparsity in the image gradient domain in the presence of Poisson-Gaussian noise, to reach further improvement of restoration performance.
4. To design a data fidelity term for coverage segmentation such that it enables removal of image blur, providing thin transition boundaries between initially blurry and noisy image components. In addition, this data term should provide an option to estimate coverage values at an increased spatial resolution, higher than that of the sensed image.
5. To explore the existing choices of sparsity promoting regularization terms and to propose new regularization terms which further improve the coverage segmentation method performances.

6. To efficiently address the numerical optimization problem for finding an argument which minimizes energy functions of different levels of complexity appearing in all developed methods.

### 1.5 Contribution of This Thesis

The main contributions of this PhD thesis are summarized as follows:

1. The thesis thoroughly explores approaches to design data fidelity term in order to treat PG noise appropriately in several restoration tasks. Novel efficient restoration methods for Poisson-Gaussian noise are proposed, Publication I-III. These suggested data fidelity terms include assumption of such noise. These methods suitably treat PG noise utilizing Variance Stabilization Transformation (VST). We also elucidate the importance of correct treatment of Poisson-Gaussian noise. The applicability of the proposed methods is demonstrated by enhancing TEM images of cilia (Publication II and Publication V).
2. Aiming to explore how restoration performance can be further improved by adequate choice of regularization term which promotes sparsity in the image gradient domain, the thesis evaluates and compares different TV based regularization terms proposed in the literature. We demonstrate that a suitable choice of edge preserving potential function in the regularization term leads to further improvement of restoration methods (Publication I and Publication IV). Huberized Total Variation (smooth version of TV) exhibits outstanding performances in edge preservation.
3. The coverage segmentation method presented in [38] is extended in Publication VI. The data term is adjusted to enable removal of image blur. Additionally, it provides an option to estimate coverage values at an increased spatial resolution, higher than that of the sensed image. This is done by combining a forward imaging model with the model of linearly mixed pure class representatives.
4. Sparsity promoting regularization terms, previously utilized in the coverage segmentation method, are improved and evaluated in different combinations. The proposed (best performing) combination includes the Huberized Total Variation (instead of the previously utilized Perimeter term in [38]), leading to less challenging optimization, and new fuzziness term, restricting the fuzziness penalty to non-edge regions and by that achieving better stability of a solution. Constraint to a correct coverage segmentation is now included in the regularization term.

5. The choice of a numerical optimization method has crucial impact on the success of minimization of all considered energy functions in both restoration and segmentation tasks. Spectral Conjugate Gradient optimization [39] is utilized in all the proposed restoration methods (Publications I-V) providing excellent solutions. For solving more complex and highly non-convex optimization problems related to blind deconvolution and coverage segmentation, more complex optimization schemes are proposed. For blind deconvolution method, an Alternating Minimization (AM) procedure is suggested, where a problem is split into convex sub-problems. For super-resolution coverage segmentation, a graduated non-convexity optimization scheme (GNC) based on gradually increasing complexity of the problem is proposed. GNC utilizes the solutions of numerically easier sub-problems as starting guesses in the optimization of the more difficult ones. Sub-problems within AM and GNC are optimized by Spectral Projected Gradient (SPG).

## 1.6 Thesis Outline

This PhD thesis is divided into 8 chapters. The first chapter explains the need for image restoration and precise segmentation and highlights the challenges associated with both tasks. It introduces the energy minimization approach which is used in this thesis for efficiently solving both problems. Most of the challenges have been addressed successfully and the results are summarized as per the Publications included in the Appendix.

Chapter 2 presents the statistical approach based on Bayes' theorem which is used to derive Maximum a Posteriori (MAP) estimators for inverse problems where images are stochastically modeled. This chapter also elucidates the link between MAP estimators and energy functions.

Chapter 3 presents various data fidelity terms for restoration and coverage segmentation tasks which exist in the literature, both the negative-log likelihood ones and their different approximations. This chapter introduces the novel data terms developed in this thesis. In particular, the VST based data fidelity term is proposed for image restoration problems in the presence of PG noise. Also, the unmixing model is combined with the blur and down-sampling model providing robust coverage segmentation in the presence of blur and noise.

Chapter 4 describes several regularization terms suitable for restoration and coverage segmentation tasks. The concept of sparsity promoting regularization is presented. Focus is on Total Variation based regularization which promotes sparsity in the gradient domain. Edge preserving potential functions are introduced.

## Chapter 1. Introduction

---

We have evaluated a hypothesis that edge preservation in image restoration can be improved if some alternative potential functions are used instead of the identity function (which reduces to classical TV). For image segmentation task, penalties of a type used in the Mumford-Shah approach [102] are elucidated. The link between TV and Mumford-Shah penalties via level sets is presented. This link motivates our choices to use TV based regularization instead of previously used Mumford-Shah inspired penalty for coverage segmentation. Additionally, a new sparsity promoting fuzziness term is introduced.

Chapter 5 highlights some important aspects of algorithms designed to efficiently compute solutions of the energy minimization problems. Different numerical optimization concepts recently used in literature for similar tasks are reviewed. In the thesis Spectral Projected Gradient method is utilized for minimization of energy functions with different data fidelity and regularization terms for both image restoration and coverage segmentation tasks. In this chapter SPG algorithm is presented and two schemes for solving non-convex energy minimization problems in blind restoration and coverage segmentation, introduced in Publications II and VI, are explained.

Chapter 6 discusses the data set and performance measures used for the evaluation of the methods proposed within this thesis. This chapter also reports on the quantitative and qualitative results obtained by the proposed methods.

Chapter 7 illustrates the applicability of restoration and coverage segmentation methods which are proposed in this thesis for TEM image enhancement, remote sensing and trabecular bone micro-structures segmentation.

Chapter 8 concludes the thesis and presents potential future research directions.



## 2 Inverse problems

Image restoration and image segmentation, which represent the focus of this thesis, belong to inverse problems [46].

In general, inverse problems aim to recover information about a physical object of interest from observed data; in super-resolution task this is a blur-free and noise-free image with increased spatial resolution, while in the coverage segmentation task the aim is to obtain the portions of pure class representatives (end-members) presence in each pixel. In real applications the observed data usually do not provide enough information for the unique and stable reconstruction of the originals and the problem is highly under-determined and therefore hard to solve. A typical example is super-resolution reconstruction where the number of pixels in the observed image (the number of known constraints) is much smaller than the number of pixels in the original image (degrees of freedom).

Hence, inverse problems belongs to the so-called ill-posed problems. An ill-posed problem is a mathematical problem which violates at least one of the properties of a well-posed problem defined by Hadamard [47]. A well-posed problem has the following three properties:

1. a solution exists;
2. the solution is unique;
3. the solution is stable/robust.

To be solved, an ill-posed problem needs to be re-formulated to impose well-posedness. Typically, this involves the incorporation of an additional assumption (constraint), often based on an appropriate a priori information about the

solution. This process is known as regularization and it aims to provide solutions with a lower sensitivity to perturbations of the input data. Regularization should provide numerical stabilization and impose desired properties to the solution.

Maximum a posteriori (MAP) estimators are classical estimators which incorporate prior information via Bayes' theorem.

## 2.1 Maximum a posteriori (MAP) estimator

A brief description of MAP estimators inference follows here; for a more detailed introduction to this topic, see [48]. Typically random variable  $\mathcal{V}$  depends on a set/vector of unknown parameter  $u$  and within this framework we also treat  $u$  as a random vector (a vector of random variables). Suppose that before the statistical experiment, our prior distribution describing the unknown parameter  $u$  is  $f_{\mathcal{U}}(u)$ . The conditional probability density function (PDF) of random variable  $\mathcal{V}|\mathcal{U} = u$  is denoted with  $f_{\mathcal{V}}(v|\mathcal{U} = u)$ . A sample of independent and identically distributed random variables (i.i.d.) for  $\mathcal{V}$  is vector  $\mathcal{V}_s = [\mathcal{V}_1, \mathcal{V}_2, \dots, \mathcal{V}_n]$  and  $v_s = [v_1, v_2, \dots, v_n]$  is its one observation/realization. The conditional probability density function of each random variable  $\mathcal{V}_1, \mathcal{V}_2, \dots, \mathcal{V}_n$  equals  $f_{\mathcal{V}}(v|\mathcal{U} = u)$ . Therefore, joint conditional distribution for the observation  $v_s$  due to their independence is the product of marginal distributions

$$f_{\mathcal{V}_s}(v_s|\mathcal{U} = u) = \prod_{i=1}^n f_{\mathcal{V}_i}(v_i|\mathcal{U} = u). \quad (2.1)$$

The function (2.1) is associated with the probability of observing the data  $v_s$  at each possible value  $u$  of an unknown parameter and it is called *likelihood* function  $\mathcal{L}(u)$ .

According to Bayes' theorem, a posterior conditional distribution of  $\mathcal{U}$  given the data  $v_s$  equals

$$f_{\mathcal{U}}(u|\mathcal{V}_s = v_s) = \frac{f_{(\mathcal{U}, \mathcal{V}_s)}(u, v_s)}{m_{\mathcal{V}_s}(v_s)}, \quad (2.2)$$

where  $f_{(\mathcal{U}, \mathcal{V}_s)}(u, v_s)$  is a joint distribution for  $(\mathcal{U}, \mathcal{V}_s)$ ,  $m_{\mathcal{V}_s}(v_s)$  is a normalizing constant (the value of marginal distribution for  $\mathcal{V}_s$  for the observation  $v_s$ )

$$m_{\mathcal{V}_s}(v_s) = \int_{u \in R_{\mathcal{U}}} f_{\mathcal{V}_s}(v_s|\mathcal{U} = u) f_{\mathcal{U}}(u) du \quad (2.3)$$

and  $R_{\mathcal{U}}$  is a set of possible values for  $u$  (support for random variable  $\mathcal{U}$ ). Joint distribution for  $(\mathcal{U}, \mathcal{V}_s)$  further equals

$$f_{(\mathcal{U}, \mathcal{V}_s)}(u, v_s) = f_{\mathcal{V}_s}(v_s|\mathcal{U} = u) f_{\mathcal{U}}(u). \quad (2.4)$$

## 2.2. MAP estimators for inverse problems and their link to energy functions

MAP estimator of the unknown parameter  $u$  is

$$\hat{u}_{MAP} = \arg \max_{u \in R_{\mathcal{U}}} f_{\mathcal{U}}(u | \mathcal{V}_s = v_s) = \arg \max_{u \in R_{\mathcal{U}}} \frac{f_{\mathcal{V}_s}(v_s | \mathcal{U} = u) f_{\mathcal{U}}(u)}{m_{\mathcal{V}_s}(v_s)}. \quad (2.5)$$

This is how Bayes' theorem modifies the prior  $f_{\mathcal{U}}(u)$  (which does not depend on the observed data at all) to the posterior by accounting for the data  $v_s$  coming from the assumed model (likelihood).

Due to the monotonicity of the logarithm function this can be rewritten as:

$$\hat{u}_{MAP} = \arg \min_{u \in R_{\mathcal{U}}} (-\log(f_{\mathcal{V}_s}(v_s | \mathcal{U} = u)) - \log(f_{\mathcal{U}}(u)) + \log(m_{\mathcal{V}_s}(v_s))) \quad (2.6)$$

which further reduces to

$$\hat{u}_{MAP} = \arg \min_{u \in R_{\mathcal{U}}} (-\log(f_{\mathcal{V}_s}(v_s | \mathcal{U} = u)) - \log(f_{\mathcal{U}}(u))) \quad (2.7)$$

as  $m_{\mathcal{V}_s}(v_s)$  is constant (does not depend on  $u$ ).

In this way, based on Bayes' theorem, MAP estimator incorporates our prior knowledge which in inverse problems imposes well-posedness.

## 2.2 MAP estimators for inverse problems and their link to energy functions

In order to derive MAP estimator of the original image, we need to stochastically model the image formation process given by Eq. (1.1), i.e., we assume that

$$\mathcal{V} = \Psi(\mathcal{U}\mathcal{U}^*) + \mathcal{N}, \quad (2.8)$$

where  $\mathcal{V}$ ,  $\mathcal{U}^*$ ,  $\Psi$  and  $\mathcal{N}$  stand for the random variables which describe the observed image, original image, signal-dependent and signal-independent noise, respectively.

We aim to find random variable  $\mathcal{U}$  which describes the recovered image such that it is as close as possible to the  $\mathcal{U}^*$  (which describes the original image), i.e.,  $\mathcal{U} \approx \mathcal{U}^*$  and which also follows the assumed image formation model

$$\mathcal{V} \approx \Psi(\mathcal{U}\mathcal{U}) + \mathcal{N}. \quad (2.9)$$

Considering that both the observed and the sought for images represented as vectors  $v = [v_1, \dots, v_n]$  and  $u = [u_1, \dots, u_n]$  are the samples of random vectors  $\mathcal{V} =$

## Chapter 2. Inverse problems

---

$[\mathcal{V}_1, \mathcal{V}_2, \dots, \mathcal{V}_n]$  and  $\mathcal{U} = [\mathcal{U}_1, \mathcal{U}_2, \dots, \mathcal{U}_n]$  where i.i.d.  $\mathcal{V}_i$  have the same distribution as  $\mathcal{V}$  for all  $i$ , respectively. An estimator of unknown image  $u$  which maximizes a posterior probability of the unknown image  $u$  given the observed data  $\mathcal{V} = v$  is:

$$\hat{u}_{MAP} = \underset{u}{\operatorname{argmax}} f_{\mathcal{U}}(u|\mathcal{V} = v), \quad (2.10)$$

or equivalently after simplifications it equals to (2.7), where  $f_{\mathcal{V}}(v|\mathcal{U} = u) = \mathcal{L}(u)$  is the *likelihood* that the outcome of the image formation process is  $v$  given  $\mathcal{U} = u$  and  $f_{\mathcal{U}}(u)$  is a prior.

Typically the latest expression is written in simpler form as

$$\hat{u}_{MAP} = \underset{u}{\operatorname{argmin}} (D(u; v) + R(u)), \quad (2.11)$$

where

- The negative log-likelihood function

$$D(u; v) = -\log(f_{\mathcal{V}}(v|\mathcal{U} = u)) = -\log(\mathcal{L}(u)) \quad (2.12)$$

is called *data fidelity term*,

- The function

$$R(u) = -\log(f_{\mathcal{U}}(u)) \quad (2.13)$$

is known as *regularization term* and it incorporates the a priori knowledge about the solution.

The balanced combination of the two functions

$$E(u) = D(u; v) + \lambda R(u) \quad (2.14)$$

with trade-off parameter  $\lambda$  is known as *energy function*.

The data fidelity term consists of the negative log-likelihood and its analytical expression corresponds to the assumed noise model, while the regularization function should ensure that the solution has the desired properties (the properties possessed by the original, non-corrupted image).

The following chapters, Chapter 3 and Chapter 4, present a number of data fidelity and regularization terms suitable for image restoration and coverage segmentation. These include both existing, and originally proposed ones.

## 3 Data fidelity terms

As discussed in Chapter 2, the data fidelity term equals the negative log-likelihood function of observing the data  $v$  given (at each possible value) unknown image  $u$  (of unknown image). Its role is to measure the distance (similarity) between the sensed image  $v$  and its reconstruction  $u$ , i.e., to keep the reconstructed image close enough to the observation so that useful information will not be discarded in the solution. The statistics of the noise give rise to a particular choice of the fidelity term. In this Chapter we describe the most common image noise models including the Gaussian, Poisson, and the PG models; nevertheless we will start by describing the basics on how to find the best estimator of an observed data corrupted with a known noise model using the criterion based on the MAP estimator. Before we dwell into the problem, we introduce notation which we use further in the thesis.

### 3.1 Preliminaries and notation

Now onwards, the most general restoration problem among problems (i)-(iii) listed at the end of the Section 1.2 is considered. This is the case when linear operator  $L$  from Eq. (1.1) models blurring and decimating process in the presence of noise.

Some notation used in this thesis follows. Gray-scale (single band) images are generally represented as column vectors, where image rows of pixels are sequentially concatenated; an image of size  $r \times c$  is represented as a vector of size  $N \times 1$ , where  $N = rc$ .

- Unknown HR image  $u$  of size  $r_h \times c_h$  is represented as a vector  $u = [u_1, \dots, u_{N_h}]^T$  of length  $N_h = r_h \times c_h$ .

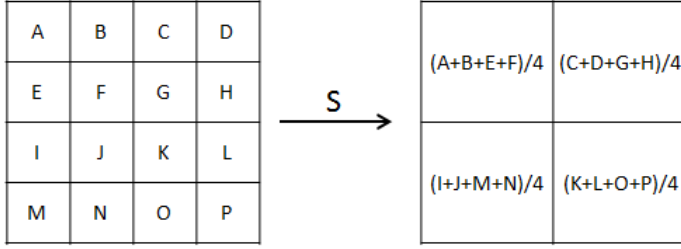


Figure 3.1: Effect of down-sampling matrix  $S$  on  $4 \times 4$  image with sampling factor  $d = 2 \times 2$ .

- Original HR image is denoted  $u^*$ . It is of size  $r_h \times c_h$  and it is represented as a vector  $u^* = [u_1^*, \dots, u_{N_h}^*]^T$  of length  $N_h = r_h \times c_h$ .
- Observed LR image is denoted  $v$ . It is of size  $r_l \times c_l$  and it is represented as a vector  $v = [v_1, \dots, v_{N_l}]^T$  of length  $N_l = r_l \times c_l$ , where  $N_l < N_h$ .

Here we aim for an increase in resolution by a factor  $d = d_r \times d_c$ , so since the observed LR image  $v$  has a size  $r_l \times c_l$ , the reconstruction  $\hat{u}$  of the HR image  $u^*$  has a size  $r_h \times c_h = d_r r_l \times d_c c_l$ , and it holds that the length of the HR image is  $N_h = d N_l$ .

The convolution with a point spread function (PSF), which models the degradation of the image with blur, is equal to the multiplication of a matrix  $H_{N_h \times N_h}$  and image  $u^*_{N_h \times 1}$ . The multiplication of a blurred HR image  $Hu^*$  with matrix  $S_{N_l \times N_h}$  corresponds to down-sampling with a factor  $d = d_r \times d_c$  where  $N_h = d N_l$ ,  $r_h = d_r r_l$  and  $c_h = d_c c_l$ . Fig. 3.1 illustrates the result of applying the down-sampling matrix  $S_{N_l \times N_h}$  on a  $4 \times 4$  image. Down-sampled blurred HR image,  $SHu^*$ , is additionally corrupted with noise. In this way, observed LR image  $v$  is formed.

## 3.2 MAP based data fidelity terms for super-resolution image reconstruction

### 3.2.1 Data fidelity term suited for Gaussian noise

As already mentioned, the two predominant sources of noise in digital image acquisition, are:

1. the stochastic nature of the photon-counting process at the detectors;

### 3.2. MAP based data fidelity terms for super-resolution image reconstruction

2. the intrinsic thermal and electronic fluctuations of the acquisition devices.

When photon-count is very high, then under standard illumination conditions, the second source of noise, which is signal-independent, dominates over the first one. This motivates the usual additive-white-Gaussian-noise assumption, while the first source of noise is neglected.

When Gaussian noise model is assumed, the degraded image  $v$  is given by:

$$v_i = (SHu^*)_i + \eta_i, \quad \forall i = 1, 2, \dots, N_l \quad (3.1)$$

where  $\eta = [\eta_1, \dots, \eta_{N_l}]$  is the sample (observation/realization) of random vector of  $N_l$  i.i.d. random variables, each following the same Gaussian distribution  $\mathcal{N}(0, \sigma_m^2)$  with density function

$$f(\eta_i) = \frac{1}{\sqrt{2\pi\sigma_m^2}} e^{-\frac{\eta_i^2}{2\sigma_m^2}}. \quad (3.2)$$

The sought for image  $u$  should be as close as possible to the original perfect noise-free and blur-free image  $u^*$  and therefore should behave approximately  $v_i \approx (SHu)_i + \eta_i, \quad \forall i = 1, 2, \dots, N_l$ . Since  $v = [v_1, \dots, v_{N_l}]$  is the sample (observation/realization) of random vector of  $N_l$  i.i.d. random variables, each following the Gaussian distribution  $\mathcal{N}((SHu)_i, \sigma_m^2)$ , we have that

$$f_{\mathcal{V}}(v_i | \mathcal{U} = u) = \frac{1}{\sqrt{2\pi\sigma_m^2}} e^{-\frac{((SHu)_i - v_i)^2}{2\sigma_m^2}}. \quad (3.3)$$

Under such assumption, the data fidelity term (2.12) equals to:

$$\begin{aligned} D_{\text{MAPG}}(u; v) &= -\log(f_{\mathcal{V}}(v | \mathcal{U} = u)) \\ &= -\log\left(\prod_{i=1}^{N_l} f_{\mathcal{V}}(v_i | \mathcal{U} = u)\right) \\ &= -\sum_{i=1}^{N_l} \log(f_{\mathcal{V}}(v_i | \mathcal{U} = u)) \\ &= -\sum_{i=1}^{N_l} \log\left(\frac{1}{\sqrt{2\pi\sigma_m^2}} e^{-\frac{((SHu)_i - v_i)^2}{2\sigma_m^2}}\right) \\ &= \sum_{i=1}^{N_l} \frac{((SHu)_i - v_i)^2}{2\sigma_m^2} + N_l \log\left(\sqrt{2\pi\sigma_m^2}\right) \\ &= \frac{1}{2\sigma_m^2} \sum_{i=1}^{N_l} ((SHu)_i - v_i)^2 + N_l \log\left(\sqrt{2\pi\sigma_m^2}\right). \end{aligned} \quad (3.4)$$

Since the aim is to find the argument which minimizes  $D_{\text{MAPG}}(u; v)$ , we can further simplify the expression by removing the constants, and define:

$$D_{\text{MAPG}}(u; v) = \frac{1}{2} \sum_{i=1}^{N_l} ((SHu)_i - v_i)^2. \quad (3.5)$$

The derived data fidelity term is quadratic  $\ell_2$  term which can be easily optimized. This makes an assumption of Gaussian noise presence in the sensed image appealing and widely used in research and applications.

#### 3.2.2 Data fidelity term suited for Poisson noise

In many applications such as fluorescence microscopy or astronomy, only relatively few photons are collected by the photo sensors, due to various physical constraints (low-power light source, short exposure time). Under these imaging conditions, the major source of noise is strongly signal-dependent. Consequently, it is more reasonable to model the output of the detectors as a Poisson-distributed random vector. Assuming Poisson noise distribution, an implicit assumption is that the stochastic nature of the photon-counting process at the detectors dominates over the intrinsic thermal and electronic fluctuations of the acquisition devices which is modeled with Gaussian noise. In this case the observed image  $v$  is given by:

$$v_i = \theta_i, \quad \forall i = 1, 2, \dots, N_l \quad (3.6)$$

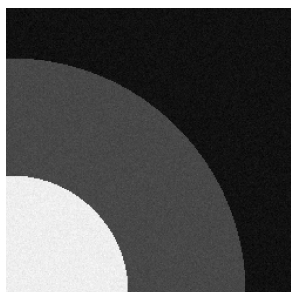
where  $\theta_i \sim \mathcal{P}((SHU^*)_i)$ . This image formation model assumes degradation by blur and subsampling. Images corrupted by different types of noise - Gaussian, Poisson and mixed PG, are presented in Fig. 3.2. An example of Poisson corrupted image is shown in Fig. 3.2 (b). Note that the noise variance is not constant and it equals to the expectation of the pixel value (the higher the pixel intensity is, the more noisy pixel is).

Such an assumption means that

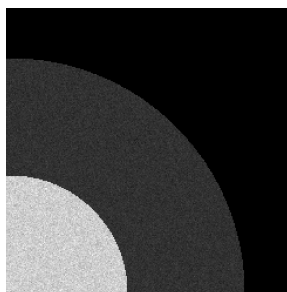
$$f_{\mathcal{V}}(v_i | \mathcal{U} = u) = \frac{((SHu)_i)^{v_i}}{v_i!} e^{-(SHu)_i}, \quad (3.7)$$



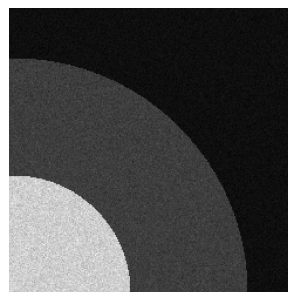
### 3.2. MAP based data fidelity terms for super-resolution image reconstruction



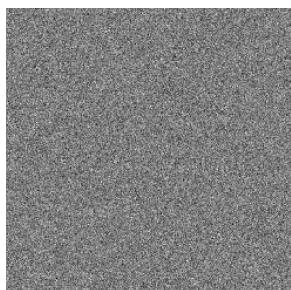
(a) Image corrupted by Gaussian noise.



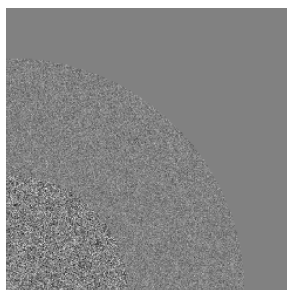
(b) Image corrupted by Poisson noise.



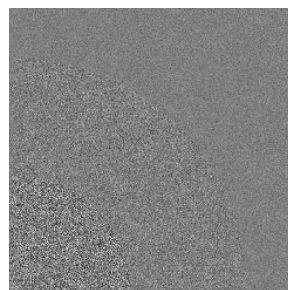
(c) Image corrupted by Poisson-Gaussian noise.



(d) Gaussian noise component.



(e) Poisson noise component.



(f) Poisson-Gaussian noise component.

Figure 3.2: Images corrupted by Gaussian, Poisson and Poisson-Gaussian noise and the corresponding residuals (difference between noisy and noise-free image which equal to noise components).

and the data fidelity term (2.12) is

$$\begin{aligned}
 D_{\text{MAP}}(u; v) &= -\log(f_{\mathcal{V}}(v|\mathcal{U} = u)) \\
 &= -\log\left(\prod_{i=1}^{N_l} f_{\mathcal{V}}(v_i|\mathcal{U} = u)\right) \\
 &= -\sum_{i=1}^{N_l} \log(f_{\mathcal{V}}(v_i|\mathcal{U} = u)) \\
 &= -\sum_{i=1}^{N_l} \log\left(\frac{((SHu)_i)^{v_i}}{v_i!} e^{-(SHu)_i}\right) \\
 &= \sum_{i=1}^{N_l} ((SHu)_i - v_i \log((SHu)_i)) + \sum_{i=1}^{N_l} \log(v_i!).
 \end{aligned} \tag{3.8}$$

After dropping the latest sum (since it does not depend on  $u$ ), the data term simplifies to the well-known Kullback-Leibler divergence:

$$D_{\text{MAP}}(u; v) = \sum_{i=1}^{N_l} ((SHu)_i - v_i \log((SHu)_i)). \tag{3.9}$$

### 3.2.3 Data fidelity term suited for Poisson-Gaussian noise

Assuming Poisson noise distribution, the signal-independent component of the noise is implicitly ignored. A more general approach is to take Gaussian noise into account, and to model noise as a mixture of both Poisson and Gaussian (PG) distributions. Assuming that the HR image  $u^*$  is degraded by blur, down-sampled and corrupted with PG noise, the intensity values  $v_i$  of the acquired LR image  $v$  are of the form

$$v_i = \theta_i + \eta_i, \quad \forall i = 1, 2, \dots, N_l \tag{3.10}$$

where  $\theta_i \sim \mathcal{P}((SHu^*)_i)$  and  $\eta_i \sim \mathcal{N}(0, \sigma_m^2)$ .

Such an assumption implies

$$f_{\mathcal{V}}(v_i|\mathcal{U} = u) = \sum_{k=0}^{+\infty} \frac{e^{-(SHu)_i} ((SHu)_i)^k}{k!} \frac{e^{-\frac{(v_i-k)^2}{2\sigma_m^2}}}{\sqrt{2\pi\sigma_m^2}}, \tag{3.11}$$

and the data fidelity term is

$$\begin{aligned}
 D_{\text{MAPPG}}(u; v) &= - \sum_{i=1}^{N_i} \log(f_{\mathcal{V}}(v_i | \mathcal{U} = u)) \\
 &= - \sum_{i=1}^{N_i} \log \left( \sum_{k=0}^{+\infty} \frac{e^{-(SHu)_i} ((SHu)_i)^k e^{-\frac{(v_i-k)^2}{2\sigma_m^2}}}{k!} \frac{1}{\sqrt{2\pi\sigma_m^2}} \right). \tag{3.12}
 \end{aligned}$$

Up till now signal-dependent PG noise model has not been widely utilized because of the theoretical and practical difficulties which arise from the fact that data fidelity term  $D_{\text{MAPPG}}$  (3.12) includes infinite sum which can not be accurately calculated. Restoration methods [5,23,41,42] treat PG noise approximating the infinite sum with finite number of summands. This way of addressing the problem related to infinite sum leads to slow and complicated algorithms, as it is demonstrated in Chapter 6.

### 3.3 Alternative data fidelity terms for non-Gaussian noise

In super-resolution studies [49–54] new, alternative data fidelity terms were proposed for non-Gaussian noise.

In [49, 52],  $\ell_1$  data fidelity term is proposed as a measure of data discrepancy

$$D_{\ell_1}(u; v) = \frac{1}{2} \sum_{i=1}^{N_i} |(SHu)_i - v_i|. \tag{3.13}$$

A combination of  $\ell_1$  and  $\ell_2$  data fidelity terms, called Huber data fidelity term, is proposed in [51, 53, 54]:

$$D_{\text{HUB}}(u; v) = \sum_{i=1}^{N_i} \Phi_{\text{HUB}} |(SHu)_i - v_i|, \tag{3.14}$$

where Huber function is

$$\Phi_{\text{HUB}}(s) = \begin{cases} \frac{s^2}{2\omega}, & t \leq \omega, \\ s - \frac{\omega}{2}, & t > \omega. \end{cases} \tag{3.15}$$

Parameter  $\omega$  controls the point of transition between  $\ell_2$  and  $\ell_1$  norm.

A Lorentzian data fidelity term is introduced in [49–51]. It is given by:

$$D_{\text{LOR}}(u; v) = \sum_{i=1}^{N_i} \Phi_{\text{LOR}} ((SHu)_i - v_i), \tag{3.16}$$

where Lorentzian function with a parameter  $\omega$  is

$$\Phi_{\text{LOR}}(s) = \log(1 + s^2/2\omega^2). \quad (3.17)$$

According to the results presented in super-resolution studies [49, 51], Huber and Lorentzian data fidelity terms outperform pure  $\ell_2$  ( $D_{\text{MAPG}}$ ) and  $\ell_1$  data fidelity terms when an image is degraded by signal-dependent Poisson noise. However, neither Huber nor Lorentzian data fidelity terms are theoretically derived for PG noise. Therefore, when they are applied for image restoration in the presence of PG noise, they lead to reduced performance of restoration. It has been confirmed by experiments conducted in our super-resolution study (Publication III), that these two data fidelity terms lead to reduced performance of super-resolution in the presence of PG noise and these results are presented in Chapter 6.

### 3.3.1 Data fidelity term based on Variance Stabilization Transformation

Another way to overcome practical difficulties related to  $D_{\text{MAPPG}}$  is to use a variance stabilizing transformation [55–59](VST) based approach. In an image corrupted by signal-dependent noise, the noise variance is typically not constant and varies with the expectation of the pixel value. For instance, the variance of a Poisson variable equals its mean and this value in general differs in each image location. VST can be used to remove signal-dependency of noise variance by making it constant throughout the whole image. The transformed data typically have an approximately Gaussian noise distribution with a known constant variance.

More formally, Variance Stabilization Transformation  $\varphi$  is a mapping which replaces the random variable  $\mathcal{V}$  by  $\mathcal{Z} = \varphi(\mathcal{V})$  such that a variance of  $\mathcal{Z}$  does not depend anymore on underlying signal  $\mathcal{U}^*$ . For instance, in the case of Poisson noise, mapping  $\varphi$  is chosen with an aim that the variance of  $\mathcal{Z}$  is a constant i.e., that it equals to 1:

$$\mathcal{V}_i \sim \mathcal{P}((SH\mathcal{U}^*)_i) \iff \mathcal{Z}_i = \varphi(\mathcal{V}_i) \approx \mathcal{N}(\varphi((SH\mathcal{U}^*)_i), 1), \quad \forall i = 1, \dots, N_l. \quad (3.18)$$

Numerous VST for the stabilization of a Poisson variable are proposed:

1. the root transformation [55]  $\varphi(v) = 2\sqrt{v}$
2. the Bartlett transformation [56]  $\varphi(v) = 2\sqrt{v + \frac{1}{2}}$ ,

### 3.3. Alternative data fidelity terms for non-Gaussian noise

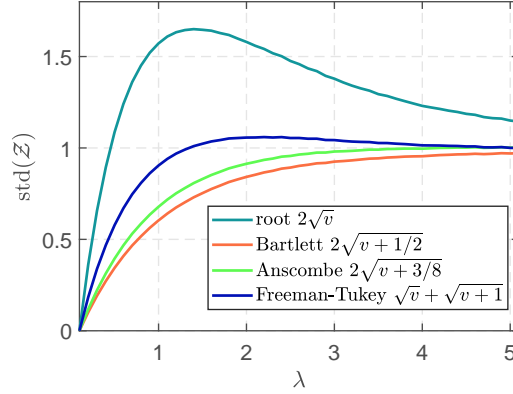


Figure 3.3: Variance stabilization transformations for Poisson random variable  $\mathcal{V} : \mathcal{P}(\lambda)$  with mean value (expectation)  $\lambda$ . The standard deviation of the stabilized Poisson variables  $\mathcal{Z} = \varphi(\mathcal{V})$  is presented for  $0 \leq \lambda \leq 5$ .

3. the Anscombe transformation [57]  $\varphi(v) = 2\sqrt{v + \frac{3}{8}}$ ,
4. the Freeman-Tukey transformation [58]  $\varphi(v) = \sqrt{v} + \sqrt{v+1}$ ,
5. the state-of-the-art nonparametric optimized transformations proposed in [59, 60].

Fig. 3.3 presents standard deviation of the stabilized Poisson variables with parameter (mean value)  $0 \leq \lambda \leq 5$ , for four different VSTs. As can be observed from the graph, the stabilization obtained through root transformation is not particularly good for low values of  $\lambda$  (standard deviation is far away from 1). Improvements have been sought by investigating a transformation of the form  $\varphi(v) = 2\sqrt{v+c}$ , where  $c$  is a constant. Bartlett [56] proposed the transformation  $\varphi(v) = 2\sqrt{v + \frac{1}{2}}$ , and Anscombe [57] improved this to  $\varphi(v) = 2\sqrt{v + \frac{3}{8}}$ . In fact, setting the constant to  $c = \frac{3}{8}$  provides optimal stabilization for large values of  $\lambda$  for this type of root transformation. Anscombe transformation is one of the most popular VSTs for Poisson distributed data, although the Freeman-Tukey transformation [58]  $\varphi(v) = \sqrt{v} + \sqrt{v+1}$  provides comparable asymptotic stabilization.

From Fig. 3.3, in addition, it can be observed that the Freeman-Tukey transformation stabilizes the variance slightly better for small mean values ( $\lambda < 2.2$ ) than the Anscombe transformation. On the other hand, the Anscombe transformation provides slightly more accurate stabilization around  $2.2 < \lambda < 4.8$ .

Another noticeable difference in the stabilized variances is that there is some overshoot for the Freeman-Tukey transformation before it approaches the target variance 1 from above, whereas such oscillation does not exist with the Anscombe transformation.

Some of the VSTs for Poisson distributed data are extended for PG distributed data. This thesis focuses on Anscombe VST transformation and, in particular, on its generalized version suited for PG distributed data [61, 62], which in absence of Gaussian distribution reduces to the original Anscombe transformation. We are not aware of a Poisson-Gaussian generalization of other mentioned VSTs (the root, Bartlett, Freeman-Tukey) therefore we focus on Anscombe VST and its generalization for PG distributed data.

The image restoration methods which utilize VST proposed in the literature focus on the simplest case where image is only degraded with noise, i.e., when no blur and/or down-sampling is involved into the capturing process.

When VST is applied for denoising, these denoising methods involve a three step procedure. In the first step the noisy data is transformed by a VST specifically designed for the chosen noise model. The second step is to treat the transformed data with any algorithm designed for the removal of Gaussian noise [63]. Finally, in the third step the desired estimate of the sought for noise-free image is obtained by applying an inverse VST ([64, 65]) to the denoised data.

However, in the presence of blur and decimation operator, the inverse VST transformation applied in the third step involves inverting blur operator and leads to amplification of noise and to further decrease of image quality and is thus not suitable. If variance stabilization transformation is used in deconvolution or super-resolution case, this fact should be included in the data fidelity term.

One of the main contributions of this thesis is generalization of the three step procedure to the cases when an image  $u^*$  is down-sampled and degraded by blur and PG noise. In such case, generalized Anscombe VST transforms the observed image  $v$  into  $z$ :

$$z_i = 2\sqrt{\max\{v_i + \frac{3}{8} + \sigma_m^2, 0\}}, \quad \forall i = 1, 2, \dots, N_l, \quad (3.19)$$

where

$$z_i \approx 2\sqrt{\max\{(SHu^*)_i + \frac{3}{8} + \sigma_m^2, 0\}} + \epsilon_i, \quad \epsilon_i \sim \mathcal{N}(0, 1), \quad \forall i = 1, 2, \dots, N_l. \quad (3.20)$$

We take a different approach in comparison to the three step procedure in this thesis and propose to use restoration methods which include transformation of

### 3.4. Data fidelity term for coverage segmentation

images degraded by blur and signal-dependent PG noise and down-sampled during the acquisition process by the generalized Anscombe VST, followed by energy minimization of the derived energy function with data fidelity term adjusted to reflect the VST process. These methods, in comparison to above described denoising procedure (in three steps) do not include the third step where an inverse VST is required. Instead, both the imaging and the transformation process are simultaneously inverted by minimization of the energy function with an appropriate data fidelity term.

In this approach we use the data fidelity term of the following form:

$$D_{\text{VSTPG}}(u) = \sum_{i=1}^{N_I} \left( z_i - 2\sqrt{\max\{(SHu)_i + \frac{3}{8} + \sigma_m^2, 0\}} \right)^2. \quad (3.21)$$

Such data fidelity term is quadratic, which ensures efficient minimization.

In Chapter 6 we discuss relevance of an appropriate assumption of the noise model present in the image and importance of a suitably designed restoration approach which takes that assumption into account. We present evaluation of restoration methods and analyse improvement in performance achieved by minimization of the energy function with data fidelity term (3.21), suited for PG noise. Details are included in Publications I-III.

## 3.4 Data fidelity term for coverage segmentation

### 3.4.1 Preliminaries and notation

A coverage representation of a crisp real object with a well-defined continuous border is, ideally, characterized by the presence of homogeneous connected regions of "pure" pixels, completely covered by any of image components. Two such regions are in general separated by a layer of "mixed" pixels, partially covered by more than one image components. Pure pixels are assigned coverage vectors with all zeros and 1 at the  $i$ -th position (indicating pixels completely included to to the  $i$ -th image component), while mixed pixels are assigned vectors with elements between 0 and 1, in accordance with their respective coverage by the image components. Under assumption that imaged components are continuous non-overlapping crisp sets and that the union of all components completely fills the image space, for the coverage model, it is reasonable to require that the assigned levels of belongingness (coverage values) of a pixel to different such image components sum up to one.

More formally, let  $V: \Omega \rightarrow \mathbb{R}^b$  denote a sensed multi-channel digital image with  $b$  bands and  $r \times c = N$  pixels on a discrete domain  $\Omega \subset \mathbb{Z}^2$ . Let a pixel  $p_{(x,y)} \in \mathbb{R}^2$

denote the Voronoi region of a grid point  $(x, y) \in \mathbb{Z}^2$  (region consisting of all points closer to that grid point than to any other). Let the set of  $m$ -component segmentation vectors be

$$\mathbb{A}_m = \left\{ \alpha = (\alpha_1, \alpha_2, \dots, \alpha_m) \in [0, 1]^m \mid \sum_{j=1}^m \alpha_j = 1 \right\}. \quad (3.22)$$

A coverage segmentation of an image  $V$  into  $m$  components is a set of ordered pairs

$$\left\{ \left( (x, y), \alpha(x, y) \right) \mid (x, y) \in \Omega, \alpha(x, y) \in \mathbb{A}_m \right\}, \quad \alpha_j \approx \frac{|p_{(x,y)} \cap S_j|}{|p_{(x,y)}|}, \quad (3.23)$$

where  $S_j \subset \mathbb{R}^2$  is the extent of the  $j$ -th image component,  $j = 1, 2, \dots, m$ . The continuous sets  $S_j$  are, in general, not known and the values  $\alpha_j$  have to be estimated from the image data.

### 3.4.2 Linear unmixing

Spectral unmixing is the procedure by which the measured spectrum of a mixed pixel is decomposed into a collection of constituent spectra, or end-members, and a set of corresponding fractions, or abundances, that indicate the proportion of each end-member present in the pixel. Linear unmixing refers to the spectral unmixing under the assumption that the measured spectrum is obtained as linear mixture of the constituent spectra.

Linear unmixing of image intensities is commonly used in the field of image processing [66–70]. This technique is often used to enable analysis of remotely sensed hyperspectral data sets with subpixel precision [66, 69, 71]. In this application the end-members normally correspond to the familiar macroscopic objects in the scene, such as water, soil or forest and linear unmixing based methods estimate the fractional abundances (coverage) of the classes at a sub-pixel level. Another field where linear unmixing is frequently applied is fluorescence microscopy [66–68] where it is usually used to determine the relative contribution from each fluorophore to every pixel of the image.

### 3.4.3 Data fidelity term for coverage segmentation based on linear unmixing

Since the aim of coverage segmentation is to estimate the fractional coverage of pixels with different image components, unmixing approach is found to be suitable for this application as well.



### 3.4. Data fidelity term for coverage segmentation

Assuming that pixel intensities are obtained as linear mixtures of pure class representatives (end-members), the sensed image  $V$  can be modeled as

$$V = A \cdot C + \text{noise}, \quad (3.24)$$

where  $C = [c_{j,k}]_{m \times b}$  is the end-member matrix and  $c_{j,k}$  corresponds to the (expected) image value of class  $j$  in band  $k$  (e.g., color channels in an RGB image).

Here the sensed image  $V$  is represented as a matrix of size  $N_l \times b$ , s.t. a row of  $V$  contains intensities of one pixel in each of the observed bands, and a column represents the pixel intensities in one band, concatenated over the whole image. In other words,  $V = [v^1, v^2, \dots, v^b]$  where vector  $v^k = [v_{1,k}, \dots, v_{N_l,k}]^T$  is the  $k$ -th band; in previous sections one channel  $b = 1$  (grey-scale) images are observed; in that case matrix  $V$  reduces to vector  $v = [v_1, \dots, v_{N_l}]^T$ . The  $m$ -component coverage segmentation of  $V$  is represent by a matrix  $A = [\alpha_{i,j}]_{N_l \times m}$ , where  $\alpha_{i,j} \in [0, 1]$  is the coverage of the  $i$ -th pixel by the  $j$ -th component. Rows of  $A$  with the value 1 at the  $j$ -th position and zeros elsewhere correspond to pure pixels (completely covered by image component  $S_j$ ), while rows with two or more non-zero elements correspond to mixed pixels.

Knowing  $V$  and the end-member matrix  $C$ , the aim is to estimate  $A$ .

The linear unmixing problem can be modeled also as an energy minimization problem. If we assume that noise present in the observed image  $V$  follows Gaussian distribution, the data fidelity term

$$D_{CS}(A; V) = \frac{1}{2} \|AC - V\|^2 \quad (3.25)$$

is the squared Euclidean distance between the linear mixture  $AC$  and the sensed data  $V$ . Its minimization, constrained to  $A \in \mathbb{A}_{N_l \times m}$ , provides an unmixing segmentation.

#### 3.4.4 Data fidelity term for super-resolution coverage segmentation

The coverage segmentation method proposed in [38] with the data fidelity term given by (3.25) exhibits poor performance in the presence of blur in the acquired image. Therefore, in Publication VI the method is extended to take into account blurring and decimation process as well. This adjustment significantly increases its robustness and applicability of the method in the presence of blur in the sensed image and enables estimation of coverage values at increased spatial resolution which, in some applications, e.g., remote sensing, is a highly desirable feature.

Assuming that the sensed image  $V$  of size  $N_l \times b$  is of a form

$$V = S \cdot H \cdot A \cdot C + \text{noise}, \quad (3.26)$$

where  $A_{N_h \times m}$  is a coverage segmentation at  $d = d_r \times d_c$  times increased spatial resolution, and where  $C_{m \times b}$ ,  $S_{N_l \times N_h}$ , and  $H_{N_h \times N_h}$  are end-member, down-sampling, and blurring matrices, respectively, then, in the presence of Gaussian noise, the data fidelity term is defined as

$$D_{SRCs}(A; V) = \frac{1}{2} \|SHAC - V\|^2 = \frac{1}{2} \sum_{k=1}^b \sum_{i=1}^{N_l} ((SHAC)_{i,k} - v_{i,k})^2, \quad (3.27)$$

and it is a generalization of  $D_{CS}(A; V)$  which takes into account blur and down-sampling resulting from the acquisition process. The benefits of this adjustment is demonstrated on the several particular examples presented in Chapters 6 and 7, where we present performance evaluation of the developed methods. The methods which perform simultaneous image restoration and crisp segmentation by energy minimization and include data terms adjusted to address blur or other degradation operators can be found in [24, 72–75].

# 4 Regularization terms

## 4.1 Regularization terms for image restoration

### 4.1.1 Sparse signals

As discussed in Chapter 2, the regularization term equals the negative log-prior of the solution and its role is to improve the ill-posed nature of an inverse problem. However, it does not necessarily mean that the regularized problem will be well-posed in the sense of Hadamard; for instance, the regularization term can be non-convex or/and non-differentiable and in that case the solution will be non-unique. Hence, a general objective in such situations is that a regularization term should significantly restrict the space of many possible solutions to the desired ones and it should provide numerical stability.

Typically, as aforementioned, the regularization term incorporates our prior knowledge about image which we want to obtain. For restoration problems, the most common regularizing assumption is that the signal  $u$  which we are interested in is *sparse*, [76]. The signal is said to be sparse in a suitable basis (dictionary) if it can be represented as linear combination of atoms (basis functions) where most of its coefficients are zero or if it has a small number of large coefficients and the remaining ones take values around zero. If the signal is not sparse in the spatial domain, it is transformed to another domain where it is sparse, i.e., where it can be represented by only few non-zero coefficients. For example, the sinus signal is clearly not sparse in the spatial (time) domain but when we transform it to the Fourier domain, the signal becomes extremely sparse (Fig. 4.1).

For images, e.g., discrete gradient or higher order differences transform them into sparse signals. Many other transforms were proposed for modelling sig-

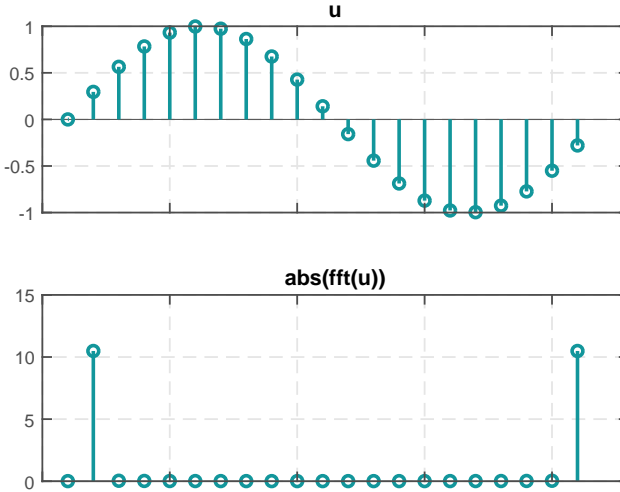


Figure 4.1: Plot of discrete sinus signal in time and Fourier domain. The dense signal in time domain (time representation) is sparse in Fourier domain (frequency representation).

nals via sparse representation in a suitable basis (dictionary); in particular, really efficient sparse representations for inverse problems in image analysis are wavelets [77–80], curvelets [81], shearlets [82–84], bandlets [85], contourlets [86], and many other. In this thesis we focus on discrete gradient transform and in relation to it Total Variation regularization which are described in more details in Section 4.1.3.

### 4.1.2 Sparsity promoting norms

An ideal measure of sparsity of the signal  $u$  is  $\ell_0$  pseudo norm

$$\ell_0(u) = \#\{i : u_i \neq 0\}, \tag{4.1}$$

which counts number of non-zero elements of the signal.

However, minimization of  $\ell_0$  norm is NP-hard problem and therefore, the  $\ell_0$  pseudo norm is replaced by  $\ell_p$  norm

$$\ell_p(u) = \left( \sum_{i=0}^{N_h} |u_i|^p \right)^{\frac{1}{p}}, \tag{4.2}$$

most frequently with the convex, although non-differentiable,  $\ell_1$  norm obtained for  $p = 1$ .

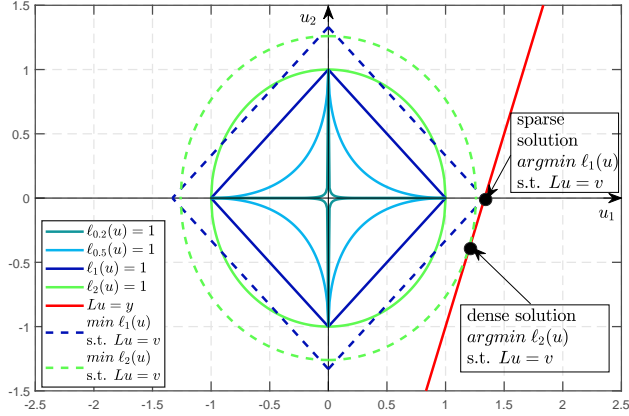


Figure 4.2: Sparsity promotion achieved by minimization of  $\ell_1$  norm. Minimization of  $\ell_p$  norm for  $0 < p \leq 1$  and  $1 < p \leq 2$  gives a sparse and a dense solution, respectively. However, for  $0 < p < 1$ , the norm is non-convex and it is challenging from the optimization point of view. Hence, convex and non-differentiable  $\ell_1$  norm is the best choice for convex approximation of  $\ell_0$  among all  $\ell_p$  norms for  $0 < p \leq 2$ .

To motivate the choice of  $\ell_1$  norm as the best choice for convex approximation of  $\ell_0$ , let us consider the following example. Suppose that we are aiming to solve an underdetermined linear system of equations  $Lu = v$ , where  $u = [u_1, u_2]^T \in \mathbb{R}^2$  is unknown,  $L = [l_{i,j}] \in \mathbb{R}^{1 \times 2}$  and  $v \in \mathbb{R}$  are known. The solution can be found by solving *basis pursuit* [87] regularized problem

$$\begin{aligned} \min_u f(u) \\ \text{such that } Lu = v, \end{aligned} \quad (4.3)$$

where  $f(u)$  is a sparsity-inducing function; typically  $f(u) = \ell_p(u)$ . In that case, the geometrical representation of the constraint  $Lu = v$  is a line in 2D space (red line in Fig. 4.2), and the solution of the problem corresponds to the intersection of the line (red) and the level set of  $\ell_p$  norm (indicated as green and blue lines for different values of  $p$ ). Since for  $0 < p \leq 1$  level sets of  $\ell_p$  norm are pointed (spiky), minimization of  $\ell_p$  for these values of  $p$  leads to sparse solutions. However, for  $p < 1$  the norm is non-convex and therefore challenging from the optimization perspective. On the other hand,  $\ell_p$  norm for  $1 < p \leq 2$  is convex, but its level set is smooth and its minimization leads to dense solution (see illustration for  $\ell_2$  minimization given in Fig. 4.2). Therefore, the optimal choice for convex approximation of  $\ell_0$  pseudo norm is sparsity promoting convex  $\ell_1$  norm, which can be efficiently minimized despite its non-differentiability [88].

In more general case, we are aiming to solve in this thesis the same under-determined linear system of equations  $Lu = v$ , where  $u \in \mathbb{R}^{N_h}$  is unknown,  $L = [L_{i,j}] \in \mathbb{R}^{N_l \times N_h}$  and  $v \in \mathbb{R}^{N_l}$  are known and  $N_h > N_l$ . This problem is a typical inverse problem - e.g., when  $v$  is a sensed low-resolution image and  $L = S$  is the down-sampling matrix, then  $u$  is the recovered high-resolution image. Another example is a sparse reconstruction problem when  $v$  is the sensed signal,  $L$  is the dictionary and  $u$  is the desired sparse representation.

### 4.1.3 Total Variation based regularization

The focus of the work presented in this thesis is on the family of regularization methods based on Total Variation (TV) [44]

$$R(u) = \sum_{i=1}^{N_h} \Phi(|\nabla(u_i)|), \quad (4.4)$$

where  $\nabla$  stands for image gradient and  $|\cdot|$  denotes  $\ell_2$  norm. The function  $\Phi$  is referred to as *potential function*. By using a potential equal to the identity function, the regularization term reduces to classic TV regularization.

Total Variation is  $\ell_1$  norm of  $\ell_2$  norm of the pixel-wise image gradients. The discrete image gradient at point  $u_i$  (at location  $i$ ),  $\nabla(u_i)$ , is computed as  $\nabla(u_i) = [u_r - u_i, u_b - u_i]^T$ , where  $r$  and  $b$  denote the indices of the edge neighbors to the right and below the pixel  $u_i$ , respectively, and its magnitude equals to  $|\nabla(u_i)| = \sqrt{(u_r - u_i)^2 + (u_b - u_i)^2}$ . TV regularization is designed with the explicit goal of preserving sharp discontinuities (edges) in images while removing noise and other unwanted fine scale detail. TV regularization term allows for discontinuities but at the same time disfavors oscillations and in that way it achieves the both goals.

Minimization of Total Variation as a regularization induces sparsity in the gradient of the image. Fig. 4.3 (b) shows an illustration of the gradient magnitude of the *Cameraman* image. As can be seen from Fig. 4.3 (b), the gradient magnitude of the image is a sparse image with limited number of non-zero values which correspond to the locations of edges in the image and zeros in big noise-free homogeneous regions. Noise in homogeneous regions (Fig. 4.3 (d)) increases TV and leads to non-sparse gradient images; at the same time, edges correspond to high TV and should be preserved. Minimization of TV results in images with preserved edges and suppressed noise in flat homogeneous regions.



(a) Original image.



(b) Gradient magnitude of the original image,  $TV = 3800$ .



(c) Noisy image.



(d) Gradient magnitude of the noisy image,  $TV = 7700$ .

Figure 4.3: Illustration of sparsity of the gradient magnitude image. Noise in homogeneous regions increase  $TV$  (sum of gradient magnitudes for all pixels).

### 4.1.4 Edge preserving potential functions

The main problem with  $TV$  regularization function is that it is not smooth around zero which makes optimization challenging. In addition to potential function that equals identity function for which regularization reduces to non-smooth  $TV$ , numerous smooth potentials are proposed and used in image restoration problems [45, 89–96]; some of potential functions considered in this thesis are listed in Table 4.1. In most cases a smooth potential function is designed s.t. small intensity changes (assumed to be noise) are additionally penalized, while large changes (assumed to be edges) are allowed to be preserved. Figure 4.4 illustrates different potential functions with analytical expressions given in Table 4.1

exhibiting different behaviours. Potential functions designed to emphasize the smoothing effect of regularization promote the equalization of the pixel intensities (e.g. Tikhonov  $\Phi_6(s) = s^2$ ). The edge preserving potentials allow smoothing in the areas with low gradient magnitudes only. That means that at locations where the variation of the pixels intensities is low, minimization of such functions leads to the smoothing effect, whereas close to edges, where the gradient magnitude is high, it keeps the intensity values and thereby preserves edges. To ensure such behavior,  $\Phi$  is usually defined to be near quadratic close to zero and near linear for larger values (e.g. Huber  $\Phi_5(s)$  regularization).

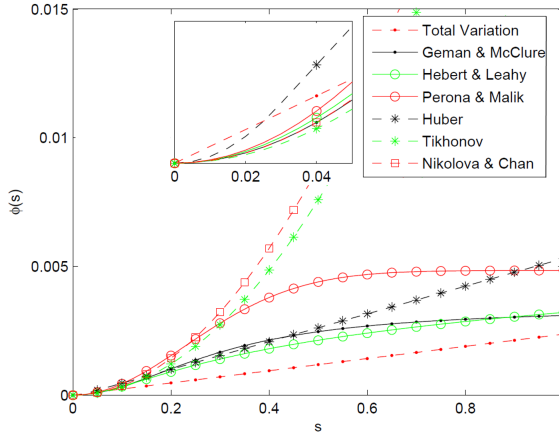


Figure 4.4: Potential functions.

In [89] the theoretical conditions for edge preserving potentials are given:

$$\lim_{s \rightarrow 0^+} \frac{\Phi'(s)}{2s} = M, \quad M > 0 \quad (4.5a)$$

$$\lim_{s \rightarrow \infty} \frac{\Phi'(s)}{2s} = 0. \quad (4.5b)$$

The condition (4.5a) regulates the behavior in homogeneous areas (gradient magnitudes are small), while (4.5b) determines the behavior of  $\Phi$  close to the edges (gradient magnitudes are large). The examples using the Huber potential for deconvolution are presented in [93], however, no explicit performance evaluation of the potentials is given. A study of the effectiveness of different potentials in image denoising is given in [45], where it is concluded that the Huber potential ( $\Phi_5$  in Table 4.1) works best overall, and that the Geman & McClure potential ( $\Phi_2$ ) shows the best performance in low noise settings.

In Publication IV we explored the utilization of seven different potential functions (Table 4.1) in the deconvolution of images degraded by both Gaussian



## 4.1. Regularization terms for image restoration

Table 4.1: Potential functions.

Potential	$\Phi(s)$	Convex
TV [44]	$\Phi_1(s) = s$	yes
Geman&McClure [90]	$\Phi_2(s) = \frac{\omega s^2}{1+\omega s^2}$	no
Hebert&Leahy [91]	$\Phi_3(s) = \ln(1 + \omega s^2)$	no
Perona&Malik [94]	$\Phi_4(s) = 1 - e^{\omega s^2}$	no
Huber [96]	$\Phi_5(s) = \begin{cases} \frac{s^2}{2\omega}, & s \leq \omega \\ s - \frac{\omega}{2}, & s > \omega \end{cases}$	yes
Tikhonov [95]	$\Phi_6(s) = s^2$	yes
Nikolova&Chan [92]	$\Phi_7(s) = \begin{cases} \sin(\omega s^2), & s \leq \sqrt{\frac{\omega}{2\pi}} \\ 1, & s > \sqrt{\frac{\omega}{2\pi}} \end{cases}$	no

noise and blur. The performed tests confirm that the utilization of potential functions in regularized image denoising and deconvolution provides a straightforward way to increase the quality of the restored images. Our conclusion was that the Huber potential performs outstandingly best, providing the best values of the used quality measures and improved edge preservation, in comparison to all the observed potentials.

Further, this study was extended to the signal-dependent types of noise, Poisson and Poisson-Gaussian case. The conducted empirical study, presented in Publication I evaluates improvement that can be achieved if Huber potential is used instead of TV regularization for such noise distributions.

The results obtained in Publication IV and Publication I are presented in Chapter 6.

It is worth mentioning, that a consequence of sparsity in the gradient domain induced by TV regularization is that the resulting images tend to be piecewise constant. This property is known as staircasing effect, which is often considered to be a drawback for certain applications. It can be avoided by non-local treatment of similar patches [97–100] or by including higher order difference operators [30, 101]. For blind deconvolution of TEM images in Publication II, in addition to the gradient, the second order difference operator is considered in order to achieve better preservation of fine textures important for the application of interest.

## 4.2 Regularization terms for segmentation

An energy minimization formulation of the segmentation problem was introduced by Mumford and Shah [102], and has, together with its numerous variations, been thoroughly studied and utilized in different applications [24, 103]. In next section we briefly explain the Mumford-Shah model, which inspired coverage segmentation method from [38]. In addition, we present also the simplified version of Mumford-Shah model introduced by Chan and Vese in order to give link between Mumford-Shah penalty (overall perimeter) and TV which motivated us to use Huberized TV instead of perimeter in Publication VI in order to improve performance of the coverage segmentation method from [38]. We at the end introduce additional fuzzy based regularization terms of coverage segmentation from [38], as well as improved ones from Publication VI.

### 4.2.1 Mumford-Shah penalty

In this section, an image  $\boldsymbol{v}$  is considered to be a continuous function  $\boldsymbol{v}: \Omega \rightarrow \mathbb{R}^b$ ,  $\Omega \subseteq \mathbb{R}^2$  (for  $b = 1$ ,  $\boldsymbol{v}$  is a gray-scale image whereas for  $b > 1$ ,  $\boldsymbol{v}$  is a multi-channel image). We represent here a digital image with a mathematical function to be able to apply Partial Differential Equations (PDE) theory for further derivations. Therefore, to distinguish between the previously used discrete image representation where for  $b = 1$  the image was a vector  $\boldsymbol{v}$  and for  $b > 1$  it is considered a matrix  $V$ , we introduce a bold symbol  $\boldsymbol{v}$  to denote its continuous representation. In continuous setting, inverse problems are usually modelled with Partial Differential Equations (PDE) and energy minimization based restoration methods are referred to as variational methods [104, 105]. The equivalence between PDE and energy minimization based image restoration is considered in several papers and books, e.g. [105, 106], as well as in the numerous references therein.

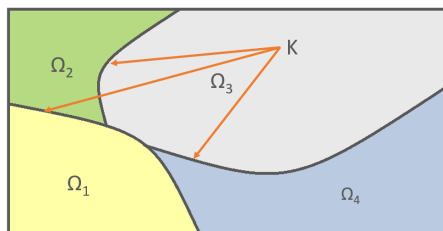


Figure 4.5: Illustration of Mumford-Shah image segmentation. Image is decomposed into regions  $\Omega_1, \Omega_2, \dots, \Omega_n$  and it varies smoothly within each  $\Omega_i$  and discontinuously across the boundary  $K$  between different  $\Omega_i$ .

## 4.2. Regularization terms for segmentation

---

A variational formulation of the crisp segmentation problem was introduced by Mumford and Shah in their seminal work [102]. As formulated by them, the segmentation problem consists of computing a decomposition (Fig. 4.5) of the domain  $\Omega \subseteq \mathbb{R}^2$  of the image  $\mathbf{v}$

$$\Omega = \Omega_1 \cup \Omega_2 \cup \dots \cup \Omega_n \cup K \quad (4.6)$$

with the properties:

- The image  $\mathbf{v}$  varies smoothly within each  $\Omega_i$ ,
- The image  $\mathbf{v}$  varies discontinuously and/or rapidly across most of the boundary  $K$  between different  $\Omega_i$ .

In essence, the problem consists in approximating the sensed image  $\mathbf{v}$  by piecewise smooth image (function)  $\mathbf{u}$  which is continuous inside each  $\Omega_i$  and may be discontinuous across  $K$ . In the continuous setting, this problem of joint estimation of a cartoon image  $\mathbf{u}$  (Fig. 4.6) and boundary set  $K$  can be written as the following optimization problem

$$\min_{\mathbf{u}, K} E(\mathbf{u}, K) = \min_{\mathbf{u}, K} \lambda \int_{\Omega} (\mathbf{u}(x) - \mathbf{v}(x))^2 dx + \int_{\Omega \setminus K} |\nabla \mathbf{u}(x)|^2 dx + \mu \text{Length}(K). \quad (4.7)$$

The first term, the data fidelity term, forces  $\mathbf{u}$  to be close to  $\mathbf{v}$ . The regularization terms, the second and the third one, are referred to as Mumford-Shah penalty. They ensure respectively, that  $\mathbf{u}$  is smooth on  $\Omega \setminus K$  and that total length of the boundary  $K$  is as small as possible i.e., that  $K$  is also smooth. The simplified version of the model given by energy functional (4.7) is piecewise constant formulation

$$\min_{\mathbf{u}, K} E(\mathbf{u}, K) = \min_{\mathbf{u}, K} \lambda \int_{\Omega} (\mathbf{u}(x) - \mathbf{v}(x))^2 dx + \mu \text{Length}(K), \quad (4.8)$$

where  $\mathbf{u}$  is constant on each  $\Omega_i$  and  $K$  is the boundary of a closed set. Here the Mumford-Shah penalty reduces to the length of the boundary  $K$  (overall perimeter). Both the original and the simplified Mumford-Shah problems are non-convex and therefore hard to solve. Various algorithms for efficient computations of the solutions are proposed in the literature; some of them are level sets approximations by Chan and Vese [107, 108], phase-field elliptic approximations by Ambrosio and Tortorelli [74, 109], finite element approximations [110, 111] and many others [24, 25].



Figure 4.6: Piecewise smooth approximation using the Mumford-Shah functional. (a) Original image  $\mathbf{v}$ , and (b) piecewise smooth approximation  $\mathbf{u}$  (source [88]).

#### 4.2.2 Link between Mumford-Shah penalty and Total Variation via level set functions

Chan and Vese in their seminal work [107] consider a further simplified version of the piecewise constant Mumford-Shah model

$$\begin{aligned} \min_{k_1, k_2, K} E(k_1, k_2, K) = & \min_{k_1, k_2, K} \lambda_1 \int_{\text{inside}(K)} (k_1 - \mathbf{v}(x))^2 dx + \\ & \lambda_2 \int_{\text{outside}(K)} (k_2 - \mathbf{v}(x))^2 dx + \mu \text{Length}(K) + \nu \text{Area}(\text{inside}(K)), \end{aligned} \quad (4.9)$$

allowing the image  $\mathbf{u}$  to have only two values

$$\mathbf{u}(x) = \begin{cases} k_1, & \text{inside } K \\ k_2, & \text{outside } K. \end{cases} \quad (4.10)$$

In addition to the Mumford-Shah penalty, Chan-Vese model penalizes the area inside  $K$ . Level set representation of the curve  $K$  facilitates numerical algorithms for solving the problem. The curve  $K$  can be represented as the zero-crossing of a level set function  $g$  by the relationship

$$K = \{x \in \Omega : g(x) = 0\}, \quad (4.11)$$

while regions outside and inside  $K$  are distinguished by sign of  $g$ . For example,

$$g(x) = R - \sqrt{x_1^2 + x_2^2} \quad (4.12)$$

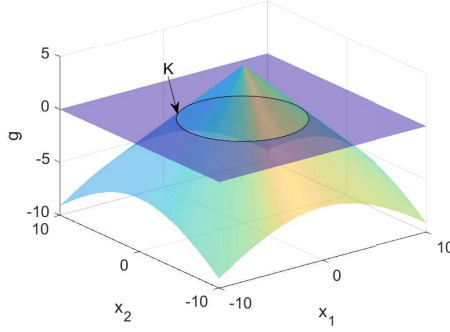


Figure 4.7: An example of a level set function of a circle  $K$  of radius 5 is function  $g(x) = 5 - \sqrt{x_1^2 + x_2^2}$  whose graph is a cone. The intersection of the cone and  $x_1 x_2$ -plane is the circle  $K$ .

is a level set function for a circle  $K$  of radius  $R$ , Fig. 4.7.

The length of the desired contour  $K$  equals the Total Variation of the Heaviside function of level set function  $g$

$$\text{Length}(K) = \int_{\Omega} |\nabla H(g(x))| dx, \quad (4.13)$$

and the area inside the curve  $K$  can be obtained as

$$\text{Area}(\text{inside}(K)) = \int_{\Omega} H(g(x)) dx, \quad (4.14)$$

where  $H$  denotes the Heaviside function

$$H(s) = \begin{cases} 1, & s \geq 0 \\ 0, & s < 0. \end{cases} \quad (4.15)$$

Chan-Vese model can be reformulated as

$$\begin{aligned} \min_{k_1, k_2, g} E(k_1, k_2, g) = & \min_{k_1, k_2, g} \lambda_1 \int_{\Omega} (k_1 - v(x))^2 H(g(x)) dx + \\ & \lambda_2 \int_{\Omega} (k_2 - v(x))^2 (1 - H(g(x))) dx + \mu \int_{\Omega} |\nabla H(g(x))| dx + \nu \int_{\Omega} H(g(x)) dx. \end{aligned} \quad (4.16)$$

The level set based representation of the boundary highlights that Mumford-Shah penalty and the Total Variation are essentially the same; they are both used alternatively for various restoration and segmentation tasks [24, 88].

### 4.2.3 Regularization terms for coverage segmentation

Inspired by Mumford-Shah model, the unmixing coverage segmentation proposed in [38] (obtained by minimization of data term defined by Eq. (3.25)) is regularized by

$$R(A) = \mu P(A) + \nu(T(A) + F(A)) \quad (4.17)$$

where  $P(A)$ ,  $T(A)$  and  $F(A)$  are overall length of boundaries (Mumford-Shah penalty), boundary thickness and total fuzziness of coverage segmentation  $A$ . The parameters  $\mu$  and  $\nu$  give relative importance to different terms.

Although the segmentation model is inspired by Mumford-Shah model, it is, at the same time, strongly determined by the idea of coverage representations, and the regularization terms are suggested in connection to that.

The *perimeter term*  $P$  is defined as the overall length of the boundaries of the  $m$  objects of the coverage segmentation  $A$  and it is computed according to the method presented in [34]:

$$P(A) = \frac{1}{2} \sum_{j=1}^m P(A_j) = \frac{1}{2} \sum_{j=1}^m \sum_{t \in \tau_{2 \times 2}(A_j)} P(t), \quad (4.18)$$

where  $A_j$  is the  $j$ th column of  $A$ . Each boundary is shared between two objects, therefore the accumulated value is divided by two. The perimeter contribution of  $2 \times 2$  tile  $t$ ,  $P(t)$ , with four assigned coverage values  $\alpha_1, \alpha_2, \alpha_3, \alpha_4$  is: For sorted coverage values s.t.  $\alpha_1 \leq \alpha_2 \leq \alpha_3 \leq \alpha_4$ ,

$$P(t) = (\alpha_2 - \alpha_1 + \alpha_4 - \alpha_3) \cdot \frac{b}{2} + (\alpha_3 - \alpha_2) \cdot w, \quad (4.19)$$

where

$$w = \begin{cases} a, & \text{if the two pixels with the smallest coverage share a common side in } t \\ b, & \text{if the two pixels with the smallest coverage are diagonally placed in } t. \end{cases} \quad (4.20)$$

The lengths  $a$  and  $b$  are optimal horizontal and diagonal unit steps. We use  $a = 0.948$  and  $b = 1.343$ . Minimization of  $P(A)$  favours smooth object boundaries and acts towards suppressing noise.

The *thickness term*  $T$  is defined as

$$T(A) = \frac{1}{2} \sum_{j=1}^m T(A_j) = \frac{1}{2} \sum_{j=1}^m \sum_{t \in \tau_{2 \times 2}(A_j)} T(t), \quad (4.21)$$

where the thickness of one tile is computed as

$$T(t) = \prod_{i=1}^4 4\alpha_i(1 - \alpha_i). \quad (4.22)$$

The *fuzziness term*  $F$  is given by

$$F(A) = \sum_{i=1}^N \sum_{j=1}^m 4\alpha_{i,j}(1 - \alpha_{i,j}). \quad (4.23)$$

Both the thickness and the fuzziness terms reach minimal values for coverage values equal to 0 or 1, and increase for non-sparse (fuzzy) solutions, i.e., coverage values within (0,1). The thickness term reaches its minimal value zero if at least one pixel in each  $2 \times 2$  configuration is crisp, thus promoting sparse solutions with thin (at most one pixel thick) fuzzy transitions between the objects, whereas the fuzziness term favours majority of (individual) pixels to be classified as pure. Combined they lead to segmentations with thin fuzzy boundaries, and favour the majority of pixels being classified as pure.

Minimization of  $P(A)$ ,  $T(A)$  and  $F(A)$  imposes restrictions to the solution according to the assumptions made about the problem: the segmented objects are expected to consist of connected regions separated by smooth, one pixel thick boundaries. This ensures that the solution has smooth boundaries and that most of the pixels are classified as pure (either zeros or ones), i.e., that the solution is sparse.

### 4.2.4 Regularization terms for super-resolution coverage segmentation

Non-smoothness of the perimeter term  $P$  makes the optimization of (1.5) challenging. Motivated by the equality of the Total Variation and Mumford-Shah overall boundary length penalty and the good performances of the Total Variation based regularizations in a number of restoration methods (Publications I-V), in order to address the issue of non-smoothness of the perimeter term, in Publication VI, it is proposed to replace the perimeter term by Huberized Total Variation.

In Publication VI, two alternatives for perimeter term are evaluated, which promote noise suppression and smooth boundaries:

- Tile-wise Huberized Perimeter term

$$P_{HUB}(A) = \frac{1}{2} \sum_{j=1}^m P_{HUB}(A_j) = \frac{1}{2} \sum_{j=1}^m \sum_{t \in \tau_{2 \times 2}(A_j)} \Phi_{HUB}(P(t)), \quad (4.24)$$

where  $A_j$  is the  $j$ th column of  $A$  and  $P(t)$  is the perimeter of a  $2 \times 2$  tile, computed as in (4.19);

- Huberized Total Variation

$$\text{TV}_{HUB}(A) = \frac{1}{2} \sum_{j=1}^m \text{TV}_{HUB}(A_j) = \frac{1}{2} \sum_{j=1}^m \sum_{i=1}^{N_h} \Phi_{HUB}(|\nabla(\alpha_{i,j})|). \quad (4.25)$$

In addition to this, in Publication VI a new sparsity promoting fuzziness term for enforcing the majority of pixels to be classified as pure is designed. New sparsity promoting fuzziness term  $\tilde{F}$  penalizes fuzzy regions, except where the contrast is high, i.e., on object boundaries:

$$\tilde{F}(A) = \sum_{i=1}^{N_h} \sum_{j=1}^m 4\alpha_{i,j}(1-\alpha_{i,j})(1-k_{i,j}), \quad (4.26)$$

where  $k_{i,j} = \max_{k \in \mathcal{N}(i)} \alpha_{k,j} - \min_{k \in \mathcal{N}(i)} \alpha_{k,j}$  and  $\mathcal{N}(i)$  is the  $3 \times 3$  neighbourhood of pixel  $i$ .  $\tilde{F}$  replaces the previously used combination of  $F$  and  $T$ . The  $\tilde{F}$  restricts the fuzziness penalty (promoting pure pixels) to non-edge regions, while not imposing this constraint along the object boundaries. An undesired competition between  $T(A)$ , allowing fuzziness along a thin boundary, and  $F(A)$ , promoting pure pixels everywhere, which was leading to instability of the optimization algorithm is, by that, avoided leading to a simpler and more numerically stable minimization of the energy function.

In Publication VI the constraint that each row of  $A$  sums to one is formulated as an additional regularization term. The term  $G$  promotes that each row of the matrix  $A$  sums up to 1:

$$G(A) = \frac{1}{2} \sum_{i=1}^{N_h} \left( \left( \sum_{j=1}^m \alpha_{i,j} \right) - 1 \right)^2. \quad (4.27)$$

This constraint is in [38] enforced by projections within the optimization. This change reduces the computational time.

Combining these terms, the following energy functions and corresponding coverage segmentation methods are defined:

- (i) CS-PTF – coverage segmentation proposed in [38]  
 $E(A) = D_{CS}(A; V) + \mu P(A) + \nu T(A) + \nu F(A),$
- (ii) CS-P $\tilde{F}$  – the improved version of CS-PTF  
 $E(A) = D_{CS}(A; V) + \mu P(A) + \nu \tilde{F}(A),$
- (iii) SRCS-P $\tilde{F}$  – super-resolution coverage segmentation  
 $E(A) = D_{SRCS}(A; V) + \mu P(A) + \nu \tilde{F}(A) + \eta G(A),$



## 4.2. Regularization terms for segmentation

---

(iv) SRCS-HP $\tilde{F}$  – super-resolution coverage segmentation

$$E(A) = D_{SRCS}(A; V) + \mu P_{HUB}(A) + \nu \tilde{F}(A) + \eta G(A),$$

(v) SRCS-HTV $\tilde{F}$  – super-resolution coverage segmentation

$$E(A) = D_{SRCS}(A; V) + \mu TV_{HUB}(A) + \nu \tilde{F}(A) + \eta G(A).$$

The introduced notation is used in the following chapters to specify the particular combination of terms used in  $E$ . In Chapter 6 a comparison of these methods is given and it is demonstrated that SRCS-HTV $\tilde{F}$  – super-resolution coverage segmentation method with newly introduced data fidelity and sparsity promoting regularization terms exhibits the best performances.



# 5 Numerical optimization

## 5.1 Minimization of energy functions

An important issue in energy minimization based imaging problems is efficient optimization of the energy function. Let us consider a problem of minimizing the energy function given by (1.2),

$$\min_u E(u). \quad (5.1)$$

In general, it is not possible to find the analytical solution of the problem, mainly because of its large scale nature. For example, for the restoration of an image of the size of  $1000 \times 1000$  pixels, an objective function includes 1 million variables. There are many challenges related to the numerical optimization of the energy function used in restoration and other tasks in image processing. In the continuation we mention the most important ones and how they can be tackled.

The energy function  $E$  is composed of the data fidelity and the regularization terms. As discussed in Chapter 3, the data fidelity terms for both restoration and coverage segmentation which are proposed within this thesis are convex quadratic functions. Therefore, the minimization of such terms is not particularly difficult. However, as explained in Chapter 4, the sparsity promoting regularization terms can be both non-smooth and non-convex which makes optimization challenging.

Non-smoothness of the regularization term entails the non-smoothness of the whole energy function and therefore the wide class of minimization methods based on the gradient or higher order derivatives such as Newton's method cannot be used. Hence the problem of numerically optimizing  $E$  is usually tackled with the first-order descent methods, which are extensions of a plain gradient descent, appropriately adapted to deal with the non-smoothness of

the objective function. Extensive and very enlightening recent surveys on the first-order gradient descent based methods for imaging problems are presented in [88, 112].

To address the non-smoothness, the problem can be “split” into simpler sub-problems which can be solved separately. This idea is commonly referred to as “proximal splitting” and, despite the fact that it relies on the ideas from the 1950s [113, 114], it has been a very active research topic in the past ten years in image and signal processing [88, 115, 116]. A variety of approaches and algorithms to minimize a regularized energy function based on proximal splitting are presented in the literature. Many of them exist under different names but refer to essentially the same algorithms. The best known are “primal-dual” algorithms such as Primal-Dual Hybrid Gradient (PDHG) [25, 117–119], the alternating direction method of multipliers (ADMM) [114, 120, 121] and the split Bregman iteration [122, 123] or “Douglas-Rachford splitting” [113]. A description of the relationships between the ADMM and similar splitting methods mentioned here can be found in [118], while many applications of such methods in image restoration and segmentation are presented in [13, 17, 23, 124–126]. An alternative way to deal with non-smooth problem is to smooth it in controlled way [127, 128].

Non-convexity of the regularization terms can cause non-convexity of the whole energy function, which can make the optimization difficult due to the existence of multiple local minima. Non-convexity prevents utilization of numerous efficient deterministic methods specifically designed for convex minimization of objective function. Some of the recently developed non-smooth and non-convex deterministic optimization methods based on proximal-splitting can be found in the studies [129–131], and in the survey [88].

The focus in this thesis is on the continuous optimization methods and not on discrete or combinatorial optimization which is another powerful approach to solve the imaging problems of similar nature. Such methods, based on graph cuts [132, 133] or network flows, are very efficient and have been extensively developed by the computer vision community to tackle most of the problems we address here with continuous optimization. Moreover, in this thesis only deterministic numerical optimization methods are taken into consideration as we want to obtain unique solutions. Very powerful stochastic optimization methods which have been driven by big data applications (mainly deep learning) are also worth mentioning. Due to the complexity of the models and very large datasets used for training, the optimization techniques for such applications are very specific and usually rely on stochastic gradient descent schemes [134].

Problem (5.1) represents an unconstrained optimization problem. However, some applications require restrictions of the search space and the problems

considered in this thesis belong to this type of problems. For instance, in the tasks of super-resolution reconstruction, as well as coverage segmentation, the search space is restricted to a Cartesian product of  $N$  intervals  $[0, 1]$  which corresponds to our assumption that each pixel has the intensity in  $[0, 1]$  or that the coverage of each pixel by object belongs to  $[0, 1]$ , respectively. In addition, in coverage segmentation, the sum of coverage values for each pixel has to be 1. Similarly, in blind restoration task, the sum of pixel intensities in the blur kernel has to be 1. In a general case, the constrained optimization of an energy function is given by

$$\min_{u \in \Theta} E(u), \quad (5.2)$$

where  $\Theta$  is a feasible set. Beside the above analyzed issues, the constraint represents an additional challenge which has to be solved. One possibility is to transform the constrained problem into a unconstrained problem, when the constraint is formulated as a new regularization term. In that way, the constraint that the coverage of each pixel by different objects has to sum up to one is included as an additional regularization term (4.27) in the energy function for coverage segmentation.

Another way to handle this challenge is the direct application of a suitable optimization method designed for constrained problems, e.g., projected gradient approach [135–138]. In Publications I–VI the constrained objective function  $E$ , with various data fidelity and regularization terms, is optimized utilizing Spectral Projected Gradient (SPG) [39, 136]. This method has all the desirable properties of an optimization method applied on an imaging problem: it is deterministic, it quickly converges to a solution, it has low memory requirements, and, most importantly, it is flexible regarding the inclusion of different types of data fidelity and regularization terms into the energy function. The last element is of particular interest in this thesis since the energy function  $E$  considered here consist of different data and regularization terms. We have observed in Publications I–VI very good performance of SPG with respect to the all desired properties mentioned. For blind deconvolution (Publication II) and super-resolution coverage segmentation (Publication VI) where energy functions are highly non-convex, specific schemes for efficient optimization are proposed. These schemes include, as the first step, solving simpler subproblems (by SPG) and as the second, gradual increase of the complexity of the main problem.

## 5.2 Spectral Projected Gradient

Spectral Projected Gradient is a deterministic gradient based method for solving a constrained optimization problem

$$\min_{x \in \Theta} f(x), \quad (5.3)$$

where  $\Theta$  is a closed convex set in  $\mathbb{R}^n$  and  $f$  is a function which has continuous partial derivatives on an open set that contains  $\Theta$ . The method is briefly outlined in Algorithm 1.

Algorithm 1.

---

### Spectral Projected Gradient

---

Choose values for parameters:  $\theta_{min}, \theta_{max}, \gamma, \sigma_1, \sigma_2, tol$

s.t.  $0 < \theta_{min} < \theta_{max}, \gamma \in (0, 1), 0 < \sigma_1 < \sigma_2 < 1, tol > 0$ .

Choose initial guess  $x_0 \in \Theta$  and  $\theta_0 = 1$ .

Compute  $x_{k+1}$  and  $\theta_{k+1}$  as follows:

$$d_k = P_{\Theta}(x_k - \theta_k \nabla f(x_k)) - x_k$$

---

#### Grippo's non-monotone line search

---

$$f_{max} = \max\{f(x^{k-j}) \mid 0 \leq j \leq \min\{k, m-1\}\};$$

$$x_{k+1} = x_k + d_k$$

$$\delta = \nabla f(x_k)^T d_k; \xi_k = 1$$

$$\mathbf{while} \ f(x_{k+1}) > f_{max} + \gamma \xi_k \delta \quad (5.4)$$

$$\xi_{temp} = -\frac{1}{2} \xi_k^2 \delta / (f(x_{k+1}) - f(x_k) - \xi_k \delta) \quad (5.5)$$

$$\mathbf{if} \ (\xi_{temp} \geq \sigma_1 \wedge \xi_{temp} \leq \sigma_2 \xi_k) \quad (5.6)$$

$$\mathbf{then} \ \xi_k = \xi_{temp}$$

$$\mathbf{else} \ \xi_k = \xi_k / 2$$

$$x_{k+1} = x_k + \xi_k d_k$$

**end while**

---

#### Barzilai-Borwein step-length selection

---

$$s_k = x_{k+1} - x_k; y_k = \nabla f(x_{k+1}) - \nabla f(x_k); \beta_k = s_k^T y_k$$

**if**  $\beta_k \leq 0$

$$\mathbf{then} \ \theta_{k+1} = \theta_{max}$$

$$\mathbf{else} \ \theta_{k+1} = \min\{\theta_{max}, \max\{\theta_{min}, \frac{s_k^T s_k}{\beta_k}\}\} \quad (5.7)$$

$$\mathbf{Repeat until:} \ |x_{k+1} - x_k|_{\infty} \leq tol. \quad (5.8)$$


---

The algorithm combines the Projected Gradient method [139] with Grippo's type non-monotone line search [140] and the Barzilai-Borwein (also known as

Spectral Gradient) step-length [141, 142]. SPG uses a line search procedure based on the Armijo [143] sufficient decrease condition (5.4). To accelerate the line search process, the trial step-length,  $\xi_{temp}$ , is calculated as the minimal value of the quadratic interpolation (5.5) of the objective function along the line search direction.

The parameters  $\theta_{min}$ ,  $\theta_{max}$ ,  $\sigma_1$  and  $\sigma_2$ , have the role to keep  $\theta_k$  (5.7) and  $\xi_{temp}$  (5.6) within the given limits. Their recommended values from [136] (which are used in Publications I-VI) are  $\theta_{min} = 10^{-3}$ ,  $\theta_{max} = 10^3$ ,  $\sigma_1 = 0.1$ ,  $\sigma_2 = 0.9$ . The parameter  $\gamma$  controls the Armijo decrease condition (5.4), and its recommended value from [136] is  $\gamma = 10^{-4}$ . In all the methods presented in Publications I-VI, Algorithm 1 terminates when the max-norm between two consecutive images is less than  $tol = 10^{-5}$  (5.8) or when the number of iterations reaches 200. The parameter  $m$  regulates the number of stored previous objective function values used by the non-monotone Grippo's line search procedure in each iteration. For  $m = 1$ , we get monotone line search, which is used in Publications I-VI. For a given arbitrary initial solution  $x_0 \in \mathbb{R}^n$  and under assumptions that

- function  $f$  is defined and has continuous partial derivatives on an open set that contains  $\Theta$ ,
- the projection,  $P_\Theta(x)$  onto the set  $\Theta$  is defined for each  $x \in \mathbb{R}^n$ ,

the algorithm converges to a constrained stationary point (more details concerning convergence analysis can be found in [39]). The SPG algorithm is particularly suited for the situations when the projection calculation is inexpensive, as in box-constrained problems. SPG has been successfully applied for solving large-scale problems in many different engineering fields including image and signal processing. For its applications we refer to the review paper [142] and references therein.

Weak requirements on the objective function, as well as its efficiency in solving large scale problems, make this tool attractive for our purpose, i.e., for solving

$$\min E(x) \quad \text{s.t. } 0 \leq x_i \leq 1, \quad i = 1, 2, \dots, N, \quad (5.9)$$

where  $x$  equals to image  $u$  in non-blind deconvolution and super-resolution case, while in coverage segmentation case  $x$  equals  $A$  where  $A$  is a coverage segmentation matrix, and the energy function  $E$  includes various data fidelity and regularization terms introduced in Chapters 3 and 4.

We define the projection  $P_\Theta$  of a vector  $x \in \mathbb{R}^n$  to the feasible set  $\Theta = [0, 1]^n$  as:  $[P_\Theta(x)]_i = \min\{1, \max\{0, x_i\}\}$ , for all  $i = 1, 2, \dots, N$ . As SPG is an iterative gradient-

Algorithm 2.

---

**Blind Restoration by Alternating Minimization**

---

Start with initial guesses for  $u$  and  $h$ .

Having  $u_k$  and  $h_k$ , estimate  $u_{k+1}$  and  $h_{k+1}$ , alternately.

*Image estimation step:*

$$u_{k+1} = \arg \min_u E(u, h_k) \text{ s.t. } 0 \leq u_i \leq 1, \quad i = 1, 2, \dots, n,$$

*PSF estimation step:*

$$h_{k+1} = \arg \min_h E(u_{k+1}, h)$$

Impose constraints on  $h_{k+1}$ :  $h_{k+1} \geq 0$  and  $\|h_{k+1}\|_1 = 1$

Repeat until convergence.

---

based method, gradient of the objective function is needed. Derivation of the gradients of  $E$  for particular method can be found in Publications I-VI.

### 5.3 Alternating minimization scheme for blind deconvolution

The discussion so far has been related to non-blind deconvolution (Publication I, IV and V), super-resolution reconstruction (Publication III) and super-resolution coverage segmentation (Publication VI) methods which assume that point spread function (PSF), which causes degradation by blur, is known in advance, or accurately estimated. However, in many applications it is very difficult, or even impossible, to accurately estimate the PSF, and blind methods may be preferable. In Publication II we have presented a novel blind deconvolution method for images degraded by PG noise, based on energy minimization, which jointly estimates the original image and the PSF from the observed data. In the blind deconvolution task an energy function  $E$  is of the form:

$$E(u, h) = D(u, h; v) + \lambda R(u), \quad (5.10)$$

where  $u$  and  $h$  are the unknown image and the PSF, respectively, which have to be recovered from the observed image  $v$ .

In the presence of PG noise, the data term  $D$  considered in Publication II is given by (3.21) ( $S$  is identity matrix) and the regularization  $R$  is given by (4.4).

The energy function (5.10) is minimized by a so-called Alternating Minimization (AM) procedure, given in Algorithm 2, where the underlying image and PSF are estimated in separate steps by utilizing SPG optimization.



## 5.4. Graduate non-convexity for super-resolution coverage segmentation

Algorithm 3.

---

**Super-resolution coverage segmentation**

---

Parameters  $\mu, \nu_0, \eta, \rho \geq 0$ . Initial segmentation  $A_0 = \lfloor \frac{1}{m} \rfloor_{N_h, m}$ .  
 $A \leftarrow \operatorname{argmin}(D_{sr}(A; V) + \eta_0 R(A))$  by SPG initialized by  $A_0$   
**For**  $k = 1, 2, \dots$   
 $\eta = \eta_0 * k$   
 $A \leftarrow \operatorname{argmin}(D_{sr}(A; V) + \mu \operatorname{TV}_{HUB}(A) + \eta R(A))$  by SPG

---

**Repeat**  
 $A \leftarrow \operatorname{argmin}^*(D_{sr}(A; V) + \mu \operatorname{TV}_{HUB}(A) + \nu \tilde{F}(A) + \eta R(A))$  by SPG  
 $\nu \leftarrow \nu(1 + \rho)$   
**Until**  $\tilde{F}(A) \leq 0.1$   
 $\nu \leftarrow \nu \cdot 0.8$   
**Until**  $\|A_{k+1} - A_k\|_\infty \leq 0.01$

---

## 5.4 Graduate non-convexity for super-resolution coverage segmentation

Coverage segmentation  $\hat{A}$  is obtained by solving the optimization problem:

$$\hat{A} = \operatorname{argmin} E(A) \quad \text{s.t.} \quad \forall \alpha_{i,j} \in [0, 1]. \quad (5.11)$$

We optimize  $E$  by utilizing SPG.

The energy function  $E$  is, in all the versions (i)-(v) of  $E$ , introduced in Section 4.2.4, non-convex, and its minimization is far from trivial. It is addressed by gradually increasing the complexity of the problem and utilizing the solutions of numerically easier sub-problems as starting guesses in the optimization of the more difficult ones. The functions (i) and (ii) are optimized as in [38]. For optimization of (iii)-(v), the process is initiated by minimizing only the terms  $D_{sr}(A; V)$  and  $R(A)$ . Then the smoothing term ( $P(A)$ ,  $P_{HUB}(A)$  or  $\operatorname{TV}_{HUB}(A)$ ) is included, while setting the relative weight for  $\tilde{F}(A)$  to zero. Finally, the non-convex  $\tilde{F}(A)$  is included and alternately switched off and on, while the weight  $\nu$  is gradually increasing. The complete optimization procedure for (super-resolution) coverage segmentation (v) is presented in Algorithm 3. The SPG optimization in the innermost loop ( $\operatorname{argmin}^*$ ) is terminated after 10 iterations, whereas other SPG optimizations are run until convergence. The process terminates when the max-norm between two consecutive estimates stays below the tolerance of 0.01. With the appropriate smoothing term, Algorithm 3 is used for (iii) and (iv) too.



# 6 Performance evaluation

## 6.1 Performance evaluation of image restoration suited for PG noise

### 6.1.1 Quantitative evaluation

#### 6.1.1.1 Data set

A set of noise-free and blur-free images is used for the evaluation. Each image from the set is degraded in a controlled way and then restored using the considered method to evaluate. The data set consists of 30 blur-free and noise-free test images (Fig. 6.1). Each test image  $u^*$  is degraded by blur with Gaussian PSFs  $H$  with 11 different standard deviations  $\sigma_p$  between 0 and 5 pixels and in that way the blurred observations  $Hu^*$  are obtained.

For the evaluation of super-resolution method the blurred observations are

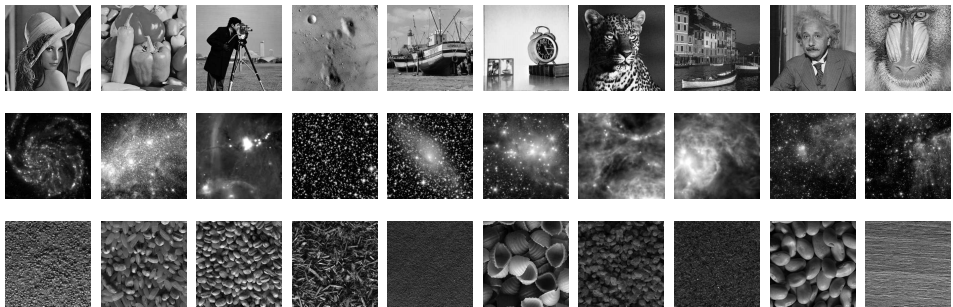


Figure 6.1: Used test images: 10 "classic" test images, 10 astronomical, and 10 texture images. All the images are  $256 \times 256$  pixels.

down-sampled with sampling factor  $d = 2 \times 2$  (by averaging pixels in  $2 \times 2$  blocks), while for the evaluation of deconvolution methods down-sampling of blurred images is not performed (matrix  $S$  is identity matrix in this case). Each (down-sampled) blurred image is further corrupted by PG noise, according to (3.10).

The images with 11 maximal pixel intensities from 100 to 10000 are considered, and in that way 11 different levels of Poisson noise are observed. For a given level of Gaussian noise the variance is chosen such that the ratio of standard deviation of Gaussian and Poisson noise,  $\sigma_m / \sqrt{peak}$ , is  $\{0.01, 0.1, 1\}$ . In this way,  $(30 \times 11 \times 11) \times 3 = 10890$  different blurred and noisy LR observations are obtained.

In some experiments the whole set of 10890 images is used while in others only a suitable subset is used.

### 6.1.1.2 Quality measures

Each degraded image used in the evaluation of specific method is restored by the considered method, and the restored image is compared with the original blur-free and noise-free image in order to see how close these two images are. The quality of the restoration (closeness between restored and original images) is measured by Peak-Signal-to-Noise Ratio (PSNR):

$$\text{PSNR} = 10 \log_{10} \left( \frac{(\max(u_i^*))^2}{\text{MSE}} \right), \quad (6.1)$$

where

$$\text{MSE} = \frac{1}{N_h} \sum_{i=1}^{N_h} (u_i^* - \hat{u}_i)^2. \quad (6.2)$$

The original and the restored images are denoted by  $u^*$  and  $\hat{u}$ , respectively, and the restored image  $\hat{u}$  is given by (1.3).

In some experiments Signal-to-Noise Ratio (SNR) is considered,

$$\text{SNR} = 10 \log_{10} \left( \frac{\frac{1}{N_h} \sum_{i=1}^{N_h} u_i^{*2}}{\text{MSE}} \right). \quad (6.3)$$

Mean squared error (MSE) measures (averaged squared) pixel-wise difference. PSNR (SNR) is inverse proportional to MSE. The larger (the smaller) PSNR (MSE) is, the closer the restored and the original images are and the higher-quality restoration is achieved. Both MSE and PSNR are easy to compute and they are commonly utilized as quality measures.

## 6.1. Performance evaluation of image restoration suited for PG noise

In general, for a given method, *the improvement in PSNR* between the noisy input and the restored output images:  $\Delta\text{PSNR}_{\text{method}} = \text{PSNR}_{\text{out}} - \text{PSNR}_{\text{in}}$  is quantified. Positive  $\Delta\text{PSNR}$  indicates that the resulting restored image is with a higher PSNR, i.e., of a higher quality, than the initial, degraded one. Different approaches are ranked by comparing the improvements in PSNR achieved using them, observing the difference in improvements reached; a positive value of  $\Delta\text{PSNR}_{\text{method1}} - \Delta\text{PSNR}_{\text{method2}}$  indicates that "method1" outperforms "method2".

The main limitation of the mentioned metrics (MSE, PSNR, SNR) is that they rely strictly on direct pixel-wise comparison of intensities and do not actually take into account any biological factors of the human vision system. For this reason, in some experiments the Structural Similarity Index Measure (SSIM) [144] is considered. It measures the similarity between two images in a manner that is more consistent with human perception.

SSIM is a measure of similarity between two images, which considers three characteristics - luminance, contrast and structure over patches of the two images. If  $u_p^*$  and  $\hat{u}_p$  are two patches extracted from an original and a restored image, respectively, then SSIM is defined as a product of luminance, contrast and structure comparison functions, each of them with relative weights  $\alpha, \beta, \gamma > 0$ :

$$\text{SSIM}(u_p^*, \hat{u}_p) = \left( \frac{2\mu_{u_p^*}\mu_{\hat{u}_p} + C_1}{\mu_{u_p^*}^2 + \mu_{\hat{u}_p}^2 + C_1} \right)^\alpha \left( \frac{2\sigma_{u_p^*}\sigma_{\hat{u}_p} + C_2}{\sigma_{u_p^*}^2 + \sigma_{\hat{u}_p}^2 + C_2} \right)^\beta \left( \frac{\sigma_{u_p^*\hat{u}_p} + C_3}{\sigma_{u_p^*}\sigma_{\hat{u}_p} + C_3} \right)^\gamma \quad (6.4)$$

where  $\mu_{u_p^*}$  and  $\mu_{\hat{u}_p}$  are mean values of  $u_p^*$  and  $\hat{u}_p$ ,  $\sigma_{u_p^*}$  and  $\sigma_{\hat{u}_p}$  are variances of  $u_p^*$  and  $\hat{u}_p$  while  $\sigma_{u_p^*\hat{u}_p}$  represents the correlation of  $u_p^*$  and  $\hat{u}_p$ . The role of constants  $C_i$ ,  $i \in \{1, 2, 3\}$  is to ensure stability by preventing division with zero. If  $\alpha = \beta = \gamma = 1$  and  $C_3 = C_2/2$ , then SSIM reduces to

$$\text{SSIM}(u_p^*, \hat{u}_p) = \left( \frac{2\mu_{u_p^*}\mu_{\hat{u}_p} + C_1}{\mu_{u_p^*}^2 + \mu_{\hat{u}_p}^2 + C_1} \right) \left( \frac{2\sigma_{u_p^*}\sigma_{\hat{u}_p} + C_2}{\sigma_{u_p^*}^2 + \sigma_{\hat{u}_p}^2 + C_2} \right) \quad (6.5)$$

and this simplified version is used in the thesis, as suggested in [144]. The SSIM between the original and the restored images is obtained by averaging the SSIMs between local sliding windows which move over the entire images. Notice that  $\text{SSIM} \leq 1$ .

Human visual system is highly adapted for extracting structural information from a scene. As SSIM compares structural changes in images imitating what human visual system does, it is nowadays a widely utilized measure of visual closeness between images.

To compare speed of the methods proposed within this thesis in a fair way, all restoration and segmentation methods are implemented in Matlab, and all experiments are executed on same computer Intel Core i7 CPU 3.40 GHz. Speed of each method is given in Appendix.

### 6.1.1.3 Optimal parameter selection

To present and compare the performances of restoration methods in a fair way, it is important to select optimal regularization parameters ( $\lambda$  for methods using TV potential and data terms suited for Gaussian/Poisson-Gaussian noise and  $(\lambda, \omega)$  for methods with other potential functions/Huber or Lorentzian data terms). Numerous parameter selection schemes are proposed in literature, e.g., SURE-based approaches [145, 146] suited for restoration in the presence of Gaussian noise, L-curve [147], generalized cross validation, discrepancy principle [148], residual based methods [149, 150], method based on no-reference measure of image content [151] which can be used for parameter estimation in the presence of blur and not-necessarily Gaussian noise. All these methods are used in applications where original image is not available. They rely on different approximations of measures like e.g., MSE, and as such they are not perfectly accurate although they are still useful when reference image is not available. Since ground truth is available in our experiments, the optimal parameters are selected empirically, and in that way the risk of bias from relying on possibly imperfect estimates from methods for parameter selection is avoided. The best performing parameters are selected for each image and each observed degradation as an argument that maximizes its  $\Delta$ PSNR. For this parameter optimization Nelder-Mead simplex search [152] is utilized. In Section 6.1.4.3 we discuss sensitivity of the best performing proposed methods w.r.t. to regularization parameters.

### 6.1.2 Importance of an appropriate treatment of Poisson-Gaussian noise

In our deconvolution (Publication I-II) and super-resolution reconstruction (Publication III) studies, we argue for an appropriate treatment of signal-dependent mixed PG noise. A novel data term (3.21), adjusted to this type of noise is suggested. To evaluate reliably and in an unbiased way its performance in comparison to other data terms, all the observed energy functions utilize TV regularization and all are minimized by SPG.

To quantify the importance of appropriate treatment of signal-dependent noise in image deconvolution, in Publication I the images degraded by mixed Poisson-Gaussian noise are restored by using deconvolution methods suited for (i) PG

## 6.1. Performance evaluation of image restoration suited for PG noise

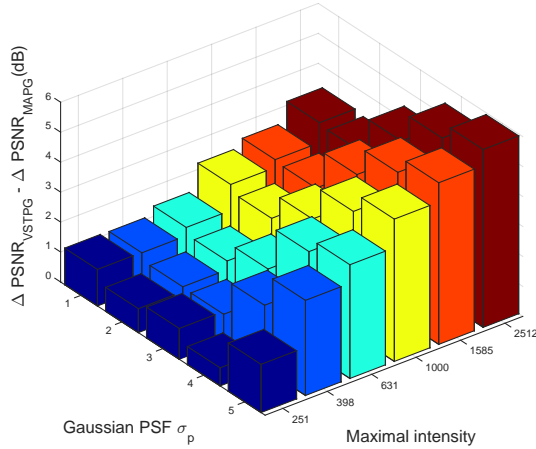


Figure 6.2: Difference in improvements in PSNR achieved due to appropriate treatment of mixed Poisson-Gaussian noise in the image deconvolution task. Improvement achieved by VSTPG is up to 6 dB higher than by MAPG.

noise (VSTPG), method with data fidelity term (3.21), and (ii) Gaussian noise (MAPG), method with quadratic data fidelity term (3.5).

The comparison of performances of VSTPG and MAPG (with TV potential) on the first 5 images in Fig. 6.1, degraded by  $4 \times 6$  different blur and PG noise levels, is presented in Fig. 6.2. It is clear that the use of an appropriate noise model is very important in image deconvolution. A consistent additional improvement in PSNR when assuming the correct noise model (indicated by positive difference in improvements achieved by VSTPG and MAPG in the plot), which goes up to 6 dB and reaches on average 2.42 dB can be observed.

To quantify the importance of appropriate treatment of signal-dependent noise in super-resolution image reconstruction, in Publication III performances of the following methods are compared:

- (i) SR-PG – the method suited for mixed PG noise with data fidelity term (3.21),
- (ii) SR-G – method suited for Gaussian noise with quadratic data fidelity term (3.5),
- (iii) SR-HUB – method with Huber data fidelity term (3.14),
- (iv) SR-LOR – method with Lorentzian data term (3.16).

A consistent additional improvement in PSNR when using the proposed SR-PG method (i) instead of any of the methods (ii)-(iv) is observed. The achieved over-

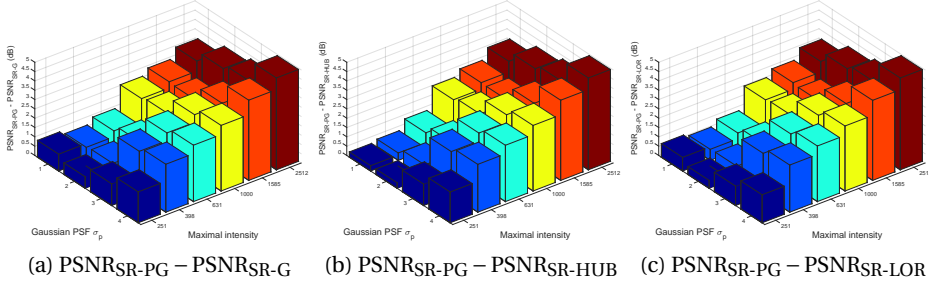


Figure 6.3: Average improvement in PSNR (dB), on the observed test set for  $\sigma_m / \sqrt{peak} = 0.01$ , achieved due to appropriate treatment of mixed Poisson-Gaussian noise, for different restoration approaches, and different levels of blur and noise (scale factor  $d = 2 \times 2$ ) in the super-resolution task. The proposed SR-PG method outperforms the other methods by up to 5 dB.

all average improvement (over 360 images) in PSNR over the three considered Gaussian-Poisson noise ratios is 2.36 dB.

The difference in the performance of SR-PG, and each of SR-G, SR-HUB and SR-LOR, on the subset of the test set with the ratio  $\sigma_m / \sqrt{peak} = 0.01$ , is presented in Fig. 6.3. It is clear that the use of the correct noise model largely improves the quality of the restored images. For this noise ratio on average (over 120 degraded images) 2.41 dB better performance when using method (i) than when using (ii)-(iv) is observed. The results for the other two considered ratios of Poisson-Gaussian noise,  $\sigma_m / \sqrt{peak} = 0.1$  and  $\sigma_m / \sqrt{peak} = 1$ , exhibit similar behavior (with average improvement of 2.34 dB and 2.34 dB, respectively).

An illustrative example of SR reconstruction with  $d = 3 \times 3$  times increased resolution is given in Fig. 6.4; a test image degraded by blur and PG noise (shown to scale in (a)) is restored by our proposed method (result shown in (b)), by methods suited for Gaussian noise (shown in (c)) and by methods using alternative data terms (shown in (d) SR-HUB and (e) SR-LOR). The PSNR reached, as well as the SSIM, are presented below each image. The proposed SR-PG method consistently reaches the best quantitative results, and subjective visual restoration performance also speaks in favour of the proposed method.

### 6.1.3 Comparison of MAP and VST approaches

In order to demonstrate that VST approach utilized in our work for PG noise treatment (3.21) does not fall behind the direct MAP approach (3.12), we have compared in Publication I the performances of the two approaches for deconv-



## 6.1. Performance evaluation of image restoration suited for PG noise



(a) Blurred and noisy LR  $85 \times 85$  image



(b) SR-PG (3.21)  
**PSNR=25.41 dB, SSIM=0.5143**  
 $(\lambda = 4.6 \cdot 10^{-4})$



(c) SR-G (3.5)  
 PSNR=21.95 dB, SSIM=0.5081  
 $(\lambda = 1.4 \cdot 10^{-3})$



(d) SR-HUB (3.14)  
 PSNR=21.97 dB, SSIM=0.5110  
 $(\lambda = 2.3 \cdot 10^{-2}, \omega = 6.4 \cdot 10^{-2})$



(e) SR-LOR (3.16)  
 PSNR=21.97 dB, SSIM=0.5118  
 $(\lambda = 2.4 \cdot 10^{-2}, T = 2.3 \cdot 10^{-1})$

Figure 6.4: SR reconstruction of a test image, degraded by blur and PG noise. Original image (Fig. 6.1) of size  $255 \times 255$  is degraded by Gaussian blur ( $\sigma_p = 1$ ), down sampled with sampling factor  $d = 3 \times 3$  (by averaging pixels in  $3 \times 3$  blocks) and corrupted by PG noise with  $peak = 1585$ ,  $\sigma_m / \sqrt{peak} = 0.01$ . Such degraded image is shown in (a). Restored HR images (b)-(e) are all  $255 \times 255$ . Below each restored image, the reached PSNR and SSIM values are given together with optimal parameters (which maximize PSNR) for each method on this particular image.

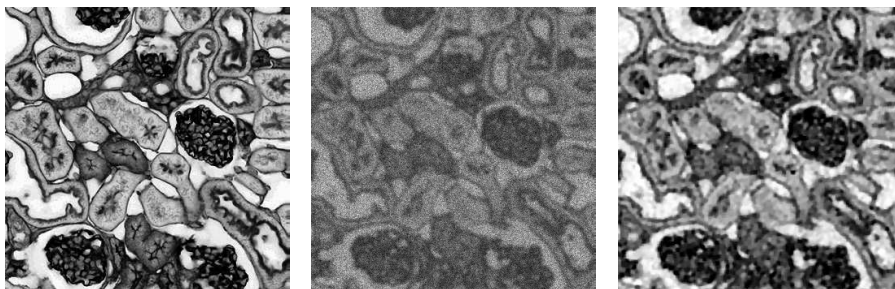


Figure 6.5: Performance of VSTPG: (a) Original image. (b) Degraded image (SNR=7.64 dB). (c) Restored image (SNR=13.87 dB).

Table 6.1: Comparison of VSTPG and MAPPG [23] for four test images used in [153]. VSTPG outperforms MAPPG in both SNR and computational time.

		VSTPG	MAPPG [23]
First image ( $350 \times 350$ ), $peak = 20$ PSF: Uniform $5 \times 5$ , $\sigma_m^2 = 9$ SNR=7.64 dB, SSIM=0.749	SNR (dB)	<b>13.87</b>	13.73
	SSIM	<b>0.934</b>	0.933
	Time (s.)	4.35	48587
Second image ( $257 \times 256$ ), $peak = 60$ PSF: Gaussian $9 \times 9$ , std 0.5, $\sigma_m^2 = 36$ SNR=9.40 dB, SSIM=0.646	SNR (dB)	<b>15.55</b>	15.43
	SSIM	<b>0.888</b>	0.880
	Time (s.)	2.43	351
Third image ( $256 \times 256$ ), $peak = 100$ PSF: Uniform $3 \times 3$ , $\sigma_m^2 = 36$ SNR=10.68 dB, SSIM=0.684	SNR (dB)	<b>14.25</b>	13.81
	SSIM	<b>0.851</b>	0.847
	Time (s.)	2.30	8322
Fourth image ( $256 \times 256$ ), $peak = 150$ PSF: Gaussian $7 \times 7$ , std 1, $\sigma_m^2 = 40$ SNR= 15.77 dB, SSIM=0.643	SNR (dB)	<b>20.57</b>	20.33
	SSIM	<b>0.875</b>	0.870
	Time (s.)	2.27	43397

lution of images degraded by PG noise.

For PG noise, the MAP approach is of much higher computational complexity than VST approach. To make a comparison of MAPPG and VSTPG feasible, their performances on a small dataset given in [153] are observed and the results of MAPPG stated there by the authors are reported here.

The data set contains four images, the first of them is shown in Fig. 6.5(a). Each test image is degraded by one PSF and one Poisson-Gaussian noise level (corresponding to very low photon count); details are given in Table 6.1. Fig. 6.5(b) shows the result of the described degradation applied to Fig. 6.5(a). VSTPG with TV potential is applied for their deconvolution. Table 6.1 presents the results. Presented deconvolution results for VSTPG in Table 6.1 are obtained with regu-

## 6.1. Performance evaluation of image restoration suited for PG noise

larization parameters which optimize SNR value. The SSIM value is calculated for the restored image obtained with parameters which maximize SNR. The SNR and SSIM values of the deconvolution results achieved by utilizing MAPPG approach [23] with TV potential, included in Table 6.1, are taken from [153].

As can be seen, the VSTPG method outperforms MAPPG in terms of SNR, SSIM and computational time. Deconvolution of one image by MAPPG takes up to 13.5h, whereas the VSTPG methods takes 4.35s on the same image and reaches a better restoration quality. Fig. 6.5(c) presents the result of restoration of Fig. 6.5(b) by VSTPG approach. The images restored by MAPPG can be found in [153].

### 6.1.4 Evaluation of different potential functions

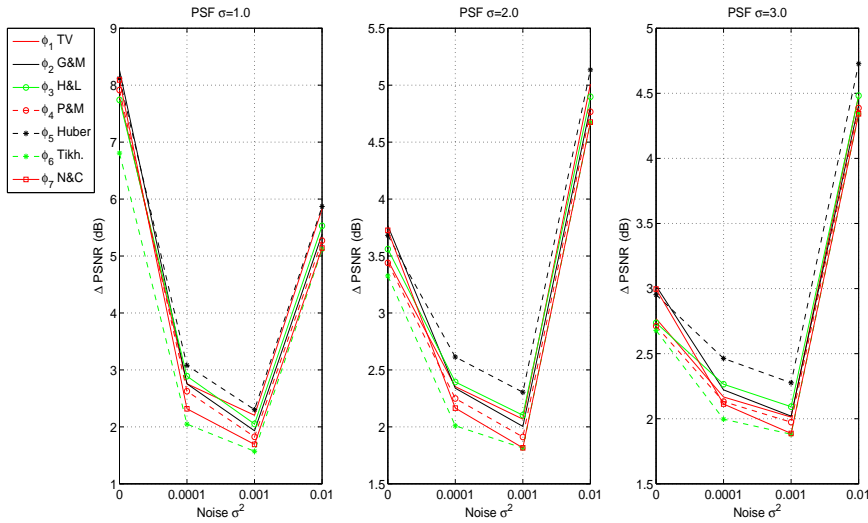
Further improvements in the considered image restoration tasks can be achieved if edge preserving potentials are utilized instead of classical TV regularization.

#### 6.1.4.1 Deconvolution and denoising in presence of Gaussian noise

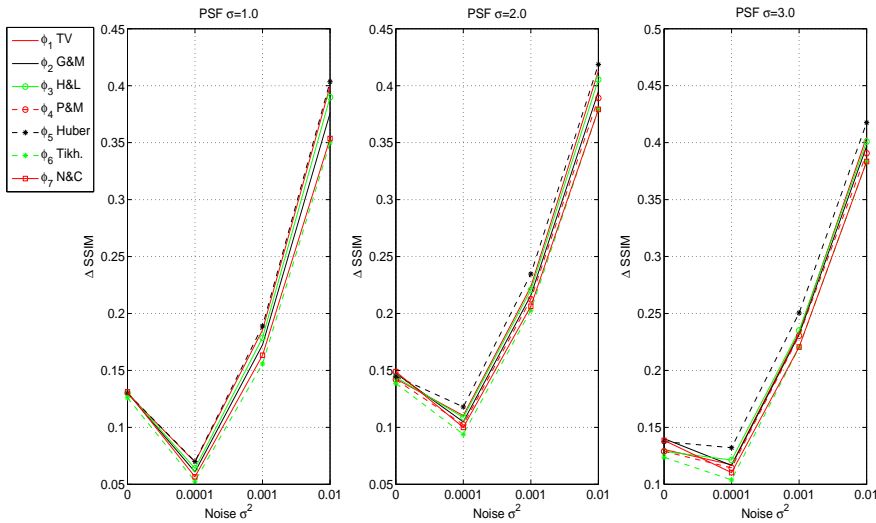
Initially the performances of seven different potential functions given in Table 4.1 are tested on images degraded by Gaussian noise and blur (Publication IV). To evaluate the performance of different potentials, first ten standard images shown in Fig. 6.1 are utilized. For every original image  $u^*$  a noisy and blurred image  $v$  is constructed by convolving it with 3 levels of PSF and adding 4 levels of white Gaussian noise according to (3.1). For each PSF and noise level, one degraded image  $v$  is obtained from which  $\hat{u}$  is restored using the data term (3.5) (with identity matrix  $S$ ) and the regularization term (4.4) with seven considered potentials from Table 4.1.

The improvement in PSNR and in SSIM for each of the seven potentials, and each of the  $3 \times 4$  blur and noise levels, is presented in Fig. 6.6. A very clear result is that the Huber potential,  $\Phi_5$ , shows superior performance in all of the evaluated settings when the quality of restoration is measured by both PSNR and SSIM. As a second runner-up comes TV based deconvolution ( $\Phi_1$ ), clearly behind in most situations. The fact that a consistent ranking of the considered potential functions for both of the used quality measures is obtained gives additional strength to our conclusion regarding the outstanding performance of the Huber potential function.

In Fig. 6.7 a zoomed-in view on the shoulder of the Cameraman is shown, to highlight the edge preservation performance of the Huber potential over the commonly used TV regularization. It is apparent that the Huber potential does



(a) PSNR improvement



(b) SSIM improvement

Figure 6.6: Average improvement of PSNR (a) and SSIM (b) over ten test images for different PSFs and different Gaussian noise levels considering all seven evaluated potentials in image deconvolution task. Huber potential outperforms other potential functions in terms of both PSNR and SSIM.

## 6.1. Performance evaluation of image restoration suited for PG noise

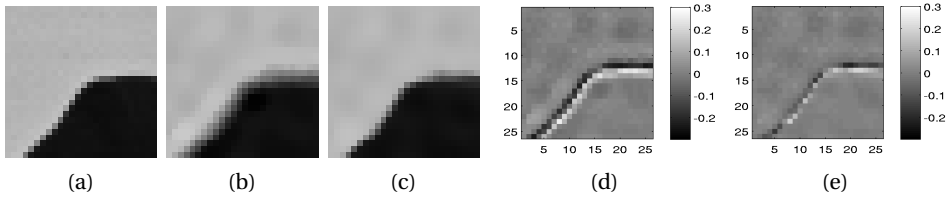


Figure 6.7: Illustration of improved edge preservation by Huber potential,  $\Phi_5$ . (a) Original image, part of Cameraman's shoulder. (b) Restored image using  $\Phi_1$  (TV). (c) Restored image using  $\Phi_5$ . (d) Residual for  $\Phi_1$ . (e) Residual for  $\Phi_5$ .

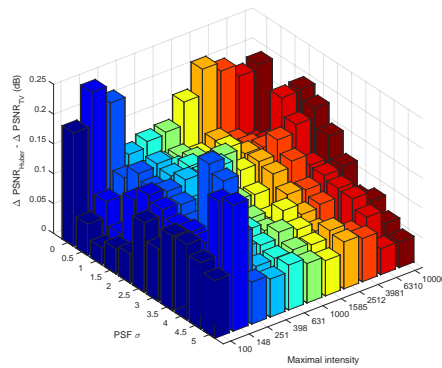


Figure 6.8: Difference in improvements achieved by VSTPG deconvolution method utilizing Huber potential, compared to TV, is shown.

a much better job in preserving the sharp edges in the image (as can also be confirmed by looking at the residual errors in Fig. 6.7(d,e)).

### 6.1.4.2 Deconvolution and denoising in presence of Poisson-Gaussian noise

Improvement achieved by the best performing potential function, Huber, utilized in deconvolution of images in presence of signal dependent PG noise is evaluated in Publication I. The difference in PSNR improvement, averaged over 30 images, obtained by using the TV and Huber potentials is evaluated.

Fig. 6.8 presents  $\Delta \text{PSNR}_{\text{Huber}} - \Delta \text{PSNR}_{\text{TV}}$  for VSTPG method for mixed PG noise with the lowest considered Gaussian noise level  $\sigma_m / \sqrt{\text{peak}} = 0.01$ . Other two Gaussian noise levels exhibit similar behaviour. The Huber potential gives on average 7.8%, 6.8% and 4.8% greater improvement in restoration quality as compared to the TV regularization for the considered Gaussian noise levels.

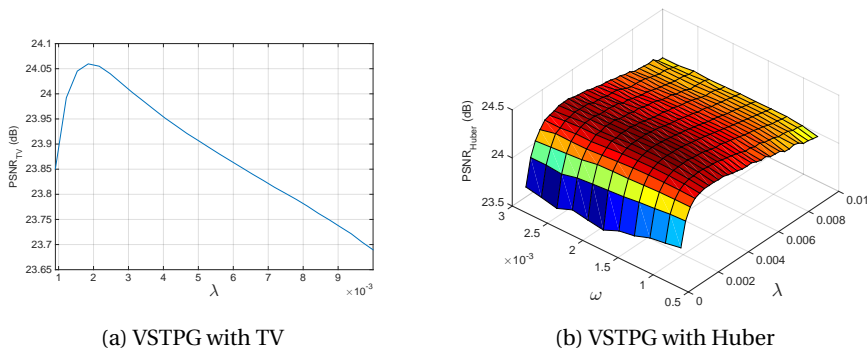


Figure 6.9: Relation between regularization parameter  $\lambda$  for TV and Huber potentials on *Peppers* image (the second in Fig. 6.1) degraded by PSF with  $\sigma_p = 1.5$  and Poisson noise with  $peak = 398$  (no Gaussian noise), when VSTPG is used for restoration. Optimal parameter  $\lambda_{\text{Huber}} = 0.0032$  is slightly larger than  $\lambda_{\text{TV}} = 0.0018$ . Quality of restoration is less sensitive to choice of  $\omega$  for the Huber potential (optimal  $\omega = 0.0016$ ).

### 6.1.4.3 On optimal parameter selection for Huber potential

The performance improvement from the Huber potential comes at the cost of one more parameter to tune. With a good optimization strategy, this cost can be kept reasonably low. Although the optimal values of  $\lambda$  and  $\omega$  vary between images, we observe some consistency. The optimal value of  $\lambda$  is generally slightly larger for Huber potential than for TV; on average, for all methods and all degraded images, optimal  $\lambda_{\text{Huber}}$  is 1.5 times larger than  $\lambda_{\text{TV}}$ . A typical example is shown in Fig. 6.9, where optimal  $\lambda_{\text{TV}} = 0.0018$  and  $\lambda_{\text{Huber}} = 0.0032$ . If a good value of  $\lambda_{\text{TV}}$  is known, this provides a good starting guess for  $\lambda_{\text{Huber}}$ . The additional parameter,  $\omega$ , controls the point of transition between  $\ell_2$ , Tikhonov regularization, and  $\ell_1$ , TV regularization. Choosing a too small value of  $\omega$  makes the Huber potential approach TV, with the above observed reduced performance. Selecting a too large value, however, leads to quadratic regularization which gives a rapid decay in performance and blurred edges as a result.

We find that a good initialization for the 2D search for optimal parameters for the Huber potential is the optimal  $\lambda$  from the 1D search for the TV potential, combined with a small value of  $\omega$  (e.g.  $10^{-5}$ , giving a behaviour of the Huber potential which is very similar to TV). Considering the logarithm of  $\lambda$  and  $\omega$ , as optimization variables, instead of  $\lambda$  and  $\omega$ , can help in addressing problems related to the difference in scale of these parameters.

6.1.5 Blind deconvolution in presence of Poisson-Gaussian noise

The performance of blind deconvolution method defined in Section 5.3 is evaluated on first five test images presented in Fig. 6.1 degraded with different levels of blur and PG noise. The obtained improvement in PSNR is compared with the improvement in PSNR achieved utilizing a non-blind approach. The average improvement over the five images in PSNR obtained with the proposed blind method for different Gaussian PSF and Poisson noise levels and for  $\sigma_m/\sqrt{peak} = 0.01$  is presented in Fig. 6.10a.

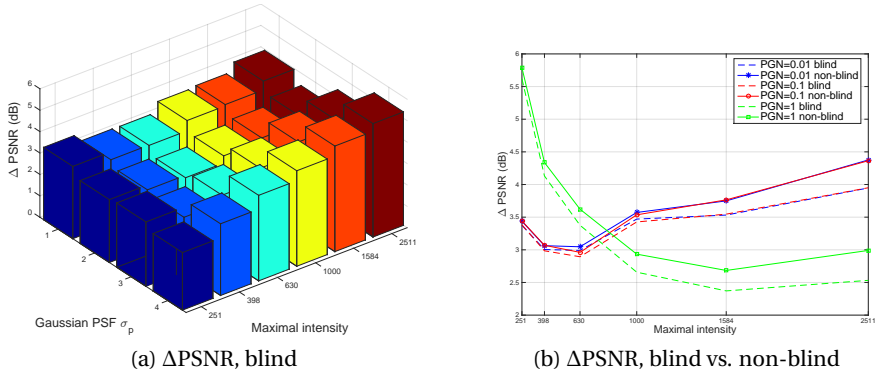


Figure 6.10: Performances of the proposed blind deconvolution method suited for PG noise. (a) The average PSNR improvement ( $\Delta$ PSNR) for blind deconvolution method for  $\sigma_m/\sqrt{peak} = 0.01$ ; (b) The average improvement in PSNR achieved with blind and non-blind (using the correct PSF) methods for Gaussian PSF with  $\sigma_p = 1$  and three PG noise (PGN) levels for  $\sigma_m/\sqrt{peak} = \{0.01, 0.1, 1\}$ .

Gaussian PSF		Motion PSF		Disk PSF	
Observed	Restored	Observed	Restored	Observed	Restored
21.41 dB	25.39 dB	22.58 dB	25.92 dB	22.40 dB	26.24 dB

Figure 6.11: Blind deconvolution of blurred and noisy images. Image Lena is convolved by  $13 \times 13$  Gaussian PSF ( $\sigma_p = 2$ ), motion blur of length 7 and angle  $\pi/4$ , and disk with radius 3, respectively, and PG noise with  $peak = 1500$ ,  $\sigma_m/\sqrt{peak} = 0.01$  is added. Reached PSNR are given. Applied and estimated PSFs are also shown.

An illustrative example presented in Fig. 6.11 shows that the proposed method reduces blur and suppresses noise in the degraded images and leads to good restoration results. Moreover, the comparison of performances of blind and non-blind method, shown in Fig. 6.10b, confirms that the results obtained by the proposed blind deconvolution method are comparable with those obtained when the blur (PSF) is known a priori.

## 6.2 Performance evaluation of super-resolution coverage segmentation

### 6.2.1 Synthetic tests

The performance of the energy functions (i)-(v) defined in Section 4.2.4, with the corresponding optimization schemes as described in Section 5.4, is evaluated on a synthetic two-channel (Red-Green) test image (Fig. 6.12(a)), which enables a direct comparison with correct coverage values. The test image of size  $100 \times 100 \times 2$  contains three geometric objects with both smooth and non-smooth boundaries. The color of the square is a linear mixture of the colors of the star and the disk (50% each). This makes it challenging to unmix the pixels on the boundary between the star and the disk, while interpreting the square as a separate pure component. By positioning the objects at different locations and in different rotations, six instances of test image are created and digitized. The coverage values are estimated by  $8 \times$  down sampling (8-sampled coverage digitization, [154]).

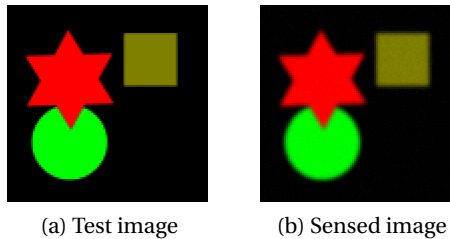


Figure 6.12: (a) Test image of size  $100 \times 100 \times 2$ ; (b) Degraded test image (a) by Gaussian blur with  $\sigma_p = 1$  and Gaussian noise with  $\sigma_n = 10^{-2}$ .

Qualitative evaluation of the observed methods in the presence of blur and noise can be made based on the results presented in Fig. 6.13. The coverage segmentations of the image shown in Fig. 6.12(b), which is a degraded version of Fig. 6.12(a), corrupted by Gaussian blur with  $\sigma_p = 1$  and Gaussian noise with



## 6.2. Performance evaluation of super-resolution coverage segmentation

$\sigma_n = 10^{-2}$ , obtained by each of the methods (i)-(v) introduced in Section 4.2.4, are presented. Each segmentation results in four coverage maps, corresponding to three objects and the background. All the maps are generated at the spatial resolution of the initial image (i.e.,  $d = 1 \times 1$  for the methods (iii)-(v)).

It can be observed that (i) fails to properly segment the square (considering its color partly a mixture of red and green, and only partly a “pure” color). While (ii) is more successful with unmixing, it does not handle well the blurred input and produces objects with highly non-smooth boundaries. Method (iii) deals with blur, but still fails with unmixing. The problem arises from the difficulty to optimize the non-convex and non-smooth perimeter term  $P$ . The Huberized perimeter term,  $P_{HUB}$ , in (iv) leads to improved, but still not perfect, segmentation. Finally, (v) utilizes  $TV_{HUB}$  and produces both a good separation of the classes and thin object boundaries with preserved sharp corners.

The robustness of methods (i)-(v) in the presence of noise and blur is assessed by observing their performance on a set of images created by degrading a synthetic test image (Fig. 6.12(a)) by increasing levels of blur and noise. The images are convolved with a Gaussian PSF with standard deviations  $\sigma_p \in \{0, 1, 2, 3, 4\}$ , and Gaussian noise with standard deviation  $\sigma_n \in \{0, 10^{-3}, 3.2 \cdot 10^{-3}, 10^{-2}, 3.2 \cdot 10^{-2}\}$  is added to the six instances of the observed test image (created as aforementioned). Each of the five segmentation methods is applied to each of the  $5 \times 5 \times 6$  images.

The performances of the methods are measured by calculating Mean Absolute Error (MAE) of the estimated coverage values:

$$\text{MAE}(A, \hat{A}) = \frac{1}{N_h m} \sum_{i=1}^{N_h} \sum_{j=1}^m |\alpha_{i,j} - \hat{\alpha}_{i,j}|, \quad (6.6)$$

where  $\alpha_{i,j}$  and  $\hat{\alpha}_{i,j}$  are true resp. estimated coverage of the  $i$ -th pixel by class  $S_j$ .

Fig. 6.14, left and center, show MAE for the five observed coverage segmentation methods for increasing levels of blur (at a fixed noise level  $\sigma_n = 10^{-2}$ ), and increasing levels of noise (blur level  $\sigma_p = 2$ ), respectively. For comparison purposes, the MAE of a theoretically optimal crisp segmentation of the corresponding *noise-free and blur-free* test image is also plotted. Clearly, SRCS-HTV $\tilde{F}$  exhibits the best performance, with the lowest average MAE, and very low variation, outperforming the ideal crisp segmentation. Consistent behaviour is observed for other combinations of blur and noise.

Methods (iii)-(v) enable segmentation at an increased spatial resolution. Their performance on images degraded by blur ( $\sigma_p = 1$ ) and noise ( $\sigma_n = 10^{-3}$ ), for increasing scale factor  $d$ , is shown in Fig. 6.14, right. The images of the spatial

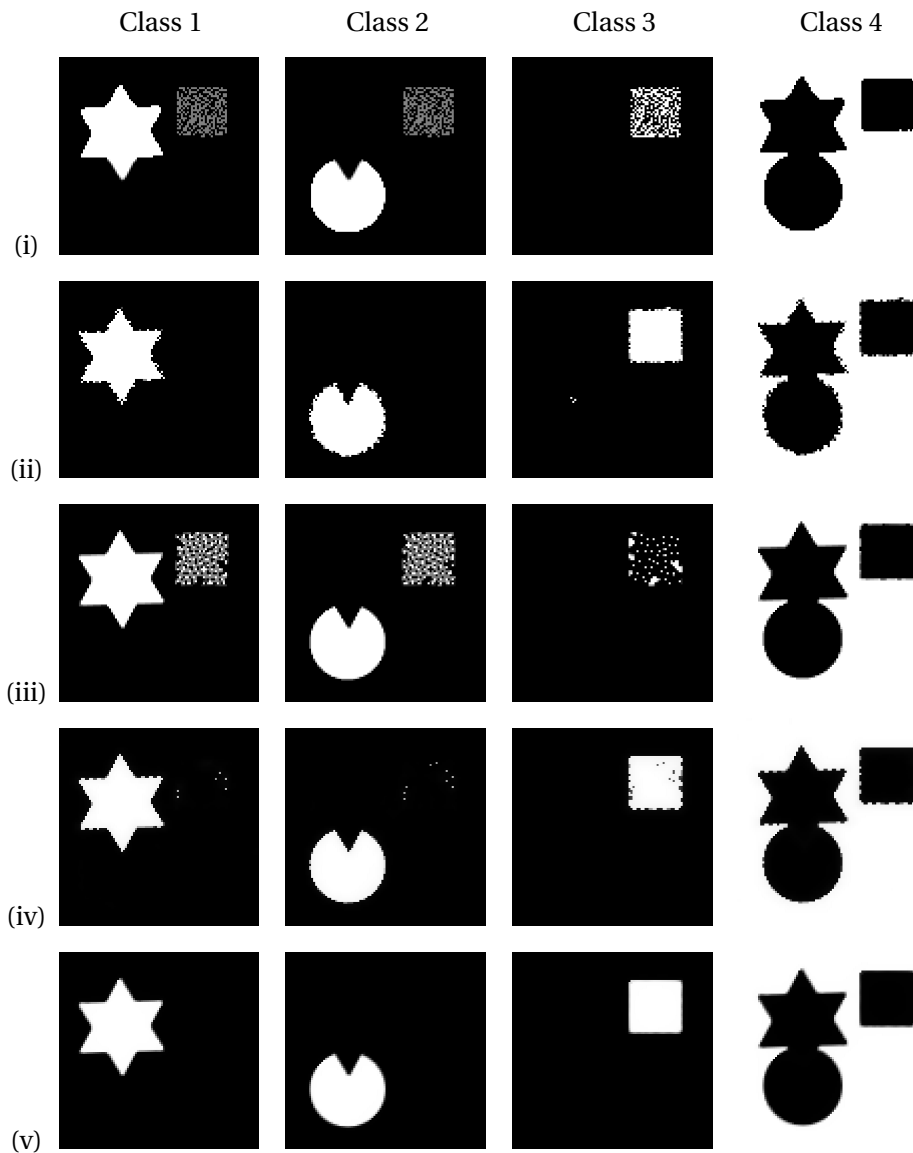


Figure 6.13: Segmentation of the image in Fig. 6.12(b). Coverage maps obtained with: (i) CS-PTE, (ii) CS-PF, (iii) SRCS-PF, (iv) SRCS-HPF, (v) SRCS-HTVF.

## 6.2. Performance evaluation of super-resolution coverage segmentation

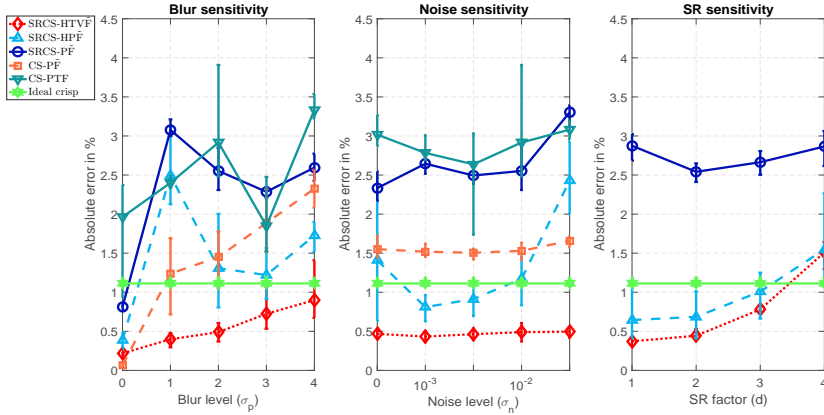


Figure 6.14: Performances of 5 considered methods for increasing levels of noise, blur and super-resolution factors measured by average MAE. Lines show averages for 6 observations and bars indicate max and min errors; *noise-free and blur-free* crisp segmentation is included. SRCS-HTVF exhibits superior performance.

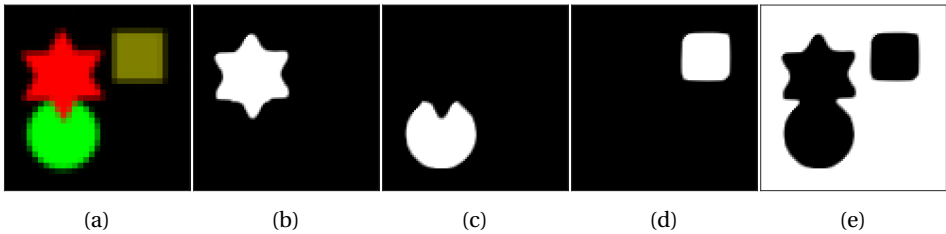


Figure 6.15: Performance of SRCS-HTVF at 3 times increased resolution. The test image of size  $99 \times 99$  from Fig. 6.13(a) is degraded by Gaussian blur ( $\sigma_p = 1$ ), down-sampled by SR factor  $d = 3 \times 3$  (by averaging pixels in  $3 \times 3$  blocks) and corrupted by Gaussian noise ( $\sigma_n = 10^{-3}$ ). Such degraded image of size  $33 \times 33$  is shown in (a). Segmented components (b)-(e) are all of size  $99 \times 99$ .

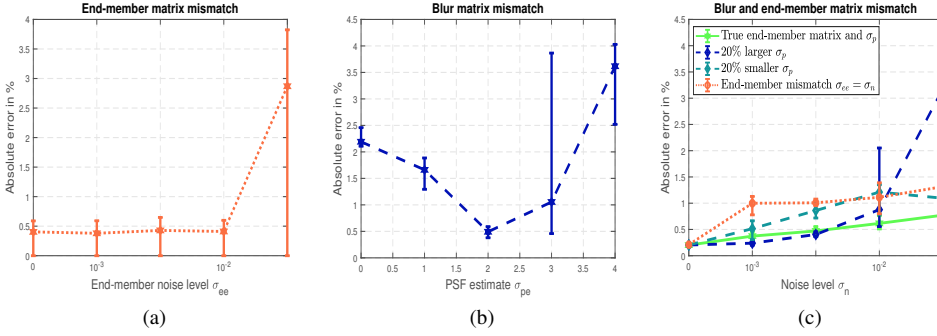


Figure 6.16: (a) Robustness of SRCS-HTV $\tilde{F}$  w.r.t. to end-member matrix mismatch for different end-member noise levels ( $\sigma_{ee}$ ) and  $\sigma_p = 2$  and  $\sigma_n = 10^{-2}$ . (b) Robustness of SRCS-HTV $\tilde{F}$  w.r.t. to blur matrix mismatch: performance for different used PSF estimates ( $\sigma_{pe}$ ) for real PSF  $\sigma_p = 2$  and  $\sigma_n = 0.01$ . (c) Mean absolute errors obtained with SRCS-HTV $\tilde{F}$  with true and wrong blur and end-member matrices for different image noise levels and  $\sigma_p = 2$ .

sizes  $100 \times 100$ ,  $50 \times 50$ ,  $33 \times 33$ , and  $25 \times 25$ , and the scale factors  $1 \times 1$ ,  $2 \times 2$ ,  $3 \times 3$ , and  $4 \times 4$  respectively, are utilized to create segmentation maps at approximately the original size of  $100 \times 100$ . Average MAE over six images confirms the superiority of SRCS-HTV $\tilde{F}$ . The coverage maps for the blurry and noisy image (a) segmented at three times higher resolution are shown in Fig. 6.15(b)-(e).

Based on this evaluation on synthetic data, the method SRCS-HTV $\tilde{F}$ , based on the energy function

$$E(A; V) = D_{SRCS}(A; V) + \mu \text{TV}_{HUB}(A) + \nu \tilde{F}(A) + \eta G(A), \quad (6.7)$$

is selected as the best performing method to be used in further tests.

The proposed method requires blur and end-member matrices as input; here we provide this information as a priori knowledge. Blind deblurring methods (e.g., [155, 156]) can be used to estimate the Point Spread Function (i.e., blur matrix) whereas end-member matrix can be estimated by methods proposed in [157–160]. We estimate the end-member matrix by random sampling of the true classes. We evaluate robustness of SRCS-HTV $\tilde{F}$  w.r.t. different levels of error in the estimated blur and end-member matrices and present the results in Fig. 6.16.

We first observe performance of the method for a range of mismatched end-member matrices, obtained by adding increasing level of noise to a true end-member matrix. Fig. 6.16(a) shows results for medium image noise and blur ( $\sigma_n = 10^{-2}$ ,  $\sigma_p = 2$ ) and with end-member noise  $\sigma_{ee} \in [0, 0.03]$ . As can be ob-

## 6.2. Performance evaluation of super-resolution coverage segmentation

served, the performance is very stable up to a point at which, in some cases, a rather different (incorrect) segmentation provides the lowest energy solution. Note that this decrease in performance occurs at a level where the end-member noise is larger than the image noise of an individual pixel (i.e., estimating the end-members from single pixel samples would provide a better result). Hence, we conclude that the method is robust w.r.t. end-member noise.

Second, we observe performance of the proposed method for a range of different Gaussian shaped PSF estimates with  $\sigma_{pe} \in [0, 4]$ . Results for the noise level  $\sigma_n = 10^{-2}$  and the case when the true blur is obtained by a Gaussian PSF with  $\sigma_p = 2$  are shown in Fig. 6.16(b). Here we observe a more proportional degradation in performance as the estimated PSF deviates from the correct one. In Fig. 6.16(c) we present performance of SRCS-HTV $\tilde{F}$  for a range of image noise levels, where either the end-member matrix has a noise magnitude equal to the image noise (orange curve) or where the used PSF either is 20% too small or 20% too large (cyan and blue curves).

Summarizing 450 = 3 × 150 tests performed with end-member noise or PSF mismatch for  $\sigma_p \in \{0, 1, 2, 3, 4\}$  and image noise with  $\sigma_n \in \{0, 10^{-3}, 3 \cdot 10^{-3}, 10^{-2}, 3 \cdot 10^{-2}\}$ , we conclude that: (i) small to medium end-member mismatch (smaller than the per pixel image noise) in general has a small impact on the segmentation, whereas larger mismatch may lead to the algorithm finding a rather different segmentation result; (ii) errors in the PSF estimate affect the performance more or less proportionally. When the solution is well defined, as in the no-noise case at the left edge of Fig. 6.16(c), the method is robust w.r.t. PSF errors.

### 6.2.2 An illustrative example on a naturally degraded image

The performance of SRCS-HTV $\tilde{F}$  is illustrated on an example image acquired under real imaging conditions. Three equal disks centered at the vertices of an equilateral triangle with a side length equal to the radii of the disks were displayed on a computer screen. The image of the screen is acquired by the 8 megapixel CMOS sensor of a Samsung Galaxy S3 cell phone camera and it is shown in Fig. 6.17(a). The black rectangular regions in Fig. 6.17(a) indicate sets of pixels used to estimate the 8 × 3 end-member matrix  $C$ . The zoomed-in portion, Fig. 6.17(b), shows details of the noisy and blurred image. The thin red lines in Fig. 6.17(b-e,g) are not parts of either sensed or segmented images and they indicate the true object boundaries. They are superimposed as reference.

For comparison, Fig. 6.17(a) is also segmented by Bayesian classification utilizing linear discriminant analysis (LDA), with the same indicated sets of pixels (black

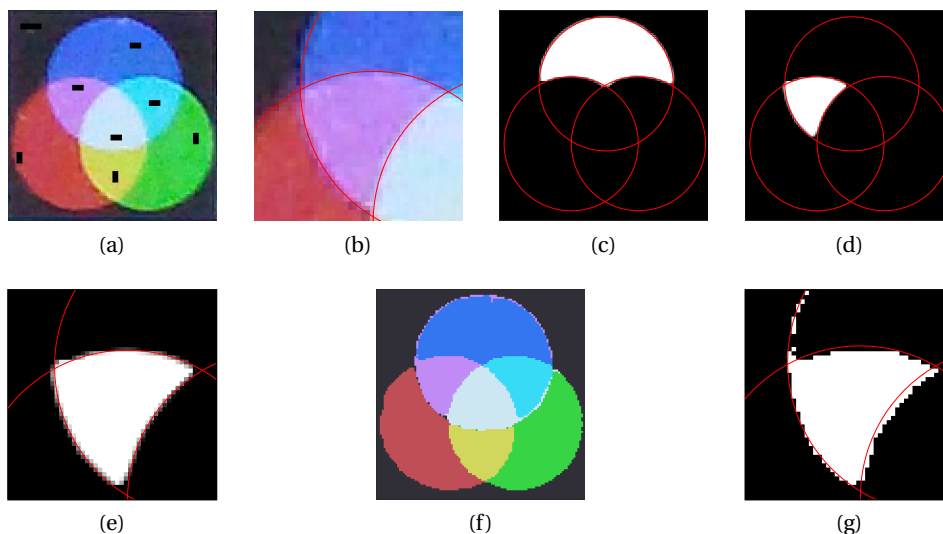


Figure 6.17: Coverage segmentation of a blurry and noisy real color image. (a) The sensed image (black regions indicate pixels used for the end-member matrix estimation). (b) Zoomed-in region from (a). (c)-(d) Two segmented components obtained by SRCS-HTV. (e) Zoomed-in segmentation of the region from (b). (f)-(g) Segmentation and zoomed-in region (b) obtained by LDA. Thin red lines indicate the correct continuous region boundaries.

regions) used for training.

Both methods perform well on the majority of pixels: the MAE of SRCS-HTV is 0.52%, and for LDA it is 0.83%. This is expected, since most of the pixels are pure, and classified as such by both methods. However, looking at the boundaries of the resulting segmentations it can be seen that the SRCS-HTV segmentation is by far superior. The coverage maps for two (out of 8, including background) components are shown in Fig. 6.17(c)-(d). Colour unmixing is successful and accurate for all the 7 appearing colours/classes. The zoomed-in portion, Fig. 6.17(e), shows the details of the segmentation. It can be observed that, in spite of the blur present in the original image, the resulting segmentation is mostly crisp, with a thin transition of partly covered pixels appropriately positioned along the true region boundaries. Noise is efficiently removed.

The segmentation obtained by LDA, Fig. 6.17(f), is crisp. As clearly visible in Fig. 6.17(g) (showing the same zoomed-in region as (e)), the noise and mixed pixels at object boundaries have a rather strong negative impact on the result, causing misclassification of a number of pixels, as well as rather unappealingly jagged segmentation boundaries.

# 7 Applications

## 7.1 Restoration of transmission electron microscopy images

### 7.1.1 Transmission electron microscopy

In biology and medicine, microscopy is essential in the understanding of the organization and function of cells and tissues. Two essentially different types of microscopy are distinguished, light and electron microscopy. Light (optical) microscope uses light to illuminate specimens and glass lenses to magnify images. It can magnify details up to  $2 \times 10^3$  times and it has a resolution (the level of details visibility) limit of about 200 nanometers (nm). A biologist can track cells over time by utilizing light microscopy since this modality can be used to image live specimens alongside with the dead ones. On the other hand, an electron microscope uses a beam of electrons to illuminate specimens and magnetic lenses to magnify images (Fig. 7.1), thus achieving much higher resolution (up to 0.01 nm) which corresponds to an extremely high magnification of details (up to  $10^6$  times). Due to such the magnification power, the information at the ultrastructural level can be obtained only by electron microscopy [161]. In addition to a superior resolution (nm range), electron microscopy also provides a comprehensive view on the sample by visualizing all organelles and ultrastructures. However, electron microscopy can only be utilized for the analysis of dead specimens due to the destructive power of electrons used in electron microscopy.

Various electron microscopy techniques exist and they differ mainly based on which type of electrons after interaction with the sample they collect. Transmission Electron Microscopy (TEM) collects transmitted electrons from the illuminated sample (Fig. 7.1) while Scanning Electron Microscopy (SEM) col-

lects backscattered electrons, i.e., electrons which have passed, or not, through the sample, respectively. Focus in this thesis is on TEM and in continuation main challenges regarding TEM imaging will be discussed.

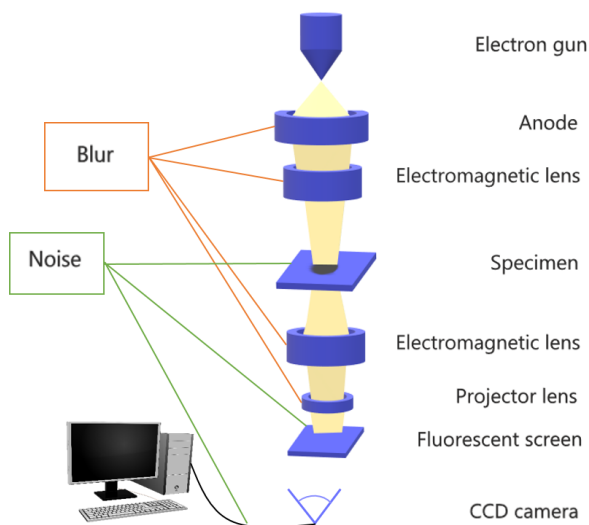


Figure 7.1: Schematic image acquisition workflow of a transmission electron microscope. Yellow rays indicate electron beams originating from the electron gun; between the gun and specimen they correspond to the emitted electrons, and below the specimen they correspond to the electrons which are transmitted by the sample. The noise and blur sources in the systems are indicated by orange and green lines, respectively.

Transmission electron microscopy suffers from various aberrations. Thus, an important goal is to improve the quality of the produced images [12]. An important step toward achieving this goal is to understand how these artefacts are created [11]. Fig. 7.1 illustrates how the conventional TEM acquires images; as mentioned above, an electron gun emits electrons which are focused on a sample by a series of electromagnetic lenses. The emitted electrons interact with the sample and the electrons transmitted by the sample are further focused by a series of lenses on an imaging device (such as a fluorescent screen, a layer of photographic film, or a sensor such as a charge-coupled device (CCD)), to form the detected image. Almost every component of the imaging device introduces artefacts in the acquired image. Due to delicate (magnetic) lens systems, statistical electron counting errors, electromagnetic interference, etc., TEM images are typically degraded by imaging artefacts such as noise and blur. The noise follows mixed Poisson-Gaussian distribution [11, 12, 162, 163]; its signal dependent



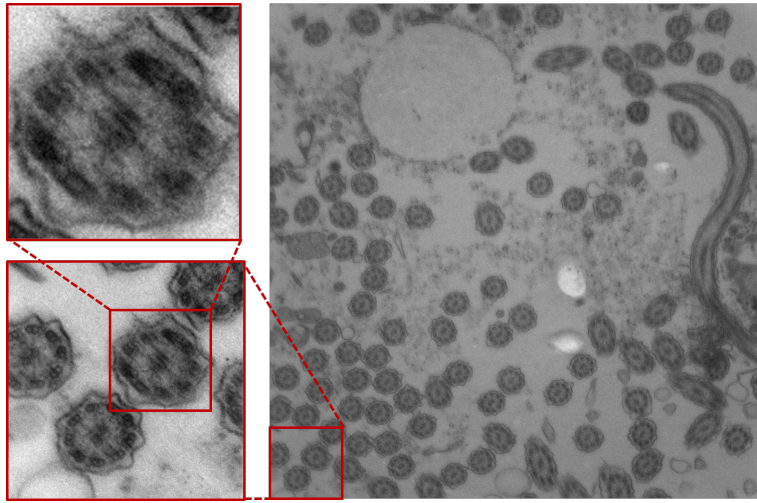


Figure 7.2: TEM image typically degraded by noise and blur.

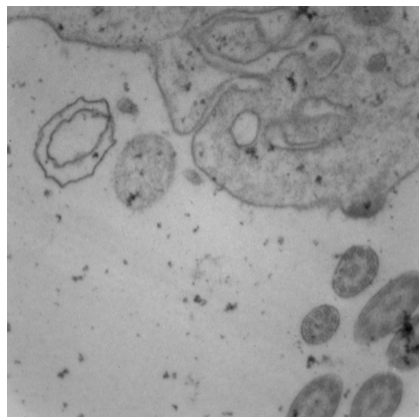
component is related to stochastic nature of electron counting process while the signal-independent component comes from electromagnetic and thermal interference. Therefore, the restoration methods suited for PG noise which are introduced in this thesis can be applied for the restoration of TEM images.

Some of the mentioned imaging artefacts, such as the motion blur artefacts due to sample drift and vibrations or the modification of the sample due to interaction with electrons, can be reduced by imaging the specimen at very low electron doses and shorter acquisition time. However, this comes at the cost of more noisy images (Fig. 7.3 (b)) which particularly increases the need for denoising.

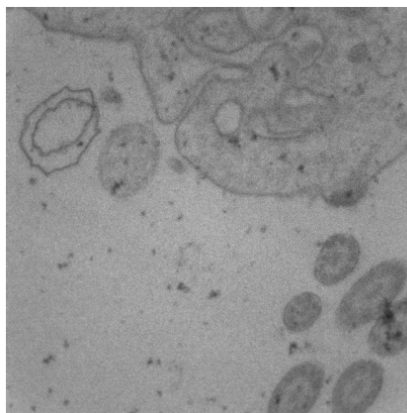
For enhancement of long exposure TEM images (Fig. 7.3 (a)) the blind deconvolution method suited for PG noise introduced in Section 5.3 is applied as shown in Publication II. In Publication V denoising of short exposure images (Fig. 7.3 (b)) is explored by both state-of-the-art traditional and newly developed deep learning based denoising methods.

### 7.1.2 Blind deconvolution of long exposure TEM images

In this section the performance of the blind deconvolution method proposed in Publication II is illustrated on the example of TEM images of cilia. Cilia are small hair-like organelles protruding from cell surface. They are used in clinical practice for diagnosis of Primary Ciliary Dyskinesia (PCD), a genetic disorder which results from dysfunction of cilia causing several diseases, such



(a) Long exposure TEM image (exposure time equals to 200ms).



(b) Short exposure TEM image (exposure time equals to 2ms).

Figure 7.3: Two TEM images of the same specimen taken at different exposure times. The short acquisition time results in more noisy image.

as airways infections and reduced fertility in both females and males [164, 165]. Transmission Electron Microscopy (TEM) is the only imaging modality that provides a resolution sufficient for the diagnosis of cilia disorders based on ultrastructural analysis.

To set a diagnosis of disorder, a pathologist has to locate and visually analyze at least 50 high quality imaged instances of perfectly perpendicularly cut cilia which can take several hours per patient. A very tedious and time consuming process is performed manually which motivates the development of automated image acquisition and analysis of cilia. Good results in automated cilia detection and rating of cilia instances have been achieved recently in [166–168]. However, no fully automated image analysis based system for setting PCD diagnosis exists, and still the best candidates which are found either automatically or manually are displayed to a pathologist who makes the final diagnosis by visually inspecting the nanostructures. Those nanostructures typically correspond to the objects of a size of a couple of pixels and they are hardly visible in TEM images corrupted by noise and blur.

As elucidated previously, the noise present in TEM images has PG distribution and therefore our restoration methods from Publications I-V which adequately tackle this type of noise are perfectly suited for the enhancement of TEM images of cilia. Those methods can ensure that the relevant structures in the cilia images appear enhanced and easier to analyze. Considering that the PSF of the imaging device is not known, an appropriate approach is to use the blind deconvolution

## 7.1. Restoration of transmission electron microscopy images

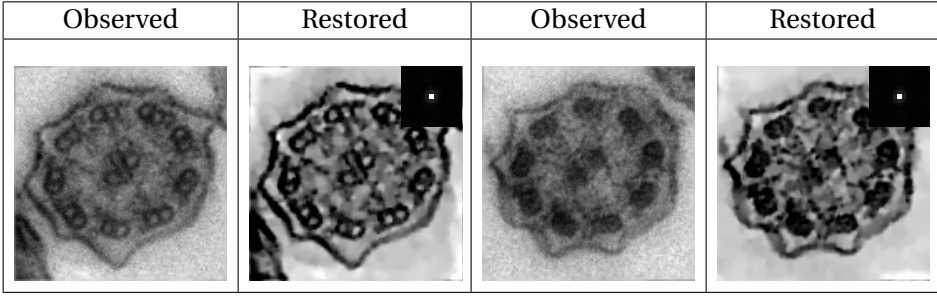


Figure 7.4: Blind deconvolution of blurred and noisy TEM images of cilia. Two examples of the observed and restored objects, as well as the estimated PSF are presented.

proposed in Publication II which simultaneously estimates blur-free and noise-free image and PSF of the imaging device.

To accurately estimate the PSF of the imaging device, a two step procedure is applied. In the first step, the PSF is estimated by applying the proposed blind deconvolution method to the original, full-size, image of the whole specimen (which contains multiple cilia instances; see Fig. 7.2, right). In the second step, this estimate of PSF is used in non-blind deconvolution (by repeated procedure, assuming known PSF) of cut-outs containing individual objects (Fig. 7.2, left upper corner) which are automatically detected in a TEM image of the whole specimen by utilizing the methods from [167, 168]. More examples of automatically detected noisy and blurry cilia are shown in Fig. 7.4.

As discussed in Chapter 4, the TV regularization transforms a smooth signal into piecewise constants and the so-called staircase effect appears in the restored images, as observed in [30, 101]. Hence, in this particular application, for better results (with less present staircase effects and better texture preservation), two regularization terms are utilized for deconvolution of cilia cut-outs in the second step; in addition to the TV-regularization, a term including second order derivatives is used. The objective function considered in this example is:

$$E(u, h) = D(u, h; v) + \lambda_1 R_1(u) + \lambda_2 R_2(u) , \quad (7.1)$$

where  $R_1$  is the TV regularization given by (4.4) and  $R_2$  is the Hessian-based penalization proposed in [30]. The regularization parameters  $\lambda_1 = 3 \times 10^{-3}$  and  $\lambda_2 = 10^{-3}$  are experimentally determined. The parameters of mixed PG noise,  $\alpha$ ,  $\sigma_m$  and  $\mu$  (Eq. (3.10)) are estimated by the method presented in [169]. In Fig. 7.4 the restored cilia cut-outs are shown alongside with the estimated PSF of the conventional TEM device<sup>1</sup> used for the acquisition of images of cilia obtained

<sup>1</sup>Vironova AB, Stockholm, Sweden

from the first step.

### 7.1.3 Denoising of short exposure TEM images

This section illustrates the results of the denoising of short exposure TEM images of cilia (Fig.7.3 (a)) presented in Publication V.

Many traditional methods for noise removal have been proposed in the literature; an exhaustive survey on denoising methods is presented in [63] while the survey [12] focuses exclusively on the application of some of the methods on electron microscopy images highlighted in [63]. Recently, deep learning based denoising methods are getting more attention due to their superior performances [170–172]. However, since these methods are data driven, the availability of ground truth, noise-free image, which is used in the training phase, is of critical importance for their good performances. Such perfect image does not exist in real application settings and obtaining such an image is the first challenge in the process of developing a powerful deep learning based denoising method.

In Publication V a novel Convolutional Neural Network (CNN) method is developed specifically for the denoising of short exposure TEM images. Fig. 7.5 illustrates CNN architecture. The network consists of two streams. The first stream is built by four convolution blocks, two transposed convolution blocks and one residual block. The convolution block encodes the image representations while removing the noise, whereas the transposed convolution block decodes these representations to restore the noise-free image content. The residual block consists of two convolution blocks. The batch normalization (BN) [173] layer is used as regularization before rectified linear unit (ReLU) [174] activation to deal with internal covariate shift. To elevate the training performance, skip connections are used and followed by a BN layer. The second stream consists of four convolution blocks, two up-sampling blocks, two max-pooling layers, and one residual block. The prediction made by the first stream restores most content with blur whereas the prediction made by the second stream contains high-frequency content, however, with an inconsistent illumination in respect to corresponding ground truth. Motivated by that observation, we performed an end-to-end training by averaging the predictions of both streams, which resulted in an improved output. The training of both streams is performed using patches of  $128 \times 128$  pixels, normalized to the range  $[0,1]$ . The patches are extracted with an overlapping stride of 16 pixels.

The CNN method is further compared with three classical methods suited for PG noise: block matching (BM3D) [98], wavelet domain (Pure-LET) [175], and

## 7.1. Restoration of transmission electron microscopy images

energy minimization (EM) from Publication I.

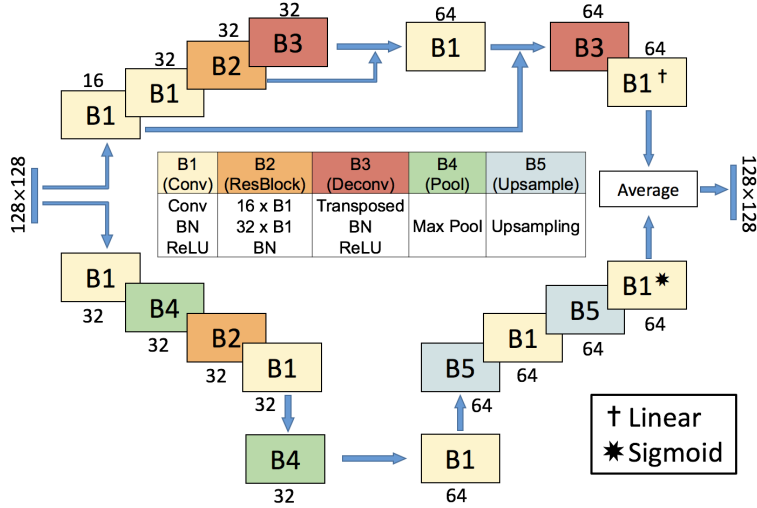


Figure 7.5: The two-stream denoising CNN (DCNN) architecture. The sizes of output feature maps of each block are shown on top of each block and generated using  $3 \times 3$  convolutions. The last  $1 \times 1$  convolution blocks of each stream use linear and sigmoid activation, respectively, instead of ReLU.

To facilitate the CNN training with noise-free image and the quantitative comparison of four considered methods, a series of 100 short exposure (2 ms) noise images is acquired at the same spatial location of the same specimen (FOV=2000 nm). All images are of the size  $2048 \times 2048$  pixels and acquired with the low-voltage MiniTEM<sup>2</sup>. Those images are aligned to the first image of the series using rigid registration and pixel-wise median is taken to form low-noise image (Fig. 7.6) which served as the ground-truth image for the CNN training and quantitative evaluation. The series of 100 short exposure images is divided into a training set (first 10 images) and test set (remaining 90 images). The images from the training set together with the ground-truth are used for the CNN training and for parameter tuning for BM3D, Pure-LET and EM. The images from the test set alongside with the ground-truth are utilized for the evaluation of all four considered methods and for their mutual comparison.

Averaging many short exposure electron microscopy images that have been mutually registered is a commonly used technique for improving the quality of short exposure electron microscopy images [176]. Inspired by this, the potential of such aggregation strategy in combination with denoising for further enhancement of ultrastructures of interest is additionally explored. Thus, apart from

<sup>2</sup>Vironova AB, Stockholm, Sweden

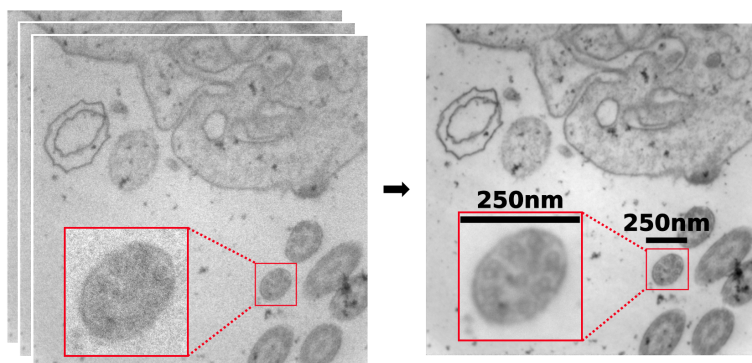


Figure 7.6: **Left:** Short exposure TEM image ( $2048 \times 2048$  pixels) from a series of 100 images. **Right:** Ground truth created by co-registration and aggregation of the stack to the left. The two insets show magnified views ( $250 \times 250$  pixels) of one cilium.

evaluating the performances on denoising single images, the performance of each method is also evaluated for two additional denoising strategies 1) *denoising of five aggregated short exposure images*, and 2) *aggregation of five denoised short exposure images*. Three traditional methods were tuned for each strategy separately, while CNN was trained only on single frames and applied without the additional tuning for two another strategies. The performance of all four methods for all three considered strategies is evaluated using the peak-signal-to-noise ratio (PSNR) and structural similarity index measure (SSIM), given by Eq. (6.1) and (6.5), respectively.

The quantitative and qualitative results for single image denoising and two aggregation strategies are presented in Table 7.1 and Fig. 7.7. As can be seen from both Table 7.1 and Fig. 7.7, denoising improved both single and aggregated short exposure images. For all three strategies, the quantitative results show that the CNN framework achieves the highest PSNR and SSIM in most of the settings, slightly outperforming the considered traditional methods which exhibit quite similar results among themselves. The CNN framework is only outperformed marginally by EM method for single image denoising in terms of SSIM. Both aggregation strategies improve the results approximately equally well.

To validate the level of agreement between the quantitative results and the visual (qualitative) results, a subjective visual evaluation conducting a two-step voting process by six of the authors of Publication V is performed. In the first step, involving only the classical methods, the authors of Publication V rated the results ( $1^{st}$ ,  $2^{nd}$ , and  $3^{rd}$  best) on the cilium subimage produced by each of the methods with different parameter settings. The seven displayed

## 7.1. Restoration of transmission electron microscopy images

Table 7.1: Results on the test data set. Average PSNR and SSIM ( $\pm$  standard deviation) over 90 single images are given in the 1st and 2nd rows. Rows 3 and 4 contain average PSNR and SSIM over 18 aggregated groups of 5 short exposure images followed by denoising. Average PSNR and SSIM over 18 images each obtained by aggregating 5 denoised short exposure images, are given in rows 5 and 6. Best performances are marked in bold.

		Initial	BM3D ( $\sigma_{bm}$ )	PURE-LET ( $\sigma_{pl}$ )	EM ( $\lambda$ )	DCNN
1	PSNR	22.25	$37.39 \pm 0.30$	$37.38 \pm 1.09$	$37.80 \pm 0.27$	<b><math>38.04 \pm 0.21</math></b>
	SSIM	0.019	$0.233 \pm 0.007$	$0.219 \pm 0.007$	<b><math>0.255 \pm 0.027</math></b>	$0.252 \pm 0.002$
2	PSNR	27.88	$40.45 \pm 1.09$	$40.19 \pm 1.06$	$40.19 \pm 0.54$	<b><math>40.86 \pm 0.37</math></b>
	SSIM	0.037	$0.270 \pm 0.019$	$0.263 \pm 0.017$	$0.277 \pm 0.017$	<b><math>0.282 \pm 0.011</math></b>
3	PSNR	22.25	$39.65 \pm 1.04$	$40.21 \pm 0.48$	$39.92 \pm 0.93$	<b><math>40.84 \pm 0.45</math></b>
	SSIM	0.019	$0.261 \pm 0.013$	$0.265 \pm 0.011$	$0.273 \pm 0.021$	<b><math>0.276 \pm 0.009</math></b>

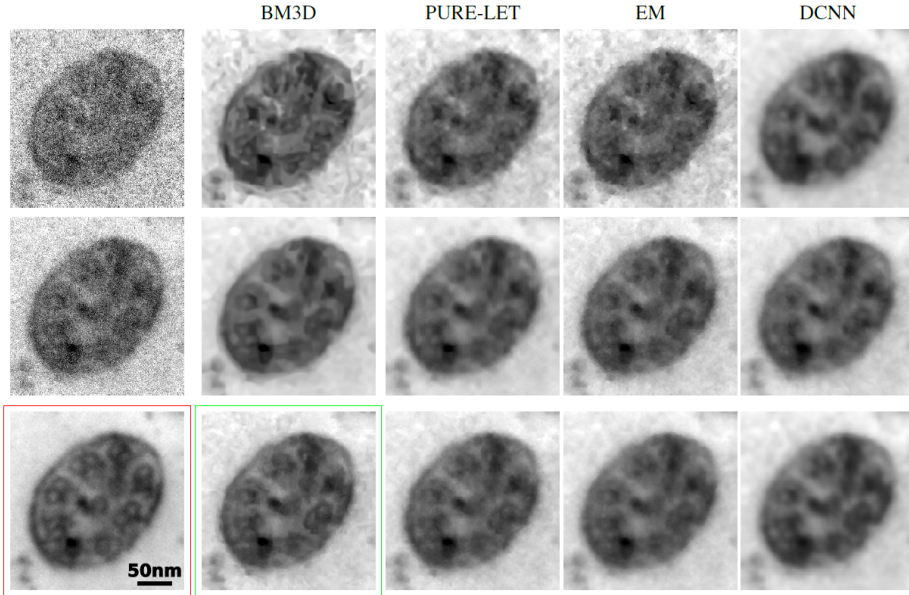


Figure 7.7: Noisy and denoised close ups of a cilium instance obtained with the considered methods. **Top**: Denoising of a single image. **Middle**: Denoising of five aggregated noisy images. **Bottom**: Aggregation of five denoised single images. The red frame (bottom left) indicates the ground truth for single noisy images. The green frame indicates the best ranked image in the two-step visual voting process.

images spanned a parameter range centered around the maximal SSIM for that

method. The procedure was repeated for the above mentioned two strategies of aggregating five short exposure images. The second step involved all four methods. The images resulting from the two aggregation strategies utilizing the tuned parameter settings as decided in the first step, together with the CNN results were displayed (in random, unknown order) and the authors of Publication V rated them again (as the 1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup> best). The denoised image with the majority of votes is highlighted in Fig. 7.7, and it is obtained by BM3D method from the aggregation of five denoised short exposure images strategy. Based on the visual assessment, BM3D method in combination with *aggregation of five denoised short exposure images* strategy produces the most appealing results.

## 7.2 Application of super-resolution coverage segmentation in medicine

In this section we illustrate performances of super-resolution coverage segmentation applied on a micro-CT image of trabecular bone. Measure of the surface of micro-structures of a trabecular bone is one of the determinants of bone strength. It is really important for the diagnosis of osteoporosis [180]. Precise segmentation, a prerequisite for accurate surface measurements, is a challenging task in blurry and noisy low-resolution CT images [40, 181–183]. We therefore find it suitable to demonstrate applicability of our super-resolution coverage segmentation method for bone segmentation.

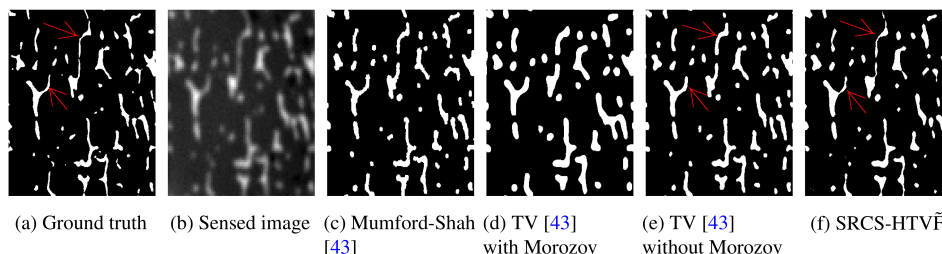


Figure 7.8: Segmentation of micro structures of a trabecular bone at 4 times increased resolution. (a) Ground truth. (b) Sensed image. (c)-(e) Segmentations obtained by variants of the method proposed in [40]. (f) Crisp segmentation derived from coverage segmentation obtained by SRCS-HTV with Morozov. Red arrows indicate some example locations where the proposed method qualitatively provides a segmentation which is more similar to the ground truth than the compared methods do. Quantitative performance comparison is given in Table 7.2.

We observe the same low resolution blurry and noisy image of trabecular bone



### 7.3. Application of super-resolution coverage segmentation in remote sensing

Table 7.2: Performance on micro-CT image segmentation.

Method	Parameter choice	DICE	Overall bone surface ref=3.94 ( $mm^2$ )
Mumford-Shah [40]	with Morozov principle	0.80	5.16
Mumford-Shah [40]	without Morozov principle	0.80	5.16
TV [40]	with Morozov principle	0.70	6.36
TV [40]	without Morozov principle	0.82	5.14
SRCS-HTV $\tilde{F}$	recommended parameters	0.84	4.32

(Fig. 7.8(b)) as is used in [40]. The image is obtained from a high resolution noise-free and blur-free ground truth image (Fig. 7.8(a)) acquired by a parallel-beam synchrotron micro-CT [184]. The image is down-sampled by a factor  $4 \times 4$  and degraded by Gaussian blur with  $\sigma_p = 4.85$  and Gaussian noise with  $\sigma_n = 1$ , as described in [40]. Fig. 7.8(c)-(e) show segmentations obtained by different versions of a joint super-resolution segmentation method suggested in [40]: Mumford-Shah model optimized by ADMM and its TV version, where regularization parameters are chosen with or without utilization of Morozov discrepancy principle [185]. The segmentation obtained by SRCS-HTV $\tilde{F}$  with recommended parameters ( $(\mu, \omega) = (0.05, 0.001)$ ,  $\nu_0 = 10^{-4}$ ,  $\rho = 0.1$ ,  $\eta_0 = 1$ ) at four times increased resolution is shown in Fig. 7.8(f). The presented crisp segmentation is obtained from the coverage segmentation by thresholding it at 0.5.

As can be seen from Fig. 7.8, the proposed method segments the micro structures of interest more precisely in comparison with all versions of the method proposed in [40]. This observation is confirmed by the quantitative evaluation presented in Table 7.2, where our method outperforms all 4 variants of the method from [40] as measured by the DICE similarity coefficient and the bone surface area.

### 7.3 Application of super-resolution coverage segmentation in remote sensing

Remote sensing is one of the application fields relevant for the proposed super-resolution coverage segmentation method, due to the relatively large pixels often being partially covered by different classes/objects imaged on the ground (e.g., water, soil, forest) [66–71], see Fig. 7.9. Due to low spatial resolution mixed pixels are commonly present in hyperspectral images restricting the processing and application of these images in practice [177]. Most state-of-the-art methods pro-



Figure 7.9: A high resolution satellite image, with a superimposed 200 meter resolution grid, illustrating the pixel size of presented satellite image. Pixel #1 is a fairly homogenous pixel, almost completely covered by urban area, whereas pixel #2 is mixed pixel, partially covered by vegetation and soil.

posed for hyper-spectral unmixing are based on the linear [178] and sparse [179] unmixing models. The coverage segmentation proposed in this thesis unifies both approaches and all of its properties are highly desirable for this application.

The method SRCS-HTV $\tilde{F}$  is evaluated on a publicly available<sup>3</sup> 220 band hyperspectral data set from an Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) image taken on NW Indiana's Indian Pine test site in June 1992. The same image is used for evaluation in [69], where linear unmixing of pixels is performed at an increased spatial resolution using simulated annealing, and in [38], where the performance of the here observed CS-PTF method is tested.

Similar to [69] and [38], the tests on the  $[31-116] \times [27-94]$  cut-out of the scene are performed. Four labelled classes: "Corn-notill" (1008 pixels), "Grass/Trees" (732 pixels), "Soybeans-notill" (727 pixels), and "Soybeans-min" (1926 pixels) are provided as the ground truth. 1455 pixels of the scene are not labelled, and are not considered in the evaluation. Since the available ground truth classification of the image does not provide abundance information, the spatial resolution of the original image is decreased by averaging blocks of  $3 \times 3$  pixels thus providing a 220 band image of  $28 \times 22$  pixels; this low resolution image was used as test image. Information from the three times higher resolved ground truth data is then used to estimate the coverage values for the created low resolved image. One band of the low resolution image and the corresponding ground truth are

---

<sup>3</sup><https://purr.purdue.edu/publications/1947/about?v=1>

### 7.3. Application of super-resolution coverage segmentation in remote sensing

shown in Fig. 7.10(a) and (b).

The 220 bands are decorrelated by a whitening transformation. For each class, 20 pure pixels are randomly selected as training data, and are used to compute the matrix  $C$ . The  $28 \times 22 \times 220$  image is then segmented at (i) the same (low) spatial resolution ( $d = 1 \times 1$ ) and (ii) at 3 times increased resolution ( $d = 3 \times 3$ ). The experiment is repeated 20 times, with a new random selection of the training sets, and the mean and standard deviation of the performance measures are computed. (The parameter values are as suggested in Publication VI). The evaluation of (i) is performed by direct comparison of the estimated coverage values and the created ground truth. The MAE of SRCS-HTV $\tilde{F}$  and CS-PTF ([38]) are given in Table 7.3. This result has to be interpreted with care, since the  $3 \times 3$  averaging, providing the ground truth, has a quantization error of up to  $\frac{1}{18} \approx 5.6\%$ .

Table 7.3: Performance on the AVIRIS data set (in %). Mean value of the measure is followed by a standard deviation, in parentheses.

	MAE	CA
Villa et al., [69]	–	90.65
CS-PTF [38]	4.53 (1.57)	[92.59(1.99), 94.74(1.32)]
SRCS-HTV $\tilde{F}$	6.10 (0.65)	93.40 (2.31)

To evaluate (ii), the coverage segmentations are thresholded at 0.5 and then compared with the full resolution ground truth. The classification accuracy (CA) (percentage of correctly classified pixels) of the crisp segmentation at the 3 times increased resolution is shown in Table 7.3. The resulting *coverage segmentation* at increased resolution from a single experiment is shown in Fig. 7.10(c). Since the CS-PTF method does not provide coverage segmentation at increased spatial resolution, its performance is presented in Table 7.3 as an interval of possible performances, where the lower bound corresponds to a nearest neighbour up-sampling and the upper bound is derived from an ideal optimal distribution of the coverage (OOA in [38]). Here the proposed SRCS-HTV $\tilde{F}$  outperforms both [69] and the naive approach of [38] and is not far from the accuracy of the “oracle” segmentation which gives the upper accuracy bound.

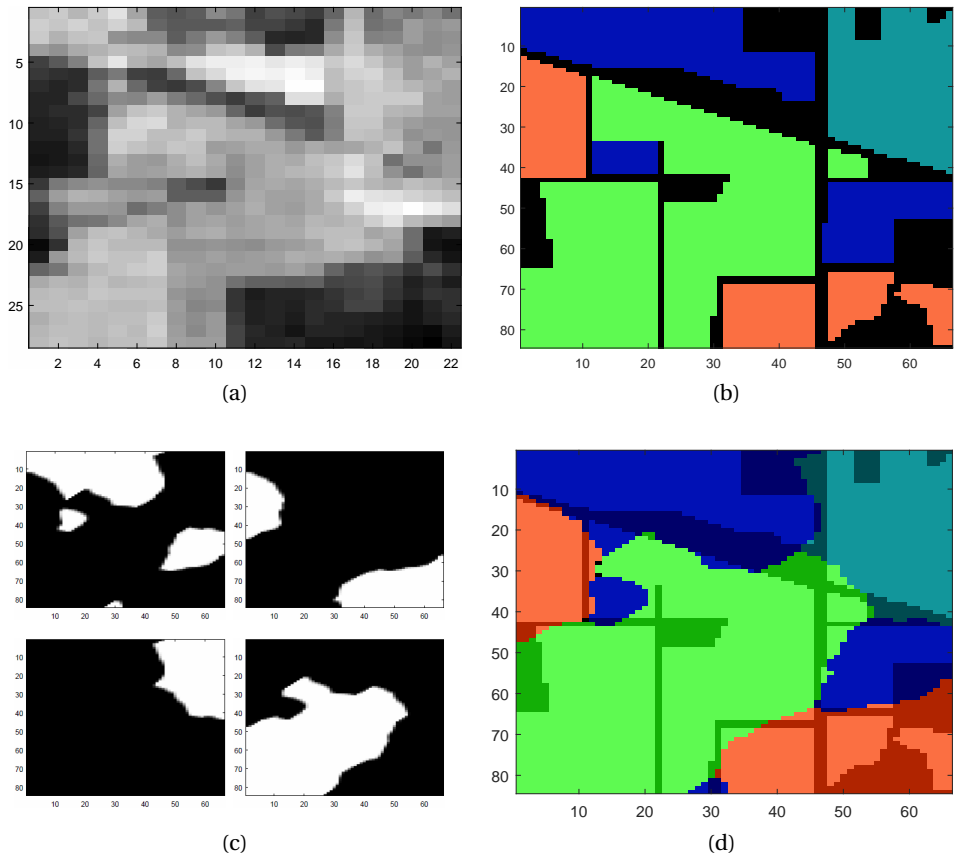


Figure 7.10: Coverage segmentation of a hyperspectral AVIRIS image. (a) Band 30 (out of 220) of the image obtained by downsampling the original image by averaging  $3 \times 3$  blocks; (b) Ground truth classification of the original image (unclassified pixels shown in black); (c) Coverage segmentation (four classes) of (a) at 3 times increased resolution; (d) Crisp segmentation derived from (c) (to compare with (b)).

# 8 Conclusion

This chapter concludes the thesis by providing a short summary of the presented work, a short reflection on the main objectives, and by suggesting some directions for future work.

## 8.1 Summary of contributions

Energy minimization methods are at the crossroad of several well established mathematical methodologies: Bayesian maximum a posteriori (MAP) theory, regularization for ill-posed inverse problems and, in relation to it, sparse representations of signals, variational methods and numerical optimization. All these methodologies and their mutual connections are briefly sketched in this thesis. More attention has been devoted to specific topics within these methodologies for which research during PhD studies is conducted. In continuation the concrete contributions of the thesis are summarized.

The thesis explores how signal-dependent Poisson-Gaussian noise which appears in many realistic imaging situations can be appropriately treated in image restoration tasks. It has been found that Variance Stabilization Transformation can be applied to the observed image in order to remove signal dependency of noise and to render it to be approximately normally distributed. Following this path, assumption of Poisson-Gaussian noise model is incorporated in the data fidelity term of the energy function. This resulted in the development of novel non-blind deconvolution (Publication I) and super-resolution reconstruction (Publication III) energy minimization based methods for images degraded by Poisson-Gaussian noise. The developed methods demonstrated excellent performances when they were utilized for restoration in the presence of Poisson-Gaussian noise in terms of both quality of restoration and computational speed.

Turning our attention to the sparsity promoting regularization terms, we explored how further improvement of performances of the methods can be achieved if edge preserving potential functions are utilized instead of classical Total Variation regularization (Publication IV).

In order to increase the applicability of the developed methods, non-blind deconvolution method is extended to the blind method (Publication II) which jointly estimates underlying image and PSF of the imaging device starting only from the observed image. The applicability of the proposed blind deconvolution method is confirmed on the long exposure Transmission Electron Microscopy images of cilia.

Exploring further how extremely noisy short exposure Transmission Electron Microscopy images can be enhanced (Publication V), CNN is developed for this specific purpose and compared with three traditional denoising methods suited for PG noise including energy minimization based method from Publication I. All four considered methods enhanced significantly the ultrastructures in short exposure TEM images of cilia. Quantitative and qualitative analysis confirmed that the CNN based denoising method exhibits outstanding performances. Energy minimization method does not fall far behind it, performing at the same time equally well as two more state-of-the-art considered traditional methods.

The coverage segmentation method from [38] is based on the assumption that partially covered pixels form one pixel thin boundary. This assumption does not hold in realistic situations whenever a sensed image is degraded by blur and noise. Observing that the coverage segmentation method performs poorly when this assumption is violated, the super-resolution blur-aware model is included in the data fidelity term of the existing energy minimization based coverage segmentation method (Publication VI). This significantly improves the performances of the method in the presence of blur and noise and, in addition, allows an estimation of coverage maps at increased resolution. New sparsity promoting regularization terms are suggested as well as a complex optimization scheme to deal with non-convexity of the considered energy function. In this way the applicability of the coverage segmentation method is increased which is demonstrated by its evaluation on real multi-channel images.

### 8.2 Future work

Several ideas for the continuation of the research presented within this thesis can be drawn.

As discussed in Section 7.1.3, three classical methods for short exposure TEM

image denoising considered in Publication V were separately tuned for single image denoising and for two other denoising strategies: aggregation of five denoised images and denoising of five aggregated noisy images. On the other hand, the CNN based denoising was trained only on single images of cilia and applied on two other denoising strategies. Considering this, it is impressive that CNN performs equally well for these two strategies which indicates huge transfer learning potential of CNN based denoising. This transfer learning potential seems worth of further exploration. It would be interesting to examine how well CNN based denoising trained on miniTEM images of cilia performs on conventional TEM images of different specimens, or on other electron microscopy images, e.g., on SEM images. Fast acquisition by SEM results also in extremely noisy images, which affects both manual and automated analysis, and increases the need for denoising. Even fast SEM acquisition is much slower than fast TEM acquisition (more than 100 times), so being able to train a CNN on short exposure TEM images and to apply it for short exposure SEM image denoising is of great practical usefulness. This will advance the state-of-the-art of SEM image denoising.

Another possible future direction would be to develop 3D CNN for denoising/restoration of Focused Ion Beam Scanning Electron Microscopy (FIB-SEM) and Serial Block Face Scanning Electron Microscopy (SBF-SEM) images. These electron microscopy modalities are used to image biological specimen in 3D by repeatedly creating a new surface and imaging this new surface with a scanning electron beam. A similar idea of super-resolution CNN based method for these imaging systems has been explored recently in [186] showing promising results.

Developing CNN restoration methods which take into account mixed Poisson-Gaussian noise assumption remains also an open research question. Combining traditional and CNN based methods can potentially give an additional boost to both CNN and traditional based methods.

The biggest challenge related to the aforementioned possible research directions is the ground truth (noise-free and blur-free image) creation for training and quantitative evaluation purposes, as experienced in Publication V.

Possible improvements of coverage segmentation method from Publication VI are multiple. Extending the existing proposed super-resolution coverage segmentation method for volumetric data would be extremely useful for practical applications. The interest in sub-pixel/voxel segmentation is not limited to 2D images. The so-called partial volume effect is particularly observed in 3D medical image analysis, referring to voxels partially covered by two or more tissue types, [187, 188]. This issue and its negative influence on the subsequent relevant measurements (e.g., volume estimation), is particularly studied in relation to

## Chapter 8. Conclusion

---

MRI and PET images of the human brain, [189]. Moreover, CT, MRI and PET images are also usually degraded by blur and noise and the need for segmenting them by a blur-aware method which allows partial belongingness of voxels at increased spatial resolution (particularly in z-direction) is of great interest.

Moreover, the super-resolution blur-aware coverage segmentation method can be extended to simultaneously estimate the PSF of the imaging device and coverage maps from the observed data, i.e., it can be extended to be “blind”. This is significant since in many applications it is very difficult or even impossible to accurately estimate the PSF and blind methods are preferable.



# Acknowledgements

This thesis is not only the fruit of the last six years of doing research; it is also the consequence of a longer life path. Lots of people crossed my road, inspired or encouraged me, surrounded me with love, provided friendship, discussions, attentive listening, or smiles. All of these contributions, regardless of their duration, were precious gems distributed along my way. To the most influential persons in my life, especially during last six years, I would like to use this opportunity to say "thank you":

- ❁ Joakim and Nataša, my advisors, for making me an independent researcher. For all the patience, energy, knowledge and support you invested in me and my work. For bringing me to Uppsala and taking care of me when I was there. For ice skating at Fjällnora, for inviting me to your home. Thank you for being great advisors and friends!
- ❁ Mila, the head of Chair for Mathematics at the Faculty of Technical Sciences, for giving me an opportunity to work at FTN and to conduct my Ph.D. studies there. For making my stays abroad possible.
- ❁ My dear colleagues and friends from FTN, Marija, Srđan, Gabika, Miroslav and Andrija, for stepping in in times of my absence and helping with my teaching obligations and endless exams without any second thought.
- ❁ Ingela Nyström, the head of Center for Image Analysis in Uppsala, for being a great host at CBA and always accepting my stays there.
- ❁ Tomas, my collaborator, for great collaboration on many topics we worked together on.
- ❁ Kika, Christophe and Damian, my non Swedish friends from Sweden, for introducing me to fikas :-). For being there for me far away from home.
- ❁ Anindya, my collaborator and my dear friend, for teaching me deep learning. For your patience and positivity. For the best summer school ever, SSIP 2015!

## Acknowledgements

---

- ✿ Marina, my sister, for your unconditional love. For coming to Đala when I was kid whenever I asked you. For our joint night life experiences. :-)
- ✿ Lacika, for taking care of me when I was kid.
- ✿ Radmila, my mother in law, for all the lunches you cooked for us in our busy times. For our endless second hand shopping. :-)
- ✿ Jela, my cousin, for being a good friend to my mother and for taking care of her.
- ✿ Ana, my best friend forever, for being always there for me since the summer 2006 in Tivat. For spending months with me in Tivat during many years. For endless joint walks and chats about everything. I am really blessed to have you as my best friend!
- ✿ Paja, for all your love and support you have given me during this whole journey. Without your support this thesis would not have been possible.
- ✿ Milka, my mother, for all your unconditional support and love. For fighting alone, since my father died, to be both mother and father to me. For being always there for me. This thesis is equally dedicated to you, too.

*Novi Sad, 1 April 2019*

Buda.

# Bibliography

- [1] W. S. Boyle and G. E. Smith, "Charge coupled semiconductor devices," *Bell System Technical Journal*, vol. 49, no. 4, pp. 587–593, 1970. (Cited on pages v and 1.)
- [2] P. J. Noble, "Self-scanned silicon image detector arrays," *IEEE Transactions on electron Devices*, vol. 15, no. 4, pp. 202–209, 1968. (Cited on pages v and 1.)
- [3] S. G. Chamberlain, "Photosensitivity and scanning of silicon image detector arrays," *IEEE Journal of Solid-State Circuits*, vol. 4, no. 6, pp. 333–342, 1969. (Cited on pages v and 1.)
- [4] L. Mandel, "Fluctuations of photon beams: the distribution of the photoelectrons," *Proceedings of the Physical Society*, vol. 74, no. 3, p. 233, 1959. (Cited on pages v and 1.)
- [5] F. Benvenuto, A. L. Camera, C. Theyes, A. Ferrari, H. Lánteri, and M. Bertero, "The study of an iterative method for the reconstruction of images corrupted by Poisson and Gaussian noise," *Inverse Problems*, vol. 24, no. 3, p. 035016, 2008. (Cited on pages v, 2, 7, and 23.)
- [6] T. E. Nichols, E. A. J. Qi, and R. M. Leahy, "Spatiotemporal reconstruction of list-mode PET data," *IEEE Transactions on Medical Imaging*, vol. 21, no. 4, pp. 396–404, 2002. (Cited on pages v and 2.)
- [7] S. Delpretti, F. Luisier, S. Ramani, T. Blu, and M. Unser, "Multiframe surelet denoising of timelapse fluorescence microscopy images," in *5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI)*, 2008, pp. 149–152. (Cited on pages v and 2.)
- [8] M. A. Vieira, P. R. Bakic, A. D. Maidment, H. Schiabel, and N. D. Mascarenhas, "Filtering of poisson noise in digital mammography using local statistics and adaptive wiener filter," in *International Workshop on Digital*

- Mammography*. Springer, 2012, pp. 268–275. (Cited on pages vii, 3, and 4.)
- [9] P. Svoboda, M. Hradiš, L. Maršík, and P. Zemčík, “CNN for license plate motion deblurring,” in *Image Processing (ICIP), 2016 IEEE International Conference on*. IEEE, 2016, pp. 3832–3836. (Cited on pages vii and 4.)
- [10] L. R. Borges, L. Azzari, P. R. Bakic, A. D. Maidment, M. A. Vieira, and A. Foi, “Restoration of low-dose digital breast tomosynthesis,” *Measurement Science and Technology*, vol. 29, no. 6, p. 064003, 2018. (Cited on pages vii and 3.)
- [11] J. Roels, J. Aelterman, J. De Vylder, H. Luong, Y. Saeys, S. Lippens, and W. Philips, “Noise analysis and removal in 3D electron microscopy,” in *Advances in Visual Computing*. Springer, 2014, pp. 31–40. (Cited on pages viii, 3, and 76.)
- [12] J. Roels, J. Aelterman, H. Luong, S. Lippens, A. Pižurica, Y. Saeys, and W. Philips, “An overview of state-of-the-art image restoration in electron microscopy,” *Journal of microscopy*, 2018. (Cited on pages viii, 3, 76, and 80.)
- [13] P. Getreuer, “Total Variation deconvolution using split Bregman,” *Image Processing On Line*, vol. 2, pp. 158–174, 2012. (Cited on pages x, 7, and 48.)
- [14] P. Rodriguez, “Total Variation regularization algorithms for images corrupted with different noise models: a review,” *Journal of Electrical and Computer Engineering*, vol. 2013, p. 10, 2013. (Cited on pages x and 7.)
- [15] M. Afonso, J. Bioucas-Dias, and M. Figueiredo, “Fast image recovery using variable splitting and constrained optimization,” *IEEE Transactions on Image Processing*, vol. 19, no. 9, pp. 234–2356, 2010. (Cited on pages x and 7.)
- [16] T. F. Chan and C. Wong, “Total Variation blind deconvolution,” *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 370–375, 1998. (Cited on pages x and 7.)
- [17] M. Almeida and M. Figueiredo, “Blind image deblurring with unknown boundaries using the alternating direction method of multipliers,” in *20th IEEE International Conference on Image Processing (ICIP)*, 2013, pp. 586–590. (Cited on pages x, 7, and 48.)
- [18] D. Perrone and P. Favaro, “Total Variation blind deconvolution: The devil is in the details,” in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014, pp. 2909–2916. (Cited on pages x and 7.)

- 
- [19] A. Levin, Y. Weiss, F. Durand, and W. Freeman, “Understanding and evaluating blind deconvolution algorithms,” in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009, pp. 1964–1971. (Cited on pages x and 7.)
- [20] M. K. Ng and N. K. Bose, “Mathematical analysis of super-resolution methodology,” *IEEE Signal Processing Magazine*, vol. 20, no. 3, pp. 6–74, 2003. (Cited on pages x and 7.)
- [21] N. Zhao, Q. Wei, A. Basarab, N. Dobigeon, D. Kouame, and J.-Y. Tourneret, “Fast single image super-resolution using a new analytical solution for  $\ell_2 - \ell_2$  problems,” *IEEE Transactions on Image Processing*, vol. 25, no. 8, pp. 3683–3697, 2016. (Cited on pages x and 7.)
- [22] A. Marquina and S. J. Osher, “Image super-resolution by TV-regularization and Bregman iteration,” *Journal of Scientific Computing*, vol. 37, no. 3, pp. 367–382, 2008. (Cited on pages x and 7.)
- [23] E. Chouzenoux, A. Jeziarska, J.-C. Pesquet, and H. Talbot, “A convex approach for image restoration with exact Poisson-Gaussian likelihood,” *SIAM Journal on Imaging Science*, vol. 8, no. 4, pp. 2662–2682, 2015. (Cited on pages x, xii, 7, 23, 48, 62, and 63.)
- [24] L. Bar, T. F. Chan, G. Chung, M. Jung, N. Kiryati, R. Mohieddine, N. Sochen, and L. A. Vese, “Mumford and Shah model and its applications to image segmentation and image restoration,” in *Handbook of Math. Methods in Imaging*, 2011, pp. 1095–1157. (Cited on pages x, 8, 30, 38, 39, and 41.)
- [25] T. Pock, D. Cremers, H. Bischof, and A. Chambolle, “An algorithm for minimizing the Mumford-Shah functional,” in *Computer Vision, 2009 IEEE 12th International Conference on*. IEEE, 2009, pp. 1133–1140. (Cited on pages x, 8, 39, and 48.)
- [26] E. Y. Sidky, C.-M. Kao, and X. Pan, “Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT,” *Journal of X-ray Science and Technology*, vol. 14, no. 2, pp. 119–139, 2006. (Cited on pages x and 8.)
- [27] J. S. Jørgensen, *Sparse image reconstruction in computed tomography*. PhD thesis, Technical University of Denmark, 2013. (Cited on pages x and 8.)
- [28] M. Benning, L. Gladden, D. Holland, C.-B. Schönlieb, and T. Valkonen, “Phase reconstruction from velocity-encoded MRI measurements—a survey of sparsity-promoting variational approaches,” *Journal of Magnetic Resonance*, vol. 238, pp. 26–43, 2014. (Cited on pages x and 8.)

## Bibliography

---

- [29] F. Knoll, K. Bredies, T. Pock, and R. Stollberger, “Second order Total Generalized Variation (TGV) for MRI,” *Magnetic resonance in medicine*, vol. 65, no. 2, pp. 480–491, 2011. (Cited on pages x and 8.)
- [30] K. Papafitsoros and C. B. Schönlieb, “A combined first and second order variational approach for image reconstruction,” *Journal of mathematical imaging and vision*, vol. 48, no. 2, pp. 308–338, 2014. (Cited on pages x, 8, 37, and 79.)
- [31] T. F. Chan and J. Shen, “Variational image inpainting,” *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, vol. 58, no. 5, pp. 579–619, 2005. (Cited on pages x and 8.)
- [32] P. Getreuer, “Total Variation inpainting using split Bregman,” *Image Processing On Line*, vol. 2, pp. 147–157, 2012. (Cited on pages x and 8.)
- [33] N. Sladoje and J. Lindblad, “Estimation of moments of digitized objects with fuzzy borders,” in *International Conference on Image Analysis and Processing*. Springer, 2005, pp. 188–195. (Cited on pages xi and 8.)
- [34] N. Sladoje, I. Nyström, and P. K. Saha, “Measurements of digitized objects with fuzzy borders in 2D and 3D,” *Image and Vision Comput.*, vol. 23, no. 2, pp. 123–132, 2005. (Cited on pages xi, 8, and 42.)
- [35] N. Sladoje and J. Lindblad, “High-precision boundary length estimation by utilizing gray-level information,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 357–363, 2009. (Cited on pages xi and 8.)
- [36] S. Dražić, N. Sladoje, and J. Lindblad, “Estimation of Feret’s diameter from pixel coverage representation of a shape,” *Pattern Recogn. Lett.*, vol. 80, pp. 37–45, 2016. (Cited on pages xi and 8.)
- [37] V. Ilić, J. Lindblad, and N. Sladoje, “Signature of a shape based on its pixel coverage representation,” in *Int. Conf. on Discrete Geom. for Computer Imagery, LNCS*, vol. 9647, 2016, pp. 181–193. (Cited on pages xi and 8.)
- [38] J. Lindblad and N. Sladoje, “Coverage segmentation based on linear unmixing and minimization of perimeter and boundary thickness,” *Pattern Recognition Letters*, vol. 33, no. 6, pp. 728–738, 2012. (Cited on pages xi, xii, xiv, 8, 9, 10, 29, 38, 42, 44, 53, 86, 87, and 90.)
- [39] E. Birgin, J. Martinez, and M. Raydan, “Nonmonotone spectral projected gradient methods on convex sets,” *SIAM Journal on Optimization*, vol. 10, no. 4, pp. 1196–1211, 2000. (Cited on pages xiii, xiv, 11, 49, and 51.)

- 
- [40] Y. Li, A. Toma, B. Sixou, and F. Peyrin, "Super-resolution/segmentation of 2d trabecular bone images by a Mumford-Shah approach and comparison to total variation," in *Signal Processing Conference (EUSIPCO), 2016 24th European*. IEEE, 2016, pp. 1693–1697. (Cited on pages 84 and 85.)
- [41] H. Lánteri and C. Theys, "Restoration of astrophysical images - the case of Poisson data with additive Gaussian noise," *EURASIP Journal on Advances in Signal Processing*, vol. 2005, pp. 2500–2513, 2005. (Cited on pages 7 and 23.)
- [42] J. Li, R. Shen, R. Yin, and X. Zhan, "A reweighted  $\ell_2$  method for image restoration with Poisson and mixed Poisson-Gaussian noise," *UCLA Computational and Applied Mathematics Reports*, pp. 12–84, 2012. (Cited on pages 7 and 23.)
- [43] J. Li, F. Luisier, and T. Blu, "PURE-LET image deconvolution," *IEEE Transactions on Image Processing*, vol. 27, no. 1, pp. 92–105, 2018. (Cited on page 7.)
- [44] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear Total Variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, pp. 259–268, 1992. (Cited on pages 7, 34, and 37.)
- [45] T. Lukić, J. Lindblad, and N. Sladoje, "Regularized image denoising based on spectral gradient optimization," *Inverse Problems*, vol. 27, p. 085010, 2011. (Cited on pages 7, 35, and 36.)
- [46] M. Bertero and P. Boccacci, *Introduction to inverse problems in imaging*. CRC press, 1998. (Cited on page 13.)
- [47] J. Hadamard, "Sur les problèmes aux dérivées partielles et leur signification physique," *Princeton university bulletin*, pp. 49–52, 1902. (Cited on page 13.)
- [48] P. H. Kvam and B. Vidakovic, *Nonparametric statistics with applications to science and engineering*. John Wiley & Sons, 2007, vol. 653. (Cited on page 14.)
- [49] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, "Advances and challenges in super-resolution," *International Journal of Imaging Systems and Technology*, vol. 14, no. 2, pp. 47–57, 2004. (Cited on pages 23 and 24.)
- [50] V. Patanavijit and S. Jitapunkul, "A Lorentzian stochastic estimation for a robust iterative multiframe super-resolution reconstruction with Lorentzian-Tikhonov regularization," *EURASIP Journal on Advances in Signal Processing*, vol. 2007, no. 2, p. 21, 2007. (Cited on page 23.)

- [51] A. Panagiotopoulou and V. Anastassopoulos, “Super-resolution image reconstruction techniques, trade-offs between the data-fidelity and regularization terms,” *Inf. Fusion*, vol. 13, no. 3, pp. 185–195, 2012. (Cited on pages 23 and 24.)
- [52] P. Purkait and B. Chanda, “Morphologic gain-controlled regularization for edge-preserving super-resolution image reconstruction,” *Signal Image Video Processing*, vol. 7, no. 5, pp. 925–938. (Cited on page 23.)
- [53] H. Song, L. Zhang, P. Wang, K. Zhang, and X. Li, “An adaptive  $\ell_1 - \ell_2$  hybrid error model to super-resolution,” in *In IEEE International Conference on Image Processing, proceedings*, 2010, pp. 2821–2824. (Cited on page 23.)
- [54] L. Yue, H. Shen, Q. Yuan, and L. Zhang, “A locally adaptive  $\ell_1 - \ell_2$  norm for multi-frame super-resolution of images with mixed noise and outliers,” *Signal Processing*, vol. 105, pp. 156–174, 2014. (Cited on page 23.)
- [55] G. Yu, “Variance stabilizing transformations of poisson, binomial and negative binomial distributions,” *Statistics & Probability Letters*, vol. 79, no. 14, pp. 1621–1629, 2009. (Cited on page 24.)
- [56] M. Bartlett, “The square root transformation in analysis of variance,” *Supplement to the Journal of the Royal Statistical Society*, vol. 3, no. 1, pp. 68–78, 1936. (Cited on pages 24 and 25.)
- [57] F. J. Anscombe, “The transformation of Poisson, binomial and negative-binomial data,” *Biometrika*, vol. 35, no. 3/4, pp. 246–254, 1948. (Cited on pages 24 and 25.)
- [58] M. F. Freeman and J. W. Tukey, “Transformations related to the angular and the square root,” *The Annals of Mathematical Statistics*, pp. 607–611, 1950. (Cited on pages 24 and 25.)
- [59] A. Foi, “Direct optimization of nonparametric variance-stabilizing transformations,” *8èmes Rencontres de Statistiques Mathématiques, CIRM, Luminy*, 2008. (Cited on pages 24 and 25.)
- [60] A. Foi, “Optimization of variance-stabilizing transformations,” *Preprint, 2009b*, vol. 94, pp. 1809–1814, 2009. (Cited on page 25.)
- [61] J.-L. Starck, F. Murtagh, and A. Bijaoui, *Image Processing and Data Analysis*. Cambridge University Press, 1998. (Cited on page 26.)
- [62] F. Murtagh, J.-L. Starck, and A. Bijaoui, “Image restoration with noise suppression using a multiresolution support,” *Astronomy and Astrophysics*, vol. 112, pp. 179–189, 1995. (Cited on page 26.)



- [63] A. Pizurica, "Image denoising algorithms: from wavelet shrinkage to non-local collaborative filtering," in *Wiley Encyclopedia of Electrical and Electronics Engineering*. Wiley, 2017, pp. 1–17. (Cited on pages 26 and 80.)
- [64] M. Makitalo and A. Foi, "Optimal inversion of the generalized Anscombe transformation for Poisson-Gaussian noise," *IEEE Transactions on Image Processing*, vol. 22, no. 1, pp. 91–103, 2013. (Cited on page 26.)
- [65] M. Mäkitalo, "Exact unbiased inverse of the Anscombe transformation and its Poisson-Gaussian generalization," *PhD thesis, Tampere University of Technology*, 2013. (Cited on page 26.)
- [66] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Processing Magazine*, vol. 19, no. 1, pp. 44–57, 2002. (Cited on pages 28 and 85.)
- [67] T. Haraguchi, T. Shimi, T. Koujin, N. Hashiguchi, and Y. Hiraoka, "Spectral imaging fluorescence microscopy," *Genes to Cells*, vol. 7, no. 9, pp. 881–887, 2002. (Cited on pages 28 and 85.)
- [68] T. Zimmermann, J. Marrison, K. Hogg, and P. O' Toole, "Clearing up the signal: spectral imaging and linear unmixing in fluorescence microscopy," *Confocal Microscopy: Methods and Protoc.*, pp. 129–148, 2014. (Cited on pages 28 and 85.)
- [69] A. Villa, J. Chanussot, J. A. Benediktsson, and C. Jutten, "Spectral unmixing for the classification of hyperspectral images at a finer spatial resolution," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 3, pp. 521–533, 2011. (Cited on pages 28, 85, 86, and 87.)
- [70] S. Verbeiren, H. Eerens, I. Piccard, I. Bauwens, and J. Van Orshoven, "Sub-pixel classification of spot-vegetation time series for the assessment of regional crop areas in Belgium," *Int. J. of Appl. Earth Obs. and Geoinf.*, vol. 10, no. 4, pp. 486–497, 2008. (Cited on pages 28 and 85.)
- [71] R. Wang, H.-C. Li, A. Pizurica, J. Li, A. Plaza, and W. J. Emery, "Hyperspectral unmixing using double reweighted sparse regression and Total Variation," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 7, pp. 1146–1150, 2017. (Cited on pages 28 and 85.)
- [72] L. Bar, N. Sochen, and N. Kiryati, "Variational pairing of image segmentation and blind restoration," in *European Conference on Computer Vision*. Springer, 2004, pp. 166–177. (Cited on page 30.)
- [73] B. M. Williams, J. A. Spencer, K. Chen, Y. Zheng, and S. Harding, "An effective variational model for simultaneous reconstruction and segmentation

- of blurred images,” *Journal of Algorithms & Computational Technology*, vol. 10, no. 4, pp. 244–264, 2016. (Cited on page 30.)
- [74] M. Foare, N. Pustelnik, and L. Condat, “A new proximal method for joint image restoration and edge detection with the Mumford-Shah model,” in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2018. (Cited on pages 30 and 39.)
- [75] M. Storath, A. Weinmann, J. Frikel, and M. Unser, “Joint image reconstruction and segmentation using the Potts model,” *Inverse Problems*, vol. 31, no. 2, p. 025003, 2015. (Cited on page 30.)
- [76] Z. Zhang, Y. Xu, J. Yang, X. Li, and D. Zhang, “A survey of sparse representation: algorithms and applications,” *IEEE access*, vol. 3, pp. 490–530, 2015. (Cited on page 31.)
- [77] S. Mallat, *A wavelet tour of signal processing: the sparse way*. Academic press, 2008. (Cited on page 32.)
- [78] J.-L. Starck, F. Murtagh, and J. M. Fadili, *Sparse image and signal processing: wavelets, curvelets, morphological diversity*. Cambridge university press, 2010. (Cited on page 32.)
- [79] A. Chambolle, R. A. DeVore, N.-Y. Lee, and B. J. Lucier, “Nonlinear wavelet image processing: variational problems, compression, and noise removal through wavelet shrinkage,” *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 319–335, 1998. (Cited on page 32.)
- [80] M. A. Figueiredo and R. D. Nowak, “An em algorithm for wavelet-based image restoration,” *IEEE Transactions on Image Processing*, vol. 12, no. 8, pp. 906–916, 2003. (Cited on page 32.)
- [81] J.-L. Starck, E. J. Candès, and D. L. Donoho, “The curvelet transform for image denoising,” *IEEE Transactions on image processing*, vol. 11, no. 6, pp. 670–684, 2002. (Cited on page 32.)
- [82] G. Kutyniok and D. Labate, “Introduction to shearlets,” in *Shearlets*. Springer, 2012, pp. 1–38. (Cited on page 32.)
- [83] V. M. Patel, G. R. Easley, and D. M. Healy, “Shearlet-based deconvolution,” *IEEE Transactions on Image Processing*, vol. 18, no. 12, pp. 2673–2685, 2009. (Cited on page 32.)
- [84] J. Qin, X. Yi, S. Weiss, and S. Osher, “Shearlet-TGV based fluorescence microscopy image deconvolution,” *CAM Report. University of California, Los Angeles (UCLA)*, pp. 14–32, 2014. (Cited on page 32.)

- [85] E. Le Pennec and S. Mallat, “Sparse geometric image representations with bandelets,” *IEEE transactions on image processing*, vol. 14, no. 4, pp. 423–438, 2005. (Cited on page 32.)
- [86] M. N. Do and M. Vetterli, “Contourlets: a directional multiresolution image representation,” in *Image Processing. 2002. Proceedings. 2002 International Conference on*, vol. 1. IEEE, 2002, pp. I–I. (Cited on page 32.)
- [87] S. S. Chen, D. L. Donoho, and M. A. Saunders, “Atomic decomposition by basis pursuit,” *SIAM review*, vol. 43, no. 1, pp. 129–159, 2001. (Cited on page 33.)
- [88] A. Chambolle and T. Pock, “An introduction to continuous optimization for imaging,” *Acta Numerica*, vol. 25, pp. 161–319, 2016. (Cited on pages 33, 40, 41, and 48.)
- [89] P. Charbonnier, L. Blanc-Féraud, G. Aubert, and M. Barlaud, “Deterministic edge-preserving regularization in computed imaging,” *IEEE Transactions on Image Processing*, vol. 6, no. 2, pp. 298–311, 1997. (Cited on pages 35 and 36.)
- [90] S. Geman and D. McClure, “Bayesian image analysis: an application to single photon emission tomography,” *Proc. Statistical Computation Section, Am. Statistical Assoc. (Washington, DC)*, pp. 12–18, 1985. (Cited on pages 35 and 37.)
- [91] T. Hebert and R. Leahy, “A generalized EM algorithm for 3-D Bayesian reconstruction from Poisson data using Gibbs priors,” *IEEE Transactions on Medical Imaging*, vol. 8, no. 2, pp. 194–202, 1989. (Cited on pages 35 and 37.)
- [92] M. Nikolova and R. Chan, “The equivalence of half-quadratic minimization and the gradient linearization iteration,” *IEEE Transactions on Image Processing*, vol. 16, pp. 1623–1627, 2007. (Cited on pages 35 and 37.)
- [93] M. Nikolova and M. K. Ng, “Analysis of half-quadratic minimization methods for signal and image recovery,” *SIAM Journal on Scientific computing*, vol. 27, no. 3, pp. 937–966, 2005. (Cited on pages 35 and 36.)
- [94] P. Perona and J. Malik, “Scale-space and edge detection using anisotropic diffusion,” *IEEE Transactions on pattern analysis and machine intelligence*, vol. 12, no. 7, pp. 629–639, 1990. (Cited on pages 35 and 37.)
- [95] A. Tikhonov and V. Arsenin, “Solutions of ill-posed problems,” *Washington, DC: Winston and Wiley*, 1977. (Cited on pages 35 and 37.)

- [96] R. Schultz and R. Stevenson, “Stochastic modeling and estimation of multispectral image data,” *IEEE Transactions on Image Processing*, vol. 4, no. 8, pp. 1109–1119, 1995. (Cited on pages 35 and 37.)
- [97] G. Gilboa and S. Osher, “Nonlocal operators with applications to image processing,” *Multiscale Modeling & Simulation*, vol. 7, no. 3, pp. 1005–1028, 2008. (Cited on page 37.)
- [98] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, “Image denoising by sparse 3-D transform-domain collaborative filtering,” *IEEE Transactions on image processing*, vol. 16, no. 8, pp. 2080–2095, 2007. (Cited on pages 37 and 80.)
- [99] A. Buades, B. Coll, and J.-M. Morel, “A review of image denoising algorithms, with a new one,” *Multiscale Modeling & Simulation*, vol. 4, no. 2, pp. 490–530, 2005. (Cited on page 37.)
- [100] G. Peyré, S. Bougleux, and L. D. Cohen, “Non-local regularization of inverse problems,” *Inverse Problems and Imaging*, vol. 5, no. 2, pp. 511–530, 2011. (Cited on page 37.)
- [101] K. Bredies, K. Kunisch, and T. Pock, “Total Generalized Variation,” *SIAM Journal on Imaging Sciences*, vol. 3, no. 3, pp. 492–526, 2010. (Cited on pages 37 and 79.)
- [102] D. Mumford and J. Shah, “Optimal approximations by piecewise smooth functions and associated variational problems,” *Commun. on Pure and Appl. Math.*, vol. 42, no. 5, pp. 577–685, 1989. (Cited on pages 12, 38, and 39.)
- [103] J.-M. Morel and S. Solimini, *Variational methods in image segmentation: with seven image processing experiments*. Springer Science & Business Media, 2012, vol. 14. (Cited on page 38.)
- [104] A. Haddad and Y. Meyer, “Variational methods in image processing,” *UCLA CAM Report*, pp. 04–52, 2004. (Cited on page 38.)
- [105] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen, *Variational methods in imaging*. Springer, 2009. (Cited on page 38.)
- [106] M. Nikolova, “Energy minimization methods,” in *Handbook of mathematical methods in imaging*. Springer, 2011, pp. 139–185. (Cited on page 38.)
- [107] T. F. Chan and L. A. Vese, “Active contours without edges,” *IEEE Transactions on image processing*, vol. 10, no. 2, pp. 266–277, 2001. (Cited on pages 39 and 40.)

- 
- [108] P. Getreuer, “Chan-Vese segmentation,” *Image Processing On Line*, vol. 2, pp. 214–224, 2012. (Cited on page 39.)
- [109] L. Ambrosio and V. M. Tortorelli, “Approximation of functional depending on jumps by elliptic functional via t-convergence,” *Communications on Pure and Applied Mathematics*, vol. 43, no. 8, pp. 999–1036, 1990. (Cited on page 39.)
- [110] A. Chambolle, “Finite-differences discretizations of the Mumford-Shah functional,” *ESAIM: Mathematical Modelling and Numerical Analysis*, vol. 33, no. 2, pp. 261–288, 1999. (Cited on page 39.)
- [111] B. Bourdin and A. Chambolle, “Implementation of an adaptive finite-element approximation of the Mumford-Shah functional,” *Numerische Mathematik*, vol. 85, no. 4, pp. 609–646, 2000. (Cited on page 39.)
- [112] M. Burger, A. Sawatzky, and G. Steidl, “First order algorithms in variational image processing,” in *Splitting Methods in Communication, Imaging, Science, and Engineering*. Springer, 2016, pp. 345–407. (Cited on page 48.)
- [113] J. Douglas and H. H. Rachford, “On the numerical solution of heat conduction problems in two and three space variables,” *Transactions of the American mathematical Society*, vol. 82, no. 2, pp. 421–439, 1956. (Cited on page 48.)
- [114] R. Glowinski and A. Marroco, “Sur l’approximation, par éléments finis d’ordre un, et la résolution, par pénalisation-dualité d’une classe de problèmes de dirichlet non linéaires,” *Revue française d’automatique, informatique, recherche opérationnelle. Analyse numérique*, vol. 9, no. R2, pp. 41–76, 1975. (Cited on page 48.)
- [115] P. L. Combettes and J.-C. Pesquet, “Proximal splitting methods in signal processing,” in *Fixed-point algorithms for inverse problems in science and engineering*. Springer, 2011, pp. 185–212. (Cited on page 48.)
- [116] N. Parikh, S. Boyd *et al.*, “Proximal algorithms,” *Foundations and Trends in Optimization*, vol. 1, no. 3, pp. 127–239, 2014. (Cited on page 48.)
- [117] M. Zhu and T. Chan, “An efficient primal-dual hybrid gradient algorithm for Total Variation image restoration,” *UCLA CAM Report*, vol. 34, 2008. (Cited on page 48.)
- [118] E. Esser, “Applications of Lagrangian-based alternating direction methods and connections to split Bregman,” *CAM report*, vol. 9, p. 31, 2009. (Cited on page 48.)

- [119] J. Eckstein and D. P. Bertsekas, “On the Douglas–Rachford splitting method and the proximal point algorithm for maximal monotone operators,” *Mathematical Programming*, vol. 55, no. 1-3, pp. 293–318, 1992. (Cited on page 48.)
- [120] D. Gabay and B. Mercier, *A dual algorithm for the solution of non linear variational problems via finite element approximation*. Institut de recherche d’informatique et d’automatique, 1975. (Cited on page 48.)
- [121] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein *et al.*, “Distributed optimization and statistical learning via the alternating direction method of multipliers,” *Foundations and Trends in Machine learning*, vol. 3, no. 1, pp. 1–122, 2011. (Cited on page 48.)
- [122] T. Goldstein and S. Osher, “The split Bregman method for  $\ell_1$ -regularized problems,” *SIAM journal on imaging sciences*, vol. 2, no. 2, pp. 323–343, 2009. (Cited on page 48.)
- [123] X. Zhang, M. Burger, X. Bresson, and S. Osher, “Bregmanized nonlocal regularization for deconvolution and sparse reconstruction,” *SIAM Journal on Imaging Sciences*, vol. 3, no. 3, pp. 253–276, 2010. (Cited on page 48.)
- [124] M. Figueiredo and J. Bioucas-Dias, “Restoration of Poissonian images using alternating direction optimization,” *IEEE Transactions on Image Processing*, vol. 19, no. 12, pp. 3133–3145, 2010. (Cited on page 48.)
- [125] F.-X. Dupe, M. Fadili, and J.-L. Starck, “A proximal iteration for deconvolving Poisson images using sparse representations,” *IEEE Transactions on Image Processing*, vol. 18, no. 2, pp. 310–321, 2009. (Cited on page 48.)
- [126] J. Bioucas-Dias, F. Condessa, and J. Kovačević, “Alternating direction optimization for image segmentation using hidden Markov measure field models,” in *Image Processing: Algorithms and Systems XII*, vol. 9019. International Society for Optics and Photonics, 2014, p. 90190P. (Cited on page 48.)
- [127] Y. Nesterov, “Smooth minimization of non-smooth functions,” *Mathematical programming*, vol. 103, no. 1, pp. 127–152, 2005. (Cited on page 48.)
- [128] S. Becker, J. Bobin, and E. J. Candès, “NESTA: A fast and accurate first-order method for sparse recovery,” *SIAM Journal on Imaging Sciences*, vol. 4, no. 1, pp. 1–39, 2011. (Cited on page 48.)
- [129] T. Pock and S. Sabach, “Inertial proximal alternating linearized minimization (iPALM) for nonconvex and nonsmooth problems,” *SIAM Journal on Imaging Sciences*, vol. 9, no. 4, pp. 1756–1787, 2016. (Cited on page 48.)

- [130] M. Nikolova and P. Tan, “Alternating proximal gradient descent for non-convex regularised problems with multiconvex coupling terms,” 2017. (Cited on page 48.)
- [131] Y. Wang, W. Yin, and J. Zeng, “Global convergence of ADMM in nonconvex nonsmooth optimization,” *Journal of Scientific Computing*, pp. 1–35, 2015. (Cited on page 48.)
- [132] Y. Boykov and V. Kolmogorov, “An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 26, no. 9, pp. 1124–1137, 2004. (Cited on page 48.)
- [133] Y. Boykov, O. Veksler, and R. Zabih, “Fast approximate energy minimization via graph cuts,” *IEEE Transactions on pattern analysis and machine intelligence*, vol. 23, no. 11, pp. 1222–1239, 2001. (Cited on page 48.)
- [134] L. Bottou, F. E. Curtis, and J. Nocedal, “Optimization methods for large-scale machine learning,” *SIAM Review*, vol. 60, no. 2, pp. 223–311, 2018. (Cited on page 48.)
- [135] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004. (Cited on page 49.)
- [136] E. Birgin, J. Martinez, and M. Raydan, “Algorithm: 813: SPG - software for convex-constrained optimization,” *ACM. Trans. Mathematical Software*, vol. 27, no. 3, pp. 340–349, 2001. (Cited on pages 49 and 51.)
- [137] M. Bertero, P. Boccacci, G. Desidera, and G. Vicidomini, “Image deblurring with Poisson data: from cells to galaxies,” *Inverse Problems*, vol. 25, no. 12, p. 123006, 2009. (Cited on page 49.)
- [138] E. G. Birgin and J. M. Mart\_nez, *Practical augmented Lagrangian methods for constrained optimization*. SIAM, 2014, vol. 10. (Cited on page 49.)
- [139] D. Bertsekas, “On the Goldstein-Levitin-Polyak gradient projection method,” *IEEE Transactions on Automatic Control*, vol. 21, no. 2, pp. 174–184, 1976. (Cited on page 50.)
- [140] L. Grippo, F. Lampariello, and S. Lucidi, “A nonmonotone line search technique for Newton’s method,” *SIAM Journal on Numerical Analysis*, vol. 23, no. 4, pp. 707–716, 1986. (Cited on page 50.)
- [141] J. Barzilai and J. M. Borwein, “Two-point step size gradient methods,” *IMA journal of numerical analysis*, vol. 8, no. 1, pp. 141–148, 1988. (Cited on page 51.)

## Bibliography

---

- [142] E. G. Birgin, J. M. Martínez, M. Raydan *et al.*, “Spectral projected gradient methods: review and perspectives,” *J. Stat. Softw*, vol. 60, no. 3, pp. 1–21, 2014. (Cited on page 51.)
- [143] L. Armijo, “Minimization of functions having Lipschitz continuous first partial derivatives,” *Pacific Journal of mathematics*, vol. 16, no. 1, pp. 1–3, 1966. (Cited on page 51.)
- [144] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004. (Cited on page 57.)
- [145] S. Ramani, T. Blu, and M. Unser, “Monte-Carlo SURE: A black-box optimization of regularization parameters for general denoising algorithms,” *IEEE Transactions on Image Processing*, vol. 17, no. 9, pp. 1540–1554, 2008. (Cited on page 58.)
- [146] Y. C. Eldar, “Generalized SURE for exponential families: applications to regularization,” *IEEE Transactions on Image Processing*, vol. 57, no. 2, pp. 471–481, 2009. (Cited on page 58.)
- [147] P. C. Hansen and D. P. O’Leary, “The use of the L-curve in the regularization of discrete ill-posed problems,” *SIAM Journal on Scientific Computing*, vol. 14, no. 6, pp. 1487–1503, 1993. (Cited on page 58.)
- [148] R. Giryes, M. Elad, and Y. C. Eldar, “The projected GSURE for automatic parameter tuning in iterative shrinkage methods,” *Applied and Computational Harmonic Analysis*, vol. 30, no. 3, pp. 407–422, 2011. (Cited on page 58.)
- [149] M. Almeida and M. Figueiredo, “Parameter estimation for blind and non-blind deblurring using residual whiteness measures,” *IEEE Transactions on Image Processing*, vol. 20, no. 7, pp. 2751–2763, 2013. (Cited on page 58.)
- [150] P. C. Hansen, M. Kilmer, and R. H. Kjeldsen, “Exploiting residual information in the parameter choice for discrete ill-posed problems,” *BIT Numerical Mathematics*, vol. 46, no. 1, pp. 41–59, 2006. (Cited on page 58.)
- [151] X. Zhu and P. Milanfar, “Automatic parameter selection for denoising algorithms using a no-reference measure of image content,” *IEEE Transactions on Image Processing*, vol. 19, no. 12, pp. 3116–3132, 2010. (Cited on page 58.)
- [152] J. C. Lagarias, J. Reeds, M. H. Wright, and P. Wright, “Convergence properties of the Nelder-Mead Simplex method in low dimensions,” *SIAM Journal of Optimization*, vol. 9, no. 1, pp. 112–147, 1998. (Cited on page 58.)



- [153] Y. Marnissi, Y. Zheng, and J.-C. Pesquet, “Fast Variational Bayesian signal recovery in the presence of Poisson-Gaussian noise,” in *41st IEEE International Conference on Acoustics, Speech and Signal Processing ICASSP 2016*, 2016, pp. 3964–3968. (Cited on pages 62 and 63.)
- [154] N. Sladoje and J. Lindblad, “The coverage model and its use in image processing,” *Zbornik Radova*, no. 23, pp. 39–117, 2012. (Cited on page 68.)
- [155] P. Campisi and K. Egiazarian, *Blind image deconvolution: theory and applications*. CRC press, 2016. (Cited on page 72.)
- [156] B. Bajić, J. Lindblad, and N. Sladoje, “Blind restoration of images degraded with mixed Poisson-Gaussian noise with application in transmission electron microscopy,” in *In IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI 2016)*, 2016, pp. 123–127. (Cited on page 72.)
- [157] E. M. Hendrix, I. García, J. Plaza, G. Martín, and A. Plaza, “A new minimum-volume enclosing algorithm for endmember identification and abundance estimation in hyperspectral data,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 7, pp. 2744–2757, 2012. (Cited on page 72.)
- [158] J. M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. Nasrabadi, and J. Chanussot, “Hyperspectral remote sensing data analysis and future challenges,” *IEEE Geoscience and remote sensing magazine*, vol. 1, no. 2, pp. 6–36, 2013. (Cited on page 72.)
- [159] J. M. Nascimento and J. M. Dias, “Vertex component analysis: A fast algorithm to unmix hyperspectral data,” *IEEE transactions on Geoscience and Remote Sensing*, vol. 43, no. 4, pp. 898–910, 2005. (Cited on page 72.)
- [160] M. E. Winter, “N-FINDR: An algorithm for fast autonomous spectral end-member determination in hyperspectral data,” in *Imaging Spectrometry V*, vol. 3753. International Society for Optics and Photonics, 1999, pp. 266–276. (Cited on page 72.)
- [161] J. J. Bozzola and L. D. Russell, *Electron microscopy: principles and techniques for biologists*. Jones & Bartlett Learning, 1999. (Cited on page 75.)
- [162] M. Vulović, E. Franken, R. B. G. Ravelli, L. J. van Vliet, and B. Rieger, “Precise and unbiased estimation of astigmatism and defocus in transmission electron microscopy,” *Ultramicroscopy*, vol. 116, pp. 115–134, 2012. (Cited on page 76.)

## Bibliography

---

- [163] B. Berkels and B. Wirth, “Joint denoising and distortion correction of atomic scale scanning transmission electron microscopy images,” *Inverse Problems*, vol. 33, no. 9, p. 095002, 2017. (Cited on page 76.)
- [164] J. M. Brown and G. B. Witman, “Cilia and diseases,” *Bioscience*, vol. 64, no. 12, pp. 1126–1137, 2014. (Cited on page 78.)
- [165] A. Barbato, T. Frischer, C. Kuehni, D. Snijders, I. Azevedo, G. Baktai, L. Bartoloni, E. Eber, A. Escribano, E. Haarman *et al.*, “Primary ciliary dyskinesia: a consensus statement on diagnostic and treatment approaches in children,” *European Respiratory Journal*, vol. 34, no. 6, pp. 1264–1276, 2009. (Cited on page 78.)
- [166] A. Gupta, A. Suveer, J. Lindblad, A. Dragomir, I.-M. Sintorn, and N. Sladoje, “Convolutional neural networks for false positive reduction of automatically detected cilia in low magnification TEM images,” in *Scandinavian Conference on Image Analysis*. Springer, 2017, pp. 407–418. (Cited on page 78.)
- [167] J. Lindblad, N. Sladoje, A. Suveer, A. Dragomir, and I.-M. Sintorn, “High-resolution reconstruction by feature distance minimization from multiple views of an object,” in *International Conference on Image Processing Theory, Tools and Applications (IPTA)*, 2015, pp. 29–34. (Cited on pages 78 and 79.)
- [168] A. Suveer, N. Sladoje, J. Lindblad, A. Dragomir, and I.-M. Sintorn, “Automated detection of cilia in low magnification transmission electron microscopy images using template matching,” in *13th IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI)*, 2016, pp. 386–390. (Cited on pages 78 and 79.)
- [169] X. Liu, M. Tanaka, and M. Okutomi, “Practical signal dependent noise parameter estimation from a single noisy image,” *IEEE Transactions on Image Processing*, vol. 23, no. 10, pp. 4361–4371, 2014. (Cited on page 79.)
- [170] C. Cruz, A. Foi, V. Katkovnik, and K. Egiazarian, “Nonlocality-reinforced convolutional neural networks for image denoising,” *arXiv preprint arXiv:1803.02112*, 2018. (Cited on page 80.)
- [171] M. Weigert, U. Schmidt, T. Boothe, M. Andreas, A. Dibrov, A. Jain, B. Wilhelm, D. Schmidt, C. Broaddus, S. Culley *et al.*, “Content-aware image restoration: pushing the limits of fluorescence microscopy,” *bioRxiv*, p. 236463, 2017. (Cited on page 80.)

- [172] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, “Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising,” *IEEE Transactions on Image Processing*, vol. 26, no. 7, pp. 3142–3155, 2017. (Cited on page 80.)
- [173] S. Ioffe and C. Szegedy, “Batch normalization: Accelerating deep network training by reducing internal covariate shift,” in *in Proceedings of the International conference on machine learning*, 2015, pp. 448–456. (Cited on page 80.)
- [174] V. Nair and G. E. Hinton, “Rectified linear units improve restricted boltzmann machines,” in *Proceedings of the International conference on machine learning (ICML)*, 2010, pp. 807–814. (Cited on page 80.)
- [175] F. Luisier, T. Blu, and M. Unser, “Image denoising in mixed Poisson–Gaussian noise,” *IEEE Transactions on Image Processing*, vol. 20, no. 3, pp. 696–708, 2011. (Cited on page 80.)
- [176] A. B. Yankovich, B. Berkels, W. Dahmen, P. Binev, and P. M. Voyles, “High-precision scanning transmission electron microscopy at coarse pixel sampling for reduced electron dose,” *Advanced Structural and Chemical Imaging*, vol. 1, no. 1, p. 2, 2015. (Cited on page 81.)
- [177] A. Plaza, Q. Du, J. M. Bioucas-Dias, X. Jia, and F. A. Kruse, “Foreword to the special issue on spectral unmixing of remotely sensed data,” *IEEE transactions on geoscience and remote sensing*, vol. 49, no. 11, pp. 4103–4110, 2011. (Cited on page 85.)
- [178] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, “Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches,” *IEEE journal of selected topics in applied earth observations and remote sensing*, vol. 5, no. 2, pp. 354–379, 2012. (Cited on page 86.)
- [179] M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, “Sparse unmixing of hyperspectral data,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 6, pp. 2014–2039, 2011. (Cited on page 86.)
- [180] E. Seeman and P. D. Delmas, “Bone quality—the material and structural basis of bone strength and fragility,” *New England journal of medicine*, vol. 354, no. 21, pp. 2250–2261, 2006. (Cited on page 84.)
- [181] C. Chen, R. E. Amelon, A. Heiner, and P. K. Saha, “Assessment of trabecular bone strength at in vivo CT imaging with space-variant hysteresis and finite element modelling,” in *Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on*. IEEE, 2016, pp. 872–875. (Cited on page 84.)

- [182] C. Chen, D. Jin, X. Zhang, S. M. Levy, and P. K. Saha, "Robust segmentation of trabecular bone for in vivo CT imaging using anisotropic diffusion and multi-scale morphological reconstruction," in *Medical Imaging 2017: Biomedical Applications in Molecular, Structural, and Functional Imaging*, vol. 10137. International Society for Optics and Photonics, 2017, p. 101371T. (Cited on page 84.)
- [183] J. H. Waarsing, J. S. Day, and H. Weinans, "An improved segmentation method for in vivo  $\mu$ CT imaging," *Journal of bone and mineral research*, vol. 19, no. 10, pp. 1640–1650, 2004. (Cited on page 84.)
- [184] M. Salomé, F. Peyrin, P. Cloetens, C. Odet, A.-M. Laval-Jeantet, J. Baruchel, and P. Spanne, "A synchrotron radiation microtomography system for the analysis of trabecular bone samples," *Medical Physics*, vol. 26, no. 10, pp. 2194–2204, 1999. (Cited on page 85.)
- [185] A. Toma, B. Sixou, and F. Peyrin, "Iterative choice of the optimal regularization parameter in TV image restoration." *Inverse Problems & Imaging*, vol. 9, no. 4, 2015. (Cited on page 85.)
- [186] L. Heinrich, J. A. Bogovic, and S. Saalfeld, "Deep learning for isotropic super-resolution from non-isotropic 3D electron microscopy," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2017, pp. 135–143. (Cited on page 91.)
- [187] J. Tohka, "Partial volume effect modeling for segmentation and tissue classification of brain magnetic resonance images: A review," *World J. of Radiol.*, vol. 6, no. 11, p. 855, 2014. (Cited on page 91.)
- [188] A. Souza, J. K. Udupa, and P. K. Saha, "Volume rendering in the presence of partial volume effects," *IEEE Trans. Med. Imaging*, vol. 24, no. 2, pp. 223–235, 2005. (Cited on page 91.)
- [189] W. J. Niessen, K. L. Vincken, J. Weickert, B. M. ter Haar Romeny, and M. A. Viergever, "Multiscale segmentation of three-dimensional MR brain images," *Int. J. of Computer Vision*, vol. 31, no. 2/3, pp. 185–202, 1999. (Cited on page 92.)



# Appendix

**Publication I**

**Appeared in:**

**Bajić B**, Lindblad J and Sladoje N.

Restoration of images degraded with signal-dependent noise based on energy minimisation: an empirical study. *SPIE Journal of Electronic Imaging*, 2016, 25(4), 043020.

# Restoration of images degraded by signal-dependent noise based on energy minimization: an empirical study

Buda Bajić,<sup>a,\*</sup> Joakim Lindblad,<sup>b,c</sup> and Nataša Sladoje<sup>b,c</sup>

<sup>a</sup>University of Novi Sad, Faculty of Technical Sciences, Trg D. Obradovića 6, 21000 Novi Sad, Serbia

<sup>b</sup>Uppsala University, Centre for Image Analysis, Department of Information Technology, Box 337, 751 05, Uppsala, Sweden

<sup>c</sup>Mathematical Institute, Serbian Academy of Sciences and Arts, Kneza Mihaila 36, Belgrade 11001, Serbia

**Abstract.** Most energy minimization-based restoration methods are developed for signal-independent Gaussian noise. The assumption of Gaussian noise distribution leads to a quadratic data fidelity term, which is appealing in optimization. When an image is acquired with a photon counting device, it contains signal-dependent Poisson or mixed Poisson–Gaussian noise. We quantify the loss in performance that occurs when a restoration method suited for Gaussian noise is utilized for mixed noise. Signal-dependent noise can be treated by methods based on either classical maximum *a posteriori* (MAP) probability approach or on a variance stabilization approach (VST). We compare performances of these approaches on a large image material and observe that VST-based methods outperform those based on MAP in both quality of restoration and in computational efficiency. We quantify improvement achieved by utilizing Huber regularization instead of classical total variation regularization. The conclusion from our study is a recommendation to utilize a VST-based approach combined with regularization by Huber potential for restoration of images degraded by blur and signal-dependent noise. This combination provides a robust and flexible method with good performance and high speed. © 2016 SPIE and IS&T [DOI: 10.1117/1.JEI.25.4.043020]

Keywords: image restoration; Poisson noise; mixed Poisson–Gaussian noise; variance stabilizing transform; total variation; Huber potential function.

Paper 16231 received Mar. 23, 2016; accepted for publication Jul. 12, 2016; published online Aug. 4, 2016.

## 1 Introduction

During acquisition, images are degraded by blur and different types of noise. Most imaging is performed with photon counting devices, e.g., CCD or CMOS sensors, which introduce signal-dependent Poisson noise in combination with signal-independent noise from the electronics. Image restoration aims to reverse the effects of imperfect imaging and to recover noise-free and blur-free images. Image restoration is commonly performed by regularized energy minimization due to its simplicity and generally good performance. The objective energy function consists of a data fidelity term, which models the image formation process and a regularization term, which imposes *a priori* knowledge on the solution.

If a Gaussian noise model is assumed, the data fidelity term is quadratic and easy to optimize. This makes such an assumption appealing and popular. However, in the presence of signal-dependent noise, this model is inappropriate and leads to significantly reduced performance. In this paper, we confirm the importance of a proper choice of noise model, also hinted at in Ref. 1, by empirical tests on a large image material. We show that a correct treatment of both signal-dependent and signal-independent components of the noise is very important, even when one of them constitutes as little as one percent of the overall noise.

The data fidelity term of the energy function should either be adjusted to the correct noise model, or alternatively, the image should be transformed to make the Gaussian model applicable. This latter approach is known as variance stabilization. For pure Poisson noise, the data fidelity term, which

maximizes *a posteriori* probability (MAP), is Kullback–Leibler divergence. Following the alternative path, a variance stabilizing transform (VST), such as Anscombe's,<sup>2</sup> is incorporated in the data term. A limited study, presented in Ref. 3, indicates that the VST approach does not fall behind the direct one. Our presented study here shows that the VST approach actually outperforms the direct approach. This is supported by a thorough evaluation on more than 3500 noisy and blurred images.

For the more general case, when considering a mixture of both Poisson and Gaussian noise components, the MAP approach leads to practical difficulties since the log-likelihood function involves infinite summation. This imposes a need for approximate solutions and leads to complicated and slow algorithms. An often used alternative is to ignore one of the noise components; this comes at a cost of reduced performance, as confirmed by our tests. The generalized VST approach is, on the other hand, straightforward; we demonstrate in this paper that it provides fast and efficient restoration of images degraded by mixed noise.

The regularization part of the energy term is, by definition, image dependent. Several options are proposed, including various sparsity promoting approaches. A popular regularizer is total variation (TV),<sup>4,5</sup> imposing sparsity in the gradient domain. Additional improvement can be achieved by the use of a potential function, which modulates the regularization component. Potential functions are designed to enhance/preserve particular image features; preservation of sharp edges is typically targeted. Our previous study<sup>6</sup>

\*Address all correspondence to: Buda Bajić, E-mail: [buda.bajic@uns.ac.rs](mailto:buda.bajic@uns.ac.rs)

shows that the Huber potential improves performance and outperforms other potential functions for the case of signal-independent Gaussian noise. Here, we evaluate, through a thorough empirical study, how large an improvement is achieved by using the Huber potential for signal-dependent noise.

The main message of our paper is signal-dependent and signal-independent components of image noise have to be treated appropriately in image restoration (Sec. 3.1). A VST-based approach, where the noise is transformed to be approximately signal independent, can successfully be used to handle Poisson and mixed noise. Our thorough empirical study (Sec. 3.2) demonstrates that this approach outperforms direct modeling of the signal-dependent noise.

## 2 Image Restoration by Energy Minimization

Energy minimization is commonly used to address inverse problems in image processing, such as image denoising, restoration, and inpainting. Typically, the energy function of regularized restoration is of the form

$$E(u) = D(u) + \lambda R(u), \tag{1}$$

where  $u$  is an (unknown) image. The energy function consists of a “data fidelity term,”  $D(u)$ , which drives the solution toward the observed data, and a “regularization term,”  $R(u)$ , which provides noise suppression. The “regularization parameter”  $\lambda$  controls the trade-off between the two terms, i. e., the level of smoothing versus faithful recovery of the (possibly noisy) image detail.

We consider that the unknown image  $u$  of size  $r \times c$  is represented as a vector  $u = [u_1, \dots, u_n]^T$  of length  $n = r \times c$ , where image rows are sequentially concatenated. With  $x$  and  $y$ , we denote original and degraded (observed) images, respectively. The convolution with a point spread function (PSF), which models degradation of the image with blur, is equal to multiplication of matrix  $H_{n \times n}$  and image  $x_{n \times 1}$ . Blurred original image,  $Hx$ , is additionally corrupted with noise. In this way, degraded image  $y$  is formed. An argument which minimizes the energy function,

$$\hat{x} = \arg \min_u E(u), \tag{2}$$

is considered to be an estimate of the original image  $x$ .

### 2.1 Data Fidelity Term

When an image  $x$  is degraded by Poisson noise, the observed image  $y$  is described as

$$y_i \sim \mathcal{P}[(Hx)_i], \quad \forall i = 1, 2, \dots, n. \tag{3}$$

The corresponding data fidelity term, derived based on MAP approach, is Kullback–Leibler divergence:

$$D_{\text{MAPP}}(u) = \sum_{i=1}^n \{(Hu)_i - y_i \cdot \log[(Hu)_i]\}. \tag{4}$$

A review of restoration methods for images corrupted with Poisson noise, based on MAP approach, is given in Ref. 7; two more recently developed methods are presented in Refs. 3 and 8.

A more general approach is to also take into account the signal-independent noise sources, modeling the noise with a mixed Poisson–Gaussian distribution. If image  $x$  is degraded by blur and signal-dependent mixed Poisson–Gaussian noise, the acquired image is such that

$$y_i = \theta_i + \eta_i, \quad \forall i = 1, 2, \dots, n, \tag{5}$$

where  $\theta_i \sim \mathcal{P}[(Hx)_i]$  and  $\eta_i \sim \mathcal{N}(0, \sigma_m^2)$ . The corresponding data fidelity term, based on MAP approach, is

$$D_{\text{MAPPG}}(u) = - \sum_{i=1}^n \log \left( \sum_{k=0}^{+\infty} \frac{[(Hu)_i]^k e^{-(Hu)_i}}{k!} \cdot \frac{e^{-\left(\frac{y_i-k}{\sqrt{2}\sigma_m}\right)^2}}{\sqrt{2\pi}\sigma_m} \right). \tag{6}$$

Since this data fidelity term includes an infinite sum, some approximation is required. One way to address this problem is to approximate the infinite sum with a finite number of summands, such as in the methods presented in Refs. 9–12. This usually leads to slow and complicated algorithms.

Another way to overcome practical difficulties related to  $D_{\text{MAPPG}}$  is to use a VST-based approach. For signal-dependent noise, VST can be used to remove signal dependency and render the noise approximately Gaussian. For such a transformed image, the corresponding data fidelity term is quadratic, which is appealing for minimization. When an image  $x$  is degraded by mixed Poisson–Gaussian noise, generalized Anscombe VST<sup>13,14</sup> is used to transform the observed image  $y$  into  $z$ :

$$z_i = 2\sqrt{\max\left\{y_i + \frac{3}{8} + \sigma_m^2, 0\right\}}, \quad \forall i = 1, 2, \dots, n, \tag{7}$$

where

$$z_i \approx 2\sqrt{\max\left\{(Hx)_i + \frac{3}{8} + \sigma_m^2, 0\right\}} + \epsilon, \quad \epsilon \sim \mathcal{N}(0,1), \tag{8}$$

$$\forall i = 1, 2, \dots, n.$$

In this case, the data fidelity term is of the form:

$$D_{\text{VST}}(u) = \sum_{i=1}^n \left( z_i - 2\sqrt{\max\left\{(Hu)_i + \frac{3}{8} + \sigma_m^2, 0\right\}} \right)^2. \tag{9}$$

When  $\sigma_m = 0$ , the data fidelity term [Eq. (9)] treats pure Poisson noise, such as in the method presented in Ref. 1.

We introduce the following abbreviations: VSTP denotes a combination of the VST approach and pure Poisson noise, VSTPG stands for a combination of the VST approach and Poisson–Gaussian noise, while MAPP and MAPPG denote combinations of the MAP approach with Poisson and Poisson–Gaussian noise, respectively.

### 2.2 Regularization

The role of the regularization term is to provide numerical stability and impose desired properties on the solution. Here,



we focus on the family of regularization methods based on TV,<sup>4</sup> i.e., we observe

$$R(u) = \sum_{i=1}^n \Phi(|\nabla(u_i)|), \tag{10}$$

where  $\nabla$  stands for image gradient and  $|\cdot|$  denotes  $\ell_2$  norm. The function  $\Phi$  is referred to as “potential function.” Classical TV regularization is obtained when  $\Phi$  is the identity function,  $\Phi_{TV}(t) = t$ .

In most cases, the potential function is designed so that small intensity changes (assumed to be noise) are penalized while large changes (assumed to be edges) are preserved. A number of potentials are studied and used in image restoration (see Refs. 6 and 15 and references therein). In Ref. 16, theoretical conditions for edge preserving potentials are given.

We have, in Ref. 6, evaluated the effectiveness of seven different potential functions in restoration of images degraded by Gaussian noise. This analysis showed that the largest improvement in image quality is achieved when the Huber<sup>17</sup> potential function

$$\Phi_{\text{Huber}}(t) = \begin{cases} \frac{t^2}{2\omega}, & t \leq \omega \\ t - \frac{\omega}{2}, & t > \omega \end{cases} \tag{11}$$

is utilized. This is a convex function. When the parameter  $\omega$  tends to zero, this function tends to classical TV. The regularization term [Eq. (10)] is differentiable at points where  $|\nabla(u_i)| \neq 0$  for both Huber [Eq. (11)] and TV potential functions. However, at points where  $|\nabla(u_i)| = 0$ , Eq. (10) is differentiable only for Eq. (11).

### 2.3 Optimization

We consider grayscale images and represent them as vectors with intensity values from [0,1] (for images in the range [0, peak] intensity is divided by peak). All considered objective functions are of the form of Eq. (1), where data fidelity terms are given by Eqs. (4) or (9), and a regularization term is given by Eq. (10). Minimization of each objective function is seen as a constrained optimization problem:

$$\min_u E(u) \text{ s.t. } 0 \leq u_i \leq 1, \quad i = 1, 2, \dots, n, \tag{12}$$

which we solve utilizing spectral projected gradient (SPG) optimization.<sup>18</sup> SPG is an efficient method for solving a constrained optimization problem  $\min_{x \in \Omega} f(x)$ , where  $\Omega$  is a closed convex set in  $\mathbb{R}^n$  and the objective function  $f$  has continuous partial derivatives on an open set containing  $\Omega$ .

A wide variety of approaches and different algorithms for minimization of regularized energy functions are presented in the literature; a number of references on the topic are given in Ref. 19. We utilize SPG for the optimization of energy functions since we have previously experienced excellent performance of SPG on restoration of images degraded by Gaussian noise.<sup>6</sup> This flexible method allows a variety of data fidelity terms and potentials to be used in the energy function. We appreciate the simplicity, flexibility, and robustness of SPG.

Pseudocode of the method is given in Algorithm 1. Equations for computing the gradients for all objective

---

#### Algorithm 1 Spectral Projected Gradient.

---

Choose values for parameters:  $\theta_{\min}, \theta_{\max}, \gamma, \sigma_1, \sigma_2, \text{tol}$  s.t.

$$0 < \theta_{\min} < \theta_{\max}, \gamma \in (0,1), 0 < \sigma_1 < \sigma_2 < 1, \text{tol} > 0.$$

Choose initial guess  $x_0 \in \Omega$  and  $\theta_0 = 1$ . Compute  $x_{k+1}$  and  $\theta_{k+1}$  as follows:

$$d_k = P_{\Omega}[x_k - \theta_k \nabla f(x_k)] - x_k; \quad x_{k+1} = x_k + d_k$$

$$\delta = \nabla f(x_k)^T d_k; \quad \lambda_k = 1$$

**while**  $f(x_{k+1}) > f(x_k) + \gamma \lambda_k \delta$

$$\lambda_{\text{temp}} = -\frac{1}{2} \lambda_k^2 \delta / [f(x_{k+1}) - f(x_k) - \lambda_k \delta]$$

**if**  $(\lambda_{\text{temp}} \geq \sigma_1 \wedge \lambda_{\text{temp}} \leq \sigma_2 \lambda_k)$

**then**  $\lambda_k = \lambda_{\text{temp}}$

**else**  $\lambda_k = \lambda_k / 2; \quad x_{k+1} = x_k + \lambda_k d_k$

**end while**

$$s_k = x_{k+1} - x_k; \quad y_k = \nabla f(x_{k+1}) - \nabla f(x_k); \quad \beta_k = s_k^T y_k$$

**if**  $\beta_k \leq 0$

**then**  $\theta_{k+1} = \theta_{\max}$

**else**  $\theta_{k+1} = \min \left\{ \theta_{\max}, \max \left\{ \theta_{\min}, \frac{s_k^T s_k}{\beta_k} \right\} \right\}$

Repeat until:  $\|x_{k+1} - x_k\|_{\infty} \leq \text{tol}$ .

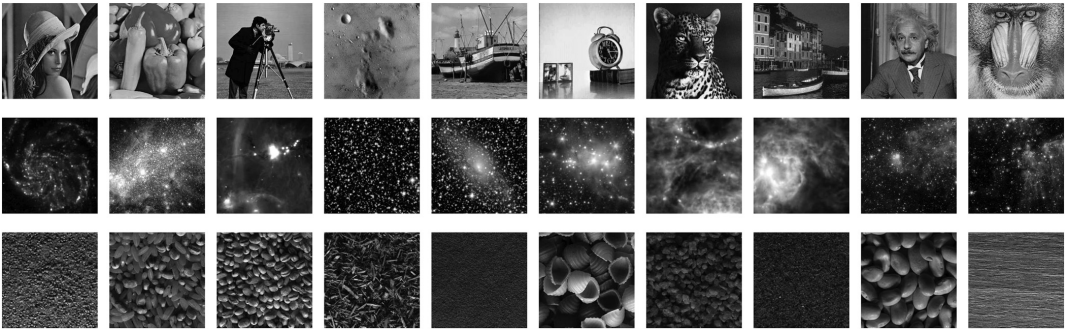
---

functions, required in SPG optimization, are given in the Appendix A.

### 3 Evaluation

For evaluation, we use a dataset consisting of 30 nonblurred and noise-free test images (Fig. 1). We blur each test image  $x$  with 11 Gaussian PSFs  $H$  and obtain blurred observations  $Hx$ , which we further degrade with 11 levels of Poisson noise [Eq. (3)] and with  $11 \times 3$  levels of mixed Poisson–Gaussian noise [Eq. (5)]. We obtain 30 degraded images  $y$  for each PSF and noise level (in total 3630 degraded images for Poisson noise and 10890 for mixed noise). Details about considered degradation levels can be found in Appendix B.

We measure the quality of restoration in terms of PSNR =  $10 \log_{10} \left( \frac{[\max(x_i)]^2}{\text{MSE}} \right)$ , where  $\text{MSE} = \frac{1}{n} \sum_{i=1}^n (x_i - \hat{x}_i)^2$ . Original and restored images are denoted by  $x$  and  $\hat{x}$ , respectively, and restored image  $\hat{x}$  is given by Eq. (2). In general, for a given method, we quantify the improvement in PSNR between noisy input and restored output images:  $\Delta \text{PSNR}_{\text{method}} = \text{PSNR}_{\text{out}} - \text{PSNR}_{\text{in}}$ . Positive  $\Delta \text{PSNR}$  indicates that the resulting restored image has a higher PSNR, i.e., is of a higher quality, than the starting degraded one. We compare different approaches by comparing the improvements in PSNR achieved by them, observing the difference in improvements reached; positive value of  $\Delta \text{PSNR}_{\text{method1}} - \Delta \text{PSNR}_{\text{method2}}$  indicates that “method1” outperforms “method2.” In some



**Fig. 1** Used test images: 10 “classic” test images, 10 astronomical, and 10 texture images. All the images are  $256 \times 256$  pixels.

experiments, we also consider signal-to-noise ratio (SNR),  $SNR = 10 \log_{10} \left( \frac{\frac{1}{N} \sum_{i=1}^N I_i^2}{MSE} \right)$ , as well as structural similarity index measure (SSIM)<sup>20</sup> as image quality measures.

To present and compare the performances of restoration methods in a fair way, it is important to select optimal regularization parameters [ $\lambda$  for methods using TV potential and  $(\lambda, \omega)$  for methods with Huber potential]. Numerous parameter selection schemes are proposed in literature, e.g., SURE-based approaches<sup>21,22</sup> suited for restoration in presence of Gaussian noise, L-curve,<sup>23</sup> generalized cross validation, discrepancy principle,<sup>24</sup> residual-based methods,<sup>25,26</sup> method based on no-reference measure of image content,<sup>27</sup> which can be used for parameter estimation in presence of blur and not-necessarily Gaussian noise. Having ground truth available, we select optimal parameters empirically, and by that avoid the risk of bias from relying on possibly imperfect estimates from methods for parameter selection. We select the best performing parameters for each image and each observed degradation as an argument that maximizes its  $\Delta PSNR$ . For this parameter optimization, we utilize Nelder–Mead simplex search.<sup>28</sup> To verify the generality of our conclusions, as a secondary measure of image quality, we compute the SSIM index of the images restored using parameters that maximize  $\Delta PSNR$ .

### 3.1 Importance of Utilizing the Correct Noise Model

To quantify the importance of appropriate treatment of signal-dependent noise, we restore images degraded by mixed Poisson–Gaussian noise using restoration methods suited for (1) mixed noise (VSTPG) and (2) Gaussian noise (MAPG).<sup>6</sup>

Comparison of performances of VSTPG and MAPG (with TV potential) on the first five images in Fig. 1, degraded by  $4 \times 6$  different blur and mixed noise levels, is presented in Fig. 2. It is clear that the use of an appropriate noise model is very important. We observe a consistent additional improvement in PSNR when assuming the correct noise model (indicated by positive difference in improvements achieved by VSTPG and MAPG in the plot), which goes up to 6 dB and reaches on average 2.42 dB.

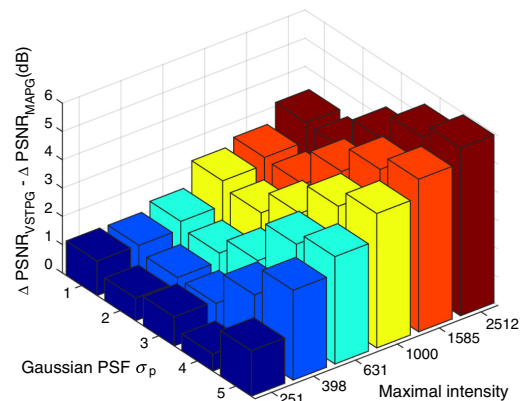
Restoration methods require estimation of the noise present in the image. A number of methods for estimation of noise parameters exist in the literature, e.g., Refs. 29–31 We evaluate the sensitivity of VSTPG with respect to the

parameter  $\sigma_m$  and conclude that an inaccurate estimate of this value affects the restoration result far less than selection of the wrong noise model. On the same data as above, using a value of  $\sigma_m$  that is three times larger than the correct one leads to a reduction in performance by only 0.18 dB.

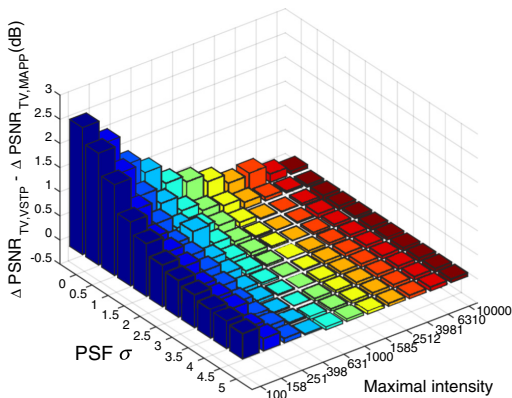
### 3.2 Comparison of Maximum A Posteriori and Variance Stabilization Approaches

For pure Poisson noise, the complexity of the MAP approach, with KL divergence as data term, is comparable with the complexity of the VST approach, which leads to a quadratic data term. It is interesting to study which of these two approaches gives a better performance in practice.

We compare the performances of MAPP and VSTP methods (used with TV potential) on 3630 images degraded by blur and pure Poisson noise. Figure 3 shows the difference in improvement in PSNR between the methods, i.e.,  $\Delta PSNR_{VSTP} - \Delta PSNR_{MAPP}$ . As can be seen from the plot, VSTP method outperforms MAPP, especially for lower Poisson counts (corresponding to higher noise levels). VSTP gives on average 42.1% and 1.1% greater improvement in restoration quality as compared to MAPP, for lower



**Fig. 2** Difference in improvements in PSNR achieved due to appropriate treatment of mixed Poisson–Gaussian noise. Improvement achieved by VSTPG is up to 6 dB higher than by MAPG.



**Fig. 3** Difference in improvements achieved by VSTP and MAPP is presented. Its positive values indicate that VSTP outperforms MAPP.

(maximal pixel intensity  $\leq 1000$ ) and higher counting (maximal pixel intensity  $> 1000$ ) regimes, respectively. The average improvement in PSNR achieved by the VST approach for all degraded images is 0.29 dB greater than that reached by the MAP approach. Our evaluation confirms that the two methods are of similar speed; the average time for recovery of one  $256 \times 256$  pixels image is 1.84 s for VSTP and 1.88 s for the MAPP approach.

For mixed Poisson–Gaussian noise, the MAP approach is of much higher computational complexity. To make comparison of MAPPG and VSTPG feasible, we observe their performances on a small dataset given in Ref. 32; we are reporting the results of MAPPG stated there by the authors.

The dataset contains four images, the first of them is shown in Fig. 4(a). Each test image is degraded by one PSF and one Poisson–Gaussian noise level (corresponding to very low photon count); details are given in Table 1. Figure 4(b) shows the result of the described degradation applied to Fig. 4(a). We apply VSTPG with TV potential for their restoration. Table 1 presents the results. Presented restoration results for VSTPG in Table 1 are obtained with regularization parameters, which optimize the SNR value. The SSIM value is calculated for the restored image obtained with parameters which maximize SNR. The SNR and SSIM values of the restoration results achieved by utilizing MAPPG approach<sup>12</sup> with TV potential, included in Table 1, are taken from Ref. 32.

**Table 1** Comparison of VSTPG and MAPPG<sup>12</sup> for four test images used in Ref. 32. VSTPG outperforms MAPPG in SNR, SSIM, and computational time.

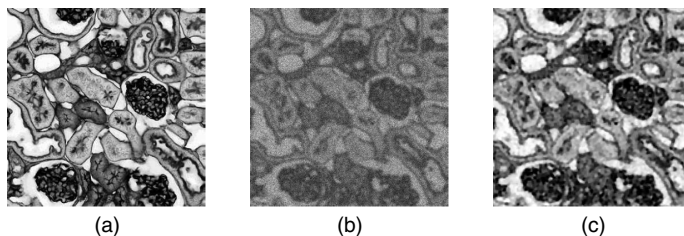
	VSTPG	MAPPG <sup>12</sup>
First image ( $350 \times 350$ ), peak = 20 PSF: Uniform $5 \times 5$ , $\sigma_m^2 = 9$ SNR = 7.64 dB, SSIM = 0.749	SNR (dB) 13.87 SSIM 0.934 Time (s) 4.35	13.73 0.933 48587
Second image ( $257 \times 256$ ), peak = 60 PSF: Gaussian $9 \times 9$ , std 0.5, $\sigma_m^2 = 36$ SNR = 9.40 dB, SSIM = 0.646	SNR (dB) 15.55 SSIM 0.888 Time (s) 2.43	15.43 0.880 351
Third image ( $256 \times 256$ ), peak = 100 PSF: Uniform $3 \times 3$ , $\sigma_m^2 = 36$ SNR = 10.68 dB, SSIM = 0.684	SNR (dB) 14.25 SSIM 0.851 Time (s) 2.30	13.81 0.847 8322
Fourth image ( $256 \times 256$ ), peak = 150 PSF: Gaussian $7 \times 7$ , std 1, $\sigma_m^2 = 40$ SNR = 15.77 dB, SSIM = 0.643	SNR (dB) 20.57 SSIM 0.875 Time (s) 2.27	20.33 0.870 43397

As can be seen, the VSTPG method outperforms MAPPG in terms of SNR, SSIM, and computational time. Restoration of one image by MAPPG takes up to 13.5 h, whereas the VSTPG method takes 4.35 s on the same image and reaches a better restoration quality. Figure 4(c) presents the result of restoration of Fig. 4(b) by VSTPG approach. The images restored by MAPPG can be found in Ref. 32.

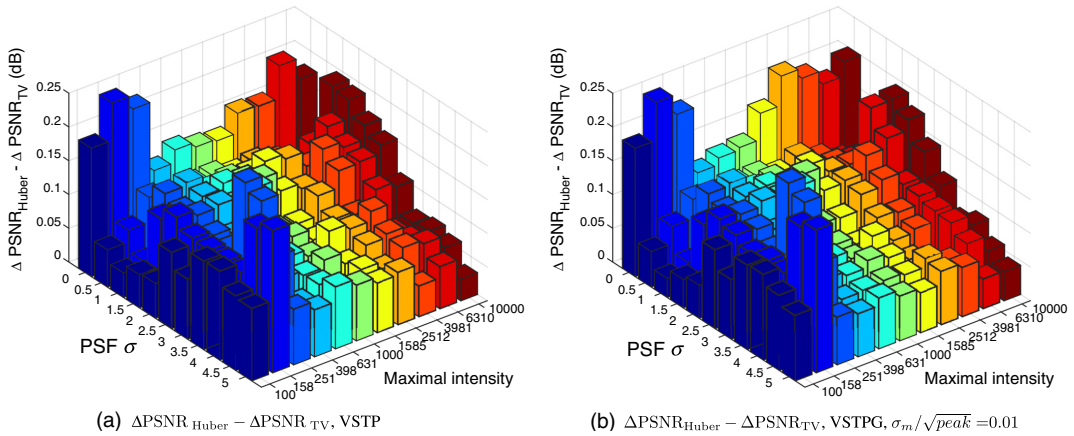
### 3.3 Improvement Achieved from Utilizing Huber Potential Function

An additional way to improve performance of restoration methods is to utilize Huber potential function. We evaluate the difference in PSNR improvement, averaged over 30 images, obtained by using TV and Huber potentials for MAPP, VSTP, and VSTPG. For details about the data used, see Appendix C.

Figure 5(a) presents  $\Delta\text{PSNR}_{\text{Huber}} - \Delta\text{PSNR}_{\text{TV}}$  obtained when VSTP is used for restoration of images degraded by Poisson noise. MAPP exhibits a similar behavior. Huber



**Fig. 4** Performance of VSTPG: (a) original image, (b) degraded image (SNR = 7.64 dB, SSIM = 0.749), and (c) restored image (SNR = 13.87 dB, SSIM = 0.943).



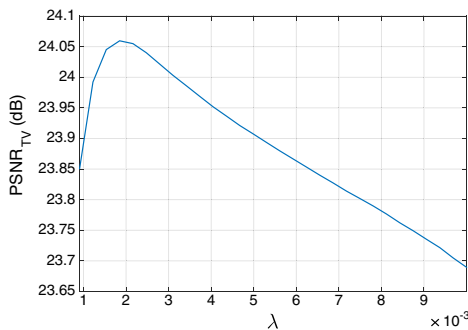
**Fig. 5** Difference in improvements achieved by the method utilizing Huber potential, compared to TV, is shown for (a) VSTP approach and (b) VSTPG approach.

potential gives on average 7.6% and 8.0% greater improvement in restoration quality as compared to TV regularization when used with VSTP and MAPP methods, respectively.

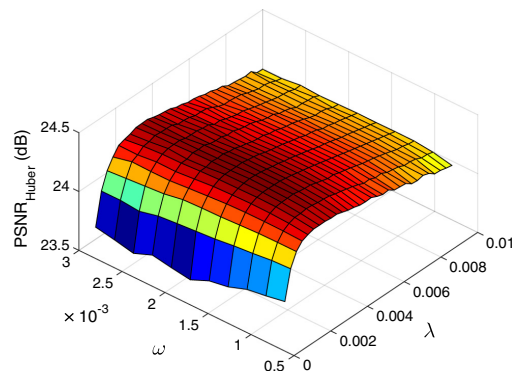
Figure 5(b) presents  $\Delta\text{PSNR}_{\text{Huber}} - \Delta\text{PSNR}_{\text{TV}}$  for VSTPG method for mixed noise with the lowest considered Gaussian noise level. The other two Gaussian noise levels exhibit similar behaviors. Huber potential gives on average 7.8%, 6.8%, and 4.8% greater improvement in restoration quality as compared to TV regularization for considered Gaussian noise levels.

The improvement from using the Huber potential is consistent with the results of the study performed in Ref. 6, where a comparison was made for pure Gaussian noise on a smaller evaluation set (120 degraded images). We have extended this study to 3630 degraded images and observe that Huber gives on average 4.6% greater improvement in restoration quality, as compared to TV, for Gaussian noise.

The performance improvement from the Huber potential comes at the cost of one more parameter to tune. With a good optimization strategy, this cost can be kept reasonably low. Although the optimal values of  $\lambda$  and  $\omega$  vary between images, we observe some consistency. The optimal value of  $\lambda$  is generally slightly larger for the Huber potential than for TV; on average, for all methods and all degraded images, optimal  $\lambda_{\text{Huber}}$  is 1.5 times larger than  $\lambda_{\text{TV}}$ . A typical example is shown in Fig. 6, where optimal  $\lambda_{\text{TV}} = 0.0018$  and  $\lambda_{\text{Huber}} = 0.0032$ . If a good value of  $\lambda_{\text{TV}}$  is known, this provides a good starting guess for  $\lambda_{\text{Huber}}$ . The additional parameter,  $\omega$ , controls the point of transition between  $\ell_2$ , Tikhonov regularization, and  $\ell_1$ , TV regularization. Choosing a too small value of  $\omega$  makes the Huber potential approach TV, with the above observed reduced performance. Selecting a too large value, however, leads to quadratic regularization, which gives a rapid decay in performance and blurred edges as a result.



(a) VSTP with TV



(b) VSTP with Huber

**Fig. 6** Relation between regularization parameter  $\lambda$  for (a) TV and (b) Huber potentials on Peppers image (the second in Fig. 1) degraded by PSF with  $\sigma_p = 1.5$  and Poisson noise with peak = 398, when VSTP is used for restoration. Optimal parameter  $\lambda_{\text{Huber}} = 0.0032$  is slightly larger than  $\lambda_{\text{TV}} = 0.0018$ . Quality of restoration is less sensitive to choice of  $\omega$  for the Huber potential (optimal  $\omega = 0.0016$ ).

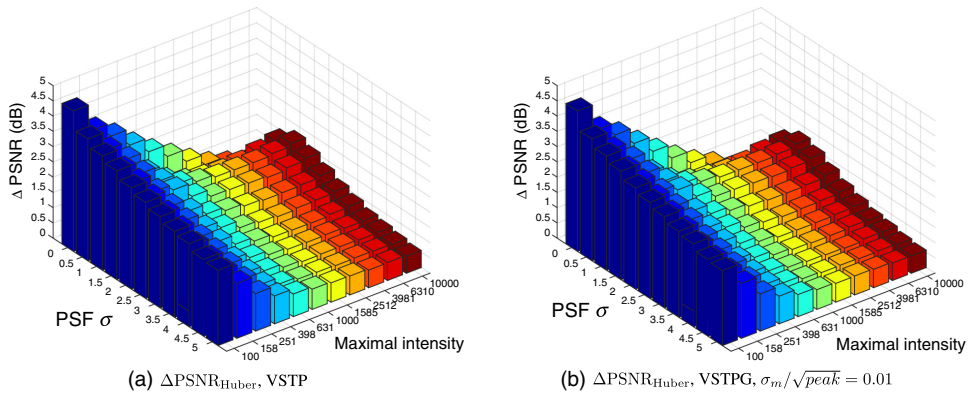


Fig. 7 Improvement in PSNR for (a) VSTP and (b) VSTPG when Huber potential is used, for different levels of blur and noise.

We find that a good initialization for the two-dimensional (2-D) search for optimal parameters for the Huber potential is the optimal  $\lambda$  from the one-dimensional search for the TV potential, combined with a small value of  $\omega$  (e.g.,  $10^{-5}$ , giving a behavior of the Huber potential which is very similar to that of TV). Considering the logarithm of  $\lambda$  and  $\omega$  as optimization variables, instead of  $\lambda$  and  $\omega$ , can help in addressing problems related to the difference in scale of these parameters.

### 3.4 Summary of Results

The performed evaluation shows that the VST-based approach together with the Huber potential function provides a very good combination for restoration of images degraded by signal-dependent noise. Figure 7 illustrates the level of improvement in PSNR obtained when VSTP and VSTPG methods are used in combination with Huber potential function. Table 2 summarizes the average improvement obtained when VST-based approach is applied with Huber potential for restoration of images degraded by signal-dependent noise and different levels of signal-independent noise.

A qualitative evaluation is given in Fig. 8, where restoration of three degraded images from the used dataset are presented. The images are degraded by a PSF with  $\sigma_p = 2$  and Poisson and mixed Poisson–Gaussian noise. The restored images, obtained utilizing the VST method with Huber potential function, are presented together with the reached PSNR and SSIM values. The optimal parameters  $\lambda$  and  $\omega$  that maximize PSNR are also given. SSIM values are calculated for restored images obtained with parameters which maximize PSNR. As can be seen, the presented method

reduces blur and suppresses noise and leads to good restoration results. Time for recovery of one  $256 \times 256$  pixels image in MATLAB on an Intel Core i7-2600 CPU 3.40 GHz is on average 1.8 s. On a limited dataset, we observe that the execution time scales linearly with the image size. Parallel implementation can provide additional speeding-up.













### 3.5 Illustrative Example on a Naturally Degraded Image

To verify applicability of the recommended VSTPG method with the Huber potential function for restoration of images degraded by signal-dependent noise, we utilize it on a naturally degraded image, as shown in Fig. 9(a). We perform restoration on a raw-data  $840 \times 840$  pixels image [Fig. 9(a)] from the CCD sensor of a Samsung digital camera WB550 at ISO 1600. It is generally considered that the main contributions to CCD sensor noise are photon noise (“shot noise”), which is Poisson distributed, and thermal noise, which is well modeled by a Gaussian distribution; our visual inspection confirms that a mixed Poisson–Gaussian noise model is appropriate.

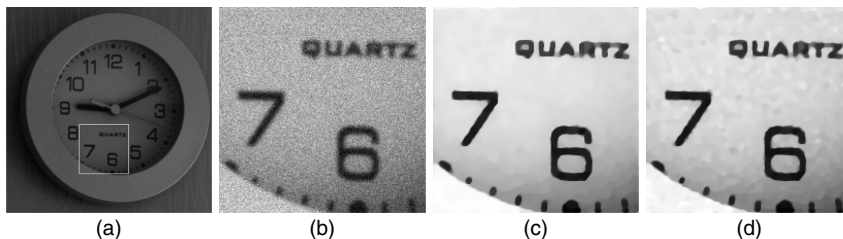
We estimate PSF utilizing the blind restoration method proposed in Ref. 33. This method is suitable for images degraded by mixed noise and enables simultaneous estimation of unknown PSF and the high quality image. The parameter of mixed noise,  $\sigma_m$ , is estimated by a method presented in Ref. 29. This method estimates parameters of noise by using the selected weak textured patches. Since ground truth is now not available, we cannot select regularization parameters by optimizing PSNR. We estimate regularization parameters  $\lambda_{\text{Huber}}$  and  $\omega$  as arguments which maximize the

Table 2 Average improvement in PSNR over all degraded images obtained by restoration using VST approach with Huber potential in regularization for low (peak  $\leq 1000$ ) and high (peak  $> 1000$ ) photon counting regimes.

	Poisson noise		Mixed noise $\sigma_m/\sqrt{\text{peak}} = 0.01$		Mixed noise $\sigma_m/\sqrt{\text{peak}} = 0.1$		Mixed noise $\sigma_m/\sqrt{\text{peak}} = 1$	
	Low	High	Low	High	Low	High	Low	High
Photon count								
VST + Huber	2.21 dB	1.08 dB	2.21 dB	1.06 dB	2.23 dB	1.06 dB	3.69 dB	1.21 dB

	Poisson noise	Mixed noise	Poisson noise	Mixed noise	Poisson noise	Mixed noise
						
PSNR	21.77 dB	20.97 dB	19.86 dB	19.22 dB	22.44 dB	22.18 dB
SSIM	0.308	0.187	0.400	0.339	0.596	0.470
						
PSNR	23.90 dB	23.84 dB	23.36 dB	21.82 dB	24.32 dB	25.68 dB
SSIM	0.492	0.540	0.701	0.591	0.783	0.730
$\lambda$	$5.3 \cdot 10^{-4}$	$5.1 \cdot 10^{-4}$	$1.2 \cdot 10^{-3}$	$1.3 \cdot 10^{-4}$	$1.8 \cdot 10^{-3}$	$5.0 \cdot 10^{-4}$
$\omega$	$5.7 \cdot 10^{-3}$	$7.4 \cdot 10^{-3}$	$2.5 \cdot 10^{-6}$	$6.9 \cdot 10^{-8}$	$1.3 \cdot 10^{-3}$	$1.1 \cdot 10^{-3}$

**Fig. 8** Restoration of blurred and noisy images using VST method with Huber potential. First row contains images degraded by PSF ( $\sigma_p = 2$ ) and two types of noise: Poisson (peak = 631) and mixed (peak = 631,  $\sigma_m/\sqrt{\text{peak}} = 1$ ). Second row contains restored images. Reached PSNR and SSIM values together with optimal parameters  $\lambda$  and  $\omega$  that maximize PSNR are given below each restored image.



**Fig. 9** Performance of VSTPG and MAPG with Huber potential on a part of a raw grayscale image acquired by a digital camera and naturally degraded by mixed Poisson–Gaussian noise: (a) Original  $840 \times 840$  pixels raw-data image. (b) Zoomed in part of size  $200 \times 200$  pixels. (c) Result of restoration of (b) by utilizing VSTPG approach. (d) Result of restoration of (b) by utilizing MAPG approach.

no-reference image content measure  $Q$  proposed in Ref. 27.  $Q$ -measure drops when the amount of noise or/and blur increases in the image; its maximum indicates optimal parameter values. To select optimal parameters in our example, we maximize  $Q$ -measure utilizing Nelder–Mead simplex search.<sup>28</sup>

Results of restoration of Fig. 9(b) are shown in Fig. 9(c), when the VSTPG approach is applied, and in Fig. 9(d), when the MAPG approach is applied. It is clearly visible that VSTPG more efficiently removes noise than MAPG. This example once more confirms that signal dependency of noise should not be neglected in restoration processes.

#### 4 Conclusion

We evaluate performance of energy minimization-based restoration methods for images degraded by signal-dependent noise on a large image material. We show the importance of using the appropriate noise model; if the signal-dependent component is neglected, we observe an average loss of performance of 2.42 dB. Signal-dependent noise can be treated following a direct MAP-based approach, or using a VST approach, which transforms the problem to a Gaussian case. We compare methods derived utilizing VST and MAP and show that the VST-based method outperforms the method based on MAP for Poisson noise with on

average 0.29 dB. For mixed Poisson–Gaussian noise, the difference is 0.24 dB in favor of the presented VSTPG method. A large difference in speed, in favor of the same method, is one more advantage of the VSTPG, which we, for all of these reasons, recommend in this paper.

An additional way to improve performances of TV-based restoration methods for signal-dependent noise is to utilize the Huber potential function in the regularization term. We show that an average improvement of 0.11 dB in restoration performances is achieved when the Huber potential is utilized instead of TV.

The presented VST-based method utilizing the Huber potential function and SPG for optimization is a fast and efficient method for restoration of images degraded by blur and Poisson or Poisson–Gaussian noise.

#### Appendix A: Gradients of Objective Functions

The gradient of the objective function  $E(u)$  is

$$\nabla E(u) = [\nabla E(u)]_{i=1}^n = [\nabla D(u)]_{i=1}^n + \lambda [\nabla R(u)]_{i=1}^n. \tag{13}$$

The gradient of the data fidelity term of MAPP Eq. (4) is

$$\nabla D_{\text{MAPP}}(u) = H^T(1 - y./Hu). \tag{14}$$

Here, the vector whose elements are all equal to 1, is denoted by  $\mathbf{1}$  and  $/$  is element-wise division.

The gradient of the data fidelity term of VST [Eq. (9)] is

$$\nabla D_{\text{VSTPG}}(u) = H^T G(u), \quad (15)$$

$$G(u)_i = \begin{cases} 2 - y_i / \sqrt{(Hu)_i + \frac{3}{8} + \sigma_m^2}, & (Hu)_i + \frac{3}{8} + \sigma_m^2 > 0 \\ 0, & (Hu)_i + \frac{3}{8} + \sigma_m^2 \leq 0 \end{cases}. \quad (16)$$

The gradient of the regularization term [Eq. (10)] is

$$\begin{aligned} \nabla R(u)_i &= \Phi'(|\nabla(u_i)|) \frac{2u_i - u_r - u_b}{|\nabla(u_i)|} + \Phi'(|\nabla(u_i)|) \frac{u_i - u_l}{|\nabla(u_i)|} \\ &\quad + \Phi'(|\nabla(u_a)|) \frac{u_i - u_a}{|\nabla(u_a)|}, \end{aligned} \quad (17)$$

where  $u_a$  and  $u_l$  denote the edge neighbors above and left of the pixel  $u_i$ , respectively. We compute the discrete image gradient at point  $u_i$ ,  $\nabla(u_i)$ , as  $\nabla(u_i) = (u_r - u_i, u_b - u_i)$ , where  $r$  and  $b$  denote indices of the edge neighbors to the right and below the pixel  $u_i$ , respectively. The image edges are handled using a periodic boundary condition.

The gradient of the regularization term with TV potential is nondifferentiable at points where  $|\nabla u_i| = 0$ . To meet the requirements of SPG, we consider a smooth version of Eq. (17) in the case of TV, where  $|\nabla u_i|$  is replaced with  $\sqrt{|\nabla(u_i)|^2 + \varepsilon^2}$  and where  $\varepsilon$  is a small positive number (we use  $\varepsilon = 10^{-5}$  throughout). The use of a relaxed gradient for TV could possibly lead to a less accurate solution; it was observed in Ref. 15 that the differences are negligible.

We assume periodic boundary conditions. All multiplications of vectors with  $H$  are performed in the frequency domain using 2-D fast Fourier transform.

## Appendix B: Spectral Projected Gradient Parameters

We perform optimization using SPG with settings recommended in Ref. 34:  $\theta_{\min} = 10^{-3}$ ,  $\theta_{\max} = 10^3$ ,  $\gamma = 10^{-4}$ ,  $\sigma_1 = 0.1$ , and  $\sigma_2 = 0.9$ . Algorithm 1 terminates when the max-norm between two consecutive images is less than  $\text{tol} = 10^{-5}$  or when the number of iterations reaches 200. We define the projection  $P_\Omega$  of a vector  $x \in \mathbb{R}^n$  to the feasible set  $\Omega = [0, 1]^n$  as  $[P_\Omega(x)]_i = \min\{1, \max\{0, x_i\}\}$ , for all  $i = 1, 2, \dots, n$ .

## Appendix C: Dataset Description

The dataset consists of 30 nonblurred and noise-free test images presented in Fig. 1. We blur each test image with Gaussian PSF, since they closely resemble real PSFs in many imaging systems. We observe PSFs with 11 different standard deviations  $\sigma_p$  between 0 and 5 pixels. We consider images with 11 maximal pixel intensities from 100 to 10,000 i.e., 11 different levels of Poisson noise. Low intensity/photon count corresponds to a high noise level. In addition, for Gaussian noise in the mixed noise model, we choose variance such that the ratio  $\sigma_m / \sqrt{\text{peak}}$  is equal to  $\{0.01, 0.1, 1\}$ .

## Acknowledgments

The authors would like to thank the anonymous reviewers for their helpful comments. The work was jointly supported by the Swedish Research Council, Vetenskapsrådet, and by the Ministry of Education, Science and Technological Development of the Republic of Serbia through Projects ON174008 and III44006 of MI-SANU. Nataša Sladoje is supported by the Swedish Governmental Agency for Innovation Systems (VINNOVA).

## References

1. F.-X. Dupe, M. Fadili, and J.-L. Starck, "A proximal iteration for deconvolving Poisson images using sparse representations," *IEEE Trans. Image Process.* **18**(2), 310–321 (2009).
2. F. J. Anscombe, "The transformation of Poisson, binomial and negative binomial data," *Biometrika* **35**(3/4), 246–254 (1948).
3. F.-X. Dupe, M. Fadili, and J.-L. Starck, "Deconvolution under Poisson noise using exact data fidelity and synthesis or analysis priors," *Statist. Methodol.* **9**(1), 4–18 (2012).
4. L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D* **60**, 259–268 (1992).
5. P. Rodriguez, "Total variation regularization algorithms for images corrupted with different noise models: a review," *J. Electr. Comput. Eng.* **2013**, 10 (2013).
6. B. Bajić, J. Lindblad, and N. Sladoje, "An evaluation of potential functions for regularized image deblurring," *Lect. Notes Comput. Sci.* **8814**, 150–158 (2014).
7. M. Bertero et al., "Image deblurring with Poisson data: from cells to galaxies," *Inverse Probl.* **25**(12), 123006 (2009).
8. M. Figueiredo and J. Bioucas-Dias, "Restoration of Poissonian images using alternating direction optimization," *IEEE Trans. Image Process.* **19**(12), 3133–3145 (2010).
9. H. Lánteri and C. Theys, "Restoration of astrophysical images - the case of Poisson data with additive Gaussian noise," *EURASIP J. Adv. Signal Process.* **2005**, 2500–2513 (2005).
10. F. Benvenuto et al., "The study of an iterative method for the reconstruction of images corrupted by Poisson and Gaussian noise," *Inverse Probl.* **24**(3), 035016 (2008).
11. J. Li et al., "A reweighted l2 method for image restoration with Poisson and mixed Poisson-Gaussian noise," *UCLA Comput. Appl. Math. Rep.* **12–84** (2012).
12. E. Chouzenoux et al., "A convex approach for image restoration with exact Poisson-Gaussian likelihood," *SIAM J. Imaging Sci.* **8**(4), 2662–2682 (2015).
13. J.-L. Starck, F. Murtagh, and A. Bijaoui, *Image Processing and Data Analysis: The Multiscale Approach*, Cambridge University Press, Cambridge (GB) (1998).
14. F. Murtagh, J.-L. Starck, and A. Bijaoui, "Image restoration with noise suppression using a multiresolution support," *Astron. Astrophys.* **112**, 179–189 (1995).
15. T. Lukić, J. Lindblad, and N. Sladoje, "Regularized image denoising based on spectral gradient optimization," *Inverse Probl.* **27**, 085010 (2011).
16. P. Charbonnier et al., "Deterministic edge-preserving regularization in computed imaging," *IEEE Trans. Image Process.* **6**(2), 298–311 (1997).
17. R. Schultz and R. Stevenson, "Stochastic modeling and estimation of multispectral image data," *IEEE Trans. Image Process.* **8**(8), 1109–1119 (1995).
18. E. Birgin, J. Martínez, and M. Raydan, "Nonmonotone spectral projected gradient methods on convex sets," *SIAM J. Optim.* **10**(4), 1196–1211 (2000).
19. P. Rodriguez and B. Wohlberg, "Efficient minimization method for generalized total variation functional," *IEEE Trans. Image Process.* **18**(2), 322–332 (2009).
20. Z. Wang et al., "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Process.* **13**(4), 600–612 (2004).
21. S. Ramani, T. Blu, and M. Unser, "Monte-Carlo SURE: a black-box optimization of regularization parameters for general denoising algorithms," *IEEE Trans. Image Process.* **17**(9), 1540–1554 (2008).
22. Y. C. Eldar, "Generalized SURE for exponential families: applications to regularization," *IEEE Trans. Signal Process.* **57**(2), 471–481 (2009).
23. P. C. Hansen and D. P. O'Leary, "The use of the L-curve in the regularization of discrete ill-posed problems," *SIAM J. Sci. Comput.* **14**(6), 1487–1503 (1993).
24. R. Giryes, M. Elad, and Y. C. Eldar, "The projected GSURE for automatic parameter tuning in iterative shrinkage methods," *Appl. Comput. Harmon. Anal.* **30**(3), 407–422 (2011).

25. M. Almeida and M. Figueiredo, "Parameter estimation for blind and non-blind deblurring using residual whiteness measures," *IEEE Trans. Image Process.* **22**(7), 2751–2763 (2013).
26. P. C. Hansen, M. Kilmer, and R. H. Kjeldsen, "Exploiting residual information in the parameter choice for discrete ill-posed problems," *Bit Numer. Math.* **46**(1), 41–59 (2006).
27. X. Zhu and P. Milanfar, "Automatic parameter selection for denoising algorithms using a no-reference measure of image content," *IEEE Trans. Image Process.* **19**(12), 3116–3132 (2010).
28. J. C. Lagarias et al., "Convergence properties of the Nelder-Mead simplex method in low dimensions," *SIAM J. Optim.* **9**(1), 112–147 (1998).
29. X. Liu, M. Tanaka, and M. Okutomi, "Practical signal dependent noise parameter estimation from a single noisy image," *IEEE Trans. Image Process.* **23**(10), 4361–4371 (2014).
30. A. Foi et al., "Practical Poissonian-Gaussian noise modeling and fitting for single-image raw data," *IEEE Trans. Image Process.* **17**(10), 1737–1754 (2008).
31. A. Jezierska et al., "An EM approach for Poisson-Gaussian noise modeling," in *Proc. 19th IEEE European Signal Processing Conf.*, pp. 2244–2248 (2011).
32. Y. Marnissi, Y. Zheng, and J.-C. Pesquet, "Fast variational Bayesian signal recovery in the presence of Poisson-Gaussian noise," in *IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, pp. 3964–3968 (2016).
33. B. Bajić, J. Lindblad, and N. Sladoje, "Blind restoration of images degraded with mixed Poisson-Gaussian noise with application in transmission electron microscopy," in *Proc. IEEE Int. Symp. on Biomedical Imaging: from Nano to Macro*, pp. 123–127 (2016).
34. E. Birgin, J. Martinez, and M. Raydan, "Algorithm: 813: SPG - software for convex-constrained optimization," *ACM Trans. Math. Softw.* **27**(3), 340–349 (2001).

**Buda Bajić** is currently pursuing her PhD at the Faculty of Technical Sciences, University of Novi Sad, Serbia. She received her BSc and MSc degrees in applied mathematics from the Faculty of Science, University of Novi Sad, Serbia, in 2011 and 2012. Her research interests include image processing, inverse problems, and image restoration.

**Joakim Lindblad** received his MSc degree in engineering physics and PhD in computerized image analysis from Uppsala University, Sweden, in 1997 and 2003, respectively. He is currently affiliated with Protracer AB, Stockholm (head of research), Mathematical Institute SANU, Belgrade (research associate professor), University of Novi Sad (associate professor), and the Centre for Image Analysis, Uppsala University (researcher). His research interests include development of general methods for image segmentation, feature extraction, and classification.

**Nataša Sladoje** received her BSc and MSc degrees in mathematics from the Faculty of Science, University of Novi Sad, Serbia, and PhD in computerized image analysis from the Centre for Image Analysis, SLU, Uppsala, Sweden. Her research interests include development of general image analysis methods, digital geometry, and precise shape analysis. She is currently an associate professor at the Centre for Image Analysis, Uppsala University, at Mathematical Institute SANU, and at the University of Novi Sad.



## Publication II

### Appeared in:

**Bajić B**, Lindblad J and Sladoje N.

Blind restoration of images degraded with mixed Poisson-Gaussian noise with application in Transmission Electron Microscopy. In *Proceedings of the 13<sup>th</sup> IEEE International Symposium on Biomedical Imaging (ISBI)*, Prague, April 2016, pp. 123–127.

# BLIND RESTORATION OF IMAGES DEGRADED WITH MIXED POISSON-GAUSSIAN NOISE WITH APPLICATION IN TRANSMISSION ELECTRON MICROSCOPY

Buda Bajić\*, Joakim Lindblad<sup>†‡</sup>, Nataša Sladoje<sup>†‡</sup>

\* Faculty of Technical Sciences, University of Novi Sad, Serbia

<sup>†</sup> Centre for Image Analysis, Department of Information Technology, Uppsala University, Sweden

<sup>‡</sup> Mathematical Institute, Serbian Academy of Sciences and Arts, Serbia

## ABSTRACT

Noise and blur, present in images after acquisition, negatively affect their further analysis. For image enhancement when the Point Spread Function (PSF) is unknown, blind deblurring is suitable, where both the PSF and the original image are simultaneously reconstructed. In many realistic imaging conditions, noise is modelled as a mixture of Poisson (signal-dependent) and Gaussian (signal independent) noise. In this paper we propose a blind deconvolution method for images degraded by such mixed noise. The method is based on regularized energy minimization. We evaluate its performance on synthetic images, for different blur kernels and different levels of noise, and compare with non-blind restoration. We illustrate the performance of the method on Transmission Electron Microscopy images of cilia, used in clinical practice for diagnosis of a particular type of genetic disorders.

*Index Terms*— deblurring, denoising, Total variation, energy minimization, cilia diseases

## 1. INTRODUCTION

During the acquisition process images are usually degraded by blur and noise. Most imaging devices, like CMOS and CCD cameras, are photon counting devices. The resulting noise is non-additive and signal-dependent and can be modelled by a mixed Poisson-Gaussian (PG) distribution, often encountered in astronomy [1, 2], biology [3], medicine [4].

Image restoration methods aim at estimating an original image from blurred and noisy observation. Restoration methods can be non-blind and blind, depending on whether the Point Spread Function (PSF) is known or not. Examples of blind restoration methods for images degraded with Gaussian noise are presented in [5, 6, 7, 8] whereas Poisson noise is studied in [9]. To the best of our knowledge, no general blind

restoration method for images degraded with PG noise exists in the literature. The recently published method in [10] does assume PG noise, however it is specifically designed for restoration of astronomical images.

Existing restoration methods [1, 2, 11, 12] for images degraded with PG noise are non-blind, i.e., they rely on knowledge, or a good estimate, of the PSF. However, in many situations it is very difficult to accurately estimate the PSF, and blind methods may be preferable. In this paper we present a novel restoration method for images degraded with PG noise, based on energy minimization, which jointly estimates the original image and the PSF from the observed data. We evaluate its performance on synthetic data, and we illustrate its applicability in Transmission Electron Microscopy (TEM).

## 2. BLIND RESTORATION BASED ON ENERGY MINIMIZATION

### 2.1. Energy function

The task of blind restoration is to simultaneously recover an original image and PSF from a blurred and noisy observation. The problem is usually solved through minimization of an energy function of the form:

$$E(u, h) = D(u, h) + \lambda R(u) , \quad (1)$$

where  $u$  and  $h$  are the unknown image and the PSF, respectively. The energy function consists of a *data fidelity term*,  $D(u, h)$ , which drives the solution towards the observed data, and a *regularization term*,  $R(u)$ , which provides noise suppression. The regularization parameter  $\lambda$  controls the trade-off between the two terms, i.e., the level of smoothing vs. faithful recovery of the (possibly noisy) image detail.

#### 2.1.1. Data fidelity term

When the original image  $x$  is degraded with blur and PG noise, the intensity values  $y_i$  of the acquired image  $y$  are of the form

$$y_i = \alpha \theta_i + \eta_i, \quad \forall i = 1, 2, \dots, n \quad (2)$$

B. Bajić, J. Lindblad and N. Sladoje are supported by the Swedish Research Council and by the Ministry of Education, Science and Technological Development of the Republic of Serbia through Projects ON174008 and III44006 of MI-SANU. Sladoje is supported by the Swedish Governmental Agency for Innovation Systems (VINNOVA).

where  $\theta_i \sim \mathcal{P}((Hx)_i)$ ,  $\eta_i \sim \mathcal{N}(\mu, \sigma_m^2)$  and  $\alpha$  is a positive scaling factor. We consider that the image  $x$  of size  $r \times c$  is represented as a vector  $x = [x_1, \dots, x_n]^T$  of length  $n = r \times c$ , where image rows are sequentially concatenated. The convolution, which models degradation of the image with blur, is expressed as multiplication of the matrix  $H_{n \times n}$  and the image  $x_{n \times 1}$ .

When noise is signal-dependent Poisson-Gaussian, the generalized Anscombe stabilizing transformation [13] can be used to remove this signal-dependency and render the noise variance constant throughout the image. Generalized Anscombe transformation of the observed image  $y$  is image  $z$ :

$$z_i = \frac{2}{\alpha} \sqrt{\max\{\alpha y_i + l, 0\}}, \quad \forall i = 1, 2, \dots, n, \quad (3)$$

where  $l = \frac{3}{8}\alpha^2 + \sigma_m^2 - \alpha\mu$ . For  $z_i, \forall i = 1, 2, \dots, n$ , holds

$$z_i \approx \frac{2}{\alpha} \sqrt{\max\{\alpha(Hx)_i + l, 0\}} + \epsilon, \quad \epsilon \sim \mathcal{N}(0, 1). \quad (4)$$

For a signal-independent (Gaussian) noise model, the data fidelity term is quadratic, so the data fidelity term of the observed energy function can, after the transformation, be expressed as

$$D(u, h) = \frac{1}{2} \sum_{i=1}^n \left( z_i - \frac{2}{\alpha} \sqrt{\max\{\alpha(Hu)_i + l, 0\}} \right)^2. \quad (5)$$

A similar approach is followed for the non-blind case in [12].

### 2.1.2. Regularization term

The role of the regularization term is to provide numerical stability and to impose desired properties to the solution. The regularization term of the energy function (1) may be designed based on the desired characteristics of the unknown image  $u$ , such as wavelet-based sparsity, smoothness, small total variation. Here we focus on Total Variation (TV) regularization based methods [14], i.e., we observe

$$R(u) = \sum_{i=1}^n |\nabla(u_i)|, \quad (6)$$

where  $\nabla$  stands for gradient and  $|\cdot|$  denotes  $\ell_2$  norm. We compute the discrete image gradient at point  $u_i$ ,  $\nabla(u_i)$ , as  $\nabla(u_i) = (u_r - u_i, u_b - u_i)$ , where  $r$  and  $b$  denote indices of the edge neighbors to the right and below the pixel  $u_i$ , respectively.

## 2.2. Alternating minimization

It was observed in [5] that the energy function  $E(u, h)$  is not jointly convex, but it is convex w.r.t. one variable when another one is fixed. It is proposed, [5], to minimize the energy function by a so-called Alternating Minimization (AM)

**Table 1:** Algorithm 1.

<b>Alternating Minimization</b>
Start with initial guesses for $u$ and $h$ . Having $u_k$ and $h_k$ , estimate $u_{k+1}$ and $h_{k+1}$ , alternatingly.
<i>Image estimation step:</i> $u_{k+1} = \arg \min_u E(u, h_k)$ s.t. $0 \leq u_i \leq 1, i = 1, 2, \dots, n$ ,
<i>PSF estimation step:</i> $h_{k+1} = \arg \min_h E(u_{k+1}, h)$
Impose constraints on $h_{k+1}$ : $h_{k+1} \geq 0$ and $\ h_{k+1}\ _1 = 1$
Repeat until convergence.

procedure, given in Algorithm 1. Restoration methods presented in [5, 6, 7, 9] estimate the underlying image and PSF in the same alternating fashion. Non-convexity of the function  $E(u, h)$  implies possibility of more than one local minimum. Consequently, a solution of deblurring is not unique; as stated in [5], if  $(u, h)$  is a solution, then, e.g.,  $(-u, -h)$  is also a solution. To get a desired solution and to remove a source of non-uniqueness, some additional constraints are usually imposed. Such are, e.g., non-negativity constraints on both image and PSF, and mass preservation constraint on PSF  $\|h\|_1 = 1$ .

The method we present also utilizes alternating minimization. The underlying image and PSF are estimated in separate steps. Algorithm 1 describes the procedure, as well as the constraints we impose on the PSF estimate. We utilize Spectral Projected Gradient (SPG) optimization method [15]. SPG is a numerical optimization tool for solving a constrained optimization problem  $\min_{x \in \Omega} f(x)$ , where  $\Omega$  is a closed convex set in  $\mathbb{R}^n$  and  $f$  is a function which has continuous partial derivatives on an open set that contains  $\Omega$ . Weak requirements on the objective function, as well as efficiency in solving large scale problems, make SPG attractive for our purpose. SPG algorithm can be found in [16].

## 3. BLIND RESTORATION BY SPG MINIMIZATION

### 3.1. Image estimation

During the image estimation step, minimization of the objective function  $E(u, h)$  is seen as a constrained optimization problem

$$\min_u E(u, h) \quad \text{s.t. } 0 \leq u_i \leq 1, \quad i = 1, 2, \dots, n, \quad (7)$$

that we optimize utilizing SPG.

SPG is an iterative gradient-based method. The gradient of the objective function  $E(u, h)$  w.r.t.  $u$  is

$$[\nabla_u E(u, h)]_{i=1}^n = [\nabla_u D(u, h)]_{i=1}^n + \lambda [\nabla_u R(u)]_{i=1}^n, \quad (8)$$

where

$$\nabla_u D(u, h)_i = \begin{cases} (H^T((2/\alpha)\mathbb{1} - y / \sqrt{\alpha H u + l\mathbb{1}}))_i, & g_i > 0 \\ 0, & g_i \leq 0 \end{cases} \quad (9)$$



**Fig. 1:** Used test images, all are  $256 \times 256$  pixels.

and  $g_i = \alpha(Hu)_i + l$ . The vector whose elements are all equal to 1 is denoted by  $\mathbf{1}$ , while  $\ ./$  and  $\sqrt{\cdot}$  denote element-wise division and element-wise square root, respectively.

The gradient of the regularization term is

$$\nabla_u R(u)_i = \frac{2u_i - u_r - u_b}{|\nabla(u_i)|} + \frac{u_i - u_l}{|\nabla(u)|} + \frac{u_i - u_a}{|\nabla(u_a)|} \quad (10)$$

where  $u_a$  and  $u_l$  denote the edge neighbours above and left of the pixel  $u_i$ , respectively.

The gradient of the regularization term defined by (10) is non-differentiable at points where  $|\nabla u_i| = 0$ . To meet requirements of SPG, we consider a smoothed version of (10), where  $|\nabla u_i|$  is replaced with  $\sqrt{|\nabla(u_i)|^2 + \varepsilon^2}$  and where  $\varepsilon$  is a small positive number (we used  $\varepsilon = 10^{-5}$  throughout).

### 3.2. PSF estimation

The gradient of the objective function  $E(u, h)$  w.r.t. to  $h$  is

$$[\nabla_h E(u, h)]_{i=1}^m = [\nabla_h D(u, h)]_{i=1}^m, \quad (11)$$

where

$$\nabla_h D(u, h)_i = \begin{cases} (U^T((2/\alpha)\mathbf{1} - y ./ \sqrt{\alpha U h + l\mathbf{1}}))_i, & g_i > 0 \\ 0, & g_i \leq 0 \end{cases} \quad (12)$$

$g_i = \alpha(Uh)_i + l$  and  $m$  is the length of the PSF, represented in a vector form. Multiplication with the matrix  $U_{m \times m}$  corresponds to the linear operation of convolving an image (here the PSF) with  $u$ ; the matrix  $U$  is such that  $Uh = Hu$ .

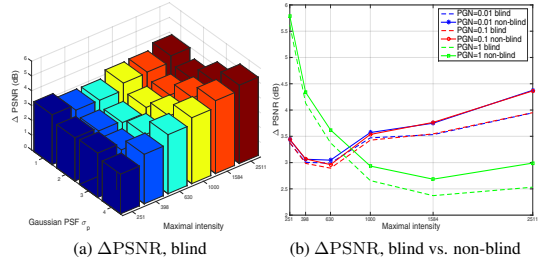
The image edges are handled using a periodic boundary condition. All multiplications with  $H$  and  $U$  are performed in the frequency domain using the 2D fast Fourier transform.

After the PSF is estimated, positivity and mass preservation constraints are imposed using Dykstra's projection algorithm [17]. As was shown in [7], imposing constraints on PSF in such way instead of having additional regularization term for the PSF in the energy function, gives the desired solution to the PSF estimation problem, without getting stuck in unwanted local optima of the energy function.

## 4. EVALUATION

### 4.1. Synthetic degradations

We evaluate the performance of the proposed algorithm on a set of synthetically degraded images, observing the improvement in Peak-Signal-To-Noise-Ratio (PSNR). We compare obtained improvement in PSNR with improvement in PSNR



**Fig. 2:** Performances of the proposed blind method. (a) The average PSNR improvement ( $\Delta$ PSNR) for blind method for  $\sigma_m/\sqrt{peak} = 0.01$ ; (b) The average improvement in PSNR achieved with blind and non-blind (using the correct PSF) methods for Gaussian PSF with  $\sigma_p = 1$  and three PG noise (PGN) levels for  $\sigma_m/\sqrt{peak} = \{0.01, 0.1, 1\}$ .

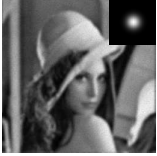
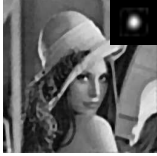
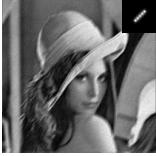
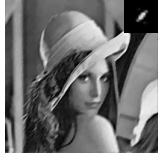
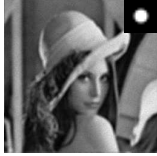
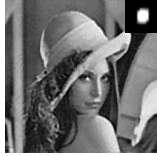
achieved utilizing a non-blind approach. The data set consists of 5 test images presented in Fig. 1. We blur each test image with Gaussian PSF with 4 different standard deviations  $\sigma_p \in \{1, 2, 3, 4\}$ , and we add mixed noise, according to (2), to each of the blurred images. We consider 6 different peak intensity levels, between 250 and 2500, and in that way we observe 6 different levels of Poisson noise. For Gaussian noise we choose  $\mu = 0$  and variance such that ratio  $\sigma_m/\sqrt{peak} \in \{0.01, 0.1, 1\}$ . The scaling factor is  $\alpha = 1$ . By this, we obtain  $5 \times 4 \times 6 \times 3 = 360$  different blurred and noisy observations which we restore using the proposed blind method and its non-blind version. The non-blind restoration is performed according to the image estimation step of Algorithm 1, assuming that the PSF is known, fixed and given as an input.

Evaluation by comparison with known original (ground truth) allows selection of optimal parameters for each individual observed image. We utilize Nelder-Mead simplex search to maximize output PSNR. By this, we avoid mixing evaluation of the proposed method with possibly imperfect parameter selection, which would otherwise make the results more difficult to interpret.

The average improvement over 5 images in PSNR obtained with the proposed blind method for different Gaussian PSF and Poisson noise levels and for  $\sigma_m/\sqrt{peak} = 0.01$  is presented in Fig. 2a. The result for other two considered ratios,  $\sigma_m/\sqrt{peak} = 0.1$  and  $\sigma_m/\sqrt{peak} = 1$ , exhibits similar behavior.

An illustrative example presented in Fig. 3 shows that the proposed method reduces blur and suppresses noise in the degraded images and leads to good restoration results. Moreover, comparison of performances of blind and non-blind method, shown in Fig. 2b, confirms that the results obtained by the proposed blind restoration method are comparable with those obtained when the blur (PSF) is known a priori.

Recovery of one  $256 \times 256$  image in Matlab on an Intel

Gaussian PSF		Motion PSF		Disk PSF	
Observed	Restored	Observed	Restored	Observed	Restored
					
PSNR=21.41 dB	PSNR=25.39 dB $\lambda = 1.4 \cdot 10^{-3}$	PSNR=22.58 dB	PSNR=25.92 dB $\lambda = 5.4 \cdot 10^{-3}$	PSNR=22.40 dB	PSNR=26.24 dB $\lambda = 8.7 \cdot 10^{-4}$

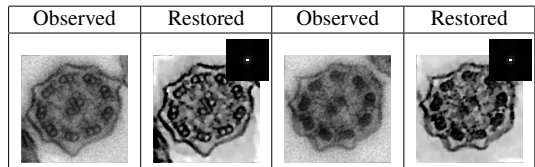
**Fig. 3:** Restoration of blurred and noisy images. Image Lena is convolved by  $13 \times 13$  Gaussian PSF ( $\sigma_p = 2$ ), motion blur of length 7 and angle  $\pi/4$ , and disk with radius 3, respectively, and PG noise with  $peak = 1500$ ,  $\sigma_m/\sqrt{peak} = 0.01$  is added. Reached PSNR and optimal parameter  $\lambda$  that maximizes PSNR, are given. Applied and estimated PSFs are also shown.

Core i7 3.40 GHz CPU takes around 200 seconds.

#### 4.2. Transmission Electron Microscopy images

Primary Ciliary Dyskinesia is a genetic disorder which causes several severe diseases. It results from dysfunction of small hair-like organelles (cilia) which, among others, clean our airways. Transmission Electron Microscopy (TEM) is the only method that provides a resolution sufficient for diagnosis of cilia disorders based on structural information in the samples. To set a diagnosis, at least 50 high quality imaged instances of perfectly cut cilia need to be located and visually analyzed by an expert pathologist; a very tedious and time consuming process if performed manually. We have recently achieved good results in automated cilia detection and rating of cilia instances, [18, 19]. Following detection, the best found candidates can be displayed to a pathologist for final diagnosis. However, structures observed by the pathologist are at the single nanometer size, shown as objects of only a couple of pixels, and are difficult to see in typically noisy TEM images. Image restoration by the method proposed in this paper can provide that the relevant structures in the cilia images appear enhanced and easier to analyse. Considering that the PSF of the imaging device is not known, an appropriate approach, is by a blind restoration. Moreover, noise present in TEM images has PG distribution [20]. Cilia are automatically detected in a TEM image and restored by the method proposed in this paper. The PSF is estimated by applying the proposed method to the original, full-size, image and this estimate is then used in non-blind restoration (by repeated procedure, assuming known PSF) of cut-outs containing individual objects. For better results (with less stair case effects), two regularization terms are utilized; in addition to the TV-regularization, a term including second order derivatives is used. The objective function considered in this example is:

$$E(u, h) = D(u, h) + \lambda_1 R_1(u) + \lambda_2 R_2(u) , \quad (13)$$



**Fig. 4:** Restoration of blurred and noisy TEM images of cilia. Two examples of observed and restored objects, as well as the estimated PSF.

where  $R_1(u)$  is TV-regularization given by (6) and  $R_2(u)$  is Hessian based penalization proposed in [21]. The regularization parameters  $\lambda_1 = 3 \times 10^{-3}$  and  $\lambda_2 = 10^{-3}$  are experimentally determined. The parameters of mixed noise,  $\alpha$ ,  $\sigma_m$  and  $\mu$  are estimated by method presented in [22]. Examples of detected cilia and their restored counterparts are shown in Fig. 4.

## 5. CONCLUSION

We present a novel blind restoration method for images degraded with blur and mixed Poisson-Gaussian noise. The method simultaneously restores the unknown image and the PSF. It is applicable in many real imaging conditions, when the PSF is not known and/or is difficult to estimate. We evaluate method performance on artificially degraded images with different types and sizes of blur and noise. Results show that the proposed method leads to good restoration results, comparable to those when the PSF is known. We confirm the applicability of the proposed method on Transmission Electron Microscopy images of cilia. The relevant structures to be inspected by a pathologist are clearly enhanced in the resulting images.

**Acknowledgments:** We thank Dr Ida-Maria Sintorn (Vironova AB, Stockholm) for providing the TEM images of cilia.

## 6. REFERENCES

- [1] F. Benvenuto, A. La Camera, C. Theys, A. Ferrari, H. Lantéri, and M. Bertero, “The study of an iterative method for the reconstruction of images corrupted by poisson and gaussian noise,” *Inverse Problems*, vol. 24, no. 3, pp. 035016, 2008.
- [2] D. L. Snyder, A. M. Hammoud, and R. L. White, “Image recovery from data acquired with a charge-coupled-device camera,” *JOSA A*, vol. 10, no. 5, pp. 1014–1023, 1993.
- [3] T. E. Nichols, E. Asma J. Qi, and R. M. Leahy, “Spatiotemporal reconstruction of list-mode pet data,” *IEEE Transactions on Medical Imaging*, vol. 21, no. 4, pp. 396–404, 2002.
- [4] S. Delpretti, F. Luisier, S. Ramani, T. Blu, and M. Unser, “Multiframe sure-let denoising of timelapse fluorescence microscopy images,” in *5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI)*, 2008, pp. 149–152.
- [5] T. F. Chan and C. Wong, “Total variation blind deconvolution,” *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 370–375, 1998.
- [6] M. Almeida and M. Figueiredo, “Blind image deblurring with unknown boundaries using the alternating direction method of multipliers,” in *20th IEEE International Conference on Image Processing (ICIP)*, 2013, pp. 586–590.
- [7] D. Perrone and P. Favaro, “Total variation blind deconvolution: The devil is in the details,” in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014, pp. 2909–2916.
- [8] A. Levin, Y. Weiss, F. Durand, and W. Freeman, “Understanding and evaluating blind deconvolution algorithms,” in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009, pp. 1964–1971.
- [9] M. Prato, A. La Camera, S. Bonettini, and M. Bertero, “A convergent blind deconvolution method for post-adaptive-optics astronomical imaging,” *Inverse Problems*, vol. 29, no. 6, pp. 065017, 2013.
- [10] R. Mourya, L. Denis, J.-M. Becker, and E. Thiébaud, “A blind deblurring and image decomposition approach for astronomical image restoration,” in *23rd European Signal Processing Conference (EUSIPCO)*, 2015, pp. 1636–1640.
- [11] J. Li, Z. Shen, R. Yin, and X. Zhan, “A reweighted l2 method for image restoration with poisson and mixed poisson-gaussian noise,” *Inverse Problems and Imaging*, vol. 9, no. 3, pp. 875–894, 2015.
- [12] E. Chouzenoux, A. Jeziarska, J.-C. Pesquet, and H. Talbot, “A convex approach for image restoration with exact poisson-gaussian likelihood,” *SIAM Journal on Imaging Science*, vol. 8, no. 4, pp. 2662–2682, 2015.
- [13] F. Murtagh, J.-L. Starck, and A. Bijaoui, “Image restoration with noise suppression using a multiresolution support,” *Astronomy and Astrophysics Supplement Series*, vol. 112, pp. 179–189, 1995.
- [14] L. I. Rudin, S. Osher, and E. Fatemi, “Nonlinear total variation based noise removal algorithms,” *Physica D: Nonlinear Phenomena*, vol. 60, no. 1, pp. 259–268, 1992.
- [15] E. Birgin, J. Martinez, and M. Raydan, “Nonmonotone spectral projected gradient methods on convex sets,” *SIAM Journal on Optimization*, vol. 10, no. 4, pp. 1196–1211, 2000.
- [16] B. Bajić, J. Lindblad, and N. Sladoje, “An evaluation of potential functions for regularized image deblurring,” in *International Conference on Image Analysis and Recognition (ICIAR), Lecture Notes in Computer Science*, 2014, pp. 150–158.
- [17] J. P. Boyle and R. L. Dykstra, “A method for finding projections onto the intersection of convex sets in hilbert spaces,” in *Lecture Notes in Statistics*, 1986, pp. 28–47.
- [18] J. Lindblad, N. Sladoje, A. Suveer, A. Dragomir, and I.-M. Sintorn, “High-resolution reconstruction by feature distance minimization from multiple views of an object,” in *International Conference on Image Processing Theory, Tools and Applications (IPTA)*, 2015, pp. 29–34.
- [19] A. Suveer, N. Sladoje, J. Lindblad, A. Dragomir, and I.-M. Sintorn, “Automated detection of cilia in low magnification transmission electron microscopy images using template matching,” in *13th IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI)*, 2016, To appear.
- [20] M. Vulović, E. Franken, R. B. G. Ravelli, L. J. van Vliet, and B. Rieger, “Precise and unbiased estimation of astigmatism and defocus in transmission electron microscopy,” *Ultramicroscopy*, vol. 116, pp. 115–134, 2012.
- [21] K. Papafitsoros and C. B. Schönlieb, “A combined first and second order variational approach for image reconstruction,” *Journal of mathematical imaging and vision*, vol. 48, no. 2, pp. 308–338, 2014.
- [22] X. Liu, M. Tanaka, and M. Okutomi, “Practical signal-dependent noise parameter estimation from a single noisy image,” *IEEE Transactions on Image Processing*, vol. 23, no. 10, pp. 4361–4371, 2014.

## Publication III

### Appeared in:

**Bajić B**, Lindblad J and Sladoje N.

Single image super-resolution reconstruction in presence of mixed Poisson-Gaussian noise. In *Proceedings of the 6th IEEE International Conference on Image Processing Theory, Tools and Applications (IPTA)*, Oulu, Finland, December 2016, pp. 1–6.





# Single image super-resolution reconstruction in presence of mixed Poisson-Gaussian noise

Buda Bajić<sup>1</sup>, Joakim Lindblad<sup>2,3</sup> and Nataša Sladoje<sup>2,3</sup>

<sup>1</sup> Faculty of Technical Sciences, University of Novi Sad, Serbia

<sup>2</sup> Centre for Image Analysis, Department of Information Technology, Uppsala University, Sweden

<sup>3</sup> Mathematical Institute, Serbian Academy of Sciences and Arts, Serbia

e-mail: buda.bajic@uns.ac.rs, joakim@cb.uu.se, natasa.sladoje@it.uu.se

**Abstract**—Single image super-resolution (SR) reconstruction aims to estimate a noise-free and blur-free high resolution image from a single blurred and noisy lower resolution observation. Most existing SR reconstruction methods assume that noise in the image is white Gaussian. Noise resulting from photon counting devices, as commonly used in image acquisition, is, however, better modelled with a mixed Poisson-Gaussian distribution. In this study we propose a single image SR reconstruction method based on energy minimization for images degraded by mixed Poisson-Gaussian noise. We evaluate performance of the proposed method on synthetic images, for different levels of blur and noise, and compare it with recent methods for non-Gaussian noise. Analysis shows that the appropriate treatment of signal-dependent noise, provided by our proposed method, leads to significant improvement in reconstruction performance.

**Keywords**—super-resolution, image zooming, signal dependent noise, energy minimization, variance stabilizing transform, total variation.

## I. INTRODUCTION

Imaging is often performed by using photon counting devices like CCD and CMOS image sensors. Due to the physical properties of light, optics, and the sensors, resolution of acquired image is limited and acquired low-resolution (LR) images are degraded by blur and noise. The resulting noise is non-additive and signal-dependent and can be accurately modelled by a mixed Poisson-Gaussian (PG) distribution, as often encountered in astronomy [1], biology [2], and medicine [3].

In most imaging applications, digital images of high resolution (HR) and good quality are desired. This motivates the development of image processing techniques, so-called super-resolution (SR) reconstruction methods, which aim to obtain blur-free and noise-free HR images from the observed (one or multiple) blurred and noisy LR images [4]–[6]. Here we focus on single image SR, where a HR image is reconstructed from one LR observation. SR reconstruction is commonly performed by inverse modelling and minimization of a regularized energy function, due to simplicity and generally good performance of this approach. The objective energy function consists of a data fidelity term, which models the image formation process, and a regularization term, which imposes a priori knowledge on the solution.

Most of existing SR reconstruction methods are developed based on an assumption that noise in acquired images is additive zero-mean white and Gaussian [6]–[8], due to its

appealing properties from standpoint of optimization. Since the resulting noise of photon counting devices cannot be accurately represented by Gaussian white noise, the more correct assumption of signal dependent mixed PG noise should be incorporated in the data fidelity term of the energy function. In this paper we present a novel energy minimization based single image SR reconstruction method which jointly removes blur and PG noise and estimates an increased resolution version of the observed image.

## II. BACKGROUND AND RELATED WORK

The task of single image SR reconstruction is to estimate a HR image from a single blurred and noisy LR image. Since details are missing, one LR image may correspond to many HR images and, consequently, the SR problem is ill-posed. To obtain a reliable solution, the problem is usually solved through minimization of a regularized energy function of the form:

$$E(u) = D(u) + \lambda R(u), \quad (1)$$

where  $u$  is the unknown HR image. The energy function consists of a *data fidelity term*,  $D(u)$ , which drives the solution towards the observed LR data, and a *regularization term*,  $R(u)$ , which provides noise suppression. The regularization parameter  $\lambda$  controls the trade-off between the two terms, i.e., the level of smoothing vs. faithful recovery of the (possibly noisy) image detail. An argument which minimizes the energy function,

$$\hat{x} = \arg \min_u E(u) \quad (2)$$

is considered to be an estimate of the HR image  $x$ .

We generally represent images as column vectors, where image rows of pixels are sequentially concatenated; an image of size  $r \times c$  is represented as a vector of size  $N \times 1$ , where  $N = rc$ . If aiming for an increase in resolution by a factor  $d = d_r \times d_c$ , and where the observed LR image  $y$  has a size  $r_l \times c_l$ , the reconstruction  $\hat{x}$  of the HR image  $x$  will have a size  $r_h \times c_h = d_r r_l \times d_c c_l$ , and be represented by a vector of length  $N_h = dN_l$ . The convolution with a point spread function (PSF), which models degradation of the image with blur, is equal to multiplication of matrix  $H_{N_h \times N_h}$  and image  $x_{N_h \times 1}$ . Multiplication of blurred HR image  $Hx$  with matrix  $S_{N_l \times N_h}$  corresponds to down-sampling with a factor  $d = d_r \times d_c$  where  $N_h = dN_l$ ,  $r_h = d_r r_l$  and  $c_h = d_c c_l$ .

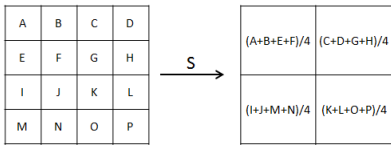


Fig. 1: Effect of down-sampling matrix  $S$  on  $4 \times 4$  image with sampling factor  $d = 2 \times 2$ .

Fig. 1 illustrates the result of applying the down-sampling matrix  $S_{N_i \times N_i}$  on a  $4 \times 4$  image. Down-sampled blurred HR image,  $SHx$ , is additionally corrupted with noise. In this way, observed LR image  $y$  is formed.

An assumption that noise in acquired images is additive zero-mean white Gaussian leads to a quadratic  $\ell_2$  data fidelity term:

$$D_G(u) = \frac{1}{2} \sum_{i=1}^{N_i} ((SHu)_i - y_i)^2. \quad (3)$$

This approach is taken in, e.g., [6]–[8]; quadratic data term is easy to optimize and different optimization strategies are observed, such as preconditioned conjugate gradient [6], Bregman iterative refinement procedure along with gradient-descent method [7], augmented Lagrangian (AL) scheme [8].

On convex problems, the particular optimization method only affects the speed of convergence. In this paper, for optimization we use Spectral Projected Gradient (SPG) [9] method, which is a versatile method that can handle a variety of energy functions. To verify the reliability of our SPG implementation, we have compared its performance for SR reconstruction of images degraded with additive white Gaussian noise, with the very recently proposed AL-FSR method [8]<sup>1</sup>, which incorporates  $\ell_2$  data term, Total Variation (TV) regularization, and optimization utilizing an AL scheme. Utilizing SPG optimization and the same energy function, i.e., with a quadratic data term (3), appropriate for Gaussian noise (denoted SR-G), we (as expected) observe very similar performance of the two approaches. On a set of 100 test images degraded with different levels of noise and blur, we observe, on average, slightly better performances of our SPG based SR-G implementation (0.03 dB higher peak signal-to-noise ratio (PSNR)), while the AL-FSR method is approximately 3 times faster than our Matlab implementation of SPG. Aiming for maximal speed is outside the scope of this paper, however, we note that SPG is fairly easy to parallelize and a GPU implementation would presumably be orders of magnitude faster.

It was observed in [10]–[12], that methods with  $\ell_2$  data term exhibit significantly reduced performance when noise is not normally distributed. More robust data fidelity norms were proposed for non-Gaussian noise, such as  $\ell_1$  norm suited for Laplacian noise [10], [13], combination of  $\ell_1$  and  $\ell_2$

norms [12], [14], [15], and Lorentzian norm [11], [12]. According to results presented in [10], [12], Huber (combination of  $\ell_1$  and  $\ell_2$ ) and Lorentzian data fidelity norms outperform pure  $\ell_2$  and  $\ell_1$  data fidelity norms when the image is degraded by signal-dependent Poisson noise.

Huber data fidelity term is given as a combination of  $\ell_1$  and  $\ell_2$  norms via Huber function, as observed in [12]:

$$D_{\text{HUB}}(u) = \sum_{i=1}^{N_i} \Phi_{\text{HUB}}((SHu)_i - y_i), \quad (4)$$

where Huber function is

$$\Phi_{\text{HUB}}(t) = \begin{cases} \frac{t^2}{2\omega}, & t \leq \omega, \\ t - \frac{\omega}{2}, & t > \omega. \end{cases} \quad (5)$$

Parameter  $\omega$  controls the point of transition between  $\ell_2$  and  $\ell_1$  norm.

Lorentzian data fidelity term [10], [12] is given by:

$$D_{\text{LOR}}(u) = \sum_{i=1}^{N_i} \Phi_{\text{LOR}}((SHu)_i - y_i), \quad (6)$$

where Lorentzian function with a parameter  $T$  is

$$\Phi_{\text{LOR}}(t) = \log(1 + t^2/2T^2). \quad (7)$$

We have already observed, [16], that a correct treatment of both signal-dependent and signal-independent components of the noise is very important, even when one of them is present to a very small amount. The data fidelity term adjusted to the PG noise model includes an infinite sum and is not easy to compute. One way to address this problem is to approximate the infinite sum with a finite number of summands, like in the methods presented in [1], [17]. This unfortunately leads to slow and complicated algorithms. Alternatively, the image can be transformed to make the Gaussian model applicable, utilizing a variance stabilizing transformation (VST) based approach. For signal-dependent Poisson-Gaussian noise, the generalized Anscombe stabilizing transformation [18] can be used to remove the signal-dependency and render the noise variance constant throughout the image [19]. Generalized Anscombe transformation of the observed image  $y$ , incorporating Poisson noise as well as zero-mean Gaussian noise of variance  $\sigma_m^2$ , is the image  $z$ :

$$z_i = 2\sqrt{\max\left\{y_i + \frac{3}{8} + \sigma_m^2, 0\right\}}, \quad \forall i = 1, 2, \dots, N_i. \quad (8)$$

The transformed LR image  $z$  behaves as the blurred, down-sampled and VST-transformed HR image  $x$  plus a normal distributed noise term; for  $z_i, \forall i = 1, 2, \dots, N_i$ , it holds that

$$z_i \approx 2\sqrt{\max\left\{(SHx)_i + \frac{3}{8} + \sigma_m^2, 0\right\}} + \epsilon, \quad \epsilon \sim \mathcal{N}(0, 1). \quad (9)$$

This enables the design and use of an appropriate quadratic data term also for the case of (transformed) PG noise.

<sup>1</sup>Available online <http://zhao.perso.enseeiht.fr/>

We have shown in [16], in a large empirical study, that such a VST based approach outperforms (particularly in terms of speed) the one based on approximate infinite sums, for deblurring and denoising of images degraded by mixed PG noise. In this paper we extend the VST approach to SR reconstruction of a single image degraded by blur and PG noise.

A related study, focusing on SR reconstruction and assuming PG noise, but not blur, is presented in [20]. Whereas in [20] SR reconstruction is performed from a sequence of images (high frame rate video), we here present a single image SR method which does remove blur, as well as mixed PG noise.

Methods for SR reconstruction which do not rely on modelling of any particular noise distribution also exist in the literature; a fast and simple example of such is bicubic interpolation. The method proposed in [21] suggests to utilize median filter transformation (MFT) to suppress noise in an estimated image. An advantage of this type of methods is their generality; however, it comes with inferior performance compared to approaches which rely on (appropriate) models.

### III. SINGLE IMAGE SUPER-RESOLUTION RECONSTRUCTION FOR MIXED POISSON-GAUSSIAN NOISE

The method we present is based on regularized energy minimization and, as such, incorporates three main components: data fidelity term, regularization term, and a suitable optimization method.

#### A. Data fidelity term

Assuming that the HR image  $x$  is degraded by blur, down-sampled and corrupted with PG noise, the intensity values  $y_i$  of the acquired LR image  $y$  are of the form

$$y_i = \theta_i + \eta_i, \quad \forall i = 1, 2, \dots, N_l \quad (10)$$

where  $\theta_i \sim \mathcal{P}((SHx)_i)$  and  $\eta_i \sim \mathcal{N}(0, \sigma_m^2)$ . Assuming cyclic convolution boundaries, the blur matrix  $H_{N_h \times N_h}$  is block circulant. Consequently, multiplication with  $H_{N_h \times N_h}$  can be performed in the frequency domain using the 2D fast Fourier transform.

Applying Generalized Anscombe transformation to the observed image  $y$ , we get the image  $z$ , such that (8) and (9) hold.

For a signal-independent Gaussian noise model, the data fidelity term is quadratic. The data fidelity term of the observed energy function (1) can, therefore, after the transformation, be expressed as

$$D_{\text{PG}}(u) = \frac{1}{2} \sum_{i=1}^{N_l} \left( 2\sqrt{\max\left\{(SHu)_i + \frac{3}{8} + \sigma_m^2, 0\right\}} - z_i \right)^2. \quad (11)$$

#### B. Regularization term

The role of the regularization term is to provide numerical stability and to impose desired properties to the solution. The

#### Algorithm 1: Spectral Projected Gradient

---

Choose values for parameters:  $\theta_{min}, \theta_{max}, \gamma, \sigma_1, \sigma_2, tol$   
s.t.  $0 < \theta_{min} < \theta_{max}, \gamma \in (0, 1), 0 < \sigma_1 < \sigma_2 < 1, tol > 0$ .

Choose initial guess  $x_0 \in \Omega$  and  $\theta_0 = 1$ .  
Compute  $x_{k+1}$  and  $\theta_{k+1}$  as follows:

$$d_k = P_{\Omega}(x_k - \theta_k \nabla f(x_k)) - x_k$$

$$x_{k+1} = x_k + d_k$$

$$\delta = \nabla f(x_k)^T d_k; \xi_k = 1$$

**while**  $f(x_{k+1}) > f(x_k) + \gamma \xi_k \delta$   
 $\xi_{temp} = -\frac{1}{2} \xi_k^2 \delta / (f(x_{k+1}) - f(x_k) - \xi_k \delta)$   
**if**  $(\xi_{temp} \geq \sigma_1 \wedge \xi_{temp} \leq \sigma_2 \xi_k)$   
  **then**  $\xi_k = \xi_{temp}$   
  **else**  $\xi_k = \xi_k / 2$   
 $x_{k+1} = x_k + \xi_k d_k$   
**end while**

$$s_k = x_{k+1} - x_k; y_k = \nabla f(x_{k+1}) - \nabla f(x_k); \beta_k = s_k^T y_k$$

**if**  $\beta_k \leq 0$   
  **then**  $\theta_{k+1} = \theta_{max}$   
  **else**  $\theta_{k+1} = \min\{\theta_{max}, \max\{\theta_{min}, \frac{s_k^T s_k}{\beta_k}\}\}$

**Repeat until:**  $\|x_{k+1} - x_k\|_{\infty} \leq tol$ .

---

regularization term of the energy function (1) may be designed based on the desired characteristics of the unknown image  $u$ , such as wavelet-based sparsity, smoothness, small total variation. In this study we choose TV regularization [22], due to its generally good performance in suppressing noise and preserving edges. The popularity of TV regularization in the literature simplifies comparison of methods incorporating different data terms (and all including TV regularization); this facilitates evaluation of our main hypothesis – that a data term particularly suited for PG noise leads to considerably improved performance of SR reconstruction of images corrupted by such type of noise.

Total Variation (TV) regularization is given by

$$R(u) = \sum_{i=1}^{N_h} |\nabla(u_i)|, \quad (12)$$

where  $\nabla$  stands for gradient and  $|\cdot|$  denotes  $\ell_2$  norm. We compute the discrete image gradient at point  $u_i$ ,  $\nabla(u_i)$ , as  $\nabla(u_i) = (u_e - u_i, u_s - u_i)$ , where  $e$  and  $s$  denote indices of the edge neighbors to the right (east direction) and below the pixel  $u_i$  (south direction), respectively.

#### C. Optimization

We consider grey-scale images and represent them as vectors with intensity values from  $[0, 1]$  (for images in the range  $[0, peak]$  intensity is divided by  $peak$ ). Minimization of the objective function  $E(u)$  is seen as a constrained optimization problem

$$\min E(u) \quad \text{s.t. } 0 \leq u_i \leq 1, \quad i = 1, 2, \dots, N_h. \quad (13)$$

We optimize the objective function  $E(u)$  utilizing Spectral Projected Gradient (SPG) [9]. We have utilized SPG in our previous studies [16], [23], [24], for optimization of energy



Fig. 2: Used test images, all of a size of  $256 \times 256$  pixels.

functions with various data fidelity and regularization terms and observed its very good performance.

SPG is a numerical optimization tool for solving a constrained optimization problem  $\min_{x \in \Omega} f(x)$ , where  $\Omega$  is a closed convex set in  $\mathbb{R}^n$  and  $f$  is a function which has continuous partial derivatives on an open set that contains  $\Omega$ . Weak requirements on the objective function, as well as efficiency in solving large scale problems, make this tool attractive for our purpose.

The method is briefly outlined in Algorithm 1. We define the projection  $P_\Omega$  of a vector  $x \in \mathbb{R}^n$  to the feasible set  $\Omega = [0, 1]^n$  as:  $[P_\Omega(x)]_i = \min\{1, \max\{0, x_i\}\}$ , for all  $i = 1, 2, \dots, n$ .

As SPG is an iterative gradient-based method, gradient of the objective function is needed. The gradient of  $E(u)$  is

$$[\nabla E(u)]_{i=1}^{N_h} = [\nabla D_{PG}(u)]_{i=1}^{N_h} + \lambda [\nabla R(u)]_{i=1}^{N_h}, \quad (14)$$

where

$$\nabla D_{PG}(u) = H^T S^T G(u), \quad (15)$$

$$G(u)_i = \begin{cases} 2 - y_i / \sqrt{g_i}, & g_i > 0 \\ 0, & g_i \leq 0 \end{cases} \quad (16)$$

and  $g_i = (SHu)_i + \frac{3}{8} + \sigma_m^2$ .

The gradient of the regularization term is

$$\nabla R(u)_i = \frac{2u_i - u_e - u_s}{|\nabla(u_i)|} + \frac{u_i - u_w}{|\nabla(u_w)|} + \frac{u_i - u_n}{|\nabla(u_n)|} \quad (17)$$

where  $u_n$  and  $u_w$  denote the edge neighbours above (north direction) and left (west direction) of the pixel  $u_i$ , respectively.

The gradient of the regularization term defined by (17) is non-differentiable at points where  $|\nabla u_i| = 0$ . To meet requirements of SPG, we consider a smoothed version of (17), where  $|\nabla u_i|$  is replaced with  $\sqrt{|\nabla(u_i)|^2 + \varepsilon^2}$  and where  $\varepsilon$  is a small positive number (we used  $\varepsilon = 10^{-5}$  throughout).

#### IV. EVALUATION

In this study we argue for an appropriate treatment of signal dependent mixed PG noise in SR reconstruction. We suggest a novel data term, adjusted to this type of noise. To reliably and in an unbiased way evaluate its performance in comparison with other data terms, we observe energy functions which all utilize TV regularization and are minimized by SPG.

The following methods are considered:

- (i) SR-PG – the proposed method suited for mixed PG noise with data fidelity term (11),
- (ii) SR-G – method suited for Gaussian noise with quadratic data fidelity term (3),

- (iii) SR-HUB – method with Huber data fidelity term (4),
- (iv) SR-LOR – method with Lorentzian data term (6).

We observe a data set consisting of 360 different blurred and noisy LR observations generated from 5 test images presented in Fig. 2. We blur each HR test image with Gaussian PSF (closely resembling real PSFs of many imaging systems) with 4 different standard deviations  $\sigma_p \in \{1, 2, 3, 4\}$ , and we down-sample it with sampling factor  $d = 2 \times 2$  (by averaging pixels in  $2 \times 2$  blocks). Finally, we corrupt each blurred image by mixed noise, according to (10). We consider 6 different peak intensity levels, between 250 and 2500, and in that way we observe 6 different levels of Poisson noise. For the level of Gaussian noise we choose variance such that the ratio of standard deviation of Gaussian and Poisson noise,  $\sigma_m / \sqrt{peak}$ , is  $\{0.01, 0.1, 1\}$ . By this, we obtain  $(5 \times 4 \times 6) \times 3 = 120 \times 3 = 360$  different blurred and noisy LR observations.

For each image, we compute the peak signal-to-noise ratio (PSNR) as a quality measure:

$$\text{PSNR} = 10 \log_{10} \left( \frac{(\max(x_i))^2}{\text{MSE}} \right),$$

where

$$\text{MSE} = \frac{1}{N_h} \sum_{i=1}^{N_h} (x_i - \hat{x}_i)^2.$$

The true and estimated HR images are, according to (2), denoted by  $x$  and  $\hat{x}$ , respectively. In some experiments we also consider Structural Similarity Index Measure (SSIM) [25], which to some extent better reflects subjectively perceived image quality than PSNR. The performed evaluation, which compares with a known original (ground truth) allows the selection of optimal (oracle style) regularization parameter,  $\lambda$ , as well as Huber and Lorentzian parameters,  $\omega$  and  $T$ , respectively, for each method and each individual observed image. By this approach, we avoid mixing evaluation of the proposed method with possibly imperfect parameter estimation, which would otherwise make the results more difficult to interpret. We select parameters utilizing Nelder-Mead simplex search to maximize output PSNR. SSIM is computed, as a secondary (control) measure of performance, on images reconstructed as to maximize PSNR (i.e., no parameter tuning is performed for maximizing SSIM).

We observe a consistent additional improvement in PSNR when using the proposed SR-PG method (i) instead of any of the methods (ii)-(iv). Observed overall average improvement (over 360 images) in PSNR over all the three considered Gaussian-Poisson noise ratios is 2.36 dB.

The difference in performance of SR-PG, and each of SR-G, SR-HUB and SR-LOR, on the subset of the test set with ratio  $\sigma_m / \sqrt{peak} = 0.01$ , is presented in Fig. 3. It is clear that the use of the correct noise model largely improves the quality of reconstructed images. For this noise ratio we observe on average (over 120 degraded images) 2.41 dB better performance when using method (i) than when using (ii)-(iv). The results for the other two considered ratios of Poisson-Gaussian noise,  $\sigma_m / \sqrt{peak} = 0.1$  and  $\sigma_m / \sqrt{peak} = 1$ ,

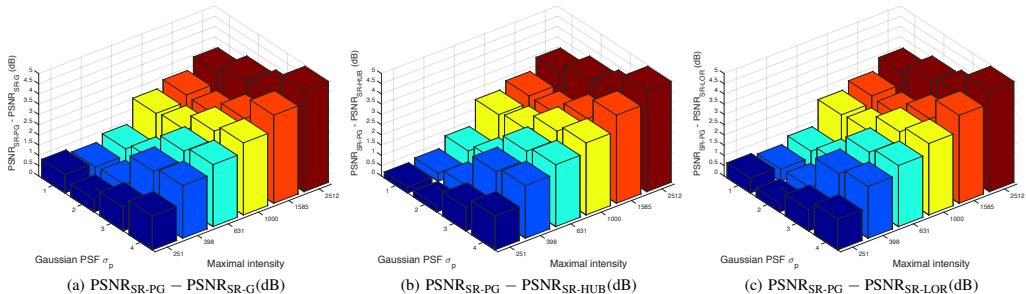


Fig. 3: Average improvement in PSNR, on the observed test set for  $\sigma_m/\sqrt{peak} = 0.01$ , achieved due to appropriate treatment of mixed Poisson-Gaussian noise, for different reconstruction approaches, and different levels of blur and noise (scale factor  $d = 2 \times 2$ ). The proposed SR-PG method outperforms the other observed methods by up to 5 dB.

exhibit similar behavior (with average improvement of 2.34 dB and 2.34 dB, respectively).

We also include comparison with two methods which do not assume any particular type of noise:

- (v) SR by bicubic interpolation,
- (vi) MFT – SR method proposed in [21]<sup>2</sup>.

Average performance (reached PSNR) over 360 degraded images and a  $2 \times 2$  increase in resolution, for all considered methods, (i)-(vi), is presented in Fig. 4. It is observed that the proposed SR-PG method clearly outperforms all the other considered methods.

An illustrative example of SR reconstruction with  $d = 3 \times 3$  times increased resolution is given in Fig. 5; a test image degraded by blur and PG noise (shown to scale in (a)) is reconstructed by our proposed method (result shown in (b)), by methods suited for Gaussian noise (shown in (c)), by using alternative data terms (shown in (d) SR-HUB and (e) SR-LOR), as well as approaches not assuming any particular noise model (in (f) bicubic interpolation, and (g) MFT). Reached PSNR, as well as SSIM, are presented below each image. The proposed SR-PG method is consistently reaching the best quantitative results, and we argue that subjective visual reconstruction performance also speaks strongly in favour of the proposed method.

To examine how reconstruction performance varies with the increase of scale factor  $d$ , we reconstruct (using SR-PG) the LR image “Lena” (first in Fig. 2) degraded by blur and PG noise from a LR size of  $256 \times 256$ ,  $128 \times 128$ ,  $85 \times 85$ ,  $64 \times 64$  and  $51 \times 51$  using a scale factor  $d$  equals to  $1 \times 1$ ,  $2 \times 2$ ,  $3 \times 3$ ,  $4 \times 4$  and  $5 \times 5$ , respectively (thus reaching approximately the original size of  $256 \times 256$ ). Table 1 shows achieved PSNR and SSIM values. Naturally, the performance drops with increasing scale factor, but at a fairly slow pace. In comparison with other methods, we conclude that the proposed SR-PG method offers

<sup>2</sup>Available online <http://www.lcc.uma.es/~ezeq1r/mft/mft.html>

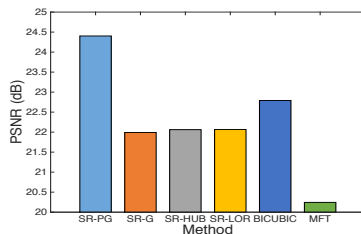


Fig. 4: Average PSNR achieved with different SR approaches (scale factor  $d = 2 \times 2$ ). The proposed SR-PG outperforms other observed methods.

Table 1: Performances of SR-PG on image “Lena” for different scale factors  $d$ . Original Lena image (the first in Fig. 2) is degraded by Gaussian blur ( $\sigma_p = 1$ ), down sampled with different sampling factors  $d$  and corrupted by PG noise with  $peak = 1585$ ,  $\sigma_m/\sqrt{peak} = 0.01$ . The degraded LR image is then reconstructed by SR-PG back to the original size.

Scale factor $d$	$1 \times 1$	$2 \times 2$	$3 \times 3$	$4 \times 4$	$5 \times 5$
PSNR	28.93 dB	27.06 dB	25.41 dB	24.16 dB	23.31 dB
SSIM	0.669	0.588	0.514	0.450	0.405

good performance even for as large scale factors as  $5 \times 5$ .

The proposed SR-PG method recovers one  $128 \times 128$  image with scale factor  $d = 2 \times 2$  (HR image is  $256 \times 256$ ) in Matlab on an Intel Core i7 3.40 GHz CPU in approximately 4 seconds.

## V. CONCLUSION

We present a novel single image super-resolution reconstruction method for images degraded by blur and mixed Poisson-Gaussian noise. The method reconstructs an unknown image from a single blurred and noisy observation of lower resolution. We evaluate the performance of the method on artificially degraded images with different levels of blur and

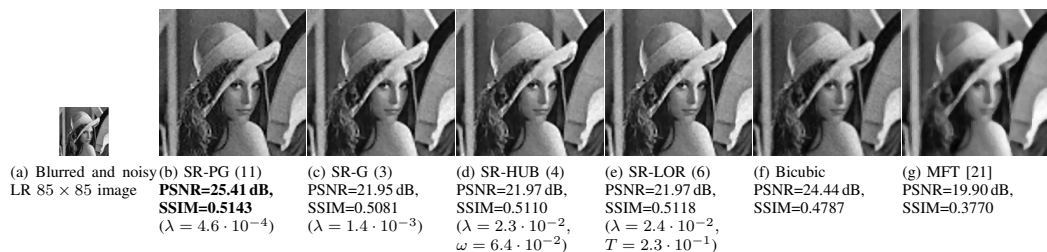


Fig. 5: SR reconstruction of a test image, degraded by blur and PG noise. Original image (Fig. 2) of size  $255 \times 255$  is degraded by Gaussian blur ( $\sigma_p = 1$ ), down sampled with sampling factor  $d = 3 \times 3$  (by averaging pixels in  $3 \times 3$  blocks) and corrupted by PG noise with  $peak = 1585$ ,  $\sigma_m/\sqrt{peak} = 0.01$ . Such degraded image is shown in (a). Reconstructed HR images (b)-(g) are all  $255 \times 255$ . Below each reconstructed image, the reached PSNR and SSIM values are given together with optimal parameters (which maximize PSNR) for each method on this particular image.

noise. Comparison with performances of other methods, which treat Gaussian as well as non-Gaussian noise, shows the importance of using the appropriate noise model in super-resolution reconstruction. Using the proposed method we observe a significant improvement in performance compared to other methods. Further improvement of the proposed method in terms of computational speed may be achieved by utilizing faster optimization strategies. This is left as future work.

#### ACKNOWLEDGEMENT

Authors are supported by the Swedish Research Council, and by the Ministry of Education, Science and Technological Development of the Republic of Serbia through Projects ON174008 and III44006 of MI-SANU. Sladoje is supported by the Swedish Governmental Agency for Innovation Systems.

#### REFERENCES

- [1] F. Benvenuto, A. La Camera, C. Theys, A. Ferrari, H. Lantéri and M. Bertero. The study of an iterative method for the reconstruction of images corrupted by Poisson and Gaussian noise. *Inverse Problems*, 24(3): 035016, 2008.
- [2] T. E. Nichols, J. Qi, E. Asma and R. M. Leahy. Spatiotemporal reconstruction of list-mode PET data. *IEEE Transactions on Medical Imaging*, 24(4): 396–404, 2002.
- [3] S. Delpretti, F. Luisier, S. Ramani, T. Blu and M. Unser. Multiframe sure-let denoising of timelapse fluorescence microscopy images. In *IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI 2008)*, proceedings, 149–152, Paris, 2008.
- [4] S. C. Park, M. K. Park and M. G. Kang. Super-resolution image reconstruction: a technical overview. *IEEE Signal Processing Magazine*, 20(3): 21–36, 2003.
- [5] K. Nasrollahi and T. B. Moeslund. Super-resolution: a comprehensive survey. *Machine Vision and Applications*, 25(5): 1423–1468, 2014.
- [6] M. K. Ng and N. K. Bose. Mathematical analysis of super-resolution methodology. *IEEE Signal Processing Magazine*, 20(3): 62–74, 2003.
- [7] A. Marquina and S. Osher. Image super-resolution by TV-regularization and Bregman iteration. *Journal of Scientific Computing*, 37(3): 367–382, 2008.
- [8] N. Zhao, Q. Wei, A. Basarab, N. Dobigeon, D. Kouamé and J.-Y. Tourneret. Fast single image super-resolution using a new analytical solution for  $l_2 - l_2$  problems. *IEEE Transactions on Image Processing*, 25(8): 3683–3697, 2016.
- [9] E. Birgin, J. Martinez and M. Raydan. Nonmonotone spectral projected gradient methods on convex sets. *SIAM Journal on Optimization*, 10(4): 1196–1211, 2000.
- [10] S. Farsiu, D. Robinson, M. Elad and P. Milanfar. Advances and challenges in super-resolution. *International Journal of Imaging Systems and Technology*, 14(2): 47–57, 2004.
- [11] V. Patanavijit and S. Jitapunkul. A Lorentzian stochastic estimation for a robust iterative multiframe super-resolution reconstruction with Lorentzian-Tikhonov regularization. *EURASIP Journal on Advances in Signal Processing*, 2007(1): 1–21, 2007.
- [12] A. Panagiotopoulou and V. Anastassopoulos. Super-resolution image reconstruction techniques: trade-offs between the data-fidelity and regularization terms. *Information Fusion*, 13(3): 185–195, 2012.
- [13] P. Purkait and B. Chanda. Morphologic gain-controlled regularization for edge-preserving super-resolution image reconstruction. *Signal Image Video Processing*, 7(5): 925–938, 2013.
- [14] H. Song, L. Zhang, P. Wang, K. Zhang and X. Li. An adaptive  $\ell_1 - \ell_2$  hybrid error model to super-resolution. In *IEEE International Conference on Image Processing*, proceedings, 2821–2824, Hong Kong, 2010.
- [15] L. Yue, H. Shen, Q. Yuan and L. Zhang. A locally adaptive  $\ell_1 - \ell_2$  norm for multi-frame super-resolution of images with mixed noise and outliers. *Signal Processing*, 105: 156–174, 2014.
- [16] B. Bajić, J. Lindblad and N. Sladoje. Restoration of images degraded by signal-dependent noise based on energy minimization: an empirical study. *SPIE Journal of Electronic Imaging*, 25(4): 043020, 2016.
- [17] E. Chouzenoux, A. Jezierska, J.-C. Pesquet and H. Talbot. A convex approach for image restoration with exact Poisson-Gaussian likelihood. *SIAM Journal on Imaging Sciences*, 8(4): 2662–2682, 2015.
- [18] F. Murtagh, J.-L. Starck and A. Bijaoui. Image restoration with noise suppression using a multiresolution support. *Astronomy and Astrophysics Supplement Series*, 112: 179–189, 1995.
- [19] M. Mäkitalo and A. Foi. Optimal inversion of the generalized Anscombe transformation for Poisson-Gaussian noise. *IEEE Transactions on Image Processing*, 22(1): 91–103, 2013.
- [20] J. Suo, Y. Deng, L. Bian and Q. Dai. Joint non-Gaussian denoising and superresolving of raw high frame rate videos. *IEEE Transactions on Image Processing*, 23(3): 1154–1168, 2014.
- [21] E. Lopez-Rubio. Super-resolution from a single noisy image by the median filter transform. *SIAM Journal on Imaging Sciences*, 9(1): 82–115, 2016.
- [22] L. I. Rudin, S. Osher and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60(1): 259–268, 1992.
- [23] B. Bajić, J. Lindblad and N. Sladoje. Blind restoration of images degraded with mixed Poisson-Gaussian noise with application in Transmission Electron Microscopy. In *IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI 2016)*, proceedings, 123–127, Prague, 2016.
- [24] B. Bajić, J. Lindblad and N. Sladoje. An evaluation of potential functions for regularized image deblurring. In *International Conference on Image Analysis and Recognition (ICIAR)*, proceedings, LNCS-8814, 150–158, Vilamoura, 2014.
- [25] Z. Wang, A.C. Bovik, H.R. Sheikh and E.P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4): 600–612, 2004.

## Publication IV

### Appeared in:

**Bajić B**, Lindblad J and Sladoje N.

An Evaluation of Potential Functions for Regularized Image Deblurring. In *Proceedings of International Conference on Image Analysis and Recognition (ICIAR)*, Algarve, Portugal, October 2014, Lecture Notes on Computer Science, 8814, pp. 150--158.

# An Evaluation of Potential Functions for Regularized Image Deblurring

Buda Bajić<sup>1</sup>, Joakim Lindblad<sup>1</sup>, and Nataša Sladoje<sup>1,2</sup>

<sup>1</sup> Faculty of Technical Sciences, University of Novi Sad, Serbia

<sup>2</sup> Centre for Image Analysis, Uppsala University, Sweden

buda.bajic@uns.ac.rs, joakim@cb.uu.se, sladoje@uns.ac.rs

**Abstract.** We explore utilization of seven different potential functions in restoration of images degraded by both noise and blur. Spectral Projected Gradient method confirms its excellent performance in terms of speed and flexibility for optimization of complex energy functions. Results obtained on images affected by different levels of Gaussian noise and different sizes of the Point Spread Functions, are presented. The Huber potential function demonstrates outstanding performance.

## 1 Introduction

Images are generally degraded in various ways in the acquisition process: by camera motion, imperfect optics, presence of noise, atmospheric turbulence, etc. Degradation is often modelled as linear and shift invariant; it is assumed that the original image is convolved by a spatially invariant Point Spread Function (PSF) and corrupted by noise. If the original image is denoted  $u$  and the acquired image  $v$ , the degradation can be expressed as

$$v = h * u + \eta, \quad (1)$$

where  $h$  is the PSF,  $\eta$  represents noise and  $*$  denotes convolution.

Image restoration methods aim at recovering the original image  $u$  from the degraded image  $v$ . However, this inverse problem is severely ill-posed and the solution is highly sensitive to noise in the observed image. Ringing effects and blurred edges are undesired consequences often appearing in restored images. A good balance between frequency recovery and noise suppression is essential for satisfactory deconvolution. A common approach is to apply some regularization, utilizing a priori knowledge when performing deconvolution. Regularization should provide numerical stabilization and impose desired properties to the solution. Total variation (TV) regularization [15] is among most popular approaches, due to its generally good performance.

Our previous studies on image denoising confirm that improved performance of TV based regularization can be achieved if potential functions are utilized. Potentials are designed to enhance/preserve particular image features during the processing; preservation of sharp edges is typically targeted. Potential functions, in general being non-convex, introduce additional complexity to the optimization



problem. We have previously experienced excellent performance of the Spectral Gradient type of optimization methods on similar tasks. These flexible methods allow a wide class of potentials to be used in the energy function, while exhibiting fast convergence.

In this paper we present an empirical evaluation of seven potential functions (listed in Table 1) when used for image deblurring/deconvolution based on regularized energy minimization utilizing the Spectral Projected Gradient (SPG) method. Image degradation includes different levels of blur (Gaussian PSF) and additive Gaussian white noise. Tests include classic TV regularization, and by that an implicit comparison with the large number of methods based on the TV model. We conclude that an appropriately chosen potential function can significantly increase the method performance at essentially no additional cost.

## 2 Background and Previous Work

### 2.1 TV regularization and potential functions

Total Variation regularization is commonly used to address inverse problems in image processing, such as image denoising, deblurring, inpainting, etc. The approach involves minimization of an energy function which incorporates a gradient based regularization term, well balanced with a data fidelity term. Ideally, minimization of the energy function provides suppression of noise while retaining true image information. One approach for improving performance of TV regularization involves the utilization of potential functions.

Typically the energy functional of regularized deblurring is of the form

$$E(u) = \frac{1}{2} \iint |h(x, y) * u(x, y) - v(x, y)|^2 dx dy + \alpha \iint \phi(|\nabla u(x, y)|) dx dy, \quad (2)$$

where  $\nabla$  stands for gradient and  $|\cdot|$  denotes  $\ell_2$  norm. The energy functional consists of a data fidelity term, which drives the solution towards the observed data (degraded image  $v$ ), and a regularization term which utilizes the image gradient to provide noise suppression. The balancing parameter  $\alpha$  controls the trade-off between the terms, i.e., the level of smoothing vs. faithful recovery of the (possibly noisy) image detail.

The function  $\phi$  is referred to as *potential function*. By using a potential equal to the identity function, the regularization term reduces to classic TV regularization. In most cases the potential function is designed s.t. small intensity changes (assumed to be noise) are penalized, while large changes (assumed to be edges) are preserved. A number of potentials are studied and used in image restoration problems [3, 4, 8–12, 16, 17]. In [3] theoretical conditions for edge preserving potentials are given. In [11] examples using the Huber potential for deblurring are presented, however no explicit performance evaluation of potentials is presented. A study of effectiveness of different potentials in image denoising is given in [9], where it is concluded that the Huber potential ( $\phi_5$  in Table 1) works best overall, and that the Geman & McClure potential ( $\phi_2$ ) shows best performance in low noise settings.

Table 1: Potential functions

Potential	Convex
TV [15] $\phi_1(s) = s$	yes
Geman&McClure [4] $\phi_2(s) = \frac{\omega s^2}{1 + \omega s^2}$	no
Hebert&Leahy [8] $\phi_3(s) = \ln(1 + \omega s^2)$	no
Perona&Malik [12] $\phi_4(s) = 1 - e^{-\omega s^2}$	no
Huber [17] $\phi_5(s) = \begin{cases} s^2, & s \leq \omega \\ 2\omega s - \omega^2, & s > \omega \end{cases}$	yes
Tikhonov [16] $\phi_6(s) = s^2$	yes
Nikolova&Chan [10] $\phi_7(s) = \begin{cases} \sin(\omega s^2), & s \leq \sqrt{\frac{\omega}{2\pi}} \\ 1, & s > \sqrt{\frac{\omega}{2\pi}} \end{cases}$	no

Table 2: Algorithm 1.

Spectral Projected Gradient
Choose values for parameters: $\theta_{min}, \theta_{max}, \gamma, \sigma_1, \sigma_2, tol$ s.t. $0 < \theta_{min} < \theta_{max}$ , $\gamma \in (0, 1)$ , $0 < \sigma_1 < \sigma_2 < 1$ , $tol > 0$ .
Choose initial guess $x_0 \in \Omega$ and $\theta_0 = 1$ . Compute $x_{k+1}$ and $\theta_{k+1}$ as follows:
$d_k = P_\Omega(x_k - \theta_k \nabla f(x_k)) - x_k$ $x_{k+1} = x_k + d_k$ ; $\delta = \nabla f(x_k)^T d_k$ $\lambda_k = 1$
<b>while</b> $f(x_{k+1}) > f(x_k) + \gamma \lambda_k \delta$ $\lambda_{temp} = -\frac{1}{2} \lambda_k^2 \delta / (f(x_{k+1}) - f(x_k) - \lambda_k \delta)$ <b>if</b> $(\lambda_{temp} \geq \sigma_1 \wedge \lambda_{temp} \leq \sigma_2 \lambda_k)$ <b>then</b> $\lambda_k = \lambda_{temp}$ <b>else</b> $\lambda_k = \lambda_k / 2$ $x_{k+1} = x_k + \lambda_k d_k$
<b>end while</b>
$s_k = x_{k+1} - x_k$ $y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$ ; $\beta_k = s_k^T y_k$
<b>if</b> $\beta_k \leq 0$ <b>then</b> $\theta_{k+1} = \theta_{max}$ <b>else</b> $\theta_{k+1} = \min \{ \theta_{max}, \max \{ \theta_{min}, \frac{s_k^T s_k}{\beta_k} \} \}$
Repeat until: $\ x_{k+1} - x_k\ _\infty \leq tol$ .

## 2.2 Optimization

An important issue in energy based image restoration is efficient optimization of the energy function. A variety of approaches and algorithms to minimize TV regularized energy function are presented in the literature; a number of references on the topic are given in [14] and some later ones can be found in [5] and [7].

Non-convexity of potentials may lead to non-convexity of the objective function (2), which makes optimization additionally challenging and excludes a number of methods specifically designed for convex minimization. Our studies presented in [9] indicated that Spectral Gradient based optimization can be successfully applied in denoising for a wide range of potential functions. Thus, we herein utilize an optimization method from the same family. SPG is an efficient tool for solving a constrained optimization problem  $\min_{x \in \Omega} f(x)$ , where  $\Omega$  is a closed convex set in  $\mathbb{R}^n$  and  $f$  is a function which has continuous partial derivatives on an open set that contains  $\Omega$ . Weak requirements on the objective function, as well as efficiency in solving large scale problems [1], make this optimization tool attractive for our purpose. The method is briefly outlined in Algorithm 1. We define the projection  $P_\Omega$  of a vector  $x \in \mathbb{R}^n$  to the feasible set  $\Omega = [0, 1]^n$  as:  $[P_\Omega(x)]_i = \min\{1, \max\{0, x_i\}\}$ , for all  $i = 1, 2, \dots, n$ .

A scaled version of SPG is used for image deblurring in [2]. However, data fidelity term is considered without regularization, and robustness of the solution is achieved by early stopping. The efficiency of the SPG method in regularized restoration of images degraded by both blur and noise is confirmed by this study.

### 3 Image deconvolution by SPG minimization of a regularized energy functional with potentials

In the observed model (1) we assume that the spatially invariant PSF is known, or can be estimated by point spread estimation techniques; the deblurring that we perform here belongs to the group of linear non-blind methods. We assume that acquired images are corrupted by additive Gaussian noise with a standard deviation  $\sigma_n$ . We consider grey scale images and represent them as vectors with intensity values from  $[0, 1]$ . Let the vector  $u = [u_1, \dots, u_n]^T$  of length  $n = r \times c$  represent an image  $u$  of size  $r \times c$ , where image rows are sequentially concatenated. Minimization of (2) can be seen as a constrained optimization problem:

$$\min_u E(u) \quad \text{s.t. } 0 \leq u_i \leq 1, \quad i = 1, 2, \dots, n. \quad (3)$$

A discrete formulation of the objective function (2) is:

$$E(u) = \frac{1}{2} \sum_{i=1}^n ((Hu - v)_i)^2 + \alpha \sum_{i=1}^n \phi(|\nabla(u_i)|), \quad (4)$$

where vector  $v$  is an observed image and  $H_{n \times n}$  is a block circulant matrix s.t.  $Hu$  is equal to convolution  $h * u$ .  $\nabla(u_i)$  is the discrete image gradient at point  $u_i$ , computed as  $\nabla(u_i) = (u_r - u_i, u_b - u_i)$ , where  $r$  and  $b$  denote indexes of the edge neighbours to the right and below the pixel  $u_i$ , respectively. The gradient of (4) is given by  $\nabla E(u) = [\nabla E(u)_i]_{i=1}^n$  and

$$\begin{aligned} \nabla E(u)_i &= (H^T(Hu - v))_i + \alpha \phi'(|\nabla(u_i)|) \frac{2u_i - u_r - u_b}{|\nabla(u_i)|} \\ &+ \alpha \phi'(|\nabla(u_i)|) \frac{u_i - u_l}{|\nabla(u_i)|} + \alpha \phi'(|\nabla(u_a)|) \frac{u_i - u_a}{|\nabla(u_a)|}, \end{aligned} \quad (5)$$

where  $u_a$  and  $u_l$  denote edge neighbours above and left of the pixel  $u_i$ , respectively. Edges are handled using periodic boundary condition.

The gradient defined by (5) is non-differentiable at points where  $|\nabla u_i| = 0$ . To meet requirements of SPG, we consider a smoothed version of (4), where  $|\nabla u_i|$  is replaced with  $\sqrt{|\nabla(u_i)|^2 + \varepsilon^2}$  and where  $\varepsilon$  is a small positive number (we used  $\varepsilon = 10^{-5}$  throughout). The use of a relaxed gradient could possibly lead to a less accurate solution. It was observed in [9] that differences are negligible.

### 4 Evaluation

To evaluate the performance of different potentials, we utilize ten standard images shown in Fig. 1. For every original image  $u^*$  we construct noisy and blurred image  $v$  by convolving it with PSF  $h$  and adding white Gaussian noise,  $v = h * u^* + \eta$ . We consider Gaussian PSFs, closely resembling real PSFs in many imaging systems. We evaluate PSFs with standard deviation  $\sigma_p \in \{1, 2, 3\}$



Fig. 1: Used test images, all  $256 \times 256$ . Intensities in  $[0, 255]$  are mapped to  $[0, 1]$ .

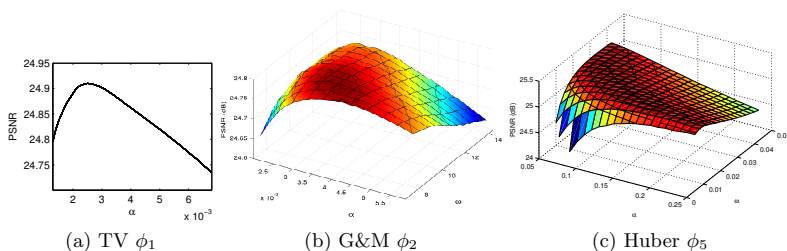


Fig. 2: Examples of average deblurring performance plotted for different parameter settings. Graphs for  $\phi_1, \phi_2, \phi_5, \sigma_p = 2$ , and  $\sigma_n^2 = 0.001$  are shown. Graphs for other potentials, PSFs, and noise levels exhibit similar characteristics.

and observe noise with variance  $\sigma_n^2 \in \{0, 0.0001, 0.001, 0.01\}$ . For each PSF and noise level, we obtain one degraded image  $v$  from which we reconstruct  $u^*$  using the seven considered potentials. Quality of reconstruction is measured with Peak Signal-to-Noise Ratio  $PSNR = 10 \log_{10} \left( \frac{(\max(u_i^*))^2}{MSE} \right)$ , where  $MSE = \frac{1}{n} \sum_{i=1}^n (u_i^* - \bar{u}_i)^2$  and  $\bar{u}$  is reconstructed image.

A number of approaches for selection of regularization parameter(s) (in our case  $\alpha$  and  $\omega$ ) exist [6, 13]. To ensure optimal selection of parameters we exhaustively explore the parameter space and selected the best performing separately for each PSF size  $\sigma_p$  and each noise level  $\sigma_n^2$  (i.e.,  $3 \times 4 \times 7$  sets of parameters). This leads to a positive bias on our results, since we perform evaluation on the training data. This bias does not favour any of the potentials. Partial evaluation on separated test and training sets, show that: (i) the bias is very limited, (ii) results are not overly sensitive to parameter tuning. Examples of typical deblurring performance (PSNR) for varying parameter settings are shown in Fig. 2.

Optimization is performed using SPG with settings recommended in [1]:  $\theta_{min} = 10^{-3}, \theta_{max} = 10^3, \gamma = 10^{-4}, \sigma_1 = 0.1, \sigma_2 = 0.9$ . Algorithm is terminated when the max-norm between two consecutive images is less than  $tol = 10^{-3}$ .

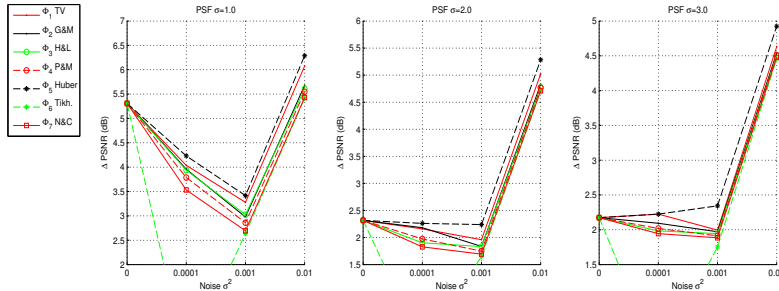


Fig. 3: Average improvement of PSNR over ten test images for different PSFs and different noise levels considering all seven evaluated potentials.

## 5 Results

The improvement in PSNR between before and after performed deblurring,  $\Delta\text{PSNR} = \text{PSNR}_{\text{out}} - \text{PSNR}_{\text{in}}$ , for each of the seven potentials, and each of the  $3 \times 4$  blur and noise levels, is presented in Fig. 3. Table 3 shows  $\Delta\text{PSNR}$ , as well as number of iterations, averaged over all images and all types of degradations. CPU time in seconds is approx. the number of iterations divided by 50 (Matlab, 3GHz Intel Core i7). A very clear result is that the Huber potential,  $\phi_5$ , shows superior performance in all of the evaluated settings. As a second runner-up comes TV based deblurring ( $\phi_1$ ), clearly behind in most situations, but providing a similar performance in the case  $\sigma_p = 3, \sigma_n^2 = 0.0001$ . On a third place comes the non-convex Geman & McClure potential ( $\phi_2$ ) which also showed to perform well in denoising [9]. The G&M potential performs slightly better than TV regularization for the case  $\sigma_p = 2, \sigma_n^2 = 0.0001$ . As opposed to the denoising study however, at no place does it outshine the Huber potential.

Visual examples of deblurring performance of TV, G&M, and Huber potentials are presented in Fig. 4. In Fig. 5 we show a zoomed-in view on the shoulder of the Cameraman, to highlight the edge preservation performance of the Huber potential over the commonly used TV regularization. It is apparent that the Huber potential does a much better job in preserving the sharp edges in the image (as also can be confirmed by looking at the residual errors in Fig. 5(d,e)).

Table 3: Average  $\Delta\text{PSNR}$  and number of iterations for the studied potentials.

Potential	$\phi_1$	$\phi_2$	$\phi_3$	$\phi_4$	$\phi_5$	$\phi_6$	$\phi_7$
$\Delta\text{PSNR}$ [dB]	3.43	3.32	3.28	3.24	<b>3.58</b>	2.55	3.17
No. iterations	30	65	26	52	40	11	30





Degraded image	$\phi_1$ (TV)	$\phi_2$ (G&M)	$\phi_5$ (Huber)
			
21.46 dB	23.75 dB	23.81 dB	24.18 dB
			
20.74 dB	22.61 dB	22.68 dB	23.27 dB

Fig. 4: First column: images degraded with PSF  $\sigma_p = 3$  and noise with variance  $\sigma_n^2 = 0.001$ . Columns 2–4: recovered images using best performing potentials  $\phi_1$ ,  $\phi_2$ , and  $\phi_5$ , respectively. PSNR is stated below each image.

## 6 Conclusions

Performed tests confirm that utilization of potential functions in regularized image denoising and deblurring provides a straightforward way to increase quality of the restored images. We have tested seven potential functions suggested in the literature, utilizing optimal parameter values for each of them, empirically found in our study. Optimization of both convex and non-convex energy functions is performed by a flexible and efficient SPG method. Our conclusion is that the Huber potential performs outstandingly best, providing best PSNR and improved edge preservation, compared to all the observed potentials.

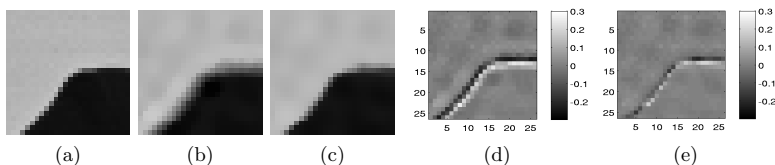


Fig. 5: Illustration of improved edge preservation by Huber potential,  $\phi_5$ . (a) Original image, part of Cameraman's shoulder. (b) Deblurred image using  $\phi_1$  (TV). (c) Deblurred image using  $\phi_5$ . (d) Residual for  $\phi_1$ . (e) Residual for  $\phi_5$ .

## Acknowledgment

Bajić, Lindblad and Sladoje are supported by the Ministry of Science of the Republic of Serbia through Projects ON174008 and III44006 of MI-SANU. Sladoje is supported by the Swedish Governmental Agency for Innovation Systems.

## References

1. Birgin, E., Martinez, J., Raydan, M.: Algorithm: 813: SPG - software for convex-constrained optimization. *ACM. Trans. Mathematical Software* 27, 340–349 (2001)
2. Bonettini, S., Zanella, R., Zanni, L.: A scaled gradient projection method for constrained image deblurring. *Inverse Problems* 25, 015002 (2009)
3. Charbonnier, P., Blanc-Féraud, L., Aubert, G., Barlaud, M.: Deterministic edge-preserving regularization in computed imaging. *IEEE Trans. Image Process.*, 6(2), 298–311 (1997)
4. Geman, S., McClure, D.: Bayesian image analysis: an application to single photon emission tomography. *Proc. Statistical Computation Section, Am. Statistical Assoc. (Washington, DC)*, 12–18 (1985)
5. Getreuer, P.: Total variation deconvolution using Split Bregman. *Image Processing On Line*, 2, 158–174 (2012)
6. Hansen, P. C., Kilmer, M. E., Kjeldsen, R. H.: Exploiting residual information in the parameter choice for discrete ill-posed problems. *BIT Numerical Mathematics*, 46(1), 41–59 (2006)
7. He, C., Hu, C., Zhang, W., Shi, B., Hu, X.: Fast total variation image deconvolution with adaptive parameter estimation via Split Bregman method. *Mathematical Problems in Engineering*, Vol. 2014, Article ID 617026 (2014)
8. Hebert, T., Leahy, R.: A generalized EM algorithm for 3D Bayesian reconstruction from Poisson data using Gibbs priors. *IEEE Trans. Med. Imaging* 8, 194–202 (1989)
9. Lukić, T., Lindblad, J., Sladoje, N.: Regularized image denoising based on spectral gradient optimization. *Inverse Problems* 27, 085010 (2011)
10. Nikolova, M., Chan, R.: The equivalence of half-quadratic minimization and the gradient linearization iteration. *IEEE Trans. Image Process.* 16, 1623–1627 (2007)
11. Nikolova, M., Ng, M.: Analysis of half-quadratic minimization methods for signal and image recovery. *SIAM J. Sci. Comput.* 27, 937–966 (2005)
12. Perona, P., Malik, J.: Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Anal. Mach. Intell.* 12, 345–362 (1990)
13. Ramani, S., Blu, T., Unser, M.: Monte-Carlo SURE: A black-box optimization of regularization parameters for general denoising algorithms. *IEEE Trans. Image Processing*, 17(9), 1540–1554 (2008)
14. Rodriguez, P., Wohlberg, B.: Efficient minimization method for generalized total variation functional. *IEEE Trans. Image Processing*, Vol. 18, No.2, 322–332 (2009)
15. Rudin, L.I., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. *Physica D*, Vol.60, No. 1-4, 259–268 (1992)
16. Tikhonov, A., Arsenin, V.: *Solutions of ill-posed problems.* (Washington, DC: Winston and Wiley) (1977)
17. Schultz, R., Stevenson, R.: Stochastic modeling and estimation of multispectral image data. *IEEE Trans. Image Process.* 4, 1109–1119 (1995)

## Publication V

### Appeared in:

**Bajić B\***, Suveer A\*, Gupta A\*, Pepić I, Lindblad J, Sladoje N, and Sintorn I-M.

Denoising of Short Exposure Transmission Electron Microscopy Images for Ultrastructural Enhancement. In *Proceedings of the 15<sup>th</sup> IEEE International Symposium on Biomedical Imaging (ISBI)*, Washington D.C., USA, April 2018, pp. 921–925.

---

\* *These authors have contributed equally.*



# DENOISING OF SHORT EXPOSURE TRANSMISSION ELECTRON MICROSCOPY IMAGES FOR ULTRASTRUCTURAL ENHANCEMENT

*Buda Bajić<sup>1\*</sup>, Amit Suveer<sup>2\*</sup>, Anindya Gupta<sup>3\*</sup>, Ivana Pepić<sup>1</sup>  
Joakim Lindblad<sup>2,4</sup>, Nataša Sladoje<sup>2,4</sup>, Ida-Maria Sintorn<sup>2,5</sup>*

<sup>1</sup> Faculty of Technical Sciences, University of Novi Sad, Serbia

<sup>2</sup> Centre for Image Analysis, Dept. of Information Technology, Uppsala University, Sweden

<sup>3</sup> T. J. Seebeck Dept. of Electronics, Tallinn University of Technology, Estonia

<sup>4</sup> Mathematical Institute of the Serbian Academy of Sciences and Arts, Belgrade, Serbia

<sup>5</sup> Vironova AB, Stockholm, Sweden

## ABSTRACT

Transmission Electron Microscopy (TEM) is commonly used for structural analysis at the nm scale in material and biological sciences. Fast acquisition and low dose are desired to minimize the influence of external factors on the acquisition as well as the interaction of electrons with the sample. However, the resulting images are very noisy, which affects both manual and automated analysis. We present a comparative study of block matching, wavelet domain, energy minimization, and deep convolutional neural network based approaches to denoise short exposure high-resolution TEM images of cilia. In addition, we evaluate the effect of denoising before or after registering multiple short exposure images of the same imaging field to further enhance fine details.

**Index Terms**— Denoising, Convolutional Neural Networks, TEM, Cilia

## 1. INTRODUCTION

Transmission Electron Microscopy (TEM) is an imaging technique providing nm resolution. It is therefore well suited and often used to analyze structural details in biological samples and tissue sections for research and clinical diagnostics. However, both manual and automated analysis of TEM images are negatively affected by a number of imaging factors, such as sample preparation artefacts, non-optimal microscope alignment and focusing, electrons interacting with and modifying the sample, and motion artefacts from e.g. sample drift and vibrations. Preprocessing with an aim to enhance the relevant details (ultrastructures) is often applied.

The imaging artefacts can be reduced by decreasing the electron dose and acquisition time. However, this results in images with more noise and increases the need for denoising. The noise induced by TEM is non-additive and signal-dependent. It can be modeled by a mixed Poisson-Gaussian

(PG) distribution [1, 2]. However, in short exposure images, the Gaussian noise dominates. We consider three classical methods suited for Gaussian and PG noise: a block matching [3], wavelet domain [4], and energy minimization [5] based method, and evaluate their performances on short exposure TEM images of cilia in a cell section sample, Fig. 1. Moreover, observing that convolutional neural networks (CNNs) have recently been shown to perform well in denoising [6, 7], we have developed a suitable denoising CNN model and included it in the comparison. To the best of our knowledge, this is the first denoising CNN evaluated on TEM noisy data.

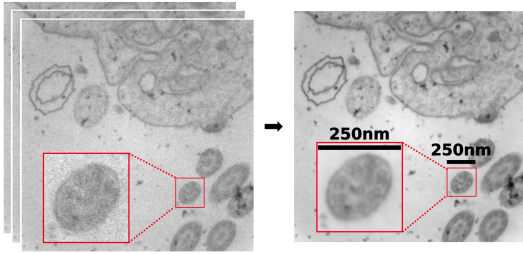
Denoising is commonly performed on a single image. However, our ultimate goal is to enhance fine details in TEM images, which, in theory, can be achieved by generating synthetic long exposure images by aggregating (median) a number of short exposure ones. We, therefore, also investigate two strategies of combining aggregation and denoising: (1) co-registration and aggregation of a number of short exposure images is performed first and followed by denoising of the aggregation; (2) short exposure images are denoised and the resulting ones are then co-registered and aggregated. Enhancement of structural information by registration and aggregation of scanned lines, images or objects, is commonly used in other biomedical imaging techniques, e.g., in scanning transmission electron microscopy (STEM) [2, 8], and cryo-EM [9].

## 2. DENOISING METHODS

### 2.1. Block-matching and 3D filtering (BM3D)

Block matching based techniques utilize self similarities present in the image. The BM3D algorithm [3] is suitable for images with structural redundancy, which is common in biological images, and also in our case. BM3D has successfully been used for denoising light microscopy images [10] and STEM images [11].

\*These authors have contributed equally.



**Fig. 1:** **Left:** Short exposure TEM image ( $2048 \times 2048$  pixels) from a series of 100 images. **Right:** Ground truth created by co-registration and aggregation of the stack to the left. The two insets show magnified views ( $250 \times 250$  pixels) of one cilium.

## 2.2. PURE-LET

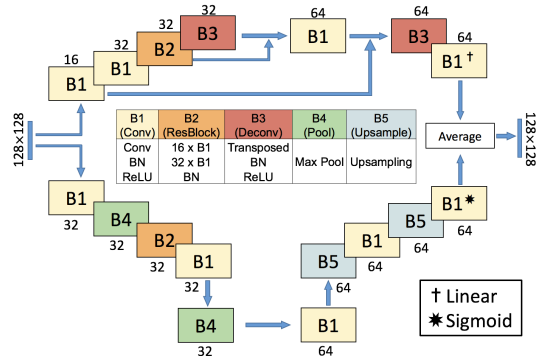
In the PURE-LET [4] method the denoising process is expressed as a linear expansion of thresholds (LET). The threshold optimization solely relies on a data-adaptive unbiased estimation of the mean squared error (MSE), derived in a non-Bayesian framework (PURE: Poisson–Gaussian unbiased risk estimate, defined in the Haar wavelet domain). The method is suitable for light microscopy images, as presented in the original paper, and it also performs well in restoring STEM images as shown in [12].

## 2.3. Energy minimization (EM)

Many denoising methods are based on solving an inverse problem through energy minimization. We perform denoising by minimizing an energy function which includes a quadratic data fidelity term, suited for Gaussian noise, and a regularization term which provides numerical stability and noise suppression. We use Total Variation (TV) regularization [13] smoothed by the Huber potential function [14], resulting in well preserved edges in images [5]. We have previously shown applicability of this approach to cilia ultrastructure enhancement in long exposure images [15], where we applied a generalized version of the method suited to PG noise and blind deblurring.

## 2.4. Denoising Convolutional Neural Network (DCNN)

Inspired by the good performance of the approaches in [16, 17], we jointly train two CNNs as an ensemble. The architecture is shown in Fig. 2. The training of both streams is performed on image patches of size  $128 \times 128$  with an overlapping stride of 16 pixels. Prior to the training, the patches are normalized to the range  $[0,1]$ . The first stream consists of four convolution blocks, two transposed convolution blocks and one residual block. The convolution block encodes the image representations while removing the noise, whereas the



**Fig. 2:** The two-stream DCNN architecture. The sizes of output feature maps of each block are shown on top of each block and generated using  $3 \times 3$  convolutions. The last  $1 \times 1$  convolution blocks of each stream use linear and sigmoid activation, respectively, instead of ReLU.

transposed convolution block decodes these representations to restore the noise-free image content. The residual block contains two convolution blocks. Batch normalization (BN) [18] is used as regularization before rectified linear unit (ReLU) activation to deal with internal covariate shift. To elevate the training performance, skip connections are used and followed by a BN layer. During experiments, we found that the prediction made by the first stream restores most content with blur. Considering that, we incorporated a second stream consisting of four convolution blocks, two up-sampling blocks, two max-pooling layers, and one residual block. The reconstructed output of the second stream contains high-frequency content, however, with an inconsistent illumination in respect to corresponding ground truth. Motivated by the above observations, we performed an end-to-end training by averaging the outputs of both streams, which resulted in an improved output. We used stochastic gradient descent (SGD) to optimize the weights in a mini-batch scheme of 16 patches. The initial learning rate was set to 0.001, and reduced to 1/10 of the current value after every epoch. We used MSE and binary cross-entropy as loss function. The DCNN is implemented using Tensorflow backend in Keras [19] and trained for 15 epochs in a five-fold cross validation scheme. The average training time is 300 s/epoch on a GPU GeForce GTX 1080.

## 3. EXPERIMENTS AND RESULTS

### 3.1. Quantitative evaluation

The dataset consists of a series of 100 noisy short exposure (2 ms) images, captured at the same spatial location in the cell section sample (FoV=2000 nm). All images are of size  $2048 \times 2048$  pixels and acquired with the low-voltage

**Table 1:** Results on the test data set. Average PSNR and SSIM ( $\pm$  SD) over 90 single images are given in the 1st and 2nd rows. Rows 3 and 4 contain average PSNR and SSIM over 18 aggregated groups of 5 short exposure images followed by denoising. Average PSNR and SSIM over 18 images each obtained by aggregating 5 denoised short exposure images, are given in rows 5 and 6. Optimal parameters (in parentheses) estimated during the training phase are used. Best performances are bolded.

		Initial	BM3D ( $\sigma_{bm}$ )	PURE-LET ( $\sigma_{pl}$ )	EM ( $\lambda$ )	DCNN
Denoising of single images	PSNR	22.25	37.39 $\pm$ 0.30 (105)	38.44 $\pm$ 1.09 (75)	37.80 $\pm$ 0.27 (0.25)	<b>38.04</b> $\pm$ 0.21
	SSIM	0.019	0.233 $\pm$ 0.007 (95)	0.219 $\pm$ 0.007 (55)	<b>0.255</b> $\pm$ 0.027 (0.20)	0.252 $\pm$ 0.002
Denoising of 5 aggregated noisy images	PSNR	27.88	40.45 $\pm$ 1.09 (95)	40.19 $\pm$ 1.06 (35)	40.19 $\pm$ 0.54 (0.125)	<b>40.86</b> $\pm$ 0.37
	SSIM	0.037	0.270 $\pm$ 0.019 (35)	0.263 $\pm$ 0.017 (25)	0.277 $\pm$ 0.017 (0.10)	<b>0.282</b> $\pm$ 0.011
Aggregation of 5 denoised single images	PSNR	22.25	39.65 $\pm$ 1.04 (95)	40.21 $\pm$ 0.48 (55)	39.92 $\pm$ 0.93 (0.10)	<b>40.84</b> $\pm$ 0.45
	SSIM	0.019	0.261 $\pm$ 0.013 (25)	0.265 $\pm$ 0.011 (45)	0.273 $\pm$ 0.021 (0.075)	<b>0.276</b> $\pm$ 0.009

MiniTEM<sup>1</sup> system. A low-noise image, used as a ground-truth, is estimated by registering each short exposure image to the first image of the series using rigid registration, followed by aggregating the information by computing the pixel-wise median value, illustrated in Fig. 1.

We utilize 10 images for the training of the DCNN<sup>2</sup> and explorative parameter tuning of each method – the regularization weight  $\lambda$  for EM, and the expected std. of Gaussian noise,  $\sigma_{bm}$  and  $\sigma_{pl}$ , for BM3D<sup>3</sup> and PURE-LET<sup>4</sup>, respectively. The tuned parameters are used to compare the performance of each method on the remaining 90 images. Apart from evaluating the performances on denoising single images, we additionally tune parameters and evaluate the performances of the methods when used to 1) denoise the resulting image after registering and aggregating (median) five short exposure images, and 2) when registering and aggregating (median) five denoised short exposure images.

The performance is evaluated using well known and often used the peak-signal-to-noise ratio (PSNR) and structural similarity index measure (SSIM) [20]. As indicated in [21], different levels of degradations applied to the same image can yield the same PSNR. We observe that PSNR performs poorly on discriminating structural content in images which plays an important role in ultrastructural analysis of TEM images. As SSIM is proposed with the aim to compare structural changes in images imitating what the human visual system does, this measure is considered a more reliable measure of visual similarity of images.

**Denoising of single short exposure images** – The average PSNR and SSIM over all 90 images from the test dataset are given in Table 1, along with the parameters tuned during the training. The EM method marginally outperforms the remaining methods in terms of SSIM. On the other hand, DCNN outperforms all classical methods in terms of PSNR. A cilium from a single noisy image and the corresponding denoised instances obtained with all 4 methods are presented in

the first row of Fig. 3.

**Denoising of 5 aggregated short exposure images** – We register groups of 5 short exposure images and aggregate them by the pixel-wise median. We denoise the resulting 18 images by all 4 considered methods. The average PSNR and SSIM (over 18 images) are given in Table 1. As confirmed by both average PSNR and SSIM, the DCNN method outperforms the other methods. A noisy cilium instance from aggregating 5 short exposure images and the corresponding denoised results obtained with all 4 methods are presented in the middle row of Fig 3.

**Aggregation of 5 denoised short exposure images** – We denoise 5 sequentially acquired short exposure images, then register them and aggregate them by the pixel-wise median. The average PSNR and SSIM for the 18 resulting images are given in Table 1. The corresponding results on the cilium subimage are shown in the bottom row of Fig. 3. Note that the first image is the ground truth, i.e., the median aggregated 100 short exposure images. In this strategy as well, the DCNN produces the highest PSNR and SSIM.

### 3.2. Qualitative evaluation

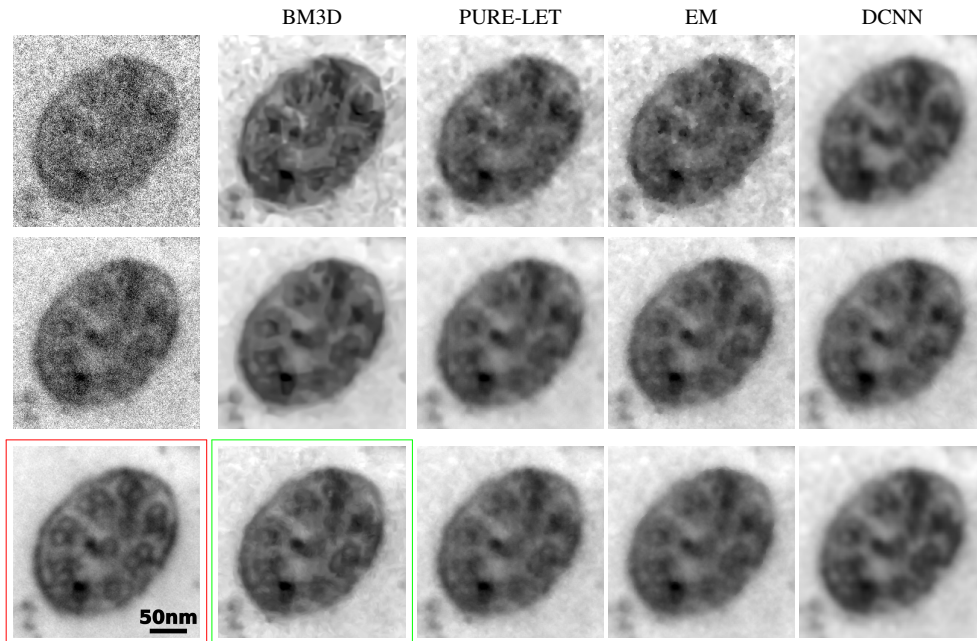
To validate the level of agreement between the quantitative results and visual (qualitative) results, we performed a subjective visual evaluation conducting a two-step voting process by six experts. In the first step, involving only the classical methods, the experts rated the results (1st, 2nd, 3rd best) on the cilium subimage produced by each of the methods with different parameter settings. The displayed images (7 for each method) spanned a parameter range centered around the maximal SSIM for that method. The procedure was repeated for the two strategies of aggregating 5 short exposure images (denoising prior or post registering and aggregation). The second step involves all four methods. The images resulting from the two aggregation strategies utilizing the tuned parameter settings as decided in Step 1, together with the DCNN results were displayed (random, unknown order) and the experts rated them again (as the 1st, 2nd, 3rd best). The denoised image with the majority of votes is highlighted in Fig. 3.

<sup>1</sup>Vironova AB, Stockholm, Sweden

<sup>2</sup>[https://bitbucket.org/anindya\\_gupta/tem-denoising/](https://bitbucket.org/anindya_gupta/tem-denoising/)

<sup>3</sup><http://www.cs.tut.fi/~foi/GCF-BM3D>

<sup>4</sup><http://bigwww.epfl.ch/algorithms/denoise/>



**Fig. 3:** Noisy and denoised close ups of a cilium instance obtained with the considered methods. **Top:** Denoising of a single image. **Middle:** Denoising of 5 aggregated noisy images. **Bottom:** Aggregation of 5 denoised single images. The red frame (bottom left) indicates the ground truth. The green frame indicates the best ranked image in the two-step visual voting process.

In the first step of the voting procedure, the experts' votes agreed well with the quantitative results based on SSIM. However, in some cases, the experts visually preferred slightly less regularized images. This is not surprising since humans prefer to see sharp details and can "ignore" noise to some degree. The results corresponding to maximal PSNR were consistently judged as over-regularized.

#### 4. DISCUSSION AND CONCLUSION

Short exposure time reduces the influence of motion blur and electron interaction with the sample. That, however, affects the image quality. We have quantitatively and qualitatively compared four different denoising methods that can be used to improve the resulting poor image quality. To additionally enhance ultrastructural information in TEM images, we have investigated two strategies i.e., denoising of aggregated series of noisy images and aggregation of several denoised short exposure images of the same view.

From the quantitative and qualitative results in Table 1 and Fig. 3 it is clear that denoising can improve both single and multiple aggregated short exposure images. Comparatively, noisy single images require more regularization. It is

also interesting to note that the optimal parameter values for the classical methods differ a lot depending on whether single short exposure or aggregated images are to be denoised. Note that the DCNN was only trained on single frames also for the strategies using aggregated images. Overall, DCNN gives the highest quantitative scores, but based on the visual assessment BM3D applied to noisy images prior to aggregation produced the most appealing result.

Both of the two aggregation strategies, denoising the registered and aggregated image or registering and aggregating after denoising the short exposure images, improve the results approximately equally well. One advantage with the former aggregation strategy is that only one denoising computations is performed instead of five.

#### 5. ACKNOWLEDGMENT

This work is supported by VINNOVA, MedTech4Health grant 2016-02329, the Ministry of Education, Science, and Techn. Development of the Rep. of Serbia (proj. ON174008 and III44006), Swedish Research Council grant 2014-4231, the IT Acad. Prog. and EU Inst. grant IUT19-11 of ERC.

## 6. REFERENCES

- [1] M. Vulović, E. Franken, R. B. G. Ravelli, L. J. van Vliet, and B. Rieger, “Precise and unbiased estimation of astigmatism and defocus in Transmission Electron Microscopy,” *Ultramicroscopy*, vol. 116, pp. 115–134, 2012.
- [2] B. Berkels and B. Wirth, “Joint denoising and distortion correction of atomic scale Scanning Transmission Electron Microscopy images,” *arXiv preprint arXiv:1612.08170*, 2016.
- [3] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, “Image denoising by sparse 3-D transform-domain collaborative filtering,” *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, 2007.
- [4] F. Luisier, T. Blu, and M. Unser, “Image denoising in mixed Poisson–Gaussian noise,” *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 696–708, 2011.
- [5] B. Bajić, J. Lindblad, and N. Sladoje, “Restoration of images degraded by signal-dependent noise based on energy minimization: an empirical study,” *J. of Electronic Imaging*, vol. 25, no. 4, pp. 043020–043020, 2016.
- [6] X. Mao, C. Shen, and Y-B. Yang, “Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections,” in *Advances in Neural Info. Process. Systems 29*, pp. 2802–2810, 2016.
- [7] J. Kim, J. Kwon Lee, and K. Mu Lee, “Accurate image super-resolution using very deep convolutional networks,” in *IEEE Conf. on Comp. Vis. and Patt. Recog. (CVPR)*, 2016.
- [8] A. B. Yankovich, C. Zhang, A. Oh, T. J. A. Slater, F. Azough, R. Freer, S. J. Haighand, R. Willett, and P. M. Voyles, “Non-rigid registration and non-local principle component analysis to improve electron microscopy spectrum images,” *Nanotechnology*, vol. 27, no. 36, pp. 364001, 2016.
- [9] Y. Cheng, N. Grigorieff, P. A. Penczek, and T. Walz, “A primer to single-particle cryo-electron microscopy,” *Cell*, vol. 161, no. 3, pp. 438–449, 2015.
- [10] J. Boulanger, C. Kervrann, P. Bouthemy, P. Elbau, J. B. Sibarita, and J. Salamero, “Patch-based nonlocal functional for denoising fluorescence microscopy image sequences,” *IEEE Trans. on Med. Imaging*, vol. 29, no. 2, pp. 442–454, 2010.
- [11] N. Mevenkamp, P. Binev, W. Dahmen, P. M. Voyles, A. B. Yankovich, and B. Berkels, “Poisson noise removal from high-resolution STEM images based on periodic block matching,” *Adv. Structural and Chemical Imaging*, vol. 1, no. 1, pp. 3, 2015.
- [12] N. Mevenkamp, A. B. Yankovich, P. M. Voyles, and B. Berkels, “Non-local means for Scanning Transmission Electron Microscopy images and Poisson noise based on adaptive periodic similarity search and patch regularization,” in *Vision Modeling and Visualization*, 2014, pp. 63–70.
- [13] L. I. Rudin, S. Osher, and E. Fatemi, “Nonlinear Total Variation based noise removal algorithms,” *Physica D: Nonlinear Phenomena*, vol. 60, pp. 259–268, 1992.
- [14] R. Schultz and R. Stevenson, “Stochastic modeling and estimation of multispectral image data,” *IEEE Trans. Image Process.*, vol. 4, no. 8, pp. 1109–1119, 1995.
- [15] B. Bajić, J. Lindblad, and N. Sladoje, “Blind restoration of images degraded with mixed Poisson-Gaussian noise with application in Transmission Electron Microscopy,” in *IEEE Int. Symp. on Biomed. Img. (ISBI 2016), proceedings*, 2016, pp. 123–127.
- [16] C. Ledig, L. Theis, F. Huszar, J. Caballero, A.P. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, “Photo-realistic single image super-resolution using a generative adversarial network,” in *IEEE Conf. on Comp. Vis. and Patt. Recog. (CVPR)*, 2017.
- [17] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with Deep Convolutional Neural Networks,” in *Advances in Neural Info. Proc. Systems 25*, pp. 1097–1105, 2012.
- [18] S. Ioffe and C. Szegedy, “Batch normalization: Accelerating deep network training by reducing internal covariate shift,” in *Int. Conf. on ML*, 2015, pp. 448–456.
- [19] F. Chollet, “Keras,” <https://github.com/fchollet/keras>, 2015.
- [20] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, 2004.
- [21] Z. Wang and A. C. Bovik, “Mean squared error: Love it or leave it? A new look at signal fidelity measures,” *IEEE Signal Process. Magazine*, vol. 26, no. 1, pp. 98–117, 2009.

## Publication VI

### Appeared in:

**Bajić B**, Lindblad J and Sladoje N.

Super-resolution coverage segmentation by linear unmixing in presence of blur and noise. *SPIE Journal of Electronic Imaging*, 2019, 28(1), 013046.

# Sparsity promoting super-resolution coverage segmentation by linear unmixing in presence of blur and noise

Buda Bajić,<sup>a,\*</sup> Joakim Lindblad,<sup>b,c</sup> and Nataša Sladoje<sup>b,c</sup>

<sup>a</sup>University of Novi Sad, Faculty of Technical Sciences, Serbia

<sup>b</sup>Uppsala University, Centre for Image Analysis, Department of Information Technology, Sweden

<sup>c</sup>Mathematical Institute of the Serbian Academy of Sciences and Arts, Belgrade, Serbia

**Abstract.** We present a segmentation method that estimates the relative coverage of each pixel in a sensed image by each image component. The proposed super-resolution blur-aware model (utilizes *a priori* knowledge of the image blur) for linear unmixing of image intensities relies on a sparsity promoting approach expressed by two main requirements: (i) minimization of Huberized total variation, providing smooth object boundaries and noise removal, and (ii) minimization of nonedge image fuzziness, responding to an assumption that imaged objects are crisp and that fuzziness is mainly due to the imaging and digitization process. Edge fuzziness due to partial coverage is allowed, enabling subpixel precise feature estimates. The segmentation is formulated as an energy minimization problem and solved by the spectral projected gradient method, utilizing a graduated nonconvexity scheme. Quantitative and qualitative evaluation on synthetic and real multichannel images confirms good performance, particularly relevant when subpixel precision in segmentation and subsequent analysis is a requirement. © 2019 SPIE and IS&T [DOI: 10.1117/1.JEI.28.1.013046]

Keywords: fuzzy segmentation; super-resolution; deconvolution; linear unmixing; total variation; energy minimization.

Paper 180835 received Sep. 25, 2018; accepted for publication Feb. 7, 2019; published online Feb. 26, 2019.

## 1 Introduction

Segmentation is generally considered one of the most important and at the same time most challenging tasks in image processing. A decision if a pixel belongs to the object of interest or not is crucial for the quality of all following analysis steps and is often very difficult to make. Numerous segmentation methods have been proposed in the literature, tuned to different applications, and dealing with challenges resulting from imaging conditions in a variety of ways. Most methods perform crisp segmentation, i.e., they assign a categorical label to each pixel in an image, indicating the single component to which the pixel belongs. No matter how sophisticated and well chosen for a particular problem, a crisp segmentation method cannot successfully address the issue of pixels partly covered by more than one object. This issue is caused by the discrete object representation itself and requires improved (noncrisp) object modeling, rather than improved crisp segmentation approaches. Coverage segmentation, a special case of fuzzy segmentation, assigns a vector with elements in the range  $[0, 1]$  to each pixel, indicating the degree of (partial) coverage of that pixel by each of the image components.

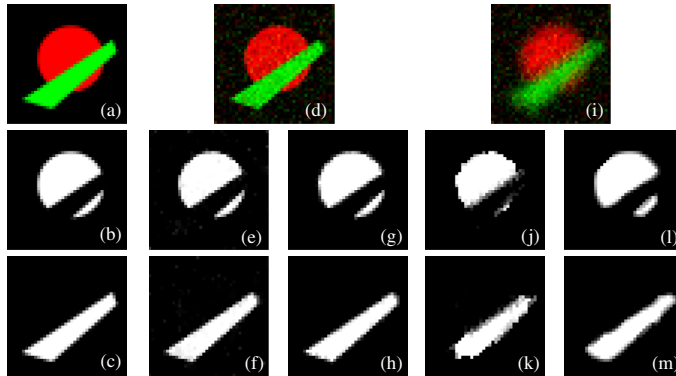
An example of an application where information about the partial coverage of pixels has shown to be beneficial is remote sensing. The relatively large pixels, characteristic for remote sensing images, are often partially covered by different classes/objects imaged on the ground (e.g., water, soil, and forest), and the subsequent image analysis can be significantly improved from subpixel precision in the segmentation. As a response to this need, subpixel proportion estimation, leading to so-called fraction images, is studied in

a number of publications. A frequently used approach is linear unmixing,<sup>1–7</sup> the measured spectrum of a mixed pixel is decomposed into a collection of constituent spectra (end-members, representing classes in the scene), accompanied by a set of corresponding fractions (abundances) that indicate the proportion of each end-member in the linear combination expressed in the pixel. When mixtures in pixels are nonlinear, more complex methods, e.g., deep neural networks,<sup>8</sup> can be utilized to estimate abundances, however, requiring significant amounts of annotated training data.

Interest for subpixel/voxel segmentation is not limited to low-resolution images. The so-called partial volume effect is particularly observed in 3-D medical image analysis, referring to voxels partially covered by two or more tissue types.<sup>9,10</sup> This issue, with its negative influence on the subsequent relevant measurements (e.g., volume estimation), is particularly studied for MR and PET images of the human brain. A study in Ref. 11 shows that consistent misplacing of the tissue borders, in a brain volume having voxels of size  $1 \text{ mm}^3$ , by only one voxel results in volume errors of ~30%, 40%, and 60% for white matter, gray matter, and cerebrospinal fluid, respectively.

Several studies show advantages of coverage segmentation and confirm that the coverage information enables increased precision and accuracy of feature estimates, as well as their decreased noisy sensitivity.<sup>12–16</sup> Methods for coverage segmentation are proposed in Refs. 17–20. Applications in medical image analysis are presented in, e.g., Refs. 21 and 22, where coverage segmentation of blood vessels at increased spatial resolution in 2-D and 3-D is proposed, and in Ref. 23, where coverage segmentation is shown to be useful in the analysis of dermoscopic images.

\*Address all correspondence to Buda Bajić, E-mail: [buda.bajic@uns.ac.rs](mailto:buda.bajic@uns.ac.rs)



**Fig. 1** An example of how the proposed method improves segmentation of a blurred test image. (a) Test image of size  $40 \times 40 \times 2$ ; (b) and (c) ground truth coverage segmentation (background component not shown). (d) Image degraded by Gaussian noise with  $\sigma_n = 0.1$ ; (e) and (f) high quality segmentation using the method in Ref. 18; and (g) and (h) high quality segmentation using the proposed method. (i) Image degraded by Gaussian blur with  $\sigma_b = 2$  and Gaussian noise ( $\sigma_n = 0.1$ ); (j) and (k) low quality segmentation using the method in Ref. 18; and (l) and (m) improved segmentation using the proposed method.

Existing coverage segmentation algorithms rely on the assumption that the mixed (partly covered) pixels form a one-pixel-thin boundary between homogeneous regions of pure pixels. However, this assumption is often violated in presence of image blur, which has a negative impact on the resulting segmentation. An illustrative example is shown in Fig. 1: the two objects in image a (individually shown in b and c) are successfully segmented by the method presented in Ref. 18 in the blur-free case, even in presence of noise (image d, segmented as e and f), whereas the method performs poorly on the blurry (and noisy) image i (segmented as j and k). The here proposed method performs well in both situations (images g and h, and l and m).

A general drawback of linear unmixing is that the number of end-members is limited by the number of spectral bands, otherwise the problem becomes underdetermined. In this paper, we present a coverage segmentation approach that, by promoting sparse unmixing and through spatial regularization, successfully addresses this, and all other above mentioned challenges. The proposed method is applicable to images degraded by blur and noise, and with possibly linearly dependent end-members. The proposed method assumes the blur matrix as input. An additional advantage of the proposed method is that it enables coverage estimation at an increased spatial resolution. Advantages and limitations

of this very versatile method, which can be applied to single- or multi-channel image data in a wide range of applications, are shown in Fig. 2, where coverage segmentation of a color licence plate, heavily blurred by motion blur, is performed at increased spatial resolution. This example uses  $\mu = 0.001$  (while other parameters are at recommended values, see Sec. 3.3) and converges to (c) after 16 iterations of Algorithm 1 (Sec. 3.3).

Sparsity promoting methods are often utilized in various image analyses and computer vision tasks (e.g., Refs. 24 and 25), showing to provide good performance; see e.g., Refs. 26 and 27 for an overview. In the context of image segmentation, most sparse methods rely on dictionary learning, e.g., in Ref. 28 dictionaries and classifiers are learned simultaneously from a set of brain atlases, which can then be used for the reconstruction and segmentation of an unseen target image. In this work, we do not rely on any learning step using annotated training data, but aim instead for a directly applicable, general, yet well performing, segmentation method.

The paper is organized as follows: in Sec. 2, we briefly review main definitions and notions, as well as the existing results on coverage segmentation by energy minimization relevant for this study. In Sec. 3, we present an algorithm for super-resolution blur-aware coverage segmentation based on minimization of our proposed energy function. Section 4



**Fig. 2** Segmentation of a heavily blurred licence plate. Original image (a) is blurred with PSF shown in the top left corner of (a), and then subsampled  $2\times$  to create the sensed image (b). Segmentation of (b) by the proposed method (SRCS-HTVF; see Sec. 3) using a three-component end-member matrix [shown as colored stripes in the top left corner of (b)] at  $2\times$  increased resolution provides the segmentation (c) (shown as a composite of the three components). The larger blue characters are well segmented, whereas the smaller red characters demonstrate limits of the method's performance.



presents performance analysis of the proposed method evaluated by experiments on simulated and real data. Section 5 concludes the paper.

## 2 Background and Previous Work

We start this section by defining the problem and related notations. Following, we summarize existing results related to coverage segmentation, deblurring, and super-resolution reconstruction by energy minimization, and we provide a brief analysis of terms utilized in previously proposed energy functions. We conclude the section by providing a list of contributions of this study.

### 2.1 Preliminaries

Let  $I: I_D \rightarrow \mathbb{R}^b$  denote a sensed multichannel digital image with  $b \in \mathbb{N}$  bands and  $r \times c = N$  pixels on a discrete domain  $I_D \subset \mathbb{Z}^2$ . Let a pixel  $p_{(x,y)} \subset \mathbb{R}^2$  denote the Voronoi region of a grid point  $(x, y) \in \mathbb{Z}^2$ . Let the set of  $m$ -component segmentation vectors be

$$\mathbb{A}_m = \left\{ \boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_m) \in [0, 1]^m \mid \sum_{j=1}^m \alpha_j = 1 \right\}. \quad (1)$$

We assume that the imaged components are continuous nonoverlapping crisp sets and the union of all components (including background) fills the image space, i.e., coverage values (a.k.a. abundances) of each pixel sum up to one.

A coverage segmentation of an image  $I$  into  $m$  components is a set of ordered pairs

$$\{(x, y), \boldsymbol{\alpha}(x, y) \mid (x, y) \in I_D, \boldsymbol{\alpha}(x, y) \in \mathbb{A}_m\},$$

$$\alpha_j \approx \frac{|p_{(x,y)} \cap S_j|}{|p_{(x,y)}|}, \quad (2)$$

where  $S_j \subset \mathbb{R}^2$  is the extent of the  $j$ 'th image component,  $j = 1, 2, \dots, m$ , and  $\boldsymbol{\alpha}(x, y)$  is the coverage vector assigned to pixel  $p_{(x,y)}$ , with coverage values  $\alpha_j$ . The continuous sets  $S_j$  are, in general, not known, and the values  $\alpha_j$  have to be estimated from the image data. Their estimation is the subject of this study.

We represent the sensed image  $I$  as a matrix of size  $N \times b$ , s.t. a row of  $I$  contains intensities of one pixel in each of the observed bands, and a column represents the pixel intensities in one band, concatenated over the whole image. Similarly, we represent the  $m$ -component coverage segmentation of  $I$  by a matrix  $A = [\alpha_{i,j}]_{N \times m}$ , where  $\alpha_{i,j} \in [0, 1]$  is the coverage of the  $i$ 'th pixel by the  $j$ 'th component. Rows of  $A$  with the value 1 at the  $j$ 'th position and zeros elsewhere correspond to pure pixels (completely covered by image component  $S_j$ ), whereas rows with two or more nonzero elements correspond to mixed pixels. When, as assumed in this study, the observed objects are crisp and with thin boundaries, the coverage segmentation matrix  $A$  is sparse; only a small number of pixels are mixed, and therefore the matrix contains relatively few nonzero elements.

Assuming that pixel intensities are obtained as linear mixtures of pure class representatives (end-members), the sensed image  $I$  can be modeled as

$$I = A \cdot C + \text{noise}, \quad (3)$$

where  $C = [c_{j,k}]_{m \times b}$  is the end-member matrix and  $c_{j,k}$  corresponds to the (expected) image intensity value of class  $j$  in band  $k$ . Knowing  $I$  and the end-member matrix  $C$ , our aim in this study is to estimate the coverage segmentation matrix  $A$ .

### 2.2 Energy Minimization-Based Coverage Segmentation

An energy minimization formulation of the segmentation problem was introduced by Mumford and Shah<sup>29</sup> and has, together with its numerous variations, been thoroughly studied and utilized in different applications.<sup>30</sup> An energy minimization approach can be used to address coverage segmentation as well. We have, in Ref. 18, proposed to compute the coverage segmentation  $\hat{A}$  of an image  $I$  as

$$\hat{A} = \arg \min_{A \in \mathbb{A}_{N \times m}} E(A; I), \quad (4)$$

and shown that the regularized energy function

$$E(A; I) = D(A; I) + \mu P(A) + \nu [T(A) + F(A)], \quad (5)$$

is a suitable choice for the task. The included terms have the following forms and roles:

#### 2.2.1 Data term

Minimization of the data fidelity term  $D(A; I)$  ensures a small distance between the linear mixture  $AC$  and the sensed data  $I$ , thereby providing a faithful segmentation of the input image. The data term is defined as  $D(A; I) = \frac{1}{2} \|AC - I\|_2^2$ , and its minimization, constrained to  $A \in \mathbb{A}_{N \times m}$ , provides an unmixing segmentation.

#### 2.2.2 Regularization terms

The remaining three terms,  $P(A)$ ,  $T(A)$ , and  $F(A)$ , are all regularization terms, imposing restrictions to the solution according to the assumptions made about the problem: the segmented objects are expected to consist of connected regions separated by smooth, one pixel thick boundaries.

Minimization of the (fuzzy) perimeter term  $P(A)$  favors smooth object boundaries and acts toward noise suppression. This term is defined in Ref. 12 as the overall length of the boundaries of the  $m$  objects of the coverage segmentation  $A$

$$P(A) = \frac{1}{2} \sum_{j=1}^m P(A_j) = \frac{1}{2} \sum_{j=1}^m \sum_{t \in \tau_{2 \times 2}(A_j)} P(t), \quad (6)$$

where  $A_j$  is the  $j$ 'th column of  $A$ . Each boundary is shared between two objects; therefore, the accumulated value is divided by two. The perimeter contribution of  $2 \times 2$  tile  $t$ ,  $P(t)$ , with four assigned coverage values  $\alpha_1, \alpha_2, \alpha_3, \alpha_4$  is computed as follows: for sorted coverage values s.t.  $\alpha_1 \leq \alpha_2 \leq \alpha_3 \leq \alpha_4$

$$P(t) = (\alpha_2 - \alpha_1 + \alpha_4 - \alpha_3) \cdot \frac{b}{2} + (\alpha_3 - \alpha_2) \cdot w, \quad (7)$$

where

$$w = \begin{cases} a, & \text{if two pixels with the smallest coverage share} \\ & \text{a common side in } t \\ b, & \text{if two pixels with the smallest coverage} \\ & \text{are diagonally placed in } t. \end{cases} \quad (8)$$

The lengths  $a$  and  $b$  are optimal horizontal and diagonal unit steps. We use  $a = 0.948$  and  $b = 1.343$ , as suggested in Ref. 31 (and used in Ref. 12).

The thickness term  $T(A)$  is computed as

$$T(A) = \frac{1}{2} \sum_{j=1}^m T(A_j) = \frac{1}{2} \sum_{j=1}^m \sum_{t \in \tau_{2 \times 2}(A_j)} T(t), \quad (9)$$

where  $\tau_{2 \times 2}(A_j)$  is the set of all  $2 \times 2$  configurations (tiles) of  $A_j$ , and thickness of one tile is computed as

$$T(t) = \prod_{i=1}^4 4\alpha_i(1 - \alpha_i). \quad (10)$$

The fuzziness term  $F(A)$  is defined as

$$F(A) = \sum_{i=1}^N \sum_{j=1}^m 4\alpha_{i,j}(1 - \alpha_{i,j}). \quad (11)$$

Both the thickness and the fuzziness term reach minimal values for coverage values equal to 0 or 1 and increase for nonsparse (fuzzy) solutions, i.e., coverage values within (0,1). The thickness term reaches its minimal value zero if at least one pixel in each  $2 \times 2$  configuration is crisp, thus promoting sparse solutions with thin (at most one pixel thick) fuzzy transitions between the objects, whereas the fuzziness term favors majority of (individual) pixels to be classified as pure. Combined they lead to segmentations with thin fuzzy boundaries.

The resulting highly nonconvex energy function [Eq. (5)] is optimized by spectral projected gradient (SPG) optimization<sup>32</sup> within a tailored graduated optimization scheme (GNC).

### 2.3 Image Deblurring and Super-Resolution Reconstruction

Image deblurring and super-resolution reconstruction are suitably performed by energy minimization; our results on the topics are presented in Refs. 33 and 34. We assume an image formation model of the form

$$I = S \cdot B \cdot X + \text{noise}, \quad (12)$$

where the sensed low-resolution image  $I$  (represented as a matrix of size  $N \times 1$ ) is obtained from a noise-free and blur-free (high-resolution) image  $X$  of size  $N_h \times 1$  ( $N_h = dN$ ) convolved with a point spread function (SPF) (represented by the convolution matrix  $B$ ) and downsampled by averaging intensities within  $d = d_r \times d_c$  blocks (represented by the rectangular downsampling matrix  $S$ ). The (estimation of the) unknown image  $X$  is obtained from  $I$  as a minimizer of the energy function

$$E(X; I) = \frac{1}{2} \|SBX - I\|_2^2 + \mu \text{TV}_H(X). \quad (13)$$

The data fidelity term  $\frac{1}{2} \|SBX - I\|_2^2$  plays its typical role of keeping the solution close to the observed data. Here, this includes handling blur and change in resolution. Noise suppression and stability of the solution are achieved by total variation (TV)-based regularization

$$\text{TV}_\Phi(X) = \sum_{i=1}^{N_h} \Phi[|\nabla(X_i)|], \quad (14)$$

where  $\nabla$  stands for image gradient in a pixel,  $|\cdot|$  denotes  $\ell_2$  norm, and  $\Phi$  is a potential function. Classical TV regularization (as in Ref. 35) is obtained for the identity potential function,  $\Phi(t) = t$ . TV is defined as the  $\ell_1$  norm of the  $\ell_2$  norm of the pixel-wise image gradients; it promotes sparsity in the gradient of the image. It has been successfully used in image restoration, providing efficient noise suppression while allowing preservation of edges. Our extensive empirical tests, presented in Ref. 33, show that the performance can be substantially improved if TV is smoothed with the Huber potential<sup>36</sup>

$$\Phi_H(t) = \begin{cases} t^2/(2\omega), & t \leq \omega \\ t - \omega/2, & t > \omega. \end{cases} \quad (15)$$

### 2.4 Contribution

Existing coverage segmentation approaches, such as those proposed in Refs. 17 and 18, have limitations when applied in realistic imaging conditions, which include blur, or when the class end-members are linearly dependent. We here propose an improved coverage segmentation method that overcomes these limitations and offers further qualitative improvements.

The data term is now formulated to enable removal of image blur, providing thin transition boundaries between initially blurry and noisy image components. We also provide the option to estimate coverage values at an increased spatial resolution, higher than that of the sensed image. These improvements increase applicability of the proposed method in realistic imaging situations.

New regularization terms are formulated and evaluated, leading to increased stability of the method and improved performance. The proposed (best performing) combination includes the Huberized TV, instead of the perimeter term. In the continuous case, TV is equivalent to the fuzzy perimeter  $P$ ; however, these two functions exhibit slightly different behaviors in the discrete case. We evaluate both options (including some modifications in an ablation style study) and reach the conclusion that the Huberized TV provides best performance combined with less challenging optimization.

We replace  $T(A)$  and  $F(A)$  by a single new term that restricts the fuzziness penalty (promoting pure pixels) to nonedge regions while not imposing this constraint along the object boundaries. An undesired competition between  $T(A)$ , allowing fuzziness along a thin boundary, and  $F(A)$ , promoting pure pixels everywhere, which was leading to instability of the optimization algorithm is, by that, avoided. The new algorithm achieves better stability of a solution,

necessary to handle the severely ill-posed problem of deblurring-segmentation. A further simplification is that the constraint to a valid coverage segmentation is now included in the regularization term, avoiding the need for the sequence of orthogonal projections in Ref. 18.

Optimization is adjusted to the new objective function, addressing its nonconvexity by a carefully redesigned graduated nonconvexity (GNC) scheme, which automatically tunes the regularization weight of the nonconvex part. This change provides improved unmixing performance in case of linearly dependent endmembers.

### 3 Method

We present details of the proposed super-resolution blur-aware coverage segmentation method, starting with definitions and motivation for a number of terms considered in the corresponding energy function. We continue with detailing the optimization approach we propose for the task; it is based on a customized GNC scheme, suitable for the proposed nonconvex objective function. We conclude the section by discussing the role and recommended values of the regularization parameters.

#### 3.1 Energy Function

We assume that the sensed image  $I$ , of size  $N \times b$ , is of a form

$$I = S \cdot B \cdot A \cdot C + \text{noise}, \tag{16}$$

where  $A_{N_r \times m}$  is a coverage segmentation at  $d = d_r \times d_c$  times increased spatial resolution, and where  $C_{m \times b}$ ,  $S_{N \times N_r}$ , and  $B_{N_h \times N_h}$  are end-member, downsampling, and blurring matrices, respectively. By this we model formation process of an image expressed as a linear mixture of its class representatives (end-members), where degradation consists of blur and decreased spatial resolution and presence of noise. Our aim is to estimate the coverage segmentation matrix  $A$ , for a given  $I$ , assuming that all three types of degradation,  $S$ ,  $B$ , and  $C$ , are known. Methods to estimate these matrices, i.e., to automatically estimate blur and class end-members, remain out of scope of this study. We refer readers interested in end-member and blur matrix estimation methods to Refs. 37–42, and references therein.

The segmentation is performed by minimization of a regularized energy function consisting of:

- Data fidelity term:

$$D_{sr}(A; I) = \frac{1}{2} \|\text{SBAC} - I\|_2^2 = \frac{1}{2} \sum_{k=1}^b \sum_{i=1}^N [(\text{SBAC})_{i,k} - I_{i,k}]^2, \tag{17}$$

which is a generalization of  $D(A; I)$  from Eq. (5).  $D_{sr}(A; I)$  incorporates information about the blur and downsampling resulting from the acquisition process.

- Huberized TV

$$\text{TV}_H(A) = \frac{1}{2} \sum_{j=1}^m \text{TV}_H(A_j) = \frac{1}{2} \sum_{j=1}^m \sum_{i=1}^{N_h} \Phi_H[\|\nabla(\alpha_{i,j})\|], \tag{18}$$

which, by promoting gradient sparsity, provides noise suppression and (spatially) smooth boundaries. As opposed to the perimeter term defined by Eq. (6), Huberized TV is a smooth function; this property contributes to a more efficient optimization process.

- New fuzziness term  $\tilde{F}(A)$

$$\tilde{F}(A) = \sum_{i=1}^{N_h} \sum_{j=1}^m 4\alpha_{i,j}(1 - \alpha_{i,j})(1 - k_{i,j}), \tag{19}$$

where

$$k_{i,j} = \max_{k \in \mathcal{N}(i)} \alpha_{k,j} - \min_{k \in \mathcal{N}(i)} \alpha_{k,j} \tag{20}$$

and  $\mathcal{N}(i)$  is the  $3 \times 3$  neighborhood of pixel  $i$ . Here,  $\tilde{F}(A)$  penalizes fuzzy regions (i.e., promotes pure regions), except where the contrast (measured by  $k_{i,j}$ ) is high, indicating object boundaries. This term replaces the previously used combination of  $F(A)$  and  $T(A)$  [as defined in Eq. (9)] and leads to a simpler and more numerically stable minimization of the energy function.

- The term  $R(A)$

$$R(A) = \frac{1}{2} \sum_{i=1}^{N_h} \left[ \left( \sum_{j=1}^m \alpha_{i,j} \right) - 1 \right]^2, \tag{21}$$

which promotes proper coverage segmentations, i.e., that each row of the matrix  $A$  sums up to 1. This constraint was in Ref. 18 enforced by projections within the optimization. The change introduced by including the term  $R(A)$  reduces the computational time.

Note that direct minimization of  $\|A\|_1$ , which often is an effective way of reaching good sparse solutions for underdetermined systems,<sup>43</sup> has no effect in this case, due to the requirement that components (rows of  $A$ ) sum up to one. We are required to include a nonconvex term  $\tilde{F}$  to drive the optimization to a sparse enough solution.

We have, in addition, evaluated

- Tilewise Huberized perimeter term

$$P_H(A) = \frac{1}{2} \sum_{j=1}^m P_H(A_j) = \frac{1}{2} \sum_{j=1}^m \sum_{t \in \tau_{2 \times 2}(A_j)} \Phi_H[P(t)], \tag{22}$$

where  $A_j$  is the  $j$ 'th column of  $A$  and  $P(t)$  is the perimeter contribution of a  $2 \times 2$  tile, computed as in Eq. (7). This is a smooth version of the perimeter term  $P$ , acting toward noise suppression and smooth object boundaries.

We here propose super-resolution blur-aware coverage segmentation as the minimizer of the energy function

$$E(A; I) = D_{sr}(A; I) + \mu TV_H(A) + \nu \tilde{F}(A) + \eta R(A). \quad (23)$$

In the continuation, we refer to this approach as SRCS-HTV $\tilde{F}$ .

To evaluate individual relevance of the terms included in the proposed energy function SRCS-HTV $\tilde{F}$ , we have conducted a systematic evaluation of variations of these terms, included in the following sequence of energy functions:

- i. CS-PTF: coverage segmentation proposed in Ref. 18

$$E(A; I) = D(A; I) + \mu P(A) + \nu T(A) + \nu F(A).$$

- ii. CS-P $\tilde{F}$ : an improved version of CS-PTF using the novel fuzziness term Eq. (19)

$$E(A; I) = D(A; I) + \mu P(A) + \nu \tilde{F}(A).$$

- iii. SRCS-P $\tilde{F}$ : super-resolution blur-aware coverage segmentation with perimeter term

$$E(A; I) = D_{sr}(A, I) + \mu P(A) + \nu \tilde{F}(A) + \eta R(A).$$

- iv. SRCS-HP $\tilde{F}$ : super-resolution blur-aware coverage segmentation with Huberized perimeter term

$$E(A; I) = D_{sr}(A, I) + \mu P_H(A) + \nu \tilde{F}(A) + \eta R(A).$$

Following the basic idea of an ablation study, we evaluate the relevance of  $TV_H(A)$  in SRCS-HTV $\tilde{F}$  by replacing it by  $P_H(A)$  (in SRCS-HP $\tilde{F}$ ) and with  $P(A)$  (in SRCS-P $\tilde{F}$ ). We evaluate the influence of the term  $R(A)$  by excluding it (and utilizing a projection instead), in CS-P $\tilde{F}$ . In the same model, we exclude the super-resolution reconstruction option. Finally, we also consider the “classic” version, CS-PTF, Ref. 18, where we replace the term  $\tilde{F}(A)$  with the combination  $T(A) + F(A)$ . Comparison of the performances of all the five methods is presented in Sec. 4.

### 3.2 Optimization

We optimize  $E(A; I)$  by utilizing SPG, a method proposed in Ref. 32 for solving a constrained optimization problem  $\min_{x \in \Omega} f(x)$ , where  $\Omega$  is a closed convex set in  $\mathbb{R}^n$  and  $f$  has continuous partial derivatives on an open set that contains  $\Omega$ . Weak requirements and efficiency in solving large-scale problems make this tool attractive for our purpose. Details related to the SPG optimization algorithm, parameters used, and the derived gradient vectors of the observed objective functions are given in Sec. 6.

The energy function  $E(A; I)$  in Eq. (23) as well as all its versions (i)–(iv) are nonconvex, and their minimization is far from trivial. We address this challenge by gradually increasing complexity of the problem and utilizing the solutions of numerically easier subproblems as initializations in optimization of the more difficult ones. The functions (i) and (ii) are optimized as in Ref. 18. For optimization of Eq. (23), as well as the functions defined by (iii)–(iv), we propose the following GNC approach: we initiate the process by minimizing only the terms  $D_{sr}(A; I)$  and  $R(A)$ .

#### Algorithm 1 Super-resolution coverage segmentation.

---

Parameters  $\mu, \nu_0, \eta_0, \rho \geq 0$ . Initial segmentation  $A_0 = \left[ \frac{1}{m} \right]_{N_y \times m}$ .

$A \leftarrow \arg \min [D_{sr}(A; I) + \eta_0 R(A)]$  by SPG initialized by  $A_0$

**For**  $k = 1, 2, \dots$

$\eta = \eta_0 \cdot k$ .

$A \leftarrow \arg \min [D_{sr}(A; I) + \mu TV_H(A) + \eta R(A)]$  by SPG

**Repeat**

$A \leftarrow \arg \min^* [D_{sr}(A; I) + \mu TV_H(A) + \nu \tilde{F}(A) + \eta R(A)]$  by SPG

$\nu \leftarrow \nu(1 + \rho)$

**Until**  $\tilde{F}(A) \leq 0.1$

$\nu \leftarrow \nu(1 - 2\rho)$

**Until**  $\|A_{k+1} - A_k\|_\infty \leq 0.01$

---

Then, we include the smoothing term  $[P(A), P_H(A), \text{ or } TV_H(A)]$ . Finally, we include the nonconvex  $\tilde{F}(A)$  and alternately switch it off and on while gradually increasing its corresponding weight  $\nu$  until we reach a crisp enough solution. The complete optimization procedure for (super-resolution blur-aware) coverage segmentation is presented in Algorithm 1. The SPG optimization in the innermost loop ( $\arg \min^*$ ) is terminated after 10 iterations, whereas other SPG optimizations are run until convergence. The process terminates when the max-norm between two consecutive estimates stays below the tolerance 0.01. With the appropriate smoothing terms, Algorithm 1 is used for (iii) and (iv) as well.

### 3.3 Regularization Parameters

Our proposed SRCS-HTV $\tilde{F}$  method includes three regularization parameters,  $\mu, \nu$ , and  $\eta$ . The parameter values are typically data dependent; here we give some general recommendations and settings. Without loss of generality we assume that the sensed image has intensity values in the range  $[0, 1]$ .

#### 3.3.1 Smoothing term weights, $\mu$ and $(\mu, \omega)$

We have experimentally confirmed stability of the method w.r.t. these parameters. Values that lead to good performances are  $\mu = 0.05$  [when  $P(A)$  is used], and  $(\mu, \omega) = (0.05, 0.001)$  (for the functions including Huberized TV and perimeter terms,  $TV_H(A)$  and  $P_H(A)$ ). In general, the level of smoothing should match the level of image noise; more image noise may require more smoothing, whereas too much smoothing leads to a risk of losing image detail.

#### 3.3.2 Fuzziness term weight $\nu$

This parameter is initialized with a small value; we recommend  $\nu_0 = 10^{-4}$ . In the proposed GNC optimization scheme,  $\nu$  is automatically tuned to a suitable level, thus making the

algorithm resilient to variations in the input data. The tuning of  $\nu$  is regulated by the parameter  $\rho$ . Too fast increase of the weight  $\nu$  may lead to an undesirably crisp solution, whereas too slow increase leads to unnecessary slow progress toward convergence. We found that a fixed value  $\rho = 0.1$  provided good performance in all our tests.

### 3.3.3 Constraint term weight $\eta$

This parameter is initially set to  $\eta_0 = 1$  and is increased linearly during the procedure to gradually impose the sum-to-one requirement of a valid solution. We found this scheme to provide a stable convergence toward a good solution in all our tests.

## 4 Evaluation and Results

We have performed three types of evaluation of the proposed method: (i) a statistical study on synthetic images provides insight in the robustness of the method and its performance with increasing levels of blur, noise, and subsampling; robustness w.r.t. the utilized end-member and blur estimates is evaluated as well. The synthetic images are generated as to enable evaluation of the method's capacity to deal with linearly dependent end-members. (ii) Qualitative and quantitative evaluation of the performance in presence of blur, on a synthetic image acquired in real imaging conditions. (iii) Qualitative and quantitative evaluation of super-resolution coverage segmentation of a hyperspectral image illustrates performance of the method on real, multichannel images. Evaluation (i) includes all five versions of the energy functions described in Sec. 3.1, together with the corresponding optimization schemes. The remaining tests (ii) and (iii) focus on the best performing approach, SRCS-HTV $\bar{F}$ .

### 4.1 Synthetic Tests

We evaluate the considered methods on a set of synthetic test images that are distorted by known levels of noise and blur. An example of a synthetic two-channel (red-green) test image used in this experiment is shown in Fig. 3(a). Correct coverage values (i.e., ground truth) are in this case known and are used in evaluation. Each test image is of size  $100 \times 100 \times 2$  and contains three geometric objects with both smooth and nonsmooth boundaries. The color of one of the end-members (square) is a linear mixture of the colors of the other two end-members (star and disk; 50% each). This makes it challenging to unmix the pixels on the (blurred) boundary between the star and the disk (appearing

in a similar color as the square) while at the same time interpreting the square as a separate pure component.

In order to capture well sensitivity of the segmentation method w.r.t. discretization, a number of variations of the test image are created and digitized where the objects are positioned at different locations and in different rotations in the digital grid. To evaluate robustness of the method in presence of noise and blur, evaluation is conducted on a set of images created by degrading the test images by increasing levels of blur and noise. We convolve each test image with Gaussian PSF with standard deviations  $\sigma_p \in \{0.1, 2, 3, 4\}$ , and add Gaussian noise with standard deviation  $\sigma_n \in \{0, 10^{-3}, 3.2 \cdot 10^{-3}, 10^{-2}, 3.2 \cdot 10^{-2}\}$ , thus creating a test set of  $5 \times 5 \times 6 = 150$  images. True coverage values (ground truth) are estimated by  $8 \times$  supersampling (eight-sampled coverage digitization, Ref. 44).

Qualitative assessment of the observed methods in presence of blur and noise can be made based on the results shown in Fig. 4. We present coverage segmentations of the image shown in Fig. 3(b), which is a degraded version of Fig. 3(a), obtained by applying to it Gaussian blur with  $\sigma_p = 1$  and Gaussian noise with  $\sigma_n = 10^{-2}$ . Each row in Fig. 4 is produced by one of the observed coverage segmentation methods, (i)–(iv), and Eq. (23), here denoted method (v). Each segmentation consists of four coverage maps, corresponding to three objects and the background. All the coverage maps are generated at the spatial resolution of the initial image [i.e., with no super-resolution reconstruction;  $d = 1 \times 1$  for methods (iii)–(v)].

We observe that method (i) fails to properly segment the square (considering its color partly a mixture of red and green and only partly as a “pure” color). While (ii) is more successful with unmixing, it does not handle well the blurred input and produces objects with highly nonsmooth boundaries. Method (iii) deals with blur but still fails with unmixing. The problem comes from the difficulty to optimize the nonsmooth perimeter term  $P$ . Huberized perimeter term,  $P_H$ , in (iv) leads to improved, but still not perfect, segmentation. Finally, (v) utilizes  $TV_H$  and produces both good separation of the classes and thin object boundaries with preserved sharp corners.

We quantify performances of the methods by computing the mean absolute error (MAE) of the estimated coverage values

$$\text{MAE}(A, \hat{A}) = \frac{1}{N_{hm}} \sum_{i=1}^{N_h} \sum_{j=1}^m |\alpha_{i,j} - \hat{\alpha}_{i,j}|, \quad (24)$$

where  $\alpha_{i,j}$  and  $\hat{\alpha}_{i,j}$  are true, resp. estimated, coverage values of the  $i$ 'th pixel by class  $S_j$ .

Figures 5(a) and 5(b) show MAE, averaged over six image displacements (as described above), for the observed coverage segmentation methods, for increasing levels of blur (at a fixed noise level  $\sigma_n = 10^{-2}$ ), and increasing levels of noise (blur level  $\sigma_p = 2$ ), respectively. For comparison, MAE of a theoretically optimal crisp segmentation of the corresponding noise-free and blur-free test image (as given by ground truth) is also plotted. It is evident that SRCS-HTV $\bar{F}$  exhibits best performance, achieving the lowest average MAE, and very low variation, consistently outperforming the ideal crisp segmentation. Consistent behavior of all the methods is observed for other combinations of blur and noise.

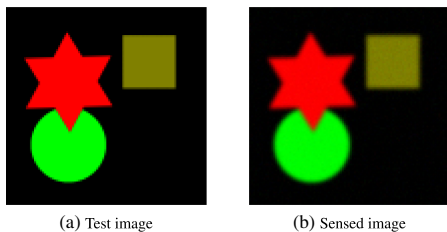
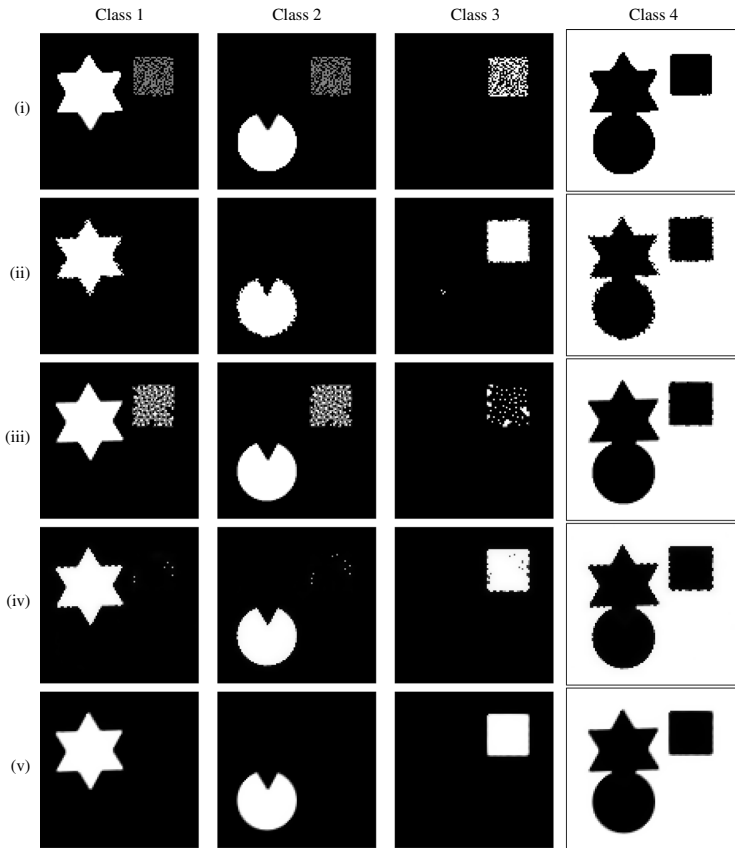


Fig. 3 (a) Test image of size  $100 \times 100 \times 2$ ; (b) degraded test image (a) by Gaussian blur with  $\sigma_p = 1$  and Gaussian noise with  $\sigma_n = 10^{-2}$ .



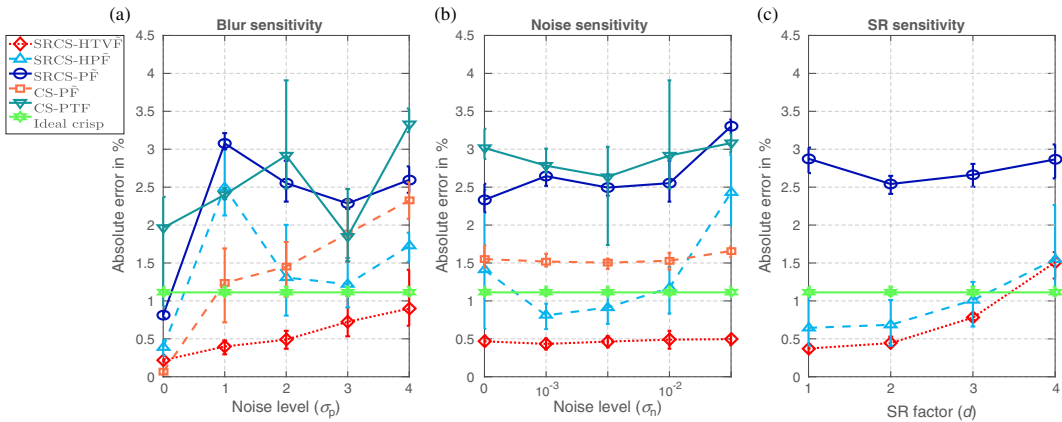
**Fig. 4** Segmentation of the image in Fig. 3(b). Coverage maps obtained with: (i) CS-PTF, (ii) CS-PF, (iii) SRCS-PF, (iv) SRCS-HPF, and (v) SRCS-HTVF.

Methods (iii)–(v) enable segmentation at increased spatial resolution. Their performance on the test set of images created from the six displacements degraded by blur with  $\sigma_p = 1$  and noise with  $\sigma_n = 10^{-3}$ , for increasing scale factor  $d$ , is shown in Fig. 5(c). We start from images of spatial sizes  $100 \times 100$ ,  $50 \times 50$ ,  $33 \times 33$ , and  $25 \times 25$  and use scale factors  $1 \times 1$ ,  $2 \times 2$ ,  $3 \times 3$ , and  $4 \times 4$ , respectively, to create segmentation maps at approximately the original size of  $100 \times 100$ . Average MAE over six image displacements confirms superiority of SRCS-HTVF. The green horizontal line in the graph shows the ideal crisp segmentation at full resolution. We observe that the proposed method provides a better segmentation (lower MAE) than what is theoretically possible using a crisp object representation even when it starts from a three times smaller, blurry, and noisy image. Figures 6(b)–6(e) show the coverage maps resulting from SRCS-HTVF applied to the blurry and noisy image (a), segmented at three times higher resolution.

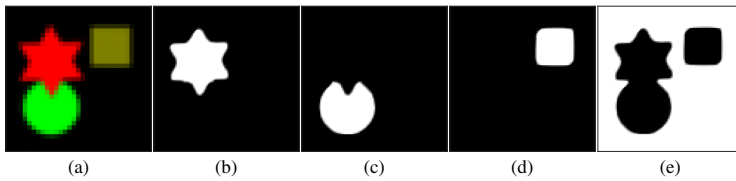
The performed evaluation on synthetic data clearly shows superiority of SRCS-HTVF (optimized according to Algorithm 1) for the task of coverage segmentation.

The proposed method requires blur and end-member matrices as input; here we provide this information as *a priori* knowledge. Blind deblurring methods (e.g., Refs. 41 and 45) can be used to estimate the PSF (i.e., blur matrix). We estimate the end-member matrix by random sampling of the true classes (Refs. 37–40 related to end-member estimation are already suggested in Sec. 3). We evaluate robustness of SRCS-HTVF w.r.t. different levels of error in the estimated blur and end-member matrices and present the results in Fig. 7.

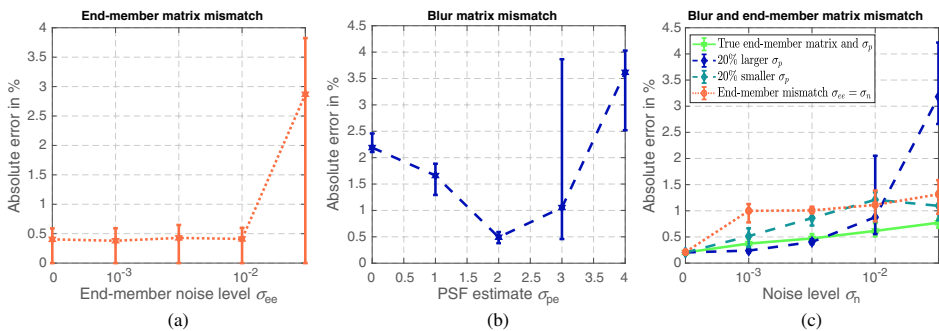
We first observe performance of the method for a range of mismatched end-member matrices, obtained by adding increasing level of noise to a true end-member matrix. Results for medium image noise and blur ( $\sigma_n = 10^{-2}$ ,  $\sigma_p = 2$ ) and with end-member noise  $\sigma_{ee} \in [0, 0.03]$  are shown in Fig. 7(a). As can be observed, the performance is very stable up to a point at which, in some cases, a rather different (incorrect) segmentation provides the lowest energy solution. Noticing that this decrease in performance occurs at a level where the end-member noise is larger than the image noise of an individual pixel (i.e., estimating the



**Fig. 5** Performances of five considered methods for increasing levels of (a) blur (at noise level  $\sigma_n = 10^{-2}$ ), (b) noise (at blur level  $\sigma_p = 2$ ), and (c) super-resolution factors, measured by average MAE. Lines show averages for six displacements with the same degradation (but different discretizations) and bars indicate max and min errors. MAE of an ideal noise-free and blur-free crisp segmentation is included as a reference (green horizontal line). The proposed SRCS-HTVF exhibits superior performance, consistently outperforming even the ideal crisp segmentation. (Lower is better).



**Fig. 6** Performance of SRCS-HTVF in coverage segmentation at three times increased resolution. The test image of size  $99 \times 99$ , Fig. 3(a), is degraded by Gaussian blur ( $\sigma_p = 1$ ), downsampled by SR factor  $d = 3 \times 3$  (by averaging pixels in  $3 \times 3$  blocks), and corrupted by Gaussian noise ( $\sigma_n = 10^{-3}$ ). The resulting degraded image of size  $33 \times 33$  is shown in (a). Images containing the resulting segmented components (b)–(e) are all of size  $99 \times 99$ .



**Fig. 7** (a) Robustness of SRCS-HTVF w.r.t. to end-member matrix mismatch for different end-member noise levels ( $\sigma_{ee}$ ) and  $\sigma_p = 2$  and  $\sigma_n = 10^{-2}$ . (b) Robustness of SRCS-HTVF w.r.t. to blur matrix mismatch: performance for different used PSF estimates ( $\sigma_{pe}$ ) for real PSF  $\sigma_p = 2$  and  $\sigma_n = 0.01$ . (c) MAEs obtained with SRCS-HTVF with true and wrong blur and end-member matrices for different image noise levels and  $\sigma_p = 2$ .

end-members from single pixel samples would provide a better result), we conclude that the method is robust w.r.t. end-member noise.

Second, we observe performance of the proposed method for a range of different Gaussian shaped PSF estimates with  $\sigma_{pe} \in [0, 4]$ . Results for the noise level  $\sigma_n = 10^{-2}$  and the case when the true blur is caused by a Gaussian PSF with  $\sigma_p = 2$  are shown in Fig. 7(b). Here, we observe a more proportional degradation in performance as the estimated PSF deviates from the correct one. In Fig. 7(c), we present performance of SRCS-HTV $\bar{F}$  for a range of image noise levels, where either the end-member matrix has a noise magnitude equal to the image noise (orange curve) or where the used PSF either is 20% too small or 20% too large (cyan and blue curves).

Summarizing 450 = 3 × 150 tests performed with end-member noise or PSF mismatch for  $\sigma_p \in \{0, 1, 2, 3, 4\}$  and image noise with  $\sigma_n \in \{0, 10^{-3}, 3 \cdot 10^{-3}, 10^{-2}, 3 \cdot 10^{-2}\}$ , we conclude that: (i) small to medium end-member mismatch (smaller than the per pixel image noise) in general has a small impact on the segmentation, whereas larger mismatch may lead to the algorithm finding a rather different segmentation result; (ii) errors in the PSF estimate affect the performance more or less proportionally. When the solution is well defined, as in the no-noise case, the left edge of Fig. 7(c), the method is robust w.r.t. PSF errors.

#### 4.2 Illustrative Example on a Naturally Degraded Image

We further illustrate performance of the SRCS-HTV $\bar{F}$  method on an example image acquired under real imaging conditions.

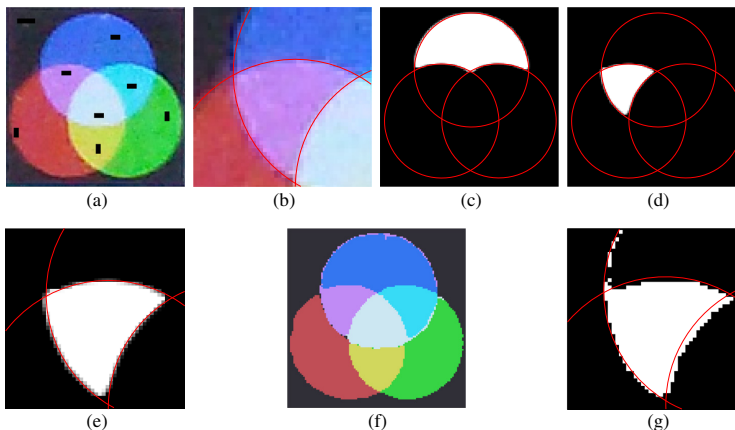
In this experiment, we display on a computer screen three disks of equal radii, centered at the vertices of an equilateral triangle with a side length equal to the radius of the disks,

and we then acquire an image of the object on the screen by the 8 megapixel CMOS sensor of a Samsung Galaxy S3 cell phone camera; the result is shown in Fig. 8(a). Black rectangular regions in Fig. 8(a) indicate sets of pixels used to estimate the  $8 \times 3$  end-member matrix  $C$ . The zoomed-in portion, Fig. 8(b), shows details of the noisy and blurred image. Red thin lines in Figs. 8(b)–8(e), and 8(g) are not parts of either sensed or segmented images. They indicate the true object boundaries and are superimposed for better assessment.

For comparison, we also segment Fig. 8(a) by Bayesian classification utilizing linear discriminant analysis (LDA), with the same indicated sets of pixels (black regions) used for training.

Both methods perform well on the majority of pixels: MAE of SRCS-HTV $\bar{F}$  is 0.52%, and for LDA it is 0.83%. This is expected since most of the pixels are pure and classified as such by both methods. However, looking at the boundaries of the resulting segments we see that the SRCS-HTV $\bar{F}$  segmentation is by far superior. In Figs. 8(c)–8(d), we show the coverage maps for 2 (out of 8, including background) components. Color unmixing is successful and accurate for all the seven appearing colors/classes. The zoomed-in portion, Fig. 8(e), shows the details of the segmentation; we observe that, in spite of the blur present in the original image, the resulting segmentation is mostly crisp, with a thin transition of partly covered pixels appropriately positioned along the true region boundaries. Noise is efficiently removed.

The segmentation resulting from LDA, Fig. 8(f), is crisp. As clearly visible in Fig. 8(g) [showing the same zoomed-in region as Fig. 8(e)], noise and mixed pixels at object boundaries have a rather strong negative impact on the result, causing misclassification of a number of pixels, as well as rather unappealingly jagged segmentation boundaries.



**Fig. 8** Coverage segmentation of a blurry and noisy color image acquired in realistic imaging conditions. (a) The sensed image (black rectangular regions indicate pixels used for the end-member matrix estimation). (b) Zoomed-in region from (a). (c)–(d) Two segmented components obtained by SRCS-HTV $\bar{F}$ . (e) Zoomed-in segmentation of the central region from (b). (f)–(g) Segmentation and zoomed-in central region from (b) obtained by LDA. Thin red lines are superimposed to indicate the correct continuous region boundaries.



### 4.3 Segmentation of a Hyperspectral Image

Remote sensing is one of relevant application fields for the proposed method. We evaluate SRCS-HTV $\bar{F}$  on a publicly available<sup>46</sup> 220 band hyperspectral data set from an airborne visible/infrared imaging spectrometer (AVIRIS) image taken on NW Indianas Indian Pine test site in June 1992.

The same image is used for evaluation in Ref. 4, where linear unmixing of pixels is performed at an increased spatial resolution using simulated annealing, and in Ref. 18, where the performance of the here observed CS-PTF method is tested.

Following the procedure of Refs. 4 and 18, we perform tests on the  $[31 - 116] \times [27 - 94]$  cut-out of the scene. Four labeled classes: “Corn-notill” (1008 pixels), “Grass/Trees” (732 pixels), “Soybeans-notill” (727 pixels), and “Soybeans-min” (1926 pixels) are provided as the ground truth. 1455 pixels of the scene are not labeled and are not considered in the evaluation. Since the available ground truth classification of the image does not provide abundance information, we decrease the spatial resolution of the original image by averaging blocks of  $3 \times 3$  pixels providing a 220 band image of  $28 \times 22$  pixels; we use this low-resolution image as test image. Information from the three times higher resolved ground truth data is then used to estimate coverage values for the created low resolved image. One band of the low-resolution image and the corresponding ground truth are shown in Figs. 9(a) and 9(b).

The 220 bands are decorrelated by a whitening transformation. For each class, 20 pure pixels are randomly selected as training data and are used to compute the matrix  $C$ . The  $28 \times 22 \times 220$  image is then segmented at (i) the same (low) spatial resolution ( $d = 1 \times 1$ ) and (ii) at three times increased resolution ( $d = 3 \times 3$ ). The experiment is repeated 20 times, with a new random selection of the training sets, and the mean and standard deviation of the performance measures are computed. The here proposed method is approximately equally fast as the method proposed in Ref. 18 when segmentation is performed at the same spatial resolution ( $d = 1 \times 1$ ). Average computational time for  $3 \times$  increased resolution segmentation of the 220 band image in MATLAB on a 2.6 GHz CPU is 224 s. Parameter values are as suggested in Sec. 3.3.

Evaluation of (i) is performed by direct comparison of the estimated coverage values with the created ground truth. The MAE of SRCS-HTV $\bar{F}$  and CS-PTF<sup>18</sup> are given in Table 1. We remark that this result has to be interpreted with care

**Table 1** Performance on the AVIRIS data set (in %). Mean value of the measure is followed by a standard deviation, in parentheses.

	MAE	CA
Villa et al. <sup>4</sup>	–	90.65
CS-PTF <sup>18</sup>	4.53 (1.57)	[92.59 (1.99), 94.74 (1.32)]
SRCS-HTV $\bar{F}$	6.10 (0.65)	93.40 (2.31)

since the  $3 \times 3$  averaging, providing the ground truth has a quantization error of up to  $\frac{1}{18} \approx 5.6\%$ .

To evaluate (ii), we make a crisp segmentation by selecting, for each pixel, the component with largest coverage and then compare with the full resolution ground truth. The classification accuracy (CA) (percentage of correctly classified pixels) of the crisp segmentation at the three times increased resolution is shown in Table 1. The resulting coverage segmentation at increased resolution from a single experiment is shown in Fig. 9(c).

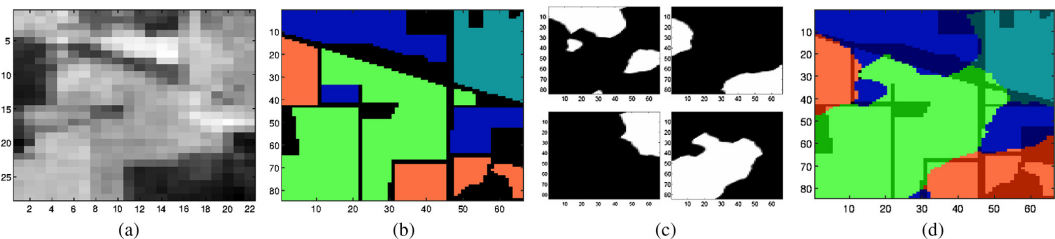
Since the CS-PTF method does not provide coverage segmentation at increased spatial resolution, we present its performance in Table 1 as an interval of possible performances, where the lower bound corresponds to a nearest neighbor upsampling and the upper bound is derived from an ideal optimal distribution of the coverage (OOA in Ref. 18).

We observe that the proposed SRCS-HTV $\bar{F}$  outperforms both Ref. 4 and the naive approach of Ref. 18 and is not far from the accuracy of the “oracle” segmentation, which gives the upper accuracy bound.

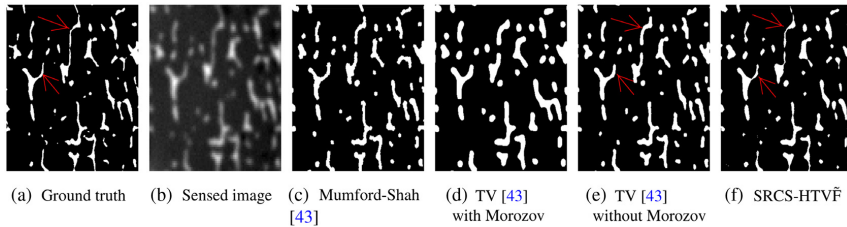
### 4.4 Segmentation of a Micro-CT Image

In this section, we illustrate performances of SRCS-HTV $\bar{F}$  applied on a micro-CT image of trabecular bone. Measurement of the surface of microstructures of a trabecular bone is one of the determinants of bone strength and is important for the diagnosis of osteoporosis.<sup>47</sup> Precise segmentation, a prerequisite for accurate surface measurements, is a challenging task in blurry and noisy low-resolution CT images.<sup>48–51</sup> We therefore find it suitable to demonstrate applicability of our method for bone segmentation.

We observe the same low resolution blurry and noisy image of trabecular bone [Fig. 10(b)] as is used in Ref. 50.



**Fig. 9** Coverage segmentation of a cut-out hyperspectral AVIRIS image. (a) Band 30 (out of 220) of the image obtained by downsampling the original image by averaging  $3 \times 3$  blocks. (b) Ground truth classification of the original image (unclassified pixels shown in black). (c) Coverage segmentation (four classes) of (a) at three times increased resolution. (d) Crisp segmentation derived from (c) [to compare with (b)].



**Fig. 10** Segmentation of microstructures of a trabecular bone at four times increased resolution. (a) Ground truth. (b) Sensed image. (c)–(e) Segmentations obtained by variants of the method proposed in Ref. 50. (f) Crisp segmentation derived from coverage segmentation obtained by SRCS-HTVĚ. Red arrows indicate some example locations where the proposed method qualitatively provides a segmentation that is more similar to the ground truth than the compared methods do. Quantitative performance comparison is given in Table 2.

**Table 2** Performance on micro-CT image segmentation.

Method	Parameter choice	DICE	Overall bone surface Ref = 3.94 (mm <sup>2</sup> )
Mumford-Shah <sup>50</sup>	With Morozov principle	0.80	5.16
Mumford-Shah <sup>50</sup>	Without Morozov principle	0.80	5.16
TV <sup>50</sup>	With Morozov principle	0.70	6.36
TV <sup>50</sup>	Without Morozov principle	0.82	5.14
SRCS-HTVĚ	Recommended parameters	0.84	4.32

The image is obtained from a high-resolution noise-free and blur-free ground truth image [Fig. 10(a)] acquired by a parallel-beam synchrotron micro-CT,<sup>52</sup> downsampled by a factor  $4 \times 4$  and degraded by Gaussian blur with  $\sigma_p = 4.85$  and Gaussian noise with  $\sigma_n = 1$ , as described in Ref. 50. Segmentations obtained by different versions of a joint super-resolution segmentation method suggested in Ref. 50 are shown in Figs. 10(c)–10(e): Mumford–Shah model optimized by ADMM and its TV version, where regularization parameters are chosen with or without utilization of Morozov discrepancy principle<sup>53</sup> are used. The segmentation obtained by SRCS-HTVĚ with recommended parameters  $[(\mu, \omega) = (0.05, 0.001), \nu_0 = 10^{-4}, \rho = 0.1, \eta_0 = 1]$  at four times increased resolution is given in Fig. 10(f). The presented crisp segmentation is obtained from the coverage segmentation by thresholding it at 0.5.

As can be seen from Fig. 10, the proposed method segments the microstructures of interest more precisely in comparison with all versions of the method proposed in Ref. 50. This observation is confirmed by the quantitative evaluation presented in Table 2, where our method outperforms all four variants of the method from Ref. 50 as measured by the DICE similarity coefficient and the bone surface area.

## 5 Conclusion

Coverage segmentation preserves valuable image information by estimating partial overlap of pixels by the image

components. This information enables subpixel precision of extracted shape features. If the imaged real objects are crisp, then mixed, partly covered pixels appear only on object boundaries. However, as a consequence of noise and blur in images, larger regions in the image may appear as partly covered, negatively affecting the segmentation result.

We present a coverage segmentation method that, by promoting sparse solutions, provides highly accurate segmentation results also in presence of blur and noise. The method allows to generate coverage maps at the same, or at an increased spatial resolution. The proposed approach successfully performs linear unmixing of classes even in case of linearly dependent pure class representatives (end-members).

The segmentation is performed by minimization of a carefully designed energy function that promotes desired segmentation properties. Optimization of the proposed non-convex objective function is challenging; we present a self-tuning graduated optimization scheme to efficiently minimize it and reliably arrive at a good segmentation. A few regularization parameters allow tuning to fit the applications at hand, more image noise typically requires stronger spatial regularization. We observe that the method is rather stable w.r.t. the parameter setting; we give recommended values that have provided good performance for a range of studied problems.

Evaluation on synthetic and real multichannel images shows that the proposed method outperforms the competitors included in the study. Robustness w.r.t. (inaccurate) end-members and blur kernel estimates, as well as w.r.t. parameter settings indicate stability of the method. Providing a good balance between flexibility and robustness, the proposed method confirms to be generally applicable in a range of real image analysis problems.

## 6 Appendix

We provide technical details for the

- I. SPG optimization algorithm;
- II. Gradient vectors of the considered versions of the objective function  $E(A; I)$  used in optimization.

### 6.1 Spectral Projected Gradient Algorithm

Pseudocode of the SPG optimization method used within the GNC scheme is given in Algorithm 2.

We perform optimization using SPG with settings recommended in Ref. 27:  $\theta_{\min} = 10^{-3}$ ,  $\theta_{\max} = 10^3$ ,  $\gamma = 10^{-4}$ ,  $\sigma_1 = 0.1$ , and  $\sigma_2 = 0.9$ . Algorithm terminates when the max-norm between two consecutive images is less than  $tol = 10^{-4}$  or when the number of iterations reaches 2000. We define the projection  $P_{\Omega}$  of a vector  $x \in \mathbb{R}^n$  to the feasible set  $\Omega = [0,1]^n$  as:  $[P_{\Omega}(x)]_i = \min\{1, \max\{0, x_i\}\}$ , for all  $i = 1, 2, \dots, n$ .

### 6.2 Gradients of the Terms Included in $E(A; I)$

The partial derivative of  $E(A; I)$  for SRCs-HTV $\bar{F}$  method w.r.t. an individual coverage value  $\alpha_{i,j}$  is

$$\begin{aligned} \frac{\partial[E(A; I)]}{\partial\alpha_{i,j}} &= \frac{\partial[D_{sr}(A; I)]}{\partial\alpha_{i,j}} + \mu \frac{\partial[TV_H(A)]}{\partial\alpha_{i,j}} + \nu \frac{\partial[\tilde{F}(A)]}{\partial\alpha_{i,j}} \\ &+ \eta \frac{\partial[R(A)]}{\partial\alpha_{i,j}}. \end{aligned} \quad (25)$$

**Algorithm 2** Spectral projected gradient.

Choose values for parameters:  $\theta_{\min}, \theta_{\max}, \gamma, \sigma_1, \sigma_2, tol$  s.t.

$0 < \theta_{\min} < \theta_{\max}, \gamma \in (0,1), 0 < \sigma_1 < \sigma_2 < 1, tol > 0$ .

Choose initial guess  $x_0 \in \Omega$  and  $\theta_0 = 1$ .

Compute  $x_{k+1}$  and  $\theta_{k+1}$  as follows:

$$\begin{aligned} d_k &= P_{\Omega}[x_k - \theta_k \nabla f(x_k)] - x_k; \quad x_{k+1} = x_k + d_k \\ \delta &= \nabla f(x_k)^T d_k; \quad \lambda_k = 1 \end{aligned}$$

**while**  $f(x_{k+1}) > f(x_k) + \gamma \lambda_k \delta$

$$\lambda_{\text{temp}} = -\frac{1}{2} \lambda_k^2 \delta / [f(x_{k+1}) - f(x_k) - \lambda_k \delta]$$

**if**  $(\lambda_{\text{temp}} \geq \sigma_1 \wedge \lambda_{\text{temp}} \leq \sigma_2 \lambda_k)$

**then**  $\lambda_k = \lambda_{\text{temp}}$

**else**  $\lambda_k = \lambda_k / 2$

$$x_{k+1} = x_k + \lambda_k d_k$$

**end while**

$$s_k = x_{k+1} - x_k; y_k = \nabla f(x_{k+1}) - \nabla f(x_k); \beta_k = s_k^T y_k$$

**if**  $\beta_k \leq 0$

**then**  $\theta_{k+1} = \theta_{\max}$

**else**  $\theta_{k+1} = \min\left\{\theta_{\max}, \max\left\{\theta_{\min}, \frac{s_k^T s_k}{\beta_k}\right\}\right\}$

**Repeat until:**  $\|x_{k+1} - x_k\|_{\infty} \leq tol$ .

a. The partial derivative of the data fidelity term  $D_{sr}(A, I)$  w.r.t.  $\alpha_{i,j}$  can be expressed as:

$$\frac{\partial[D_{sr}(A; I)]}{\partial\alpha_{i,j}} = [B^T \cdot S^T (S \cdot B \cdot A \cdot C - I) C^T]_{i,j}. \quad (26)$$

b. The partial derivative of the smoothing term  $TV_H(A)$  is

$$\begin{aligned} \frac{\partial[TV_H(A)]}{\partial\alpha_{i,j}} &= \frac{1}{2} \left\{ \Phi'[\|\nabla(\alpha_{i,j})\|] \frac{2\alpha_{i,j} - \alpha_{i,j,r} - \alpha_{i,j,b}}{|\nabla(\alpha_{i,j})|} \right. \\ &+ \Phi'[\|\nabla(u\alpha_{i,j,l})\|] \frac{\alpha_{i,j} - \alpha_{i,j,l}}{|\nabla(\alpha_{i,j,l})|} \\ &\left. + \Phi'[\|\nabla(\alpha_{i,j,a})\|] \frac{\alpha_{i,j} - \alpha_{i,j,a}}{|\nabla(\alpha_{i,j,a})|} \right\}, \end{aligned}$$

where  $\alpha_{i,j,a}$  and  $\alpha_{i,j,l}$  denote the edge neighbors above and to the left of the coverage value  $\alpha_{i,j}$ , respectively. We compute the discrete image gradient at point  $\alpha_{i,j}$ ,  $\nabla(\alpha_{i,j})$ , as  $\nabla(\alpha_{i,j}) = (\alpha_{i,j,r} - \alpha_{i,j}, \alpha_{i,j,b} - \alpha_{i,j})$ , where  $r$  and  $b$  denote indexes of the edge neighbors to the right and below the coverage value  $\alpha_{i,j}$ , respectively.

c. To differentiate the fuzziness term  $\tilde{F}(A_j)$  w.r.t.  $\alpha_{i,j}$ , we observe that  $\alpha_{i,j}$  appears in nine neighboring tiles (which form a  $5 \times 5$  neighborhood with  $\alpha_{i,j}$  as the central pixel). The part of the partial derivative of  $\tilde{F}(A)$  w.r.t.  $\alpha_{i,j}$ , which comes from the  $3 \times 3$  tile  $\mathcal{N}(i)$  with  $\alpha_{i,j}$  as the central pixel is

$$\frac{\partial\{\tilde{F}[\mathcal{N}(i)]\}}{\partial\alpha_{i,j}} = \begin{cases} c, & \text{if } k_{i,j} \neq \alpha_{i,j} \\ d, & \text{if } \max_{k \in \mathcal{N}(i)} \alpha_{k,j} = \alpha_{i,j} \\ e, & \text{if } \min_{k \in \mathcal{N}(i)} \alpha_{k,j} = \alpha_{i,j} \end{cases} \quad (27)$$

where

$$\begin{aligned} c &= 4(1 - k_{i,j})(1 - 2\alpha_{i,j}), \\ d &= 4\{1 + \min_{k \in \mathcal{N}(i)} \alpha_{k,j} \\ &\quad - \alpha_{i,j}[4 + 2 \min_{k \in \mathcal{N}(i)} \alpha_{k,j}] + 3\alpha_{i,j}^2\}, \\ e &= 4[1 - \max_{k \in \mathcal{N}(i)} \alpha_{k,j} \\ &\quad + 2\alpha_{i,j} \max_{k \in \mathcal{N}(i)} \alpha_{k,j} - 3\alpha_{i,j}^2]. \end{aligned} \quad (28)$$

$\alpha_{i,j}$  can also be maximal or minimal coverage value in  $3 \times 3$  neighborhood related to coverage values  $\alpha_{z,j}$ ,  $z \in \mathcal{N}(i) \setminus \{i\}$  (remaining eight out of nine  $3 \times 3$  tiles, which overlap with coverage value  $\alpha_{i,j}$ ). We therefore need to include the following value:

$$s_{z,j} = \begin{cases} 4\alpha_{z,j}(1 - \alpha_{z,j}), & \text{if } \min_{k \in \mathcal{N}(z)} \alpha_{k,j} = \alpha_{i,j} \\ -4\alpha_{z,j}(1 - \alpha_{z,j}), & \text{if } \max_{k \in \mathcal{N}(z)} \alpha_{k,j} = \alpha_{i,j} \\ 0, & \text{otherwise,} \end{cases} \quad (29)$$

for all  $z \in \mathcal{N}(i) \setminus \{i\}$ .

Finally, the partial derivative of  $\tilde{F}(A)$  w.r.t.  $\alpha_{i,j}$  is

$$\frac{\partial \tilde{F}(A)}{\partial \alpha_{i,j}} = \frac{\partial \{\tilde{F}[N(i)]\}}{\partial \alpha_{i,j}} + \sum_{z \in N(i) \setminus \{i\}} s_{z,j}. \quad (30)$$

d. The partial derivative of  $R(A)$  w.r.t.  $\alpha_{i,j}$  is

$$\frac{\partial R(A)}{\partial \alpha_{i,j}} = [(A \mathbb{1}_{m \times 1} - \mathbb{1}_{N_h \times 1}) \mathbb{1}_{1 \times m}]_{i,j}. \quad (31)$$

$\mathbb{1}$  denotes a vector with all elements equal to 1.

The other observed versions of  $E(A; I)$  (i) to (iv) (Sec. 3.1) include different regularization terms. Partial derivatives for the terms  $P(A)$ ,  $T(A)$ , and  $F(A)$  are given in Ref. 16. For  $P_H(A)$  partial derivatives are computed combining the derivatives for  $P(A)$  and the chain rule.

### Acknowledgments

The authors acknowledge the support from Swedish Research Council through Projects 2014-4231, 2015-05878, and 2017-04385. Ministry of Education, Science and Technological Development of the Republic of Serbia is acknowledged for the support through Projects ON174008 and III44006 of MI-SANU.

### References

- N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Process. Mag.* **19**(1), 44–57 (2002).
- T. Haraguchi et al., "Spectral imaging fluorescence microscopy," *Genes Cells* **7**(9), 881–887 (2002).
- T. Zimmermann et al., "Clearing up the signal: spectral imaging and linear unmixing in fluorescence microscopy," *Confocal Microsc. Methods Protoc.* **1075**, 129–148 (2014).
- A. Villa et al., "Classification of hyperspectral images at a finer spatial resolution," *IEEE J. Sel. Top. Signal Process.* **5**(3), 521–533 (2011).
- S. Verbeiren et al., "Sub-pixel classification of spot-vegetation time series for the assessment of regional crop areas in Belgium," *Int. J. Appl. Earth Obs. Geoinf.* **10**(4), 486–497 (2008).
- C. Lanaras, E. Baltasvias, and K. Schindler, "Hyperspectral super-resolution with spectral unmixing constraints," *Remote Sens.* **9**(11), 1196 (2017).
- S. Huang, H. Zhang, and A. Pižurica, "A robust sparse representation model for hyperspectral image classification," *Sensors* **17**(9), 2087 (2017).
- L. Zhang, L. Zhang, and B. Du, "Deep learning for remote sensing data: a technical tutorial on the state of the art," *IEEE Geosci. Remote Sens. Mag.* **4**(2), 22–40 (2016).
- J. Tohka, "Partial volume effect modeling for segmentation and tissue classification of brain magnetic resonance images: a review," *World J. Radiol.* **6**(11), 855 (2014).
- A. Souza, J. K. Udupa, and P. K. Saha, "Volume rendering in the presence of partial volume effects," *IEEE Trans. Med. Imaging* **24**(2), 223–235 (2005).
- W. J. Niessen et al., "Multiscale segmentation of three-dimensional MR brain images," *Int. J. Comput. Vision* **31**(2–3), 185–202 (1999).
- N. Sladoje and J. Lindblad, "Estimation of moments of digitized objects with fuzzy borders," *Lect. Notes Comput. Sci.* **3617**, 188–195 (2005).
- N. Sladoje, I. Nyström, and P. K. Saha, "Measurements of digitized objects with fuzzy borders in 2D and 3D," *Image Vision Comput.* **23**(2), 123–132 (2005).
- N. Sladoje and J. Lindblad, "High-precision boundary length estimation by utilizing gray-level information," *IEEE Trans. Pattern Anal. Mach. Intell.* **31**(2), 357–363 (2009).
- S. Dražić, N. Sladoje, and J. Lindblad, "Estimation of Feret's diameter from pixel coverage representation of a shape," *Pattern Recognit. Lett.* **80**, 37–45 (2016).
- V. Ilić, J. Lindblad, and N. Sladoje, "Signature of a shape based on its pixel coverage representation," *Lect. Notes Comput. Sci.* **9647**, 181–193 (2016).
- N. Sladoje and J. Lindblad, "Pixel coverage segmentation for improved feature estimation," *Lect. Notes Comput. Sci.* **5716**, 929–938 (2009).
- J. Lindblad and N. Sladoje, "Coverage segmentation based on linear unmixing and minimization of perimeter and boundary thickness," *Pattern Recognit. Lett.* **33**(6), 728–738 (2012).
- F. Malmberg et al., "A graph-based framework for sub-pixel image segmentation," *Theor. Comput. Sci.* **412**(15), 1338–1349 (2011).
- J. Liang et al., "Image coverage segmentation based on soft boundaries," *Lect. Notes Comput. Sci.* **8536**, 374–381 (2014).
- K. Lidayová et al., "Coverage segmentation of thin structures by linear unmixing and local centre of gravity attraction," in *IEEE Int. Symp. Image and Signal Process. and Anal.*, pp. 83–88 (2013).
- K. Lidayová et al., "Coverage segmentation of 3D thin structures," in *IEEE Int. Conf. Image Process. Theory, Tools and Appl.*, pp. 23–88 (2015).
- T. Majtner et al., "Improving skin lesion segmentation in dermoscopic images by thin artefacts removal methods," in *IEEE Eur. Workshop Vis. Inf. Process.*, pp. 1–6 (2016).
- X. Lan et al., "Learning common and feature-specific patterns: a novel multiple-sparse-representation-based tracker," *IEEE Trans. Image Process.* **27**(4), 2022–2037 (2018).
- Y. Li et al., "Image super-resolution with parametric sparse model learning," *IEEE Trans. Image Process.* **27**(9), 4638–4650 (2018).
- Z. Zhang et al., "A survey of sparse representation: algorithms and applications," *IEEE Access* **3**, 490–530 (2015).
- J. Wright et al., "Sparse representation for computer vision and pattern recognition," *Proc. IEEE* **98**(6), 1031–1044 (2010).
- T. Tong et al., "Segmentation of MR images via discriminative dictionary learning and sparse coding: application to hippocampus labeling," *NeuroImage* **76**, 11–23 (2013).
- D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Commun. Pure Appl. Math.* **42**(5), 577–685 (1989).
- L. Bar et al., "Mumford and Shah model and its applications to image segmentation and image restoration," in *Handbook of Mathematical Methods in Imaging*, O. Scherzer, Ed., pp. 1095–1157, Springer, New York (2011).
- L. Dorst and A. W. Smeulders, "Length estimators for digitized contours," *Comput. Vision Graphics Image Process.* **40**(3), 311–333 (1987).
- E. Birgin, J. Martinez, and M. Raydan, "Nonmonotone spectral projected gradient methods on convex sets," *SIAM J. Opt.* **10**(4), 1196–1211 (2000).
- B. Bajić, J. Lindblad, and N. Sladoje, "Restoration of images degraded by signal-dependent noise based on energy minimization: an empirical study," *J. Electron. Imaging* **25**(4), 43020 (2016).
- B. Bajić, J. Lindblad, and N. Sladoje, "Single image super-resolution reconstruction in presence of mixed Poisson-Gaussian noise," in *IEEE Int. Conf. Image Process. Theory, Tools and Appl.*, pp. 1–6 (2016).
- L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D* **60**(1–4), 259–268 (1992).
- R. Schultz and R. Stevenson, "Stochastic modeling and estimation of multispectral image data," *IEEE Trans. Image Process.* **4**(8), 1109–1119 (1995).
- E. M. Hendrix et al., "A new minimum-volume enclosing algorithm for endmember identification and abundance estimation in hyperspectral data," *IEEE Trans. Geosci. Remote Sens.* **50**(7), 2744–2757 (2012).
- J. M. Bioucas-Dias et al., "Hyperspectral remote sensing data analysis and future challenges," *IEEE Geosci. Remote Sens. Mag.* **1**(2), 6–36 (2013).
- J. M. Nascimento and J. M. Dias, "Vertex component analysis: a fast algorithm to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.* **43**(4), 898–910 (2005).
- M. E. Winter, "N-FINDR: an algorithm for fast autonomous spectral end-member determination in hyperspectral data," *Proc. SPIE* **3753**, 266–275 (1999).
- P. Campisi and K. Egiazarian, *Blind Image Deconvolution: Theory and Applications*, CRC Press, Boca Raton (2016).
- Q. Lu et al., "Robust blur kernel estimation for license plate images from fast moving vehicles," *IEEE Trans. Image Process.* **25**(5), 2311–2323 (2016).
- D. L. Donoho, "For most large underdetermined systems of linear equations the minimal  $\ell^1$ -norm solution is also the sparsest solution," *Commun. Pure Appl. Math.* **59**(6), 797–829 (2006).
- N. Sladoje and J. Lindblad, "The coverage model and its use in image processing," in *Selected Topics on Image Processing and Cryptology*, Zbornik radova (Collection of Papers), M. Mihaljević, Ed., Vol. **15**, pp. 39–117, Mathematical Institute of the Serbian Academy of Sciences and Arts, Belgrade (2012).
- B. Bajić, J. Lindblad, and N. Sladoje, "Blind restoration of images degraded with mixed Poisson–Gaussian noise with application in transmission electron microscopy," in *IEEE Int. Symp. Biomed. Imaging: From Nano to Macro*, pp. 123–127 (2016).
- M. F. Baumgardner, L. L. Biehl, and D. A. Landgrebe, "220 band AVARIS hyperspectral image data set: June 12, 1992 Indian Pine Test Site 3," 2015, <https://pur.purdue.edu/publications/1947/1> (25 October 2016).
- E. Seeman and P. D. Delmas, "Bone quality—the material and structural basis of bone strength and fragility," *N. Engl. J. Med.* **354**(21), 2250–2261 (2006).

48. C. Chen et al., "Assessment of trabecular bone strength at in vivo CT imaging with space-variant hysteresis and finite element modelling," in *IEEE 13th Int. Symp. Biomed. Imaging (ISBI)*, pp. 872–875, IEEE (2016).
49. C. Chen et al., "Robust segmentation of trabecular bone for in vivo CT imaging using anisotropic diffusion and multi-scale morphological reconstruction," *Proc. SPIE* **10137**, 101371T (2017).
50. Y. Li et al., "Super-resolution/segmentation of 2D trabecular bone images by a Mumford–Shah approach and comparison to total variation," in *24th Eur. Signal Process. Conf. (EUSIPCO)*, IEEE, pp. 1693–1697 (2016).
51. J. H. Waarsing, J. S. Day, and H. Weinans, "An improved segmentation method for in vivo  $\mu$ CT imaging," *J. Bone Miner. Res.* **19**(10), 1640–1650 (2004).
52. M. Salomé et al., "A synchrotron radiation microtomography system for the analysis of trabecular bone samples," *Med. Phys.* **26**(10), 2194–2204 (1999).
53. A. Toma, B. Sixou, and F. Peyrin, "Iterative choice of the optimal regularization parameter in TV image restoration," *Inverse Prob. Eng.* **9**(4), 1171–1191 (2015).

**Buda Bajić** is currently pursuing her PhD at the Faculty of Technical Sciences, University of Novi Sad, Serbia. She received her BSc and MSc degree in applied mathematics from the Faculty of Science, University of Novi Sad, Serbia, in 2011 and 2012. Her research

interests include image processing, inverse problems, and image restoration.

**Joakim Lindblad** received his MSc degree in engineering physics and PhD in computerized image analysis from Uppsala University, Sweden, in 1997 and 2003, respectively. He is currently affiliated with the Centre for Image Analysis, Uppsala University (researcher), Mathematical Institute SANU, Belgrade (research associate professor) and Toppolf Sweden AB, Stockholm (head of research). His research interests include development of general methods for image segmentation, feature extraction, and classification.

**Nataša Sladoje** received her BSc and MSc degrees in mathematics from the Faculty of Science, University of Novi Sad, in 1992 and 1998, respectively, and her PhD in computerized image analysis from the Centre for Image Analysis, Swedish University of Agricultural Sciences, Uppsala, in 2005. She is a senior lecturer at the Centre for Image Analysis, Uppsala University, and is also a professor at the Mathematical Institute, Serbian Academy of Sciences and Arts. Her research interests include theoretical development of image analysis and machine learning methods with applications in biomedicine and medicine.

## Buda Bajić

14.2.1988

Dože Derđa 6/8, 21000 Novi Sad Serbia

+381 64 09 18 362

[buda.bajic@uns.ac.rs](mailto:buda.bajic@uns.ac.rs)

[imft.ftn.uns.ac.rs/~buda/](http://imft.ftn.uns.ac.rs/~buda/)

[scholar.google.com/citations?user=dwxKfgYAAAAJ&hl=en](https://scholar.google.com/citations?user=dwxKfgYAAAAJ&hl=en)



## EDUCATION

<b>Faculty of Science, University of Novi Sad, Serbia</b>	<b>2007-2011</b>
<ul style="list-style-type: none"><li>- B.Sc. degree in Applied Mathematics</li><li>- Grade point average <b>10.00 of 10.00</b></li></ul>	
<b>Faculty of Science, University of Novi Sad, Serbia</b>	<b>2011-2012</b>
<ul style="list-style-type: none"><li>- M.Sc. degree in Applied Mathematics</li><li>- Grade point average <b>10.00 of 10.00</b></li></ul>	
<b>Faculty of Technical Sciences, University of Novi Sad, Serbia</b>	<b>2013-now</b>
<ul style="list-style-type: none"><li>- Ph.D. student</li><li>- Advisers: Nataša Sladoje, Ph.D, and Joakim Lindblad, Ph.D.</li><li>- Research area – applied mathematics, image processing, image restoration, numerical optimization</li></ul>	

## WORK EXPERIENCE

<b>National Bank of Serbia, Belgrade, Serbia</b>	<b>2012-2013</b>
<ul style="list-style-type: none"><li>- Directorate for Economic Research and Statistics – Macroeconomic Analyses Division</li><li>- Development of econometric model for prediction of growth rate of gross domestic product (GDP)</li></ul>	
<b>Faculty of Technical Sciences, University of Novi Sad, Serbia</b>	<b>2013-now</b>
<ul style="list-style-type: none"><li>- <b>Teaching assistant</b> – Department of Fundamental Sciences – Chair for mathematics</li><li>- Teaching various undergraduate courses in mathematics (calculus, probability and stochastic processes) at different engineering program</li></ul>	

## RESEARCH

<b>Journal publications</b>	
<ul style="list-style-type: none"><li>- Restoration of images degraded with signal-dependent noise based on energy minimization: an empirical study. <i>Journal of Electronic Imaging (SPIE)</i>, Vol. 25, No. 4, 043020, 2016. B. Bajić, J. Lindblad, and N. Sladoje.</li></ul>	<b>2016</b>
<ul style="list-style-type: none"><li>- Super-resolution coverage segmentation by linear unmixing in presence of blur and noise. <i>Journal of Electronic Imaging (SPIE)</i>, Vol. 28, No. 1, 013046, 2019. B. Bajić, J. Lindblad, and N. Sladoje.</li></ul>	<b>2019</b>
<ul style="list-style-type: none"><li>- Generalized Deep Learning Framework for HEP-2 Cell Recognition. Submitted. B. Bajić, T. Majtner, J. Lindblad and N. Sladoje.</li></ul>	<b>2018</b>
<b>Peer reviewed conference publications and participations</b>	
<b>International Conference on Image Analysis and Recognition, ICIAR 2014, Portugal</b>	<b>October 2014</b>
<ul style="list-style-type: none"><li>- An Evaluation of Potential Functions for Regularized Image Deblurring. B. Bajić, J. Lindblad, and N. Sladoje, <i>Lecture Notes in Computer Science</i>, Vol. 8814, pp. 150–158, Springer, 2014.</li></ul>	
<b>IEEE International Symposium on Biomedical Imaging, ISBI 2016, Czech Republic</b>	<b>April 2016</b>
<ul style="list-style-type: none"><li>- Blind Restoration of Images Degraded with Mixed Poisson-Gaussian Noise with Application in Transmission Electron Microscopy. B. Bajić, J. Lindblad, and N. Sladoje. pp. 123-127.</li></ul>	
<b>IEEE International Conference on Image Processing Theory, Tools and Applications, IPTA 2016, Finland</b>	<b>December 2016</b>
<ul style="list-style-type: none"><li>- Single image super-resolution reconstruction in presence of mixed Poisson-Gaussian noise. B. Bajić, J. Lindblad, and N. Sladoje. <b>Awarded as one of the "Best Reviewed Papers"</b>.</li></ul>	
<b>IEEE International Symposium on Biomedical Imaging, ISBI 2018, United States of America</b>	<b>April 2018</b>
<ul style="list-style-type: none"><li>- Denoising of Short Exposure Transmission Electron Microscopy Images for Ultrastructural Enhancement. B. Bajić, A. Suveer, A. Gupta, I. Pepić, J. Lindblad, N. Sladoje, I.-M. Sintorn. pp. 921-925.</li></ul>	
<b>Non-reviewed conference presentations</b>	
<b>Symposium of the Swedish Society for Automated Image Analysis, Sweden</b>	<b>March 2015</b>
<ul style="list-style-type: none"><li>- Performance Evaluation of Potential Functions for Regularized Image Enhancement. B. Bajić, J. Lindblad, and N. Sladoje. pp. 117-122.</li></ul>	
<b>Symposium of the Swedish Society for Automated Image Analysis, Sweden</b>	<b>March 2016</b>
<ul style="list-style-type: none"><li>- Blind deconvolution of images degraded with mixed Poisson-Gaussian noise with application in Transmission Electron Microscopy. B. Bajić, J. Lindblad, and N. Sladoje. pp. 137-141.</li></ul>	
<b>Participations in research projects</b>	
<ul style="list-style-type: none"><li>- "Advanced Techniques of Cryptology, Image Processing and Computational Topology for Information Security", grant OI 174008 of Ministry of Sciences of the Republic of Serbia. 2011-2017.</li><li>- Swedish Research Council, Vetenskapsrådet, within the Swedish Research Links program: Collaborative development of methods for robust and precise image analysis for cost effective and reliable detection of cervical cancer. 2015-2017.</li></ul>	

Summer Schools	
<b>22nd Summer School in Image Processing</b>	<b>July 2014</b>
- Faculty of Electrical Engineering and Computing, <b>University of Zagreb</b> , Croatia	
<b>23rd Summer School in Image Processing</b>	<b>July 2015</b>
- <b>University of Szeged</b> , Hungary	
<b>BMS summer school 2016 - Mathematical and Numerical Methods in Image Processing</b>	<b>July 2016</b>
- <b>Humboldt University and Technical University of Berlin</b> , Germany	
<b>Summer School on Advanced Methods in Biomedical Image Analysis (AMBIA)</b>	<b>September 2017</b>
- <b>Masaryk University, Brno</b> , Czech Republic	
Visiting researcher	
- Centre for Image Analysis, <b>Uppsala University</b> , Sweden	23 Feb - 1 Apr 2015
- Laboratoire d'Informatique Gaspard Monge, <b>Universite Paris-Est</b> , France	14 - 25 Sep 2015
- Centre for Image Analysis, <b>Uppsala University</b> , Sweden	18 Feb - 17 Mar 2016
- Centre for Image Analysis, <b>Uppsala University</b> , Sweden	14 Aug - 28 Oct 2017
- Center for Microscopy and Image Analysis, <b>University of Zurich</b> , Switzerland	19 Mar - 29 Mar 2018