HABILITATION A DIRIGER LES RECHERCHES
DE L'ECOLE NORMALE SUPERIEURE DE CACHAN

presented by
Jérôme GILLES

Speciality :
Mathematics

Title:

Contributions to image analysis and restoration

Defended in December 6th, 2012 ENS Cachan

Reviewers:

Guillermo SAPIRO	Duke University
Vicente CASELLES	Universitat Pompeu Fabra, Barcelona
Laure BLANC-FÉRAUD	INRIA - Sophia-Antipolis

Committee members:

Florence TUPIN	TELECOM ParisTech
Jean-Luc STARCK	CEA - Saclay
Patrick PEREZ	Thomson R&D
Jean-Michel MOREL	CMLA - ENS Cachan

Ecole Normale Supérieure de CACHAN
61, Avenue du Président Wilson, 94235 CACHAN Cedex, FRANCE

Acknowledgments

I wish to thank Laure Blanc-Féraud, Guillermo Sapiro and Vicent Caselles who have accepted to review this Habilitation Thesis. I wish to thank the committee members, Florence Tupin, Jean-Luc Starck and Patrick Perez for their time as well as Jean-Michel Morel who accepted to be the guarantor of this Habilitation Thesis.

I wish to especially thank Yves Meyer for the useful and enthusiastic discussions we have and to be a constant source of inspiration for me. I thank Stanley Osher and Andrea Bertozzi for welcoming me in the Applied Math group at UCLA and leaving me enough liberty to follow my own research path as well as proposing me to work on very interesting projects.
I also thank my ancient colleagues at the DGA for their good mood and with whom I gained experience in optronics (thanks Stéphane!) and the application of image processing into real problems. I wish to thank all my other collaborators: Luminita Vese, Yves van Gennip, Rustum Chotski, Mario Michelli, Yu Mao, David Frakes, Richard Espinola, Endre Repasi, Tristan Dagobert, Enric Meinhardt-Llopis, ...

A warm thank to all the students who worked with me on many interesting topics in image processing these last ten years with a special mention to Yohann Tendero who was my first Ph.D former student.

Finally, I wish to thank my wife, Agnès, for her support in both my career and my life.

Contents

Notations

The following notations are used throughout the whole document.

$\hat{u}$ or $u^\wedge$: Fourier transform of u
$\tilde{u}$: optimal u
$\star$: convolution product
$\mathbb{R}$: set of real numbers
BV: Bounded Variations function space
G: Meyer's space of oscillatory functions
$W^{s,p}$: Sobolev spaces
$B^s_{p,q}$: Besov spaces
WST, CST: Wavelet and Contourlet Soft Thresholdings, respectively
Δ_j: Littlewood-Paley filter at scale j

ENS: *Ecole Normale Supérieure*
CMLA: *Centre de Mathématiques et de Leurs Applications*
UCLA: University of California, Los Angeles
DGA: *Direction Générale pour l'Armement* (an institution of the French Ministry of Defense)
D.O.D: Department Of Defense
NATO: North Atlantic Treaty Organization
IR: Infrared
MTF: Modulation Transfert Function
PSF: Point Spread Function

Chapter 1

Introduction

In less than two decades, thanks to the constant increase of computing power, image processing has become a full-fledged field. Like any other scientific field, mathematics is now at the core of new developments meeting new problems and creating new tools (both theoretical and applied). All mathematical branches, such as partial differential equations, Fourier analysis, multiresolution analysis (wavelets and their geometric extensions), functional analysis, ... are used to solve image processing problems. Despite a huge number of new methods and algorithms, two main topics in image processing still need further investigation: image analysis and image restoration.

The goal of image analysis is to extract information from an image or a sequence of images. This kind of problem has strong connections with computer vision and covers a large spectrum of applications like measuring physical parameters (cell sizes in microscopy, matter properties, temperature of objects, ...), detecting specific objects (tumor detection in medical imaging, vehicle detection and recognition for military applications, ...) or providing some autonomy to artificial systems (robots, for example) by reconstructing their environment. The main goal of image analysis is to obtain relevant features which best represent the desired information. Such purposes generally lead to mathematical modeling problems. A well known example concerns textures characterization where different approaches were proposed (statistical tools, wavelets analysis, ...) providing interesting results in particular cases. But the construction of a general model to represent textures remains an intense and challenging research field.

Image restoration is probably one of the most important topics today beause it appears in all imaging processes. Indeed, each image is acquired via specific equipment with imperfections that consequently creates some errors (blur, distortions, ...) in the image. Moreover, independent of the type of imaging system (light, electromagnetic, ...), the media through which "objects" are observed also influences the quality of the image received by the sensor. These are the reasons why restoration techniques are widely studied in the literature.

Before my current position at UCLA, I spent nine years as an "image processing expert" at the Direction Générale pour l'Armement (DGA), an institution of the French Ministry of Defense. My main work concerned image processing for passive and active infrared imaging systems. An important military image processing application is the Automatic Target Detection/Recognition and Tracking (ATD/ATR/ATT) algorithms. The most difficult case is when targets (typically vehicles) are camouflaged in a rural environment, because advanced textures analysis tools are needed. My interest in texture modeling and analysis comes from this problem. It was the topic of my Ph.D. and continues now to drive some of my research activities. My contributions on texture modeling/decomposition are presented in chapter 3. My second major interest concerns the problem of observation through atmospheric turbulences. This question arises for ground to ground long range imaging systems where the impact of the atmosphere becomes non negligible. Turbulence in the air causes two main effects: blur and geometric distortions which need to be restored if we want to retrieve details on objects for recognition purposes. I led investigations on this topic which resulted in two important contributions; one on turbulence stabilization and another on atmospheric deblurring. Chapter 4 presents this work and shows that it is at the crossroad of many fields (mathematics, physics, optics, ...). Other contributions in infrared image processing are presented in chapter 5 while chapter 6 gives an overview of my current and future research investigations. Finally, all my publications are listed in chapter 7.

Chapter 2

Personal statement

2.1 *Curriculum Vitae*

Jérôme Gilles - *Assistant Adjunct Professor*

Born in January 07, 1974 at Langres (France).

Academic Address ✉:
Department of Mathematics
University of California Los Angeles (UCLA),
520 Portola Plaza - Office 7324,
Los Angeles, CA 90095-1555, USA

Phone number ☎: (001) 310-794-7737
Email ✍: jegilles@math.ucla.edu
Homepage ⌨: http://www.math.ucla.edu/∼jegilles

2.1.1 Professional experience

Since November 2010, I have been an Assistant Adjunct Professor at the Department of Mathematics at UCLA. In this position, I have a teaching charge and a research post on Professor Stanley Osher's team.

From September 2001 to August 2010, I occupied the position of *Professeur Agrégé* for the French Ministry of Defense in France. I was assigned to the *Direction Générale pour l'Armement* (DGA), in charge of research and studies in image processing in the Department of Space, Observation, Intelligency and UAVs. My first few years in this position were also dedicated to my Ph.D. From 2004 to my departure I was the organizer of internal seminars about robotics, signal and image processing. I was appointed Head of scientific activities and perspectives of the Department in 2007.

In addition, I am an associate member of the *Centre de Mathématiques et de Leurs Applications* (CMLA) of *Ecole Normale Supérieure* (ENS) of *Cachan*.

2.1.2 Education

2006: Ph.D. Thesis in Mathematics at the *Ecole Normale Supérieure of Cachan*, Defended in June 22, 2006
Title: *Décomposition et détection de structures géométriques en imagerie* (Decomposition and detection of geometrical structures in imaging).
Commitee:

- President: Prof. Jean-Michel Morel
- Rewiewers: Prof. Henri Maître and Prof. Gilles Aubert
- Referee: Dr. Frédéric Pradeilles
- Supervisors: Prof. Yves Meyer and Dr. Bertrand Collin

1999: Fourth year at ENS Cachan. *Diplôme d'Etudes Approfondies* (DEA) of ENS Cachan
Speciality: Automatic and signal processing,

1998: Third year at ENS Cachan. Preparation of *Agrégation* in Electrical Engineering (speciality: Electronics and Computer Sciences),

1997: Second year at ENS Cachan. *Maîtrise EEA* delivered by Paris XI University,

1996: First year at ENS Cachan. *Licence ingénierie électrique* delivered by Paris XI University,

1992-1995: *Classes préparatoires aux Grandes Ecoles*,

1992: *Baccalauréat*, speciality: electronics.

2.2 Academic involvements

2.2.1 Teaching

I currently teach the following undergraduate courses at UCLA:

- Methods of Applied Mathematics ($\sim$ 30 students): dimensional analysis, perturbation methods, variational calculus, Fourier series, Green's functions.

- Mathematical Modeling ($\sim$ 45 students): population dynamics (logistic equation, predator/prey model, Lotka-Volterra equation, phase-plane methods for nonlinear systems), traffic modeling (concepts of velocity vector fields, associated PDEs, methods of characteristics, wave propagation).

- Calculus of Several Variables - B (integral calculus) ($\sim$ 210 students): definition of multiple integrals; polar, cylindrical and spherical coordinates; curvilinear integrals; surface integrals; integrals of vector fields; divergence and curl operators; Green's, Stokes and divergence theorems.

Previously on the margins of my position at the DGA, I was Assistant Professor in different engineering schools in France:

- Lecturer for the Master at ENST: Image processing for Defense applications, from 2007 to 2009.

- EUROSAE Advanced Training at ENSTA: Image processing - Deformable models (active contours) - Image Modeling, from 2005 to 2010.

- Lecturer at ESIEE: Introduction to signal processing (third year), Advanced concepts for signal processing (fourth year), from 2006 to 2009.

- Preparatory classes at ENSTA: Initiation to C programing, from 2004 to 2009.

2.2.2 Student supervision

At the DGA, I supervised several students:

- Carlo De Franchis. "*Déconvolution aveugle d'images*" (Image blind deconvolution). First year at Ecole Polytechnique, 2007/2008.

- Tristan Dagobert. "*Amélioration d'images infrarouges déformées par la turbulence atmosphérique*" (Restoration of infrared images degraded by atmospheric turbulences). CNAM Engineer Diploma, 2007.
- Anne Favède. "*Correction des effets de la turbulence atmosphérique en imagerie passive et active*" (Correction of atmospheric turbulence effects in passive and active imaging). Fifth year at INPG Grenoble-ENSERG, 2006.
- Damien Jeandrau. "*Classification de textures par SVM*" (Texture classification by SVM). Second year at IUT de Cachan, 2005 .
- Sylvain Rousselle. "*Segmentation de textures par plongement fractal*" (Texture segmentation by fractal embeddding). Fifth year at ESIEE, 2005.
- Nicolas Widynski. "*Détection semi-automatique de réseaux routiers dans le cadre d'images aériennes et satellites*" (Semi-automatic road network detection in aerial and satellite images). Third year at EPITA, 2005.
- Laurent Pigois. "*Simulation des effets de la turbulence atmosphérique par écrans de phase*" (Simulation of atmospheric turbulence effects by phase screens). DEA, Cergy University, 2004.
- Aurélien Zelty. "*Segmentation d'images texturées par filtrage de Gabor*" (Segmentation of textured images by Gabor filtering). Third year at EPITA, 2003.
- Mélissa Pinel. "*Contrôle adaptatif de contraste local adapté aux images IR à forte dynamique*" (Local contrast adaptive control for high dynamic IR images). DESS, Orsay University, 2003.
- William Phommaly. "*Correction du bruit de non-uniformité d'une caméra matricielle*" (Non-uniformity noise correction for matrix based camera). DESS, Jussieu University, 2003.

From September 2009 to June 2012, I was involved in the supervision of the Ph.D. thesis of Yohann Tendero with Prof. Jean-Michel Morel at CMLA (ENS Cachan) on the topic of infrared image processing (Yohann Tendero defended his Ph.D in June, 22th 2012). I am also currently involved in the work of two Ph.D. candidates at UCLA: Hayden Schaeffer about image decomposition models and Melissa Tong about underwater deconvolution.

I was in the referee commitee of the following Ph.D. thesis:

- Stéphanie Bigot-Marchand. *Outils de traitement d'images adaptés au traitement d'images omnidirectionnelles* (Image processing tools adapted to omnidirectional image processing). Picardie Jules Vernes University, speciality: Mathematics. Defended in October 15, 2008.
- Nicolas Morizet. *Reconnaissance biométrique par fusion multimodale du visage et de l'iris* (Biometric recognition by face and iris multimodal fusion). Telecom ParisTech, speciality: signal and images. Defended in December 15, 2008.

2.3 Scientific community involvements

I am an active reviewer in the following journals (more than forty reviews at the date this manuscript is written):

- IEEE transactions in image processing,
- IEEE signal processing letters,
- IEEE transactions in signal processing,
- IEEE transactions on geoscience and remote sensing,
- Journal of Mathematical Imaging and Vision,
- Journal of Visual Communication and Image Representation,
- The Visual Computer Journal.
- Journal of Electronic Imaging

Besides, I am in the scientific commitee of the ACIVS (Advanced Concept for Intelligent Vision Systems) conference since 2009.

2.4 Software development

During the years I spent at the DGA, I was one of the person in charge of the development of the internal image processing library. This library, called GAMMAG, was written in C and expected to provide an access to the most used image processing algorithms (segmentation, enhancement, restoration, ...) to people of the department and to compile the different research developments.
Since I am at UCLA, I wrote a Matlab toolbox, called the Bregman Cookbook, which aims to provide denoising and deconvolution algorithms using L^1 minimization based on Bregman iterations. This toolbox is freely available on Matlab Central (a direct link can be found on my homepage).

Chapter 3

Image decomposition

This chapter mainly focuses on the work I did during my Ph.D. with Prof. Yves Meyer. We investigated the ability to decompose an image into its respective cartoon and textures parts. In the first section, I describe the mathematical models providing such decomposition and some numerical cues. The second section is devoted to the idea of seeking a cartoon + texture + noise decomposition to deal with noisy images. Finally in the last section I concentrate on my contributions on theoretical aspects, first by recalling a theorem which gives the different decompositions according to the choice of the parameters; and secondly by giving new recent results.

3.1 Structures + textures decomposition

3.1.1 Meyer's $BV - G$ model

The image decomposition story starts with the work of Y. Meyer [30] who investigated the Rudin-Osher-Fatemi (ROF) algorithm [32]. Meyer noted that this model is not efficient to capture textures. In order to fix the notations and to understand the concepts behind decomposing an image, let us recall the ROF model. Assuming f is an observed image that is the addition of the ideal scene u, the one we want to recover, and some noise b. The authors of [32] propose to minimize the following functional to denoise u:

$$F_\lambda^{ROF}(u) = J(u) + \frac{\lambda}{2}\|f - u\|_{L^2}^2. \tag{3.1}$$

Because it preserves sharp edges, this model assumes that u is in BV (the space of Bounded Variations), by setting $J(u)$ as being the total variation of u. This algorithm gives good results and is very easy to implement by the use of Chambolle's nonlinear projectors [18] or Bregman iterations.

Now if, instead of considering a denoising problem, we follow the image decomposition point of view, $f = u + v$, where u is the structure part and v the texture part, the functional in Eq.(3.1) can be rewritten as

$$F_\lambda^{ROF}(u, v) = J(u) + \frac{\lambda}{2}\|v\|_{L^2}^2. \tag{3.2}$$

Meyer shows that this model is not adapted to properly achieve this decomposition. The following example illustrates that the more a texture oscillates, the more it is removed from both u and v parts.

Example 1 *Let v be a texture created from an oscillating signal over a finite domain. Then v can be written ($x = (x_1, x_2)$) as follows:*

$$v(x) = \cos(\omega x_1)\theta(x), \tag{3.3}$$

where ω is the frequency and θ the characteristic function over the considered domain. Then we can calculate the L^2 and BV norms of v, respectively. We get

$$\|v\|_{L^2} \approx \frac{1}{\sqrt{2}}\|\theta\|_{L^2}, \tag{3.4}$$

which is constant $\forall\omega$ and does not specially capture the textures. In addition,

$$\|v\|_{BV} = \frac{\omega}{2\pi}\|\theta\|_{L^1}, \tag{3.5}$$

which grows as $\omega \to \infty$ and then clearly rejects the textures.

In order to capture textures in the v component, Meyer proposes to adapt the ROF model by replacing L^2 by another space, called G, which might correspond to a space of oscillating functions. He proves that this space is the dual space of $\mathcal{BV}$ (where $\mathcal{BV} = \{f \in L^2(\mathbb{R}^2)\ ,\ \nabla f \in L^1(\mathbb{R}^2)\}$, which is close to the BV space); see [30] for more theoretical details about these spaces.
This space G is endowed by the following norm:

$$\|v\|_G = \inf_g \left\| \left(|g_1|^2 + |g_2|^2\right)^{\frac{1}{2}} \right\|_{L^\infty}, \tag{3.6}$$

where $g = (g_1, g_2) \in L^\infty(\mathbb{R}^2) \times L^\infty(R^2)$ and $v = \operatorname{div} g$. If we calculate the G-norm of the oscillating texture in Eq.(3.3) of example 1, we get

$$\|v\|_G \leqslant \frac{C}{|\omega|}, \tag{3.7}$$

where C is a constant. Then it is obvious to see that G is well adapted to capture textures. Now, the modified functional performing the structures + textures decomposition is

$$F_{\lambda}^{YM}(u,v) = J(u) + \lambda \|v\|_G, \tag{3.8}$$

where $f = u + v$, $f \in G$, $u \in BV$, $v \in G$.

3.1.2 Numerical implementation

The main drawback of this model is the presence of an L^{∞} norm in the the expression of the G-norm (3.6) which does not allow classic variational calculus for setting up numerical schemes.

The first numerical algorithm proposed to solve Meyer's model was the one published by Vese and Osher in [34]. Their approach was based on the theorem saying that $\forall f \in L^{\infty}(\Omega), \|f\|_{L^{\infty}} = \lim_{p\to\infty} \|f\|_{L^p}$ and they wrote a slightly modified version of Meyer's functional:

$$F_{\lambda,\mu,p}^{OV}(u,g) = J(u) + \lambda \|f - (u + \text{div } g)\|_{L^2}^2 + \mu \left\| \sqrt{g_1^2 + g_2^2} \right\|_{L^p}. \tag{3.9}$$

Then classic variational calculus applies and results in a system of three connected partial differential equations. All discretization details are available in [34]. This algorithm works well but is very sensitive to the choice of its parameters, which induces many instabilities.

Another way to solve Meyer's model was proposed in [2, 4, 7] by Aujol et al. The authors proposed a dual-method approach that naturally arises because of the dual relation between G and BV spaces. The problem is assumed to be in the discrete case and defined over a finite domain Ω. They proposed a modified functional to minimize, given by:

$$F_{\lambda,\mu}^{AU}(u,v) = J(u) + J^{*}\left(\frac{v}{\mu}\right) + (2\lambda)^{-1} \|f - u - v\|_{L^2}^2 \tag{3.10}$$

and

$$(u,v) \in BV(\Omega) \times G_{\mu}(\Omega). \tag{3.11}$$

The set G_{μ} is the subset in G where $\forall v \in G_{\mu}$, $\|v\|_G \leqslant \mu$. Moreover, J^{*} is the characteristic function over G_1 having the property that J^{*} is the dual operator of J ($J^{**} = J$). Thus,

Algorithm 1 $BV-G$ Aujol's cartoon + textures decomposition algorithm.

```
  Initialization: $u_0 = v_0 = 0$
  while '$\max(|u_{n+1} - u_n|, |v_{n+1} - v_n|) \leqslant \epsilon$ or a prescribed maximal number of iterations is not reached' do
    Compute $v_{n+1} = P_{G_\mu}(f - u_n)$
    Compute $u_{n+1} = f - v_{n+1} - P_{G_\lambda}(f - v_{n+1})$
  end while
```

Algorithm 2 ROF minimization by Split Bregman iterations.

```
  $u^0 = f, d_x^0 = 0, d_y^0 = 0, b_x^0 = 0, b_y^0 = 0$
  while "Not converged" do
    Update $u^{k+1}$ by using $\hat{U} = (\lambda\hat{I} - \eta\Re(\hat{\Delta}))^{-1}\left[\lambda\hat{F} - \eta\left(\mathrm{div}\,(d^k - b^k)\right)^{\wedge}\right]$
    Compute $s^k = \sqrt{|\nabla_x u^k + b_x^k|^2 + |\nabla_y u^k + b_y^k|^2}$
    $d_x^{k+1} = max(s^k - 1/\eta), 0)\frac{\nabla_x u^k + b_x^k}{s^k}$
    $d_y^{k+1} = max(s^k - 1/\eta), 0)\frac{\nabla_y u^k + b_y^k}{s^k}$
    $b_x^{k+1} = b_x^k + \nabla_x u^{k+1} - d_x^{k+1}$
    $b_y^{k+1} = b_y^k + \nabla_y u^{k+1} - d_y^{k+1}$
  end while
```

$$J^*(v) = \begin{cases} 0 & \text{if } v \in G_1 \\ +\infty & \text{else.} \end{cases} \tag{3.12}$$

The authors proved that the minimizers $(\tilde{u}, \tilde{v})$ are also minimizers of the original Meyer functional Eq. (3.8). The precited Chambolle's projectors (denoted P_{G_μ}) are projectors over the sets G_μ. More details about these projectors can be found in [18]. Algorithm. 1 resumes Aujol's decomposition algorithm.

We recently [26] proposed another approach based on Bregman iterations [27] to numerically solve the decomposition. Our idea is also based on the duality principle but used in the reverse direction. If we write $P_{ROF}(f, \lambda)$ the minimizer of the ROF model (3.1), it can be efficiently implemented with Bregman iterations (see Algorithm. 2) then proposition 1 gives how we can compute projections onto G_μ subsets (the proof of this proposition is available in [26]). The overall Bregman based decomposition algorithm is given by Algorithm. 3.

Proposition 1 *The function $v \in G_\mu$ which minimizes*

$$J^*\left(\frac{v}{\mu}\right) + \|f - v\|_2^2 \tag{3.13}$$

is given by

$$\tilde{v} = f - P_{ROF}(f, 1/\mu) \tag{3.14}$$

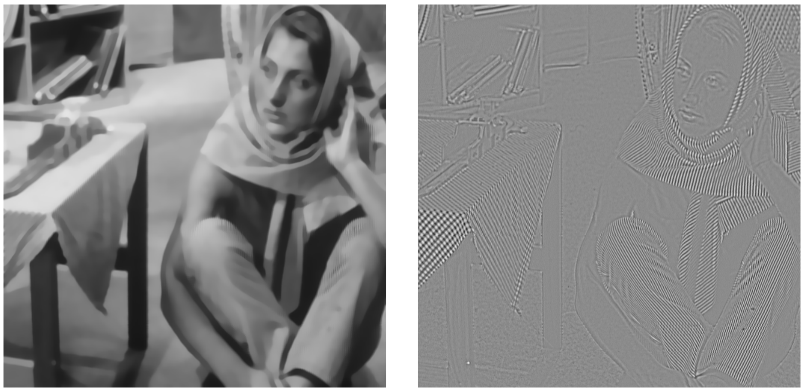
Algorithm 3 Cartoon + textures decomposition algorithm based on Bregman iterations.

```
Initialisation: $u_0 = v_0 = 0$
while " Not converged" do
    Update $u$ by $u_{n+1} = P_{ROF}(f - v_n, \lambda)$
    Update $v$ by $v_{n+1} = f - u_{n+1} - P_{ROF}(f - u_{n+1}, 1/\mu)$
end while
```

3.1.3 Decomposition examples

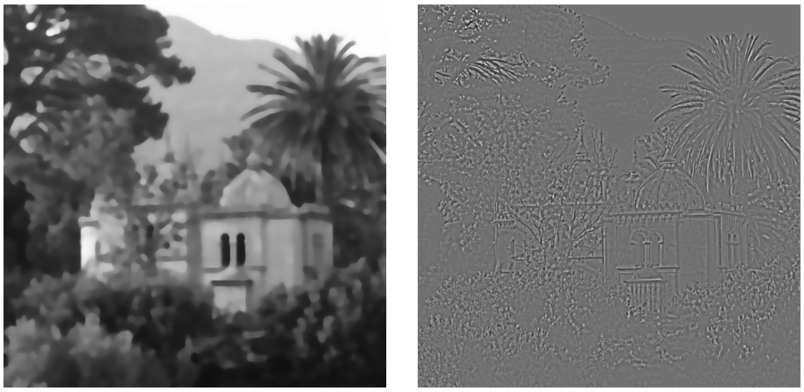
Figure 3.1 presents two original images (Barbara, Building) used in our experiments. Figures 3.2 and 3.3 illustrate the decomposition results. On each test we see that the separation between structures and textures works well. Some residual textures remain in the structures part; this can be explained by the fact the parameter λ acts as a tradeoff between the "power" of separability and too much regularization of u.

Figure 3.1: Original Barbara and Building images.



Structures Textures

Figure 3.2: *BV*-*G* structures + textures image decomposition of Barbara image.



Structures Textures

Figure 3.3: *BV*-*G* structures + textures image decomposition of Building image.

3.1.4 The $BV - \dot{B}^{\infty}_{-1,\infty}$ model

The G-norm being difficult to handle, in [30] Meyer also proposed to replace the space G by the Besov space $\dot{B}^{\infty}_{-1,\infty}$ because $G \subset \dot{B}^{\infty}_{-1,\infty}$ (in the following, for simplicity, we will denote $E = \dot{B}^{\infty}_{-1,\infty}$). The advantage is that the norm of a function v over this space can be defined from its wavelet expansion. The corresponding model proposed by Meyer is

$$F_{\lambda}^{YM2}(u,v) = J(u) + \lambda \|v\|_E \tag{3.15}$$

Algorithm 4 $BV - E_\mu$ cartoon + textures decomposition algorithm.

```
Initialization: $u_0 = v_0 = 0$
while 'max $(|u_{n+1} - u_n|, |v_{n+1} - v_n|) \leqslant \epsilon$ or a prescribed maximal number of iterations is not reached' do
    Compute $v_{n+1} = P_{E_\mu}(f - u_n) = f - u_n - WST(f - u_n, 2\mu)$
    Compute $u_{n+1} = f - v_{n+1} - P_{G_\lambda}(f - v_{n+1})$
end while
```

In [5], Aujol and Chambolle proposed a numerical algorithm that uses the space E. As previously, they reformulated the model by using a dual-method approach, where E_μ is the subset of E, where $\forall f \in E_\mu$, $\|f\|_E \leqslant \mu$ and $B^*(f)$ is the characteristic function over E_1. Then the equivalent functional to minimize is given by Eq. (3.16)

$$F^{AC}_{\lambda,\mu}(u,v) = J(u) + B^*\left(\frac{v}{\mu}\right) + (2\lambda)^{-1}\|f-u-v\|^2_{L^2}. \tag{3.16}$$

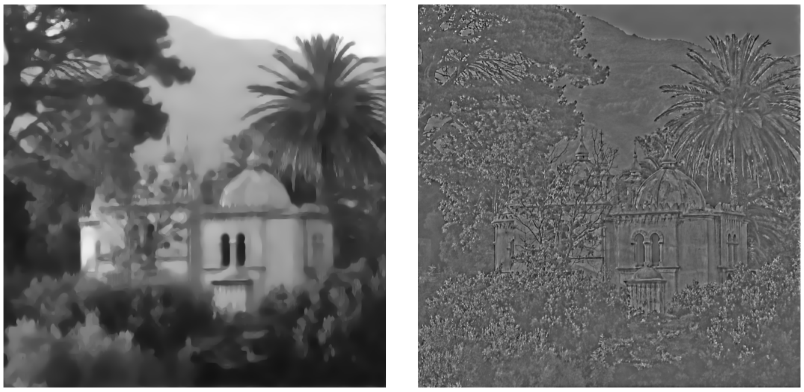
The existence of a projector on this space, denoted P_{E_μ}, was proven in [19] and is defined by

$$P_{E_\mu}(f) = f - WST(f, 2\mu), \tag{3.17}$$

where $WST(f, 2\mu)$ is the reconstructed image after a wavelet soft thresholding of f with a threshold set at 2μ. The corresponding algorithm is given in Algorithm 4.
Results obtained from this model are presented in Figures 3.4 and 3.5.

Structures Textures

Figure 3.4: BV-E_μ structures + textures image decomposition of Barbara image.



Structures Textures

Figure 3.5: BV-E_μ structures + textures image decomposition of Building image.

This algorithm also works well, but its main drawback is that the texture component captures some structures informations (like the legs of the table in the Barbara image; see Figure 3.4). This behavior arises because E is much bigger than G, in particular E contains functions that are not only textures.

3.1.5 Other models

Some other models were proposed in the literature trying different spaces to replace BV or G. In [31], the authors replace G by the Sobolev space H^{-1} while the authors of [5, 6] propose to replace BV by the smaller Besov space $B^1_{1,1}$, or G by some Hilbert spaces, which give the possibility to extract textures with a certain directionality. In [28], the Besov space $\dot{B}^1_{1,\infty}$ is used instead of BV (the norms over these two spaces are equivalent) with the L^2 norm for the v part. Duals of John and Nirenberg spaces $\operatorname{div}(BMO)$, $B\dot{M}O^{-\alpha}$, and Sobolev spaces $\dot{W}^{-\alpha,p}$ to model the texture component are studied in [29, 33].

3.2 Structures + textures + noise decomposition

The previous algorithms yield good results but are of limited interest for noisy images (we add a Gaussian noise with $\sigma = 20$ on each test image of Figure 3.1; the corresponding noisy test images can be viewed in Figure 3.6).

Indeed, noise can be viewed as a very highly oscillating function (this means that noise can be viewed as lying in G). Therefore, the algorithms extract the noise in texture component and consequently the textures are corrupted by noise (see Figure 3.7 for example).

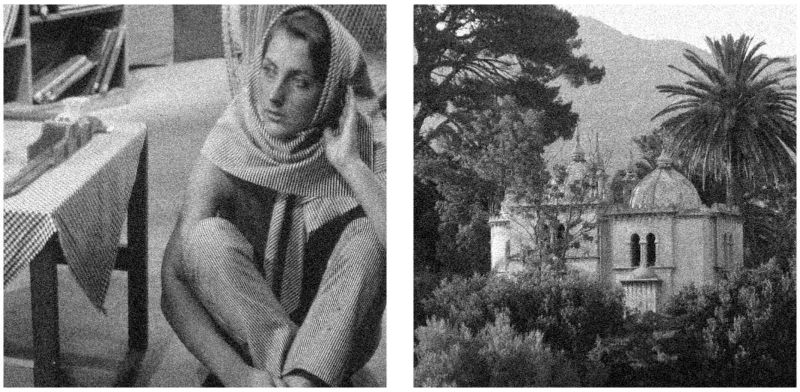


Figure 3.6: Original Barbara and Building images corrupted by Gaussian noise ($\sigma = 20$).

Structures Textures

Figure 3.7: *BV-G* structures + textures image decomposition of the noisy Barbara image.

In the following, we present some extensions of two-component models into three-component models, $f = u + v + w$, aiming to discriminate among structures (u), textures (v), and noise (w).

3.2.1 BV-G-G local adaptative model

In [22], we proposed an extension of the $BV - G$ model to decompose an image into three parts: structures (u), textures (v), and noise (w). As in the $u + v$ model, we considered that structures and textures are modeled by functions in BV and G respectively. Moreover, we assumed we have a zero mean Gaussian noise added to the image. In our approach we viewed noise as a specific high oscillating function. In virtue of Meyer's work [30], where it is shown that the more a function oscillates, the smaller its G-norm, we proposed to model w as a function in G and considered that its G-norm is much smaller than the texture norm ($\|v\|_G \gg \|w\|_G$). These assumptions can be resumed by

$$v \in G_{\mu_1}\ ,\ w \in G_{\mu_2}, \qquad \text{where} \quad \mu_1 \gg \mu_2. \tag{3.18}$$

To increase the efficiency, we proposed to add a local adaptability behavior to the algorithm following the idea proposed by Gilboa et al. in [20]. The authors investigated the ROF model given by Eq.(3.1) and proposed a modified version able to preserve textures during the denoising process. This approach consists to choose λ (the regularization parameter) as a function $\lambda(f)(x, y)$ which represents local properties of the image instead of a constant. In a cartoon-type region, the algorithm enforces the denoising process by increasing the value of λ while in a texture-type region, the algorithm decreases λ to attenuate the regularization in order to preserve texture details. Then $\lambda(f)(x, y)$ can be viewed as a smoothed partition between textured and untextured regions.

Based on this idea we propose to use the following functional to decompose an image into three parts:

$$F^{JG}_{\lambda,\mu_1,\mu_2}(u, v, w) = J(u) + J^*\left(\frac{v}{\mu_1}\right) + J^*\left(\frac{w}{\mu_2}\right) + (2\lambda)^{-1}\|f - u - \nu_1 v - \nu_2 w\|^2_{L^2}, \tag{3.19}$$

where functions ν_i represent the smoothed partition of textured and untextured regions (and play the role of λ in Gilboa's paper).
In our approach, we chose ν_1 and ν_2 as complementary, $\nu_2 = 1 - \nu_1 : \mathbb{R}^2 \to]0; 1[$. The choice of ν_1 and ν_2 is discussed after the following proposition, which characterizes the minimizers of $F^{JG}_{\lambda,\mu_1,\mu_2}(u, v, w)$.

Proposition 2 *Let $u \in BV$, $v \in G_{\mu_1}$, and $w \in G_{\mu_2}$ be the structure, texture, and noise parts, respectively, and f the original noisy image. Let the*

Algorithm 5 $BV-G-G$ cartoon + textures + noise decomposition algorithm.

```
Initialization: $u_0 = v_0 = w_0 = 0$
Compute $\nu_1$ and $\nu_2 = 1 - \nu_1$ from $f$
while 'Not converged' do
    Compute $w_{n+1} = P_{G_{\mu_2}}\left(\frac{f-u_n-\nu_1 v_n}{\nu_2+\kappa}\right)$, ($\kappa$ is a small value in order to prevent division by zero)
    Compute $v_{n+1} = P_{G_{\mu_1}}\left(\frac{f-u_n-\nu_2 w_{n+1}}{\nu_1+\kappa}\right)$
    Compute $u_{n+1} = f - \nu_1 v_{n+1} - \nu_2 w_{n+1} - P_{G_\lambda}(f - \nu_1 v_{n+1} - \nu_2 w_{n+1})$
end while
```

functions $(\nu_1(f)(.,.), \nu_2(f)(.,.))$ *be defined on* $\mathbb{R}^2 \to]0;1[$*, and assume that these functions could be considered as locally constant compared to the variation of* v *and* w*. Then the minimizers defined by*

$$(\tilde{u}, \tilde{v}, \tilde{w}) = \underset{(u,v,w)\in BV\times G_{\mu_1}\times G_{\mu_2}}{\arg} \min F^{JG}_{\lambda,\mu_1,\mu_2}(u, v, w), \tag{3.20}$$

are given by

$$\tilde{u} = f - \nu_1\tilde{v} - \nu_2\tilde{w} - P_{G_\lambda}(f - \nu_1\tilde{v} - \nu_2\tilde{w}), \tag{3.21}$$

$$\tilde{v} = P_{G_{\mu_1}}\left(\frac{f - \tilde{u} - \nu_2\tilde{w}}{\nu_1}\right), \tag{3.22}$$

$$\tilde{w} = P_{G_{\mu_2}}\left(\frac{f - \tilde{u} - \nu_1\tilde{v}}{\nu_2}\right), \tag{3.23}$$

The proof of this proposition can be found in [22]. As in the two-part BV-G decomposition model, we got an equivalent numerical scheme as depicted in Algorithm. 5.

Concerning the choice of the ν_i functions, we were inspired by the work in [20]. The authors chose to compute a local variance on the texture + noise part of the image obtained by the ROF model $(f - u)$. In our model, we use the same strategy but on the v component obtained by the two parts decomposition algorithm. This choice is guided by the fact that the additive Gaussian noise can be considered as orthogonal to textures. As a consequence, the variance of a textured region is larger than the variance of an untextured region.

In practice, we first compute the two-part decomposition of the image f. On the texture part, for all the pixels (i, j), we compute the local variance on a

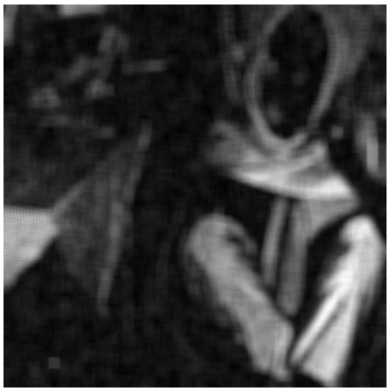


Figure 3.8: Texture partition ν_1 obtained by local variance computation.

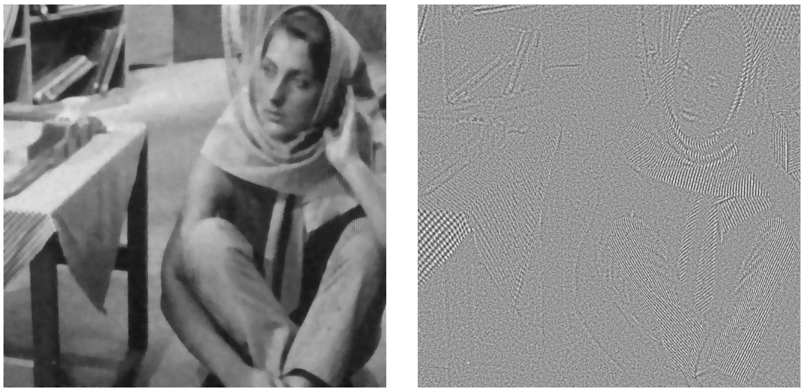
small window (odd size L) centered on (i, j). At the end, we normalize this variance image to obtain values in $]0; 1[$. All details about the computation of the ν_i functions can be found in [22]. Figure 3.8 shows an example from the noisy Barbara image. As expected, the variance is higher in the textured regions and lower in the others.
Figures 3.9 and 3.10 show the results of the $u+v+w$ decomposition obtained by the BV-G-G local adaptive model. This model can separate noise from textures. If we look more precisely, we can see that some residual noise remains in textures, and some textures are partially captured in the noise component. This is related to the choice of λ, μ_1, and μ_2 which fix the separability power of the algorithm.

3.2.2 Aujol-Chambolle BV-G-E model

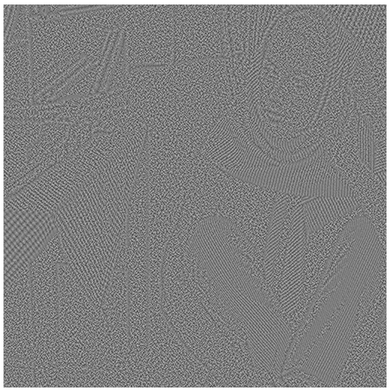
At the same time, in [5] Aujol et al. addressed the same structures + textures + noise decomposition problem. They proposed a model close to ours but with a main difference in that the noise is considered as a distribution and living in the Besov space $E = \dot{B}^{\infty}_{-1,\infty}$. Their corresponding functional is given by Eq. (3.24).

$$F^{AC2}_{\lambda,\mu,\delta}(u,v,w) = J(u)+J^*\left(\frac{v}{\mu}\right)+B^*\left(\frac{w}{\delta}\right)+(2\lambda)^{-1}\|f-u-v-w\|^2_{L^2}, \quad (3.24)$$

where $u \in BV$, $v \in G_\mu$, and $w \in E_\delta$ as defined previously. The numerical algorithm is given by Algorithm. 6.



Structures Textures



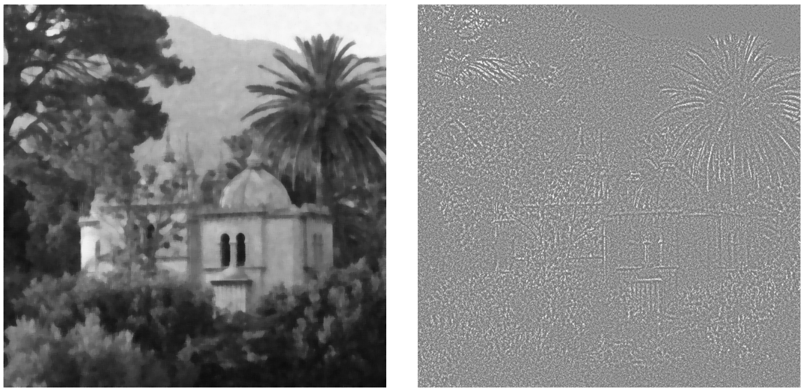
Noise

Figure 3.9: *BV*-*G*-*G* structures + textures + noise image decomposition of Barbara image.

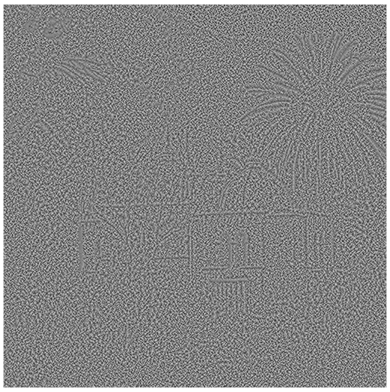
Algorithm 6 $BV - G - E$ cartoon + textures + noise decomposition algorithm.

```
Initialization: $u_0 = v_0 = w_0 = 0$
while 'Not converged' do
    Compute $w_{n+1} = f - u_n - v_n - WST(f - u_n - v_n, 2\delta)$
    Compute $v_{n+1} = P_{G_\mu}(f - u_n - w_{n+1})$
    Compute $u_{n+1} = f - v_{n+1} - w_{n+1} - P_{G_\lambda}(f - v_{n+1} - w_{n+1})$
end while
```

Results of this algorithm on our test images are shown in Figures 3.11 and 3.12, respectively. We can see that textures are better denoised by this model. This is a consequence of a better noise modeling by distributions



Structures Textures



Noise

Figure 3.10: BV-G-G structures + textures + noise image decomposition of Building image.

in the Besov space. But the residual texture is more important than the one given by our algorithm in the noise part. Another drawback appears in the structure part; the edges in the image are damaged because some important wavelet coefficients are removed. In my Ph.D., I tested to add the local adaptivity behavior of the BV-G-G model to the BV-G-E model. This modified version shows little improvements compared to the original model. We prefered to explore the replacement of wavelets by recent geometric multiresolution expansions like contourlets.

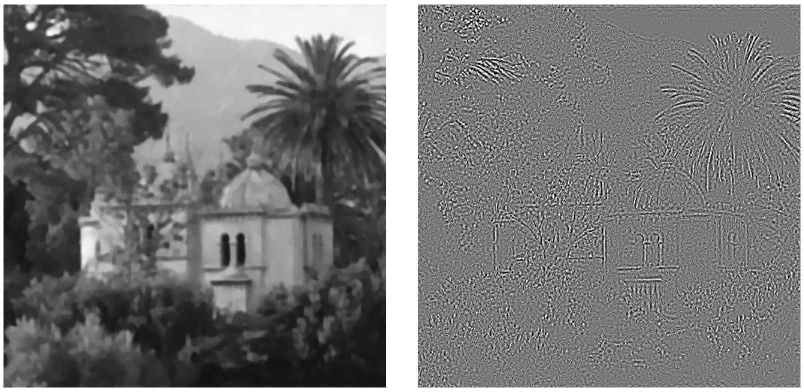
Structures Textures

Noise

Figure 3.11: *BV*-*G*-*E* structures + textures + noise image decomposition of Barbara image.

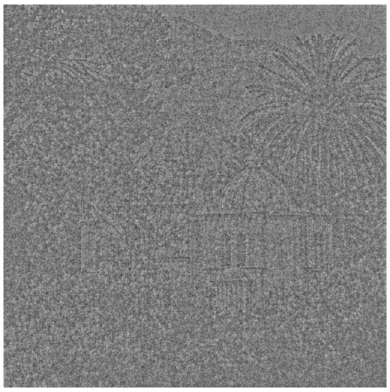
3.2.3 The BV-G-$\dot{C}o^{\infty}_{-1,\infty}$ decomposition model

Recent directional multiresolution tools, such as curvelets or contourlets, exhibit very good results in denoising. They also better reconstruct edges in an image. Then the idea to replace wavelets by curvelets or contourlets in Aujol et al. model naturally arose. We focused our choice on contourlets but in fact all directional multiresolution expansions can be used in the same way. This is equivalent to change the Besov space in the model described in the previous subsection by an homogeneous contourlet space $\dot{C}o^{\infty}_{-1,\infty}$ (in fact, it is an anisotropic Besov space). Then, the equivalent functional providing the decomposition is given in Eq.(3.25).



Structures

Textures



Noise

Figure 3.12: *BV-G-E* structures + textures + noise image decomposition of Building image.

$$F^{Co}_{\lambda,\mu,\delta}(u,v,w) = J(u) + J^*\left(\frac{v}{\mu}\right) + J^*_{Co}\left(\frac{w}{\delta}\right) + (2\lambda)^{-1}\|f-u-v-w\|^2_{L^2}, \quad (3.25)$$

where $J^*_{Co}(f)$ is the characteristic function over the set Co_1 if we denote $Co_\delta = \left\{ f \in Co^{\infty}_{-1,\infty} / \|f\|_{Co^{\infty}_{-1,\infty}} \leqslant \delta \right\}$ (norms over contourlet spaces can be defined in the same way as Besov norms from the contourlet coefficients, see [23] for more details).

The corresponding numerical scheme is the same as in the *BV-G-E* algorithm, except that we replace the wavelet expansion by the contourlet expansion in the soft thresholding. It is depicted in Algorithm. 7 (*CST* stands

Algorithm 7 $BV - G - \dot{C}o^{\infty}_{-1,\infty}$ cartoon + textures + noise decomposition algorithm.

```
Initialization: $u_0 = v_0 = w_0 = 0$
while 'Not converged' do
    Compute $w_{n+1} = f - u_n - v_n - CST(f - u_n - v_n, 2\delta)$
    Compute $v_{n+1} = P_{G_\mu}(f - u_n - w_{n+1})$
    Compute $u_{n+1} = f - v_{n+1} - w_{n+1} - P_{G_\lambda}(f - v_{n+1} - w_{n+1})$
end while
```

for the Contourlet Soft Thresholding).

Figures 3.13 and 3.14 show the obtained results by replacing wavelets by contourlets. The advantage of using geometric frames is in that it better preserves the integrity of edges and oriented textures.

3.3 Theoretical analysis

In this section I will give some theoretical results we got which permit to characterize the solutions of the decomposition model. In order to do this analysis, we rewrite the decomposition model slightly differently than in the previous sections. Let us assume that we are given an image $f \in L^2(\mathbb{R}^2)$ and two positive parameters λ et μ. Then we seek to decompose f as a sum

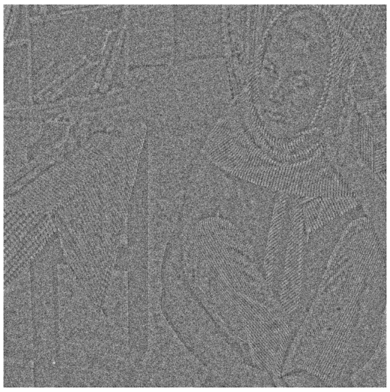
$$f = u + v + w \tag{3.26}$$

by minimizing the functional $E(u, v, w)$ defined by

$$\|u\|_{BV} + \lambda\|v\|^2_{L^2} + \mu\|w\|_G. \tag{3.27}$$

In this model, u is the cartoon part, w the textures and v the residual component. The ROF model corresponds to the case $\mu = +\infty$. As $BV \subset L^2$, we necessarily have $w \in L^2$. The existence of an optimal decomposition is given by the "Hilbert's direct method". Since BV is a dual space, from every bounded sequence $u_j \in BV$ we can extract a subsequence that converges, in the distributionnal sense, to $u \in BV$. The same argument can be used for L^2 and G. The uniqueness is not ensured except for the v part. In my Ph.D., we proved Theorem. 1 which aims to characterize minimizers of Eq. (3.27) accordingly to the choice of λ and μ.

Structures Textures



Noise

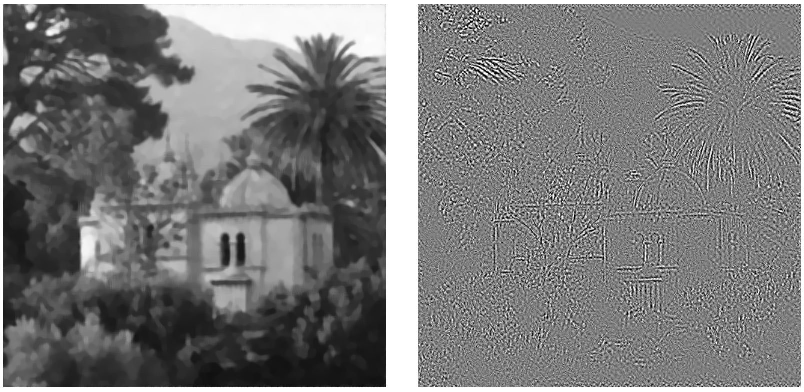
Figure 3.13: *BV-G-Co* structures + textures + noise image decomposition of Barbara image.

Theorem 1 *If* $\|f\|_G \leqslant \frac{1}{2\lambda}$ *and* $\|f\|_{BV} \leqslant \frac{\mu}{2\lambda}$, *then* $u = w = 0$ *and the optimal decomposition is* $f = 0 + f + 0$.

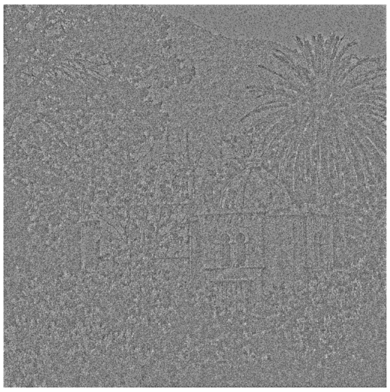
If $\|f\|_G \leqslant \frac{1}{2\lambda}$ *but* $\|f\|_{BV} > \frac{\mu}{2\lambda}$, *three cases appear for an optimal decomposition* $f = u + v + w$.

(1) $u = 0$, $\|v\|_{BV} = \frac{\mu}{2\lambda}$, $\|v\|_G < \frac{1}{2\lambda}$ *and* $\langle v, w\rangle = \frac{\mu}{2\lambda}\|w\|_G$,

(2) $w = 0$, $\|v\|_{BV} \leqslant \frac{\mu}{2\lambda}$, $\|v\|_G = \frac{1}{2\lambda}$ *and* $\langle u, v\rangle = \frac{1}{2\lambda}\|u\|_{BV}$ *and finally,*



Structures Textures



Noise

Figure 3.14: *BV-G-Co* structures + textures + noise image decomposition of Building image.

(3) $\|v\|_{BV} = \frac{\mu}{2\lambda}$, $\|v\|_G = \frac{1}{2\lambda}$, $\langle u, v\rangle = \frac{1}{2\lambda}\|u\|_{BV}$ *and* $\langle v, w\rangle = \frac{\mu}{2\lambda}\|w\|_G$.

Conversely, all triplet (u, v, w) *which fulfills (1), or (2), or (3) is optimal for* $f = u + v + w$ *and their corresponding values of* λ *and* μ.

The proof and more details about this theorem are available in [25].

An interesting application of this theorem is the following: let us assume that we deal with long and thin objects in an image. This kind of object can be modeled by

$$f(x_1,x_2)=1 \qquad \text{if} \qquad 0\leqslant x_1\leqslant L \quad \text{and} \quad 0\leqslant x_2\leqslant \epsilon \tag{3.28}$$

where $L\gg 1$ and $0<\epsilon\ll 1$. Then $\|f\|_G\leqslant\epsilon$ while $\|f\|_{BV}=2(L+\epsilon)$. It is easy to see that we are in the case 1 of Theorem. 1 if $\epsilon<\sqrt{\frac{\pi}{\lambda\mu}}$ and $\mu<4\lambda(L+\epsilon)$ i.e if L is rather large compared to μ.

By Theorem 1, we conclude that $u=0$, $\|v\|_{BV}=\frac{\mu}{2\lambda}$. Then we have $\|w\|_{BV}\geqslant\|f\|_{BV}-\|v\|_{BV}\geqslant 2(L+\epsilon)-\frac{\mu}{2\lambda}$ which is high. In this case, the w part is the most important one. This means that this kind of objects will be attracted in the w component.

This property was used in [21, 25] as a preprocessing stage in an aerial road networks detection. Indeed, road networks could be considered as long and thin objects in the image. The previous result teaches us that this kind of objects will be enhanced in the texture component (but u is not strictly equal to 0 and w does not contain only roads because the original image contains different kind of objects). So we decompose the image and then apply a detection algorithm on the w component. Figure 3.15 shows the constrast enhancement of roads in the w component on a zoomed portion of an aerial image. We clearly see that roads are the most visible objects in the texture component. Figure 3.16 shows the result we get, on a bigger image by this approach with a very simple detection algorithm (it is a combination of an *a contrario* segment detection algorithm and active contours, see [21] for more details) applied on the w component.
In the following, I present new (submitted in June 2012) results: two useful lemma which help to prove a final theorem.

Lemma 1 *Let E be a function space, $\|.\|_E$ its norm which we assumed to be translation invariant ($BV, L^p,\ldots$), let Δ_j be a Littlewood-Paley filter associated to function ψ . Then for all function $f\in E$ we have*

$$\|\Delta_j[f]\|_E\leqslant C\|f\|_E \qquad \textit{where} \quad C=\|\psi\|_{L^1} \tag{3.29}$$

This lemma only said that the energy, in the $E-$norm sense, of the filtered version of a function f is lower than the energy of f itself.

The next lemma gives an equivalency between the BV and L^1 norms of functions of compactly supported spectrum.

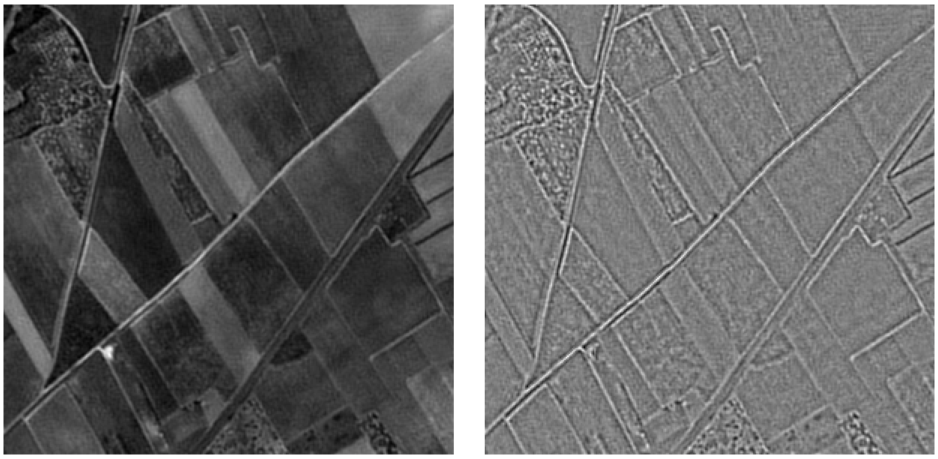


Figure 3.15: Example of a portion of an aerial image: original image on left, w component on right which lets appear enhanced roads.

Lemma 2 *If* $\hat{f}$ *is compactly supported by* $R \leqslant |\xi| \leqslant 3R$ *with* $R \gg 1$ *then* $\|f\|_{BV} \approx R\|f\|_{L^1}$.

The next theorem shows that it is possible to extract a specific texture from an image by properly choosing λ, μ and the right filter to apply on the texture component. Let us consider an image $f(x) = a(x) + b(x)\cos(\omega_1 x + \varphi_1) + c(x)\cos(\omega_2 x + \varphi_2)$ where a, b, c are $\mathcal{C}^1$ compactly supported functions.

Theorem 2 *If* $f(x) = a(x) + b(x)\cos(\omega_1 x + \varphi_1) + c(x)\cos(\omega_2 x + \varphi_2)$ *and if we assume* $1 \leqslant \lambda \ll |\omega_1| \ll \frac{\mu}{\lambda} \ll |\omega_2|$ *then, for a given scale* j *and* Δ_j *its corresponding Littlewood-Paley filter, the decomposition* $f = u + v + w$ *verifies*

$$\|\Delta_j[w](x) - c(x)\cos(\omega_2 x + \varphi_2)\|_{L^1} \leqslant \epsilon = C\frac{\mu}{\lambda|\omega_2|}. \tag{3.30}$$

This theorem is interesting in the way that it proposes some directions to optimaly separate different textures by a combination of image decomposition and band-pass filtering. In an upcoming paper [24], based on this theorem, we proposed a multiscale texture separation (MTS) algorithm: We compute the cartoon + texture decomposition and finally apply the Littlewood-Paley operator to accurately extract the most oscillating textures. Let denote $w_j = \Delta_j[w]$. The lower oscillating counterpart is obtained by substracting w_{j+1} from the input image f_j. In the sequel we denote $f_{j+1} = f_j - w_{j+1}$, then a single scale texture separation block can be depicted as in Fig. 3.17.

This process can be iterated to reach a multiscale texture separation algorithm (at each scale we set $\mu_{j+1} = \mu_j/2$) as in Fig. 3.18. We can add a directional capacity to this algorithm by considering only portions of each dyadic shell. This idea was previously used in the construction of curvelet frames by Candès et al. [10, 11, 12]. Then we can build a directional filter bank following the construction of the curvelets and, instead of applying only one filter to the texture part, we apply this bank of filters to get the different textures corresponding to each direction. This process corresponds to modify the initial single texture separation block shown in Fig. 3.17 into a directional single texture separation block as depicted in Fig.3.19 where $\Delta_j^{\theta_l}$ represents the Littlewood-Paley filter associated with direction θ_l. The decomposition obtained from the input image presented in Fig. 3.20 is given on Fig. 3.21 and Fig. 3.22.

Bibliography

[1] F. Andrieu, C. Ballester, V. Caselles, and J.M. Mazon. Minimizing total variation flow. Differential and Integral Equations, 4(3):321–360, 2001.

[2] J.F. Aujol. Contribution à l'analyse de textures en traitement d'images par méthodes variationnelles et équations aux dérivées partielles. PhD thesis, University of Nice Sophia Antipolis, 2004.

[3] J.F Aujol, G. Aubert, L. Blanc-Féraud, and A. Chambolle. Image decomposition into a bounded variation component and an oscillating component. Journal of Mathematical Imaging and Vision, 22(1), 2005.

[4] J.F. Aujol, G. Aubert, L. Blanc-Féraud, and A. Chambolle. Decomposing an image. application to textured images and SAR images. Technical report, University of Nice Sophia-Antipolis, 2003.

[5] J.F. Aujol and A. Chambolle. Dual norms and image decomposition models. International Journal of Computer Vision, 63(1):85–104, 2005.

[6] J.F. Aujol and G. Gilboa. Constrained and SNR-based solutions for TV-Hilbert space image denoising. Journal of Mathematical Imaging and Vision, 26(1-2):217–237, 2006.

[7] J.F. Aujol, G. Gilboa, T. Chan, and S. Osher. Structure-texture image decomposition-modeling, algorithms and parameter selection. International Journal of Computer Vision, 67(1):111–136, 2006.

[8] M. Bertalmio, V. Caselles, B. Rougé, and A. Solé. Total variation image restoration with local constraints. Journal of Scientific Computing, 19(1-3):95–122, 2003.

[9] L. Blanc-Féraud. Approche variationelle pour la restauration et la décomposition d'images. In ORASIS'07 Journées des jeunes chercheurs en vision par ordinateur, Obernai, France, May 2007.

[10] E. Candès, L. Demanet, D.L. Donoho, and L. Ying. Fast discrete curvelet transforms. Multiscale Modeling and Simulation, 5:861–899, 2005.

[11] E. Candès and David Donoho. Continuous curvelet transform, part I: Resolution of the wavefront set. Applied Computational Harmonic Analysis, 19:162–197, 2003.

[12] E. Candès and David Donoho. Continuous curvelet transform, part II: Discretization and frames. Applied Computational Harmonic Analysis, 19:198–222, 2003.

[13] V. Caselles. Total variation based image denoising and restoration. International Congress of Mathematicians, III:1453–1472, 2006.

[14] V. Caselles, F. Catté, B. Coll, and F. Dibos. A geometric model for active contours in image processing. Num. Mathematik, 6:1–31, 1993.

[15] V. Caselles, B. Coll, and J.M. Morel. Scale space versus topographic map for natural images. In Scale-space'97, Utrecht, July 1997.

[16] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. International Journal of Computer Vision, 22(1):61–79, 1997.

[17] V. Caselles and L. Rudin. Multiscale total variation. In Second European Conference on Image Processing, Palma, Spain, September 1995.

[18] A. Chambolle. An algorithm for total variation minimization and applications. Journal of Mathematical Imaging and Vision, 20(1-2):89–97, mars 2004.

[19] A. Chambolle, R.A. DeVore, N. Lee, and B.J. Lucier. Nonlinear wavelet image processing: variational problems, compression and noise removal through wavelet shrinkage. IEEE Trans. on Image Processing, 7:319–335, 1998.

[20] G. Gilboa, Y. Zeevi, and N. Sochen. Texture preserving variational denoising using an adaptive fidelity term. pages 137–144. Proceedings of VLSM, 2003.

[21] J. Gilles. Choix d'un espace de représentation image adpaté à la détection de réseaux routiers. Traitement et Analyse de l'Information: Méthode et Application (TAIMA) Workshop, Hammamet, Tunisia, 2007.

[22] J. Gilles. Noisy image decomposition: a new structure, texture and noise model based on local adaptivity. Journal of Mathematical Imaging and Vision, 28(3):285–295, 2007.

[23] J. Gilles. Séparation en composantes structures, textures et bruit d'une image, apport de l'utilisation des contourlettes. In GRETSI Symposium on Signal and Image Processing, Dijon, France, September 2009.

[24] J. Gilles. Multiscale texture separation. Submited to SIAM Multiscale, Modeling and Simulation Journal, 2012.

[25] J. Gilles and Y. Meyer. Properties of BV-G structures + textures decomposition models. application to road detection in satellite images. IEEE Transaction in Image Processing, 19(11):2793–2800, 2010.

[26] J. Gilles and S. Osher. Bregman implementation of Meyer's G-norm for cartoon + textures decomposition. UCLA CAM Report 11-73, 2011.

[27] T. Goldstein and S. Osher. The Split Bregman method for L1 regularized problems. SIAM Journal on Imaging Sciences, 2(2):323–343, 2009. UCLA CAM Report 08-29.

[28] A. Haddad. Méthodes variationnelles en traitement d'image. PhD thesis, Ecole Normale Supérieure de Cachan, France, 2005.

[29] M.L. Triet J.B. Garnett, P. Jones and L. Vese. Modeling oscillatory components with the homogeneous spaces $\dot{BMO}^{-\alpha}$ and $\dot{W}^{-\alpha,p}$. Technical report, UCLA CAM Report 07-21, 2005.

[30] Y. Meyer. Oscillating patterns in image processing and in some nonlinear evolution equations. The Fifteenth Dean Jacquelines B. Lewis Memorial Lectures, American Mathematical Society, 2001.

[31] S. Osher, A. Sole, and L. Vese. Image decomposition and restoration using total variation minimization and the H^{-1} norm. Multiscale Modeling and Simulation: A SIAM Interdisciplinary Journal, 1(3):349–370, 2002.

[32] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. Physica D, 60:259–268, 1992.

[33] M.L. Triet and L. Vese. Image decomposition using total variation and $div(BMO)$. Multiscale Modeling and Simulation, SIAM Interdisciplinary Journal, 4(2):390–423, 2005.

[34] L. Vese and S. Osher. Modeling textures with total variation minimization and oscillating patterns in image processing. Journal of Scientific Computing, 19(1-3):553–572, 2002.

[35] J. Wright, Y. Ma, J. Mairal, G. Sapiro, T.S. Huang, and Y. Shuicheng. Sparse representation for computer vision and pattern recognition. Proceedings of the IEEE, 98(6):1031–1044, 2010.

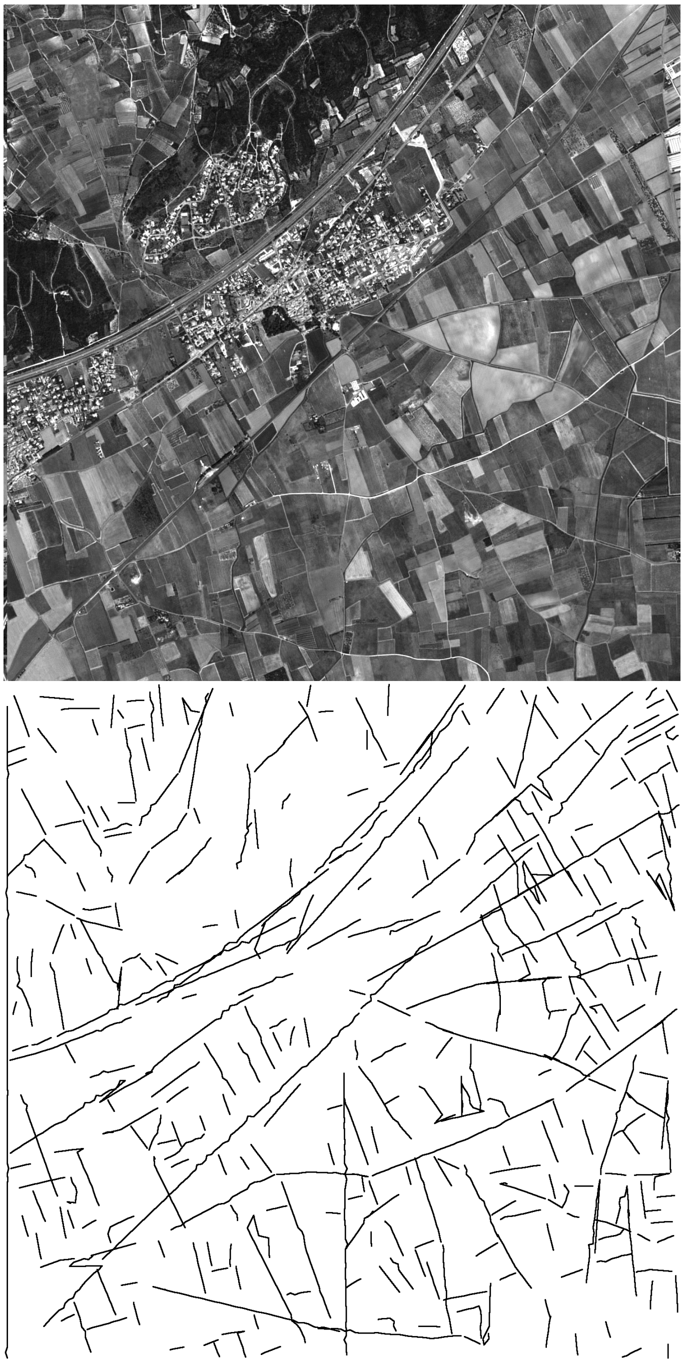


Figure 3.16: Example of road network detection: original image on top, detected roads on bottom.

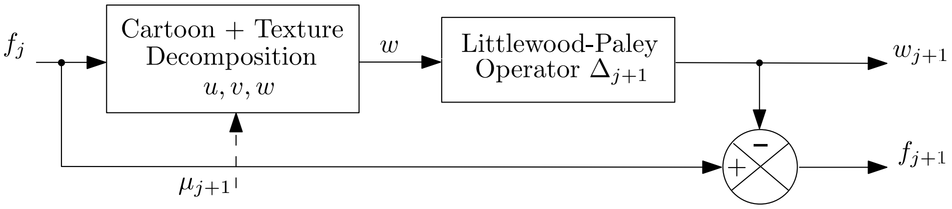


Figure 3.17: Single scale texture separation block

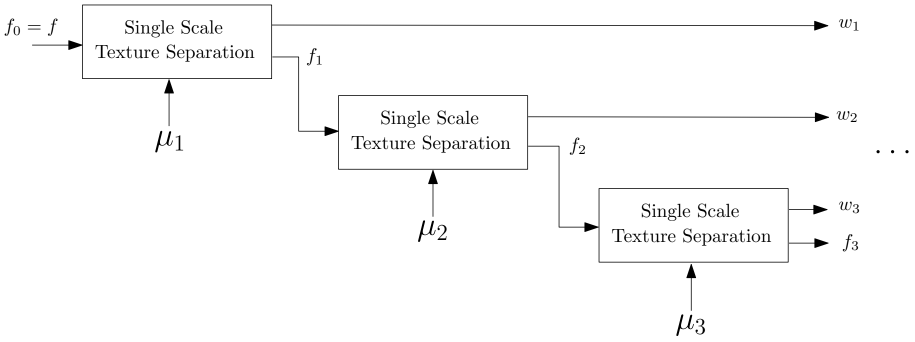


Figure 3.18: Multiscale scale texture separation

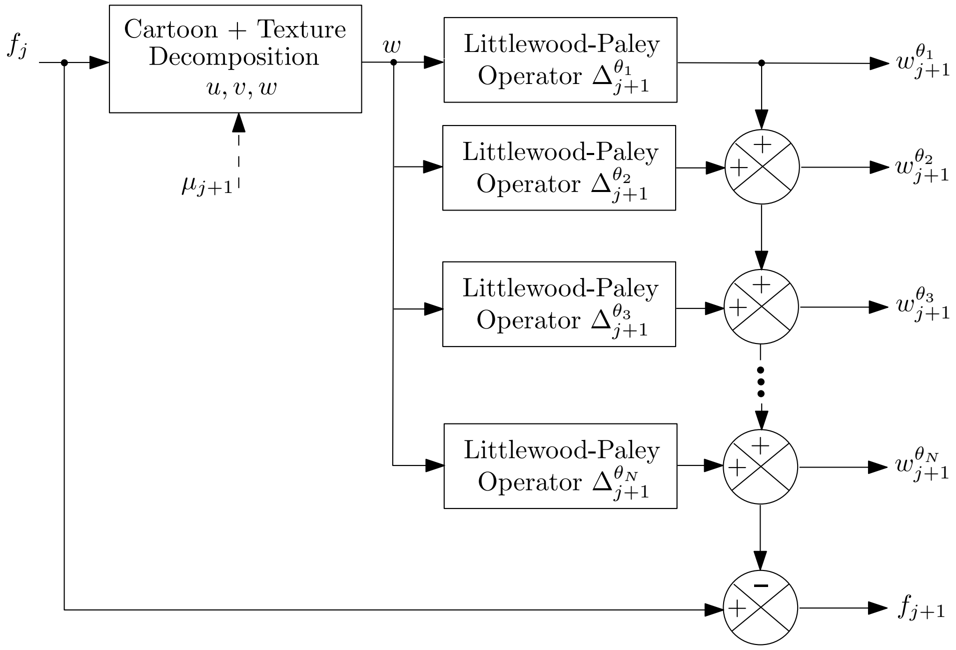


Figure 3.19: Directional Single scale texture separation block

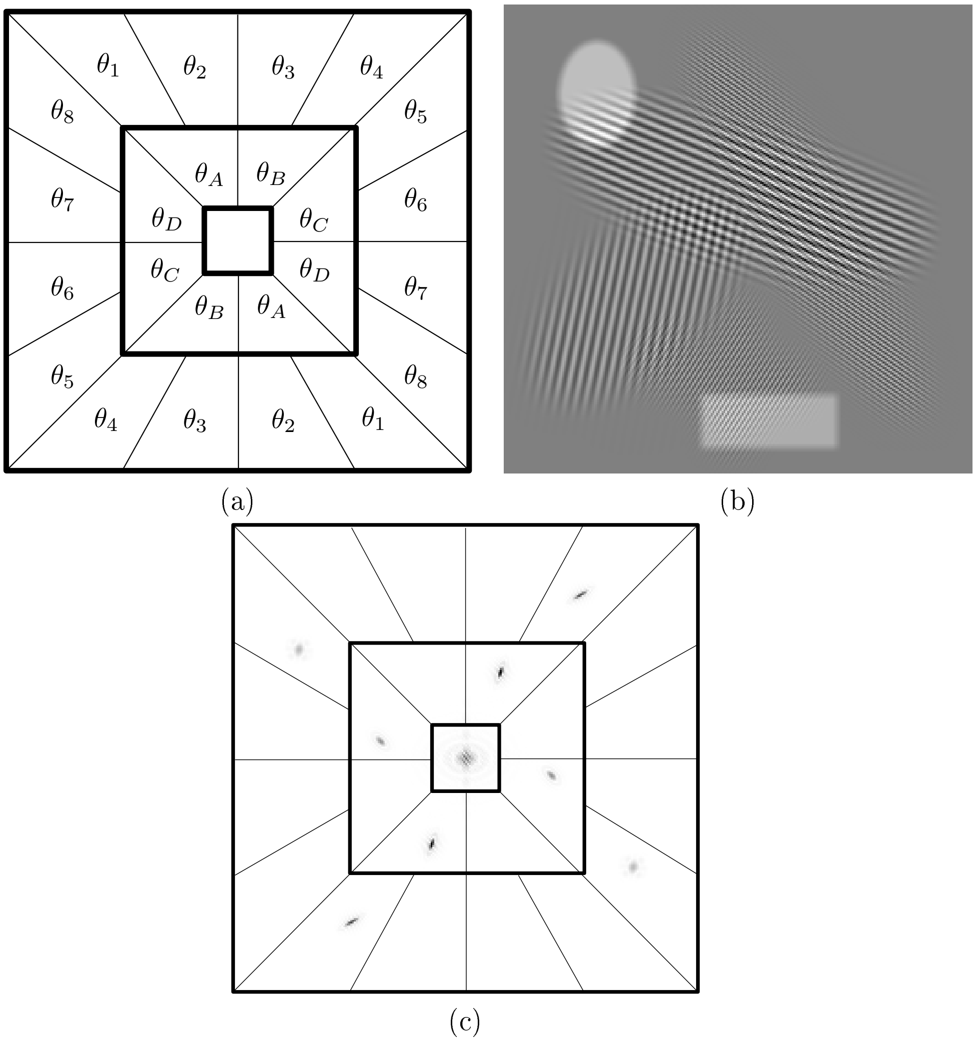
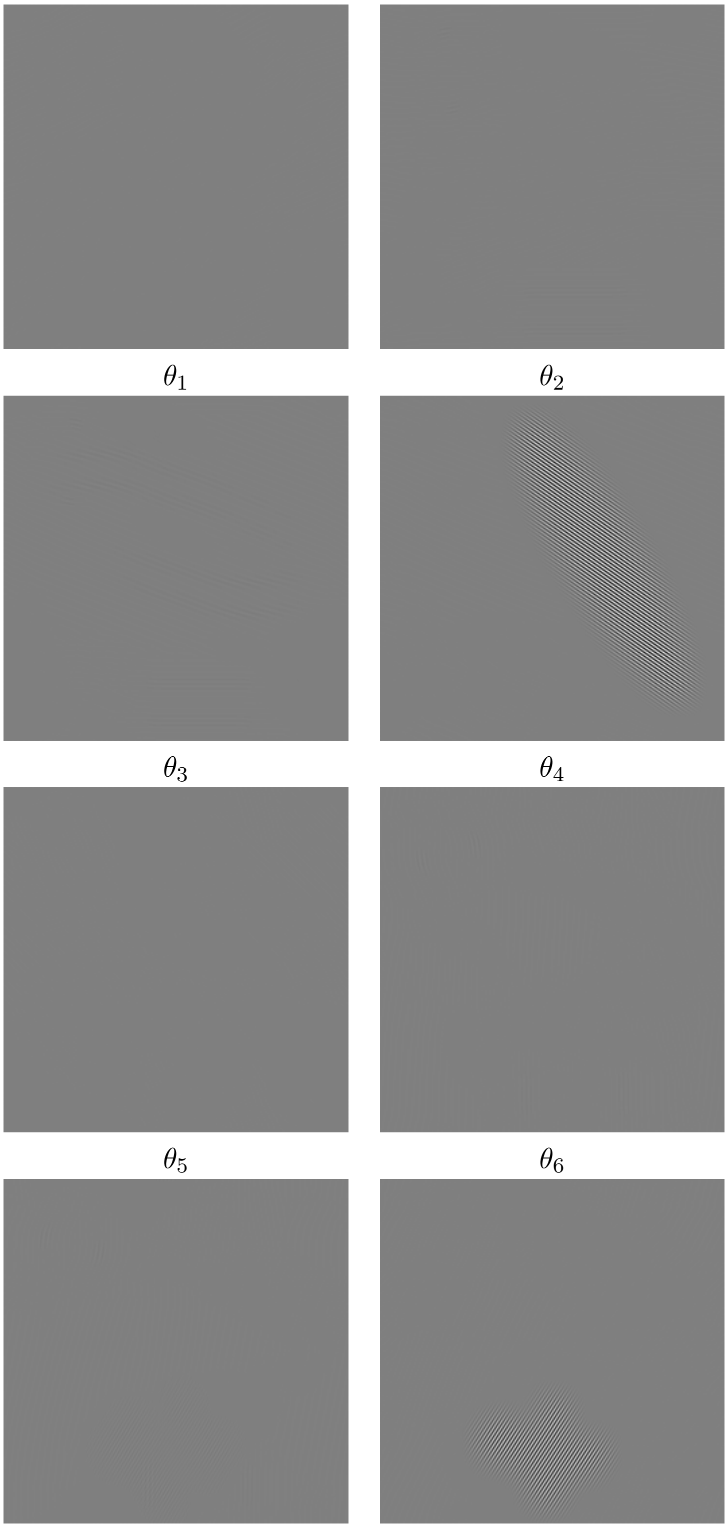


Figure 3.20: Partition of the Fourier domain corresponding to the directional filter bank (a), test image (b) and its Fourier transform (c).



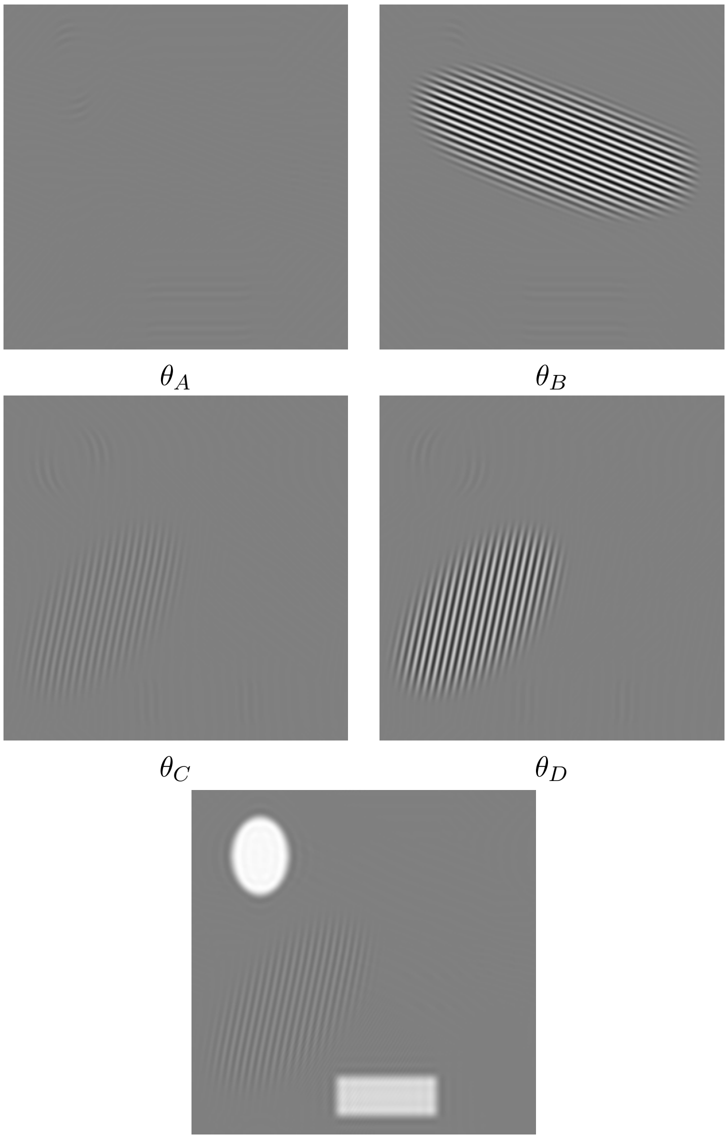
θ_1 θ_2

θ_3 θ_4

θ_5 θ_6

θ_7 θ_8

Figure 3.21: First scale directional texture components obtained from the synthetic image.



θ_A θ_B

θ_C θ_D

f_2

Figure 3.22: Second scale directional texture components and the low scale image f_2 obtained from the synthetic image.

Chapter 4

Imaging through turbulence

This chapter presents recent restoration work for imaging through turbulent atmosphere. I started to work on this topic when I was at the DGA as the french representative to a NATO Research Task Group about the modeling of active imaging systems. Such long range systems are dedicated to observation at several kilometers. The main limiting factor, in terms of resolution, is the atmospheric impact and notable turbulence. It results in two main visual defects: random geometric distortions and blur. Fig. 4.1 shows some examples obtained by a camera in real conditions. For each video we arbitrarily chose three frames to display here.
A pioneering work on turbulence modeling for mitigation algorithms was done by Frakes [10, 12]. The authors modeled the turbulence phenomenon using two operators:

$$f_i(x) = D_i(H(u(x))) + \text{noise} \tag{4.1}$$

where u is the static original scene we want to retrieve, f_i is the observed image at time i, H is a blurring kernel, and D_i is an operator which represents the geometric distortions caused by the turbulence at time i. Based on this model, in [14] we proposed a method to evaluate H^{-1} and D^{-1}. The H^{-1} operator is obtained by blind deconvolution, while the correction of the geometrical distortions D^{-1} is computed by an elastic registration algorithm based on diffeomorphic mappings (taking the inverse of D can be viewed as a stabilization problem). This approach gives good results, but it has two main drawbacks. Firstly, it is time consuming to perform the calculations due to the two iterative processes involved in the algorithm. Secondly, performances are sensitive to the choice of the parameters and this results in an unstable algorithm.

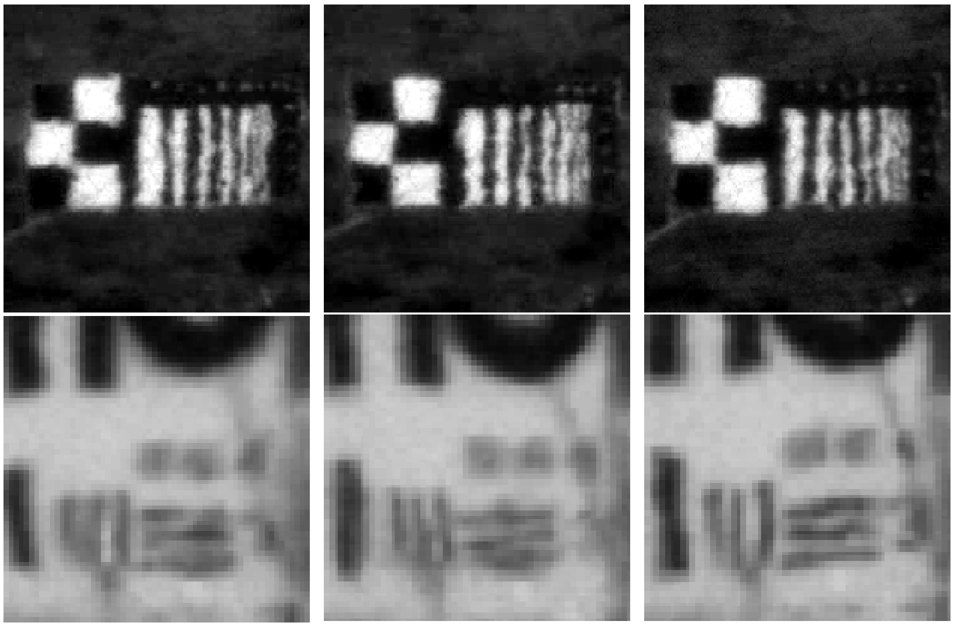


Figure 4.1: Sample images. Each row contains three arbitrary frames from different testing turbulence videos.

In the remainder of this chapter, I present new investigations on elastic stabilization and a deblurring approach based on physics electromagnetic wave propagation equations.

4.1 Elastic stabilization

4.1.1 Model and algorithm

In this section we only deal with the geometric distortions of the image and do not consider any deblurring aspects. If we denote the observed image sequence as $\{f_i\}_{i=1,\dots,N}$ and the true image that needs to be reconstructed as u, we model the deformations as

$$f_i(x) = u(\phi_i(x)) + \text{noise}, \quad \forall i \tag{4.2}$$

where ϕ_i corresponds to the geometric deformation on the i-th frame (note that the ϕ_i are the deformations between the true image and the observed

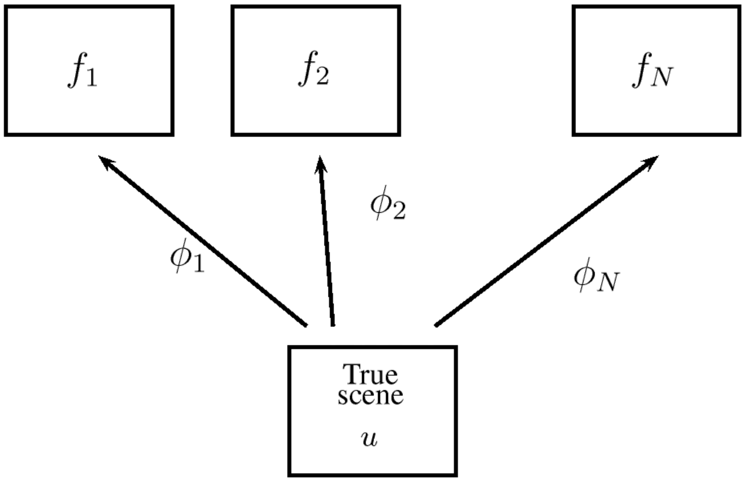


Figure 4.2: The model of deformation used in this paper.

frame i and not the continuous movement flow from frame to frame, see Fig. 4.2).

If we fix ϕ_i, $u(\cdot) \to u(\phi_i(\cdot))$ can be treated as a linear operator on u, so we can write the right hand side of (4.2) as

$$u(\phi_i(\cdot)) = (\Phi_i u)(\cdot) \tag{4.3}$$

where Φ_i is the linear operator corresponding to deformation ϕ_i, then (4.2) becomes

$$f_i = \Phi_i u + \text{noise}, \quad \forall i \tag{4.4}$$

which gives the fidelity equations in the model.

On the other hand, it is reasonable to assume that our image has certain regularization features. If we denote the regularization term of the image as $J(u)$, then we can formulate the stabilization problem as

$$\min_{u,\phi_i} J(u) + \frac{\lambda}{2} \sum_{i=1}^{N} \|\Phi_i u - f_i\|_2^2 \tag{4.5}$$

The regularization term of u, $J(u)$, has many different choices. Most notably, total variation based optimization [28] has been successful for edge preserving regularization. More recently, the nonlocal means regularization model [5, 13] has been introduced which modifies the intensity of a pixel by considering the nearby pixel values with similar patterns. Its basic assumption

is that a natural image contains repeating structures instead of repeating pixels; this makes the nonlocal approach capable of preserving important detailed features in an image. Based on these reasons, we decided to utilize the nonlocal regularization. A detailed introduction of this regularization method and its numerical implementation can be found in [13].
The kind of model (4.5) can be solved by the alternative optimization method, i.e. optimizing over different variables (u and ϕ_i) alternatively. In our model (4.5), if we have a good guess on u, then the optimal ϕ_i can be estimated by (4.2) via certain optical flow algorithms (e.g. methods developed in [3, 4, 29]). On the other hand, for fixed $\{\phi_i\}$ the model (4.5) can be efficiently solved as a constrained problem by using Bregman Iteration [27].
In our implementation we combine the updating step by optical flow for Φ_i into the Bregman updating loop, following [17]. The overall algorithm is resumed on Algorithm 8.

Algorithm 8 The Alternative Optimization Algorithm

```
Initialize: Start from some initial guess $u$. Let $\tilde{f}_i = f_i$.
while $\sum_i \|\Phi_i u - f_i\|^2$ not small enough do
    Estimate $\Phi_i$ which maps $u$ onto $f_i$ from (4.2) via optical flow scheme.
    while $\sum_i \|\Phi_i u - \tilde{f}_i\|^2$ not converge do
        $v \leftarrow u - \delta \sum_i \Phi_i^\top (\Phi_i u - \tilde{f})$
        $u \leftarrow \arg\min_u J(u) + \frac{\lambda}{2\delta}\|u - v\|^2$
    end while
    $\tilde{f}_i \leftarrow \tilde{f}_i + f_i - \Phi_i u$.
end while
```

The initial guess for u is chosen as the temporal average of the frames; Fig. 4.3 shows an example. We can see that the average of the frames is very blurry but gives a good initial guess of the rough shape of the object. In our numerical experiments we use less than 100 frames and generally we can obtain satisfactory results with only 10 frames.

There are two parameters in our algorithm: δ and λ; δ is the step size for the gradient descent of the fidelity term. As shown in [9], the step size should

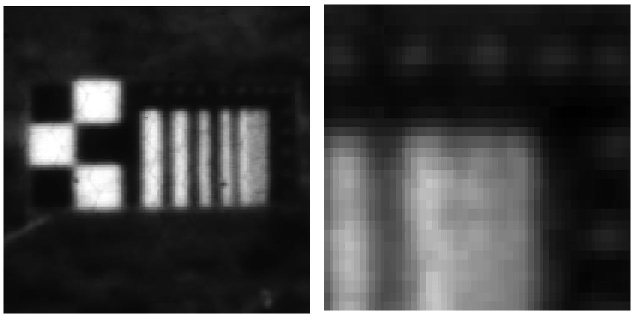


Figure 4.3: The average of the frames of one example video and its magnification of the top right corner.

be chosen such that $u \to u - \delta \sum_i \Phi_i^\top(\Phi_i u - \tilde{f})$ is contractive to achieve the convergence. The parameter λ is not a crucial factor and numerical results also indicate that the algorithm is not sensitive to the choice of λ. In practice we suggest a small initial λ to make the image regular enough at the very beginning, then increase λ gradually. This method is also used in other image reconstruction methods, e.g. [1]. More details about the numerical implementation are available in [26].

4.1.2 Experiments

In Figs. 4.4, 4.5 and 4.6 some frames taken from real videos are shown, as well as magnified details. Our reconstruction results are shown in the last column. Only five iterations are implemented in our algorithm to get this satisfactory result (especially in Fig. 4.6 where the medium sized letters on the board are much more readable in the processed image than in the original frames).

In order to evaluate the influence of the optical flow algorithm, we tried two different methods [3, 4]. The Black-Anandan approach gives accurate optical flow estimation but is time consuming while the classical Lucas-Kanade optical flow algorithm is faster. Various experiments show that the obtained results are very similar with both schemes and then our restoration algorithm is not sensitive to the choice of the optical flow method.

We also studied the impact in the choice of $J(u)$ by experimenting other regularization terms like usual TV and sparsity in some frame expansions (like framelets or curvelets). The different tests show very small differences between the use of these regularizers. NLTV has real interest in the case of textured images, otherwise classic (and faster to compute) $J(u)$ can be used.

Figure. 4.7 shows a comparison between results obtained using our method and two other state of the art methods: the algorithm based on PCA [23] and the algorithm using the Lucky-Region Fusion [2]. In these tests, we deal with short exposure sequences; it is not surprising that the PCA method failed as in this case the assumption of a Gaussian kernel for the blur is not correct. The Lucky-Region Fusion approach gives some good results but our method shows a better geometry reconstruction (see, for example the high spatial frequencies in the second row, the separation between each bar is clearer in our results). The other advantage of our algorithm is that it requires only a few frames to get good results, compared to 100 frames which were used in the Lucky-Region Fusion.

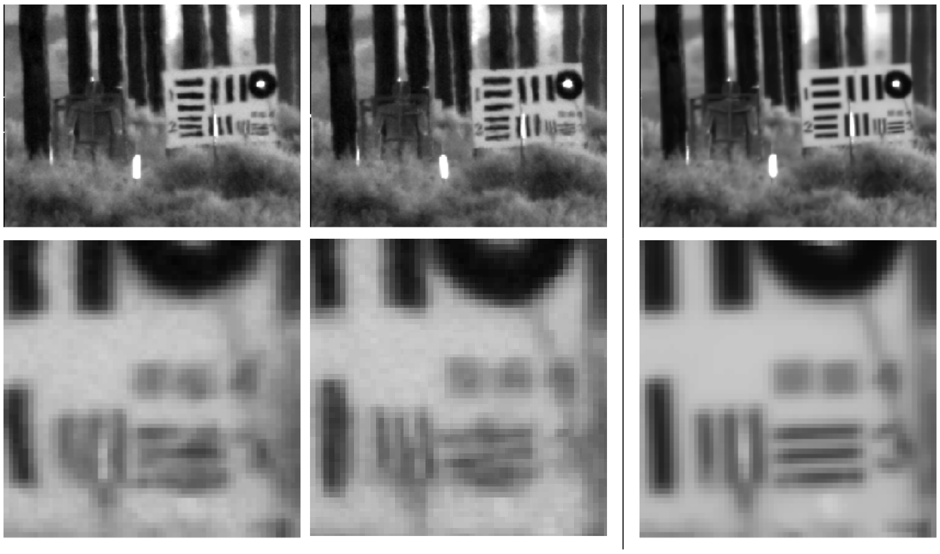


Figure 4.4: The first two columns are example input frames and magnifications of the right part of the frames. The last column shows our reconstructed result.

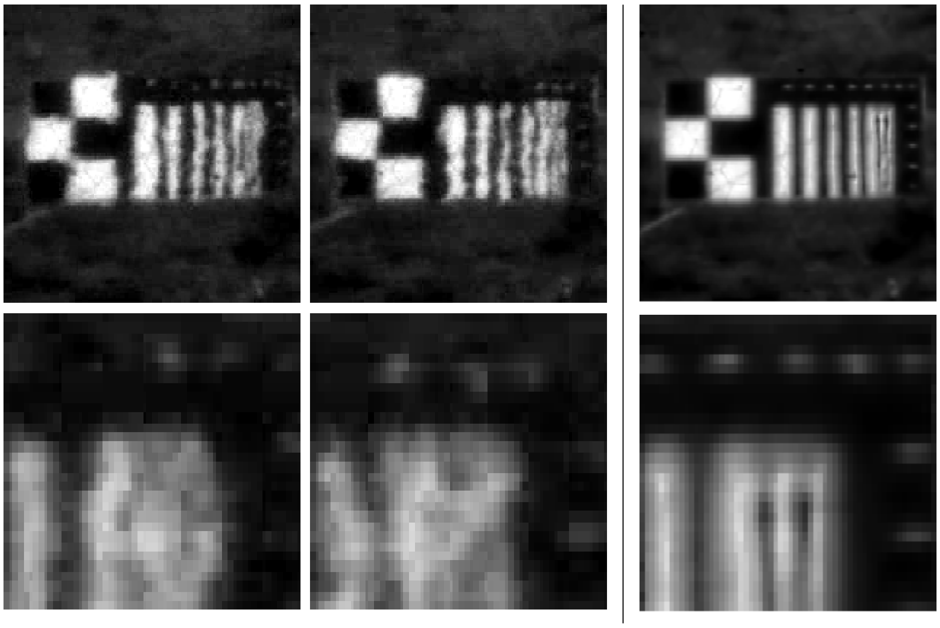


Figure 4.5: The first three rows are example input frames and the magnification of the top right part of the frames. The last row shows our reconstructed result.

Figure 4.6: The first two panels are example frames. The last panel is our reconstructed result. Only 10 frames are used in this example.

4.2 Atmospheric deblurring

The preceeding section presented the algorithm we built to correct the geometric distortions, e.g. inversing operator D_i in model (4.1). In this section

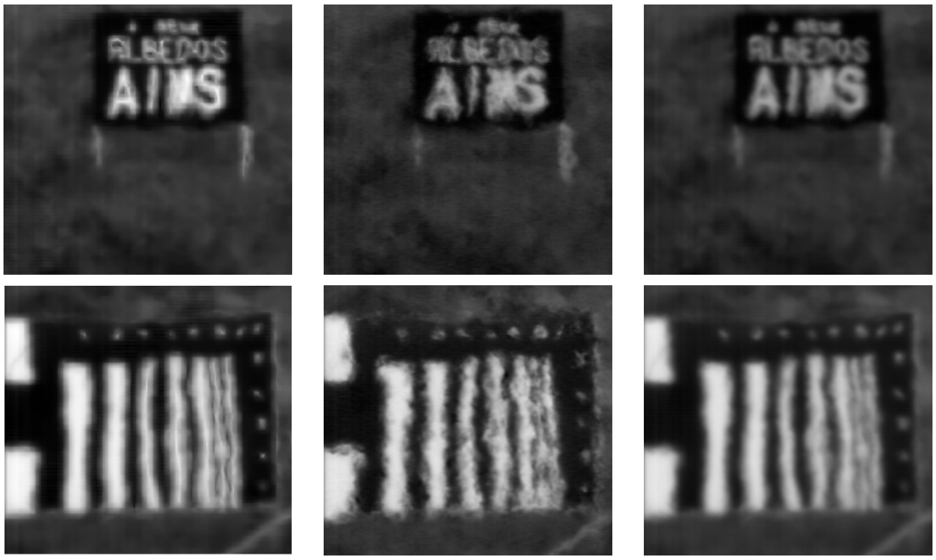


Figure 4.7: Results obtained with different algorithms: our method on left, the PCA based approach in the middle and the Lucky-Region Fusion on the right.

we address the problem of inversing H by an original deconvolution approach [15]. The originality of our approach comes from the use of an atmospheric kernel which models the blur impact of atmospheric propagation. Next we define this kernel formulation and then use it in a framelet based deconvolution algorithm.

4.2.1 Fried kernel

In order to deconvolve an image, two methods are mainly proposed in the literature. The first one is to assume that the blur kernel is unknown and to use some blind deconvolution algorithm but in the general case it is a difficult problem and the community remains very active on this approach. The second one is to find some analytic expressions which model the blurring effect of the atmosphere. Surprisingly, this model exists since … 1966! David Fried (well known in the optics community) proposed in [11] an analytical formulation to model the Modulation Transfer Function (MTF) of the atmosphere. His work was recently revisited by D. Tofsted in [30] and some experiments based on field trials showed the effectiveness of this MTF to model real phenomena [6].

Based on [11, 30], the Fried kernel can be viewed as the product of two terms. One effect, M_0, corresponds to the imaging system plus atmosphere MTFs, when turbulence is negligible. The second, M_{SA}, also called the short-term exposure MTF, models the impact (in term of blur) of phase-tilt due to turbulence. Denoting by ω the spatial frequency (in 2D we consider an isotropic kernel and ω is the frequency modulus), $M_0(\omega)$ can be expressed by

$$M_0(\omega) = \begin{cases} \frac{2}{\pi}\left(\arccos(\omega) - \omega\sqrt{1-\omega^2}\right) & \omega < 1 \\ 0 \qquad \omega > 1 \end{cases} \tag{4.6}$$

and $M_{SA}(\omega)$ is given by

$$M_{SA}(\omega) = \exp\left\{-(2.1X)^{5/3}(\omega^{5/3} - V(Q,X)\omega^2)\right\} \tag{4.7}$$

If we denote

- D: system entrance pupil diameter (we recall from geometrical optics rules that $D = f/N$ where f is the focal length and N the optics F-number),
- L: path length (distance from the sensor to the observed scene),
- λ: wavelength,
- C_n^2: refractive index structure representing turbulence magnitude of the atmosphere [31],

we can define the following quantities: $k = \frac{2\pi}{\lambda}$, the coherence diameter r_0 defined by $r_0 = 2.1\rho_0$ where $\rho_0 = 1.437(k^2LC_n^2)^{-3/5}$ (coherence length), $P = \sqrt{\lambda L}$, $Q = \frac{D}{P}$, $X = \frac{D}{r_0}$.

Finally, the quantity $V(Q,X)$ is defined as follows

$$V(Q,X) = A + \frac{B}{10}\exp\left\{-\frac{(x+1)^3}{3.5}\right\} \tag{4.8}$$

where $x = \log_{10}(X)$, $q = \log_2(Q)$ and

$$A = \begin{cases} 0.840 + 0.116\Sigma_{qa} & \text{with} \quad qa = 1.35(q+1.50) \\ & \qquad \text{if } q < -1.50 \\ 0.840 + 0.280\Sigma_{qc} & \text{with} \quad qc = 0.51(q+1.50) \\ & \qquad \text{if } q \geqslant -1.50 \end{cases} \tag{4.9}$$

and $\Sigma_q = \frac{e^q-1}{e^q+1}$. The coefficient B is defined by

$$B = 0.805 + 0.265\Sigma_{qb} \quad \text{with} \quad qb = 1.45(q - 0.15) \tag{4.10}$$

Finally, in the Fourier domain, Fried MTF ($M_F(\omega)$) is the product of $M_0(\omega)$ and $M_{SA}(\omega)$:

$$M_F(\omega) = M_0(\omega)M_{SA}(\omega) \tag{4.11}$$

Practically, this kernel only depends on four parameters: D, L, λ and C_n^2. The three first clearly depend on the acquisition system and the observed scene. The last parameter, C_n^2 represents the turbulence magnitude and, according to measurement, [31], is generally in the range $10^{-16}m^{-2/3}$ to $10^{-12}m^{-2/3}$ corresponding respectively to weak and strong turbulence.

4.2.2 Nonblind Fried deconvolution

Equiped with the Fried kernel, we can use a nonblind deconvolution algorithm. In our work [15], we proposed to use a framelet based deconvolution method proposed by Cai et al. in [7, 8]. This algorithm aims to find the image $\tilde{u}$ which has a sparse representation in a framelet expansion. We denote $\mathcal{D}$ and $\mathcal{D}^T$ the framelet decomposition and framelet reconstruction operators, respectively (see [8] for more details). Following the framelet properties we have $\mathcal{D}^T\mathcal{D} = I$ (tight frame) where I stands for the identity. Denoting g the acquired blurred image, M_F the known Fried kernel (here we assume that all parameters are known), the nonblind Fried deconvolution is achieved by finding $\tilde{u}$ which minimizes the following functional

$$\tilde{u} = \arg\min \|\mathcal{D}u\|_1 + \frac{\mu}{2}\|M_F \star u - g\|_2^2 \tag{4.12}$$

To solve this minimization problem, we set $d = \mathcal{D}u$ (the framelet expansion of u) and we use the split Bregman iteration (see [16] for details):

$$\begin{cases} u^{k+1} = \arg\min \frac{\mu}{2}\|M_F \star u - g\|_2^2 + \frac{\eta}{2}\|d^k - \mathcal{D}u - b^k\|_2^2 \\ d^{k+1} = \arg\min \|d\|_1 + \frac{\eta}{2}\|d - \mathcal{D}u^{k+1} - b^k\|_2^2 \\ b^{k+1} = b^k + \mathcal{D}u^{k+1} - d^{k+1} \end{cases} \tag{4.13}$$

Updating for u^{k+1} is a classical L^2 minimization problem while it as been shown that solving for d^{k+1} is equivalent to use the shrinkage operator:

$$d^{k+1} = shrink(\mathcal{D}u^{k+1} + b^k, 1/\eta) = sign(u)\max(0, |u| - 1/\eta) \tag{4.14}$$

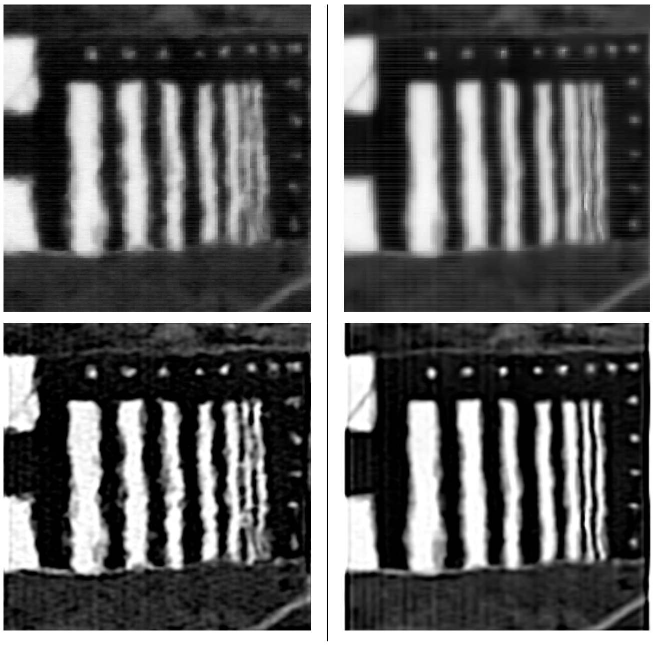


Figure 4.8: Nonblind Fried deconvolution on real barchart. The original images are on top, the deconvolved ones on bottom. The first column uses an original acquired image as input while a stabilized one is used on the second column.

In Fig. 4.8, 4.9, we test the nonblind Fried deconvolution on real images acquired during ground test field experiments made for turbulence characterization by the NATO group TG40 in 2005. The interesting fact is that the C_n^2 coefficient were measured during each image acquisition. Here, measured C_n^2 values are $1.51 \times 10^{-13} m^{-2/3}$ for Fig. 4.8 and $1.91 \times 10^{-13} m^{-2/3}$ for Fig. 4.9, the observed panels were at a distance $L = 1$km, the pupil diameter of the system was $D = 0.05$m and the system works in the visible spectrum ($\lambda \approx 700$nm). Framelet parameters are set to $\mu = 1000, \eta = 10$ and only two iterations were performed (experiments show that this choice gives the best visual improvements). The top row shows input images while the bottom one shows deconvolved images $\tilde{u}$. The first column corresponds to directly consider an observed image while the second column uses a stabilized image obtained by the method presented in the previous section. We can see great improvements, particularly in the edge sharpness and the readibility of the letter board.

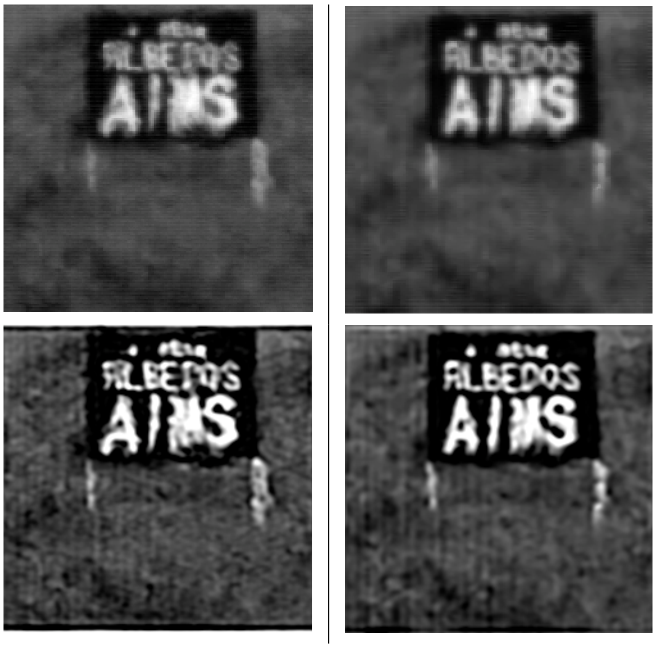


Figure 4.9: Nonblind Fried deconvolution on real letter board. The original images are on top, the deconvolved ones on bottom. The first column uses an original acquired image as input while a stabilized one is used on the second column.

In [15], we studied the influence of each parameter and we found that only large errors on the parameters reduce the restoration efficiency (details like textures are not well reconstructed). In practice, this parametric study gives us promising news because it appears that, even if the parameters are unknown, coarse approximations should be sufficient to get an improved deconvolved image.

4.2.3 Semiblind Fried deconvolution

If it seems reasonable to assume that parameters D,L and λ could be known in practice, the C_n^2 parameter is more difficult to handle as it depends on the atmospheric behavior and is not constant. In [15], we considered the case where C_n^2 is unknown and needs to be estimated from the acquired blurred image f. From the previous parametric analysis, we observe that if C_n^2 is under-estimated then the restored image $\tilde{u}$ remains blurred; if it is

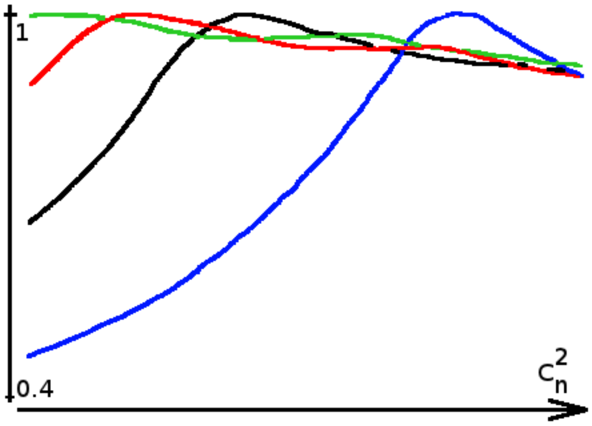


Figure 4.10: $TV(C_n^2)$ curves corresponding to different blurry cases.

over-estimated, $\tilde{u}$ is too much regularized: many details are removed and the edges are over sharpened. This behavior naturally recalls the total variation (TV) defined by $TV_u = \int_\Omega |\nabla u|$. Indeed in the case of a blurred image, TV is weak because the image gradients are over-smoothed while in the case of an over-regularized image (even if the gradients have more amplitude) the total amount of gradients decreases. Then it seems natural to expect that TV must be maximum (with respect to the choice of C_n^2) for a correct restoration. We easily verified this behavior by simulating a set of different blurred images by fixing some C_n^2 values. Then for each case, we computed the set of deconvolved images $\tilde{u}(C_n^2)$ for every C_n^2 choosen in an appropriate range linearly sampled which covers the typical physical range values. Then we can compute the normalized curves $TV_{\tilde{u}}(C_n^2)$ for each case. Fig. 4.10 shows the corresponding curves for cases $\tilde{C}_n^2 = \{1 \times 10^{-14}; 5 \times 10^{-14}; 10 \times 10^{-14}; 20 \times 10^{-14}\}$ $m^{-2/3}$. As expected, the position of $TV(C_n^2)$'s maximum corresponds to a good estimate of the real C_n^2.

In practice, it is too expensive to compute the whole $TV(C_n^2)$ curve (one deconvolution per C_n^2 value is needed). We proposed to speed up the estimation process by computing a limited, $N_{C_n^2}$, equidistant number of points on the curve and we used a polynomial approximation of the complete curve to finally deduce the optimal C_n^2. The semiblind Fried deconvolution consists to add this step to the previous nonblind algorithm and is resumed in Algorithm. 9. Figures 4.11 and 4.12 show the results we get with this semiblind Fried deconvolution. A comparison between measured and estimated C_n^2 is

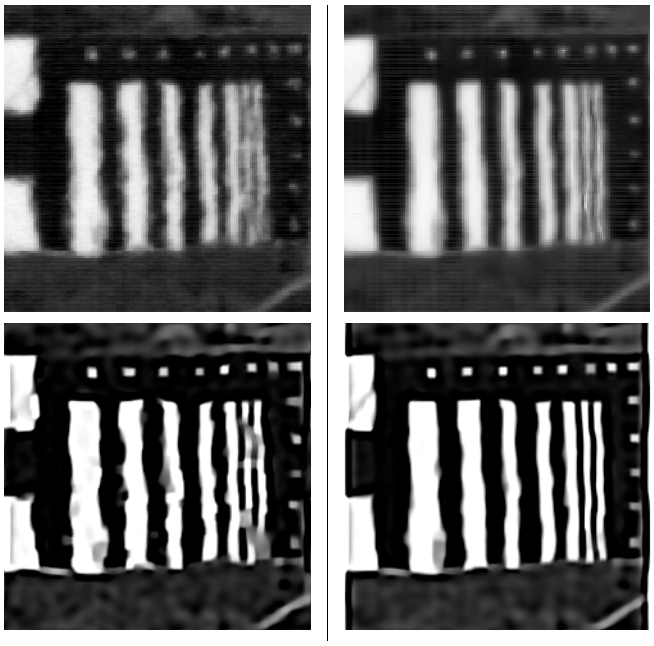


Figure 4.11: Semiblind Fried deconvolution on real barchart. The original images are on top, the deconvolved ones on bottom. The first column uses an original acquired image as input while a stabilized one is used on the second column.

provided in [15] and shows that the estimation process performs well.

Algorithm 9 Semiblind Fried deconvolution

```
- $D, L, \lambda$ are known. Fix the regularization parameters $\mu, \eta$.
- Compute the prescribed equidistant $N_{C_n^2}$ points (typically ten points) of the $TV(C_n^2)$ curve for a range $C_n^2 \in [C_{n,min}^2, C_{n,max}^2]$.
- Find the polynomial approximation by least square minimization.
- Estimate $\tilde{C}_n^2$ from the maximum of the polynomial approximation.
- Build the Fried kernel $M_F(\omega)$ corresponding to the known parameters and $\tilde{C}_n^2$.
- Use the framelet nonblind deconvolution algorithm with $M_F(\omega)$ to find the final restored image $\tilde{u}$.
```

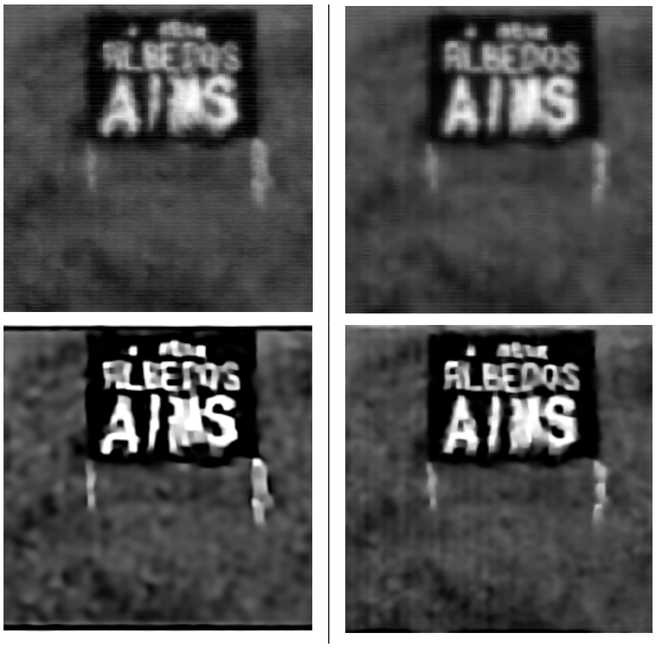


Figure 4.12: Semiblind Fried deconvolution on real letter board. The original images are on top, the deconvolved ones on bottom. The first column uses an original acquired image as input while a stabilized one is used on the second column.

Bibliography

[1] M. Almeida and L. Almeida. Blind and semi-blind deblurring of natural images. IEEE Trans on Image Processing, 19:36–52, 2010.

[2] M. Aubailly, M.A. Vorontsov, G.W. Carhat, and M.T. Valley. Automated video enhancement from a stream of atmospherically-distorted images: the lucky-region fusion approach. In Proceedings of SPIE, volume 7463, 2009.

[3] M.J. Black and P. Anandan. The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields. Comput Vision and Image Understanding, 63:75–104, 1996.

[4] J.-Y. Bouguet. Pyramidal implementation of the lucas kanade feature tracker description of the algorithm. Intel Corporation Microprocessor Research Labs, 2000.

[5] A. Buades, B. Coll, and J.M. Morel. A review of image denoising algorithms, with a new one. Multiscale Modeling and Simulation (SIAM interdisciplinary journal), 4(2):490–530, 2005.

[6] K. Buskila, S. Towito, E. Shmuel, R. Levi, N. Kopeika, K. Krapels, R.G. Driggers, R.H. Vollmerhausen, and C.E. Halford. Atmospheric modulation transfer function in the infrared. Applied Optics, 43(2):471–482, January 2004.

[7] J-F. Cai, S. Osher, and Z. Shen. Linearized bregman iterations for frame-based image deblurring. SIAM Journal on Imaging Sciences, 2(1):226–252, 2009.

[8] J-F. Cai, S. Osher, and Z. Shen. Split bregman methods and frame based image restoration. Multiscale Modeling and Simulation, 8(2):337–369, 2009.

[9] P.L. Combettes and V.R. Wajs. Signal recovery by proximal forward-backward splitting. SIAM Journal on Multiscale Modeling and Simulation, 4(4):1168–1200, 2005.

[10] D.H. Frakes, J.W. Monaco, and M.J.T. Smith. Suppression of atmospheric turbulence in video using an adaptive control grid interpolation approach. IEEE Conference Proceedings on Acoustics, Speech and Signal Processing, 3:1881–1884, 2001.

[11] D.L. Fried. Optical resolution through a randomly inhomogeneous medium for very long and very short exposures. Journal of The Optical Society of America, 56(10):1372–1379, October 1966.

[12] S. Gepshtein, A. Shtainman, and B. Fishbain. Restoration of atmospheric turbulent video containing real motion using rank filtering and elastic image registration. In Proceedings of EUSIPCO, Vienna, Austria, September 2004.

[13] G. Gilboa and S. Osher. Nonlocal operators with applications to image processing. Multiscale Model Sim, 7:1005–1028, 2008.

[14] J. Gilles, T. Dagobert, and C. De Franchis. Atmospheric turbulence restoration by diffeomorphic image registration and blind deconvolution. In Advanced Concepts for Intelligent Vision Systems (ACIVS), Juan Les Pins, France, October 2008.

[15] J. Gilles and S. Osher. Fried deconvolution. In SPIE Defense, Security and Sensing conference, Baltimore, US, April 2012.

[16] T. Goldstein and S. Osher. The Split Bregman method for L1 regularized problems. SIAM Journal on Imaging Sciences, 2(2):323–343, 2009. UCLA CAM Report 08-29.

[17] L. He, A. Marquina, and S. Osher. Blind deconvolution using TV regularization and Bregman iteration. International Journal of Imaging Systems and Technology, 15:74–83, 2005.

[18] A. Jalobeanu, L. Blanc-Féraud, and J. Zerubia. Restoration automatique d'images satellitaires par une méthode mcmc. In 17th GRETSI Colloquium, Vannes, France, September 1999.

[19] A. Jalobeanu, L. Blanc-Féraud, and J. Zerubia. Estimation of blur and noise parameters in remote sensing. In IEEE ICASSP'02, Orlando, USA, May 2002.

[20] A. Jalobeanu, L. Blanc-Féraud, and J. Zerubia. Satellite image deblurring using complex wavelet packets. International Journal of Computer Vision, 51(3):205–218, 2003.

[21] A. Jalobeanu, L. Blanc-Féraud, and J. Zerubia. An adaptive gaussian model for satellite image deconvolution. IEEE Trans. Image Processing, 13(4):613–621, 2004.

[22] M. Khoumri, L. Blanc-Féraud, and J. Zerubia. Unsupervised deconvolution of satellite images. In 5th IEEE International Conference on Image Processing, volume II, Chicago, USA, October 1998.

[23] D. Li, R.M. Mersereau, and S. Simske. Atmospheric turbulence-degraded image restoration using principal components analysis. IEEE Geoscience and Remote Sensing Letters, 4:340–344, July 2007.

[24] M. Mahmoudi and G. Sapiro. Fast image and video denoising via non-local means of similar neighborhoods. IEEE Signal Processing Letters, 12(12):839–842, 2005.

[25] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman. Non-local sparse models for image restoration. In IEEE 12th International Conference on Computer Vision, Kyoto, Japan, September - October 2009.

[26] Y. Mao and J. Gilles. Non rigid geometric distortions correction - application to atmospheric turbulence stabilization. Accepted in Inverse Problems and Imaging Journal, xx:xx, 2012.

[27] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin. An iterative regularization method for total variation-based image restoration. Multiscale Model Sim, 4:460–489, 2005.

[28] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. Physica D, 60:259–268, 1992.

[29] D. Sun, S. Roth, J. Lewis, and M.J. Black. Learning optical flow. In Computer Vision–ECCV, 2008.

[30] D.H. Tofsted. Reanalysis of turbulence effects on short-exposure passive imaging. Optical Engineering, 50(1):1–9, January 2011.

[31] A. Tunick, N. Tikhonov, M. Vorontsov, and G. Carhart. Characterization of optical turbulence C_n^2 data measured at the ARL A_LOT facility. Technical Report ARL-MR-625, Army Research Lab, September 2005.

Chapter 5

Infrared imaging

This chapter presents different investigations related to infrared (IR) imaging that began when I was at the DGA and continue today. My contributions concern two different topics: the non-uniformity correction in raw uncooled infrared cameras and the design of an evaluation methodology to assess performances of infrared Automatic Target Detection/Recognition and Tracking (ATD/ATR/ATT) algorithms.

5.1 Non-uniformity correction

Infrared (IR) imaging has proved to be a very efficient tool in a wide range of industry, medical, and military applications. IR cameras are used to measure temperatures, IR signatures, detection, etc. However, the performance of the imaging system is strongly affected by a random spatial response of each pixel sensor. Under the same illumination the readout of each sensor is different. This is due to mismatches in the fabrication process, among other issues [6]. Furthermore for uncooled cameras the problem is even worse because the sensor response non-uniformity is not stationary and slowly drifts in time. For this kind of camera a periodic update of the non-uniformity-correction (NUC) is required.

A good non-uniformity-correction is a success factor key for any post processing such as pattern recognition, image registration, etc. To get rid of the non-uniformity, two main kinds of methods have been developed:

- Calibration based techniques consist in an equalization of the response to an uniform black body source radiation. They are not convenient for real time applications, since they force to interrupt the image flow.

(This calibration is usually automatic, a shutter closing in front of the lens periodically).

- Scene based techniques, involving motion compensation or temporal accumulation. Such methods are complex and require certain observation conditions.

The perturbation model is

$$z_t(X) = f_{(X,t)}(u0_t(X)) + \eta_t(X)$$

where (X is the position and t is the time for the following) $z_t(X)$ is the observed value, $u0_t(X)$ is the ideal landscape, $f_{(X,t)}$ is the (unknown) transfer function of the sensor, and $\eta_t(X)$ is a random sensor Poisson noise. A non-uniformity correction algorithm aims at discovering $f_{(X,t)}$ or $u0_t(X)$ for each X and t. In [21], we propose a single frame based algorithm and show that motion compensation or accumulation algorithms are not necessary to achieve a good image quality. However, the proposed method can be viewed as a first step fostering the success of more sophisticated motion based correction algorithms. These are slow while the proposed algorithm is real time, and the obtained quality after a single frame correction is sufficient for many uses.

5.1.1 Anterior work

Numerous algorithms have been reported in the literature to remove the fixed-pattern-noise caused by the lack of a cross-column sensor equalization. Some algorithms estimate the sensor parameter and others attempt at recovering the true landscape. Most of them use a simplified (linear) model for the transfer function of the pixel sensor:

$$z_t(X) = u0_t(X)g_t(X) + b_t(X) + \eta_t(X),$$

where X, t, $z_t(X)$, $u0_t(X)$ are defined as previously; $g_t(X)$ and $b_t(X)$ are the gains/ offsets (in place of $f_{(X,t)}$) and $\eta_t(X)$ is the random noise (nevertheless, the true transfer function is non linear). These algorithms process a sequence of images $(z_t)_{t \in 1,\dots,N}$, not a single frame. The proposed algorithm uses no registration, hence we will focus on single frame algorithms. There are methods [14] suggesting to equalize the mean and the standard deviation of each pixel sensor by a linear transform. But this is only possible if there is a long camera sequence with enough motion where each sensor sweeps many different parts of the scene. A variant [16] adjusts the minimum and the maximum of the readout values, assuming the time histograms observed in each sensor to become equal over a long enough time sequence. This last method is called Constant Range [22]. As pointed out by several authors [13] the length N of the sequence is a crucial factor of success here. Two problems may arise:

- If N is too small then the estimation is wrong because all sensors have not seen the same landscape;
- If N is too large and because of the approximation bias and time drift of the sensor behavior, the previous images may appear as ghosts in the last ones. This undesirable effect is known as "ghost artifact".

There is a way to avoid the ghost artifacts [13], which consists in a reset of the estimation when the scene changes too much. But again, all this requires a long exposition time with a varying scene or a serious camera motion.

5.1.2 Midway infrared correction (MIRE)

The midway histogram equalization method

The midway algorithm was initially designed to correct for gain differences between cameras [7]. It permits to compare two images taken with different cameras more easily after their histograms have been equalized. This algorithm was later extended to flicker correction [8].
Consider two cumulative histograms H_1, H_2. The midway cumulative histogram of the corrected image is simply

$$Hmid^{-1} = \frac{H_1^{-1} + H_2^{-1}}{2},$$

and this average can be extended to an arbitrary number of images. Once the midway histogram is computed, a monotone contrast change is applied to image to specify H as its histogram. Thus, all images get the midway histogram, which is the best compromise between all histograms (see Fig 5.1).

Our approach for infrared images

Since many IR correction algorithms actually propose to equalize the temporal histograms of each pixel sensor, the midway is quite adapted to get a

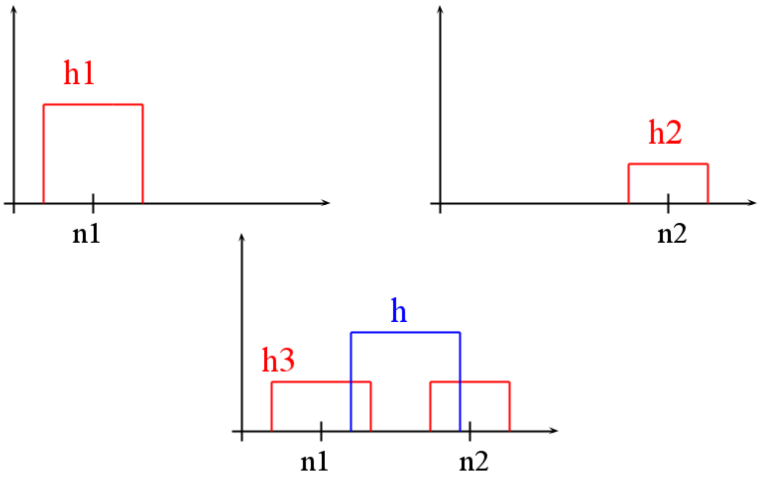


Figure 5.1: Two histograms h_1, h_2 (on top) and the corresponding midway histogram h (on bottom), compared to the direct histogram average h_3, which would create two modes and is therefore wrong.

better result than a simple equalization. We propose a still much simpler strategy. Equalization can be based on the fact that single columns (or lines, depending of the readout system) carry enough information by themselves for an equalization. The images being continuous, the difference between two adjacent columns is statistically small, implying that two neighboring histograms are nearly equal. So the proposition is to transport the histogram of each column (or line) to the midway of histograms of neighboring columns (resp. lines). In presence of strong fixed-pattern-noise (FPN) it will be useful to perform this *sliding midway method* over a little more than two columns, because the FPN is not independent in general. Assuming that the equalization is performed with columns, the proposed algorithm is resumed in Algorithm. 10.

The choice of the Gaussian standard deviation σ depends only on the camera, and not on the landscape. Thus, it can be fixed once and for ever for each camera. Since we work on separated images, the method is not affected by motions or changes of scene, which completely avoids "ghost artifacts" and any problem caused by the calibration parameters drifting over time. An automatic method to obtain a good σ is proposed hereafter.

Algorithm 10 Midway Infrared Equalization (MIRE).

```
Compute the cumulative histogram $H_i$ of each column $c_i$;
For each column $c_i$ compute a local midway histogram $\tilde{H}_{mid}(i)^{-1} := \sum_{j\in(-N,...,N)} \Phi(j)H_{i+j}^{-1}$ using Gaussian weights $\Phi = \Phi_\sigma$ with standard deviation $\sigma$ average;
Specify the histogram of the column $c_i$ onto this midway histogram $\tilde{H}_{mid}(i)$.
```

Automatically fitting the perfect parameter

The non-uniformity leads to an increased total-variation norm. Hence the smoothest image is also the one with little or no non-uniformity at all. So the simplest way to find the good parameter automatically is :

$$\sigma^* = argmin_\sigma ||I_\sigma||_{TV},$$

where I_σ is the image processed by MIRE with the parameter σ. The optimization could be done by a dichotomy on σ. We also proved [21] the following theorem which provides us the asymptotic behavior of this algorithm.

Theorem 3 *If h_i $i \in 1, ..., N$ are N histograms of the same landscape seen by N different columns of the sensor, and $H_{mid} = \sum_{j=1}^{N} \frac{H_j^{-1}}{N}$ then :*

$$||h_{mid} - h_{true}||_2 \leq \underset{i\in(1,...,N)}{max} (||h_i - h_{true}||_2)$$

Moreover if the h_i $\forall i \in (1, ..., N)$ from the N columns of the sensor are i.i.d. and centered on h_{true} then

$$||hmid - h_{true}||_2 \underset{N\to\infty}{\to} 0$$

Figures 5.2 and 5.3 show some results obtained on a simulated FPN and a real infrared image. More results are available in [21].

In a recent work [20], we proposed a locally adaptable version of this algorithm (named ADMIRE) where the parameter σ depends on the processed position in the image. One obtained result is given on figure. 5.4.

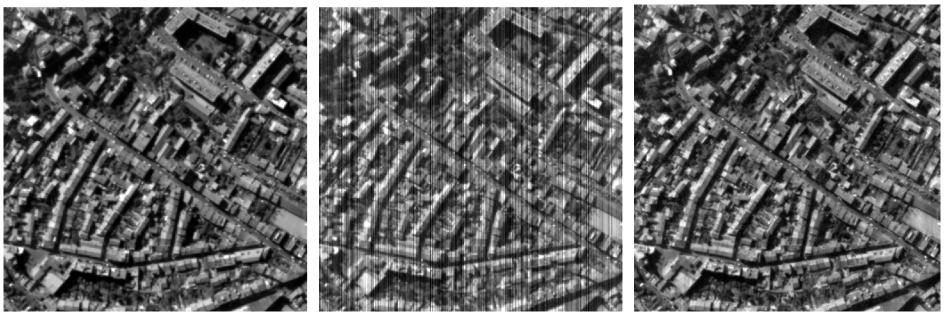


Figure 5.2: Image 1 : The groundtruth (left), the simulated FPN (middle) and MIRE restoration (right).

Figure 5.3: Real infrared image : RAW (left), MIRE restoration (right).

Figure 5.4: Adaptive MIRE result obtain on a real infrared image.

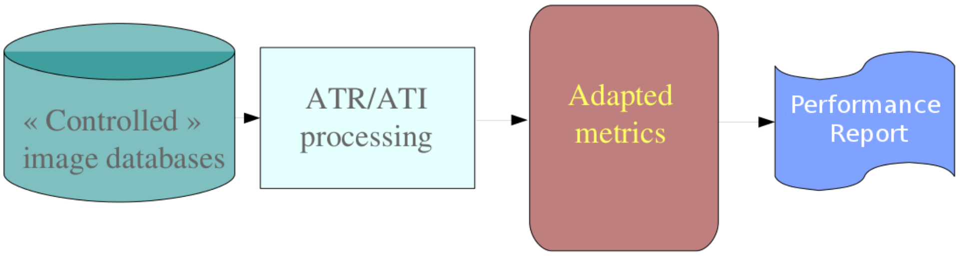


Figure 5.5: ATR evaluation principle

5.2 ATD/ATR/ATT performance evaluation

For an institution like DGA which conducts projects to develop new equipments for military forces, it is important to have the capabilities to evaluate the proposed industrial or academic solutions. A major topic in infrared is the development of Automatic Target Detection/Recognition and Tracking (ATD/ATR/ATT) algorithms. During the years I spent at the DGA, I led the scientific part of a project involving academic laboratories and industrial departments working on ATD/ATR/ATT algorithms. Beside driving the development of new algorithms, our main contribution was to propose a complete methodology to evaluate such algorithms. The basic principle for evaluation is depicted in Fig. 5.5. The evaluation chain has two important parts: the image databases (which must correspond to the wanted domain of applicability) and the metrics definitions (which must provide fair scores). Next I present our contributions concerning these two parts.

5.2.1 Image databases

Bigger the size of the test image set, more confident the evaluation is. If creating image datasets in the visible spectrum is easy because the needed equipment is widespread and cheap, in infrared imaging the needed ressources are more difficult to obtain and very expensive. Another binding in infrared images for Defense is the data confidentiality. A solution holds in the use of scenes simulators, but those remain expensive in terms of computational time and it is especially difficult to select the various parameters in order to sweep a maximum of operational scenarios in an exhaustive way. In order to dispose of these constraints, we proposed [10, 11] to generate hybrid databases by superimposition of targets and occultant in front of a background under constraint of image quality metrics. Entry parameters of these constraints are most effective to describe realistic operational scenarios. Moreover, one important aspect of the IR imagery is the intrinsic thermal variability of the target signature. We proposed an original method allowing to take into account of this variability during the scene generation. For that, we use real images of each vehicle acquired in their "extreme" operating processes: target at ambient temperature, target with all its potentially hot elements to the maximum of their temperatures. These signatures are then plated on a 3D model of the vehicle segmented under subelements of signature considered to have an homogeneous and independent thermal behavior. It is thus possible to selectively parameterize the temperature of these under-parts, to build alternatives of the signature of the same target. We used a 2D projection of this model under the wished visual angle for finally superimposing it under constraint and applying the wanted sensor effect to it.

Hybrid scenes generation

This scene generation is called "hybrid" because it consists to superimpose a real target signature, eventually with different kind of occultants (like trees, rocks,...), in a real background. The interest of the method is that it is possible to control the output image quality by some metrics [9, 23]. The used metrics are: local contrast RSS, "detectability" quantity Q_D, signal to clutter ratio SCR, occultation ratio R_x and internal target contrast K. These quantities are defined by

$$RSS = \frac{1}{\nu_k}\sqrt{(\mu_C - \mu_{F_1})^2 + \sigma_C^2} \tag{5.1}$$

$$Q_D = RSS.S_C \tag{5.2}$$

$$SCR = \frac{\nu_k RSS}{\sigma_F} \tag{5.3}$$

$$R_x = \frac{S_{\text{occluded target area}}}{S_{\text{full target area}}} \tag{5.4}$$

$$K = \frac{\mu_{F_1} - \mu_C}{\nu_k RSS} = \frac{\Delta\mu}{\nu_k RSS} \tag{5.5}$$

where C is the target, F_1 the local background over C and F_2 the remaining background (we denote the global background $F = F_1 \cup F_2$), see Figure 5.6. The quantities S_x, μ_x, σ_x are the surface, mean and standard deviation of area x where x is C, F_1 or F_2, respectively. The coefficient ν_k is the

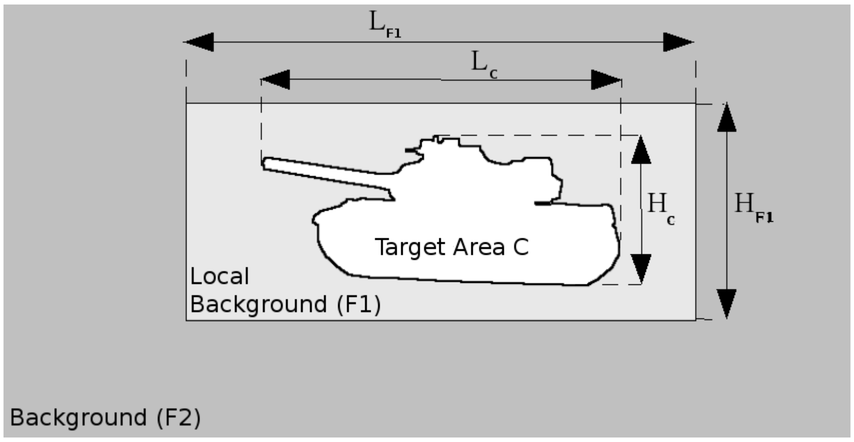


Figure 5.6: Definition of different areas for target superimposition over a choosen background.

coefficient which permits to do the conversion between pixel gray levels and temperature in Kelvin. The choice of these parameters fixes some gains and offsets to apply on the pixels of both the target and background in order to obtain the resulting image. Finally, the sensor effect (MTF and noise) is applied. The hybrid scene generation process is summarized in Figure 5.7. We start, Ⓐ, by positioning the occultant, then, Ⓑ, the positioning of the target inside the background. We apply the calculated gains and offsets to histograms of each region, Ⓒ. We finish by applying the sensor effect, Ⓓ. More details and expressions of the different gains and offsets to apply can be found in [19]. However, one aspect is not taken into account in this algorithm: the intrinsic thermal variability of targets. In the next section, we propose an approach to deal with this aspect in the hybrid scene generation.

Intrinsic thermal variability of targets

In [10, 11], we proposed a new method which permits to deal with the intrinsic thermal variability of a target in IR imagery. Indeed, a same target can have many different thermal aspects according to its activity. For example, tires of a vehicle which stopped since a long time are colder than a vehicle which is moving for a long time. But their engine are quite identical. However, the current vehicle recognition algorithms use some training to be efficient. It is obvious that this kind of algorithm will have less performances if they never "learn" the different aspects of concerned targets.

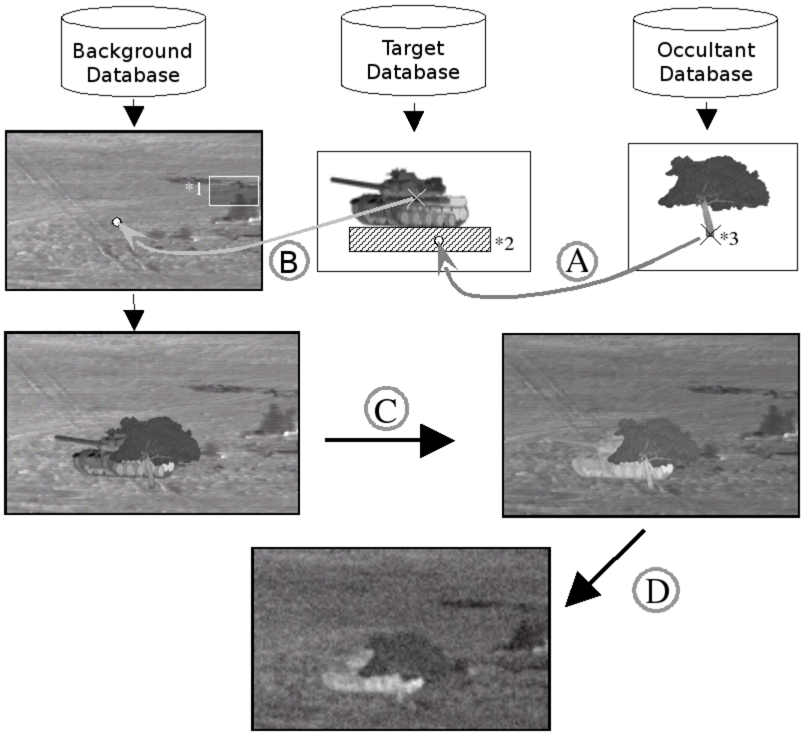


Figure 5.7: Hybrid scene generation principle.

This variability is function of the vehicle operation, this means that it is sufficient to modify the signature of the vehicle. As it is too complex, from a practical point of view, to use an accurate thermal physical models for different targets, we propose to create intermediate signatures by interpolation from ambient (TA) and operationnal (TF) temperatures, taken from real radiometric images. For that, we lay out 3D models of vehicles on which we plate infrared textures. These textures are available for the TA and TF temperatures. We propose to segment the surface of the vehicle into homogeneous thermal behavior areas which are dependent on the different operationnal vehicle's areas. The relevant choosen areas are: the engine, the main body, the muffler, windows, tires/caterpillar (see Figure 5.8).

An intermediate thermal state of an area R (TI_R) relevant to the wanted variability, is generated by mixing the states TA and TF, according to equation 5.6.

$$TI_R = (1 - \lambda_R)TA_R + \lambda_R TF_R, \tag{5.6}$$

where $\lambda_R \in [0; 1]$ represent the variability rate of area R. We can define three different behaviors:

1. ambient temperature: $\lambda_R \in [0; 0.1]$,
2. intermediate temperature: $\lambda_R \in]0.1; 0.9[$,
3. in operation temperature: $\lambda_R \in [0.9; 1]$.

Figure 5.8: Map of homogeneous thermal behavior of a given vehicle.

The final choice of λ_R is done by random drawings according to Gaussian laws (or half-Gaussian at extremities, see Figure 5.9). The standard deviation of each Gaussian is choosen in order to have 99% of its surface inside the intervals considered above. This is equivalent to $3\sigma_{TA} = 3\sigma_{TF} = 0.1$ and $3\sigma_{TI} = 0.4$, this give us $\sigma_{TA} = \sigma_{TF} = 0.33$ and $\sigma_{TI} = 0.133$ respectively. Then the different laws are given by equations (5.7), (5.8) and (5.9) (for all λ_R taken in the previous intervals).

$$P_{TA}(\lambda) = \frac{1}{\sqrt{2\pi\sigma_{TA}^2}} \exp(-\frac{\lambda_R^2}{2\sigma_{TA}^2}) \tag{5.7}$$

$$P_{TF}(\lambda) = \frac{1}{\sqrt{2\pi\sigma_{TF}^2}} \exp(-\frac{(1-\lambda_R)^2}{2\sigma_{TF}^2}) \tag{5.8}$$

$$P_{TI}(\lambda) = \frac{1}{\sqrt{2\pi\sigma_{TI}^2}} \exp(-\frac{(\lambda_R - 0.5)^2}{2\sigma_{TI}^2}) \tag{5.9}$$

By selecting different thermal configurations (for example a vehicle in standby where its engine and muffler are hot, its body, windows and tires at ambient temperature), we can generate the intermediate texture to plate on the 3D model. Thus, we can make 2D views, at different angles, by projection in order to increase the number of signatures in the database used for the hybrid scene generation.

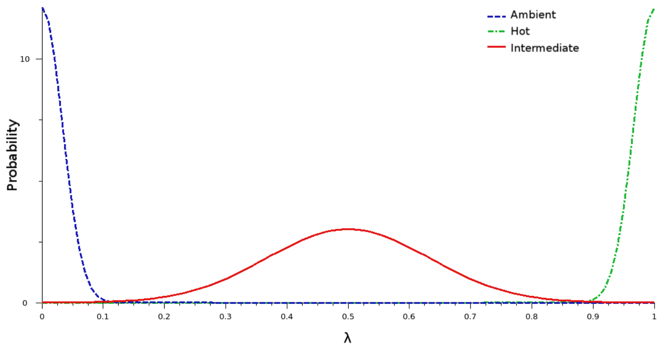


Figure 5.9: Probability laws of λ for the different operationnal mode.

Simulation examples

Here, we presente some examples we have got via the previously described method. Firstly, Figure 5.10 shows different thermal configurations of a same vehicle presented in the same point of view. We can see that it is possible to create realistic IR signatures which correspond to predefined operationnal states (vehicle completely motionless, vehicle in motion, ...). In conclusion, the method enables us the generation of all needed views of a vehicle. Secondly, these new signatures are added to a new target database which will be used by the hybrid scene generator. This allows us to generate scenes which hold account the intrinsic thermal variability of targets by superimposing the wanted target taken in this new database. Figure 5.11 shows an example of a same scene, generated with the same image quality constraints, containing the same vehicle with different thermal configurations.

The extension to generate a sequence of simulated frames with a moving target is given in [12].

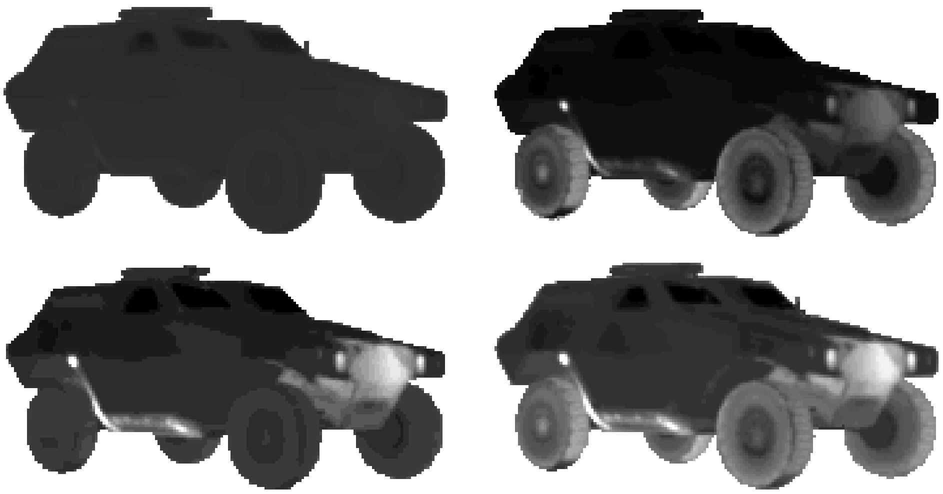


Figure 5.10: Example of different thermal configurations generated by the proposed methode.

5.2.2 Metrics definitions

In this section, we address the question of which metrics are relevant to measure performances of ATD/ATR/ATT algorithms. In fact, as each kind of algorithm has its own characteristics it is impossible to define a way to compute a unique performance score. Each part (detection, recognition and tracking) must have adapted metrics. In [12], we gave the criteria on which the different type of algorithm will be evaluated. We assume that for each test images, their corresponding ground truth are available.

Detection

A detection algorithm is said efficient when the target is detected, well localized and its size is well estimated. Let us define some notations based on figure 5.12. We assume that the assessed algorithm outputs the bounding box (BBox) around the detected target. The reference target is denoted Z^* and the detected one by Z. The variables $X_{ref}, Y_{ref}, W_{ref}, H_{ref}$ and S_{ref} are respectively the coordinates of the BBox's center, W_{ref}, H_{ref} and S_{ref} are the width, height and surface of the reference target's BBox. The same definitions are used for the detected target (D is the corresponding subscript).

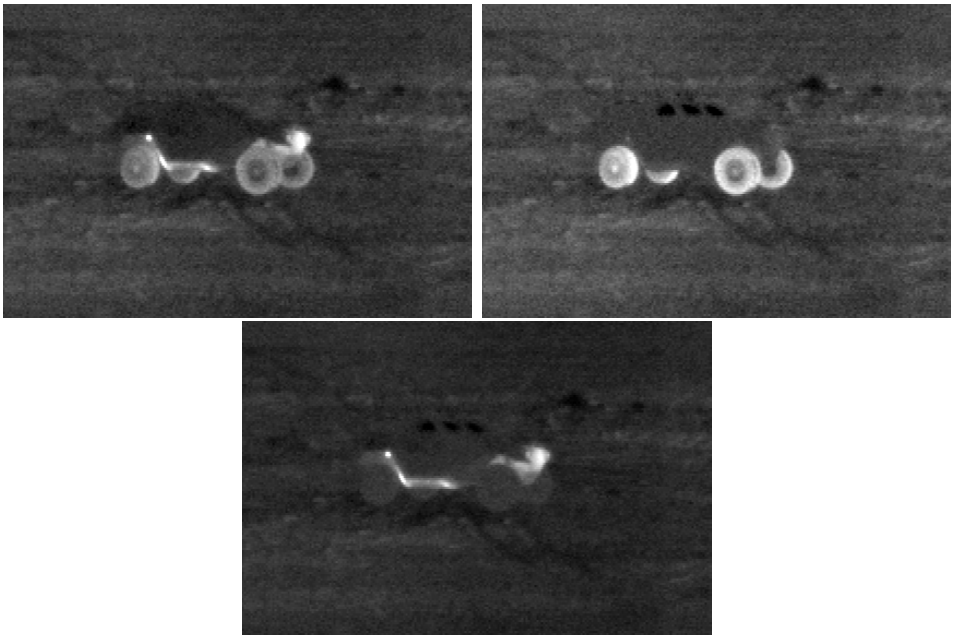


Figure 5.11: Example of a same scene with different thermal configurations of a vehicle.

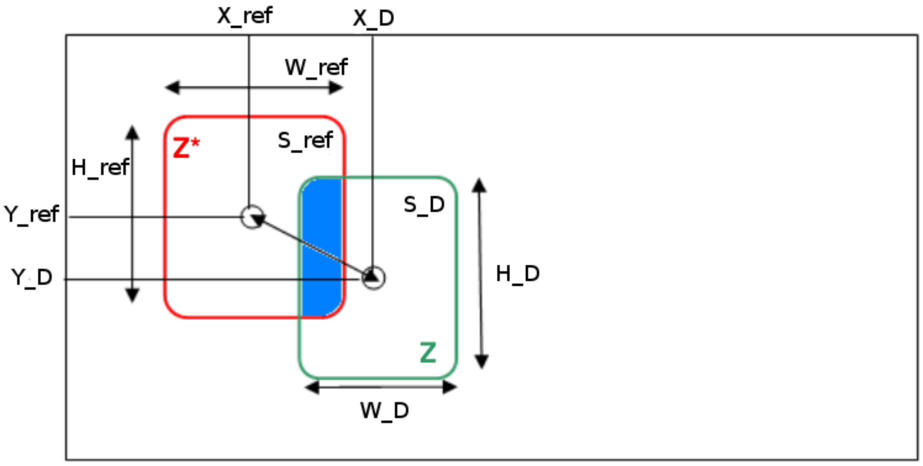


Figure 5.12: Detection's notations.

In practice, we have two ways to declare a detection as good or not. The first one is to use Jaccard's criterion [1, 15] (Eq.5.10).

$$m_0(Z, Z^*) = \frac{|\bigcap(Z, Z^*)|}{|\bigcup(Z, Z^*)|} > \epsilon_0 = 0.5 \tag{5.10}$$

The second one, inspired from the ROBIN competition [4], is to use a combinaison of three criteria: a localization criterion m_1, a scale accuracy criterion m_2 and a segmentation accuracy criterion m_3, respectively defined by equations (5.11) to (5.13).

$$m_1(Z, Z^*) = \frac{2}{\pi} \arctan\left(\max\left(\frac{|X_D - X_{ref}|}{W_{ref}}, \frac{|Y_D - Y_{ref}|}{H_{ref}}\right)\right) \tag{5.11}$$

$$m_2(Z, Z^*) = \frac{|S_D - S_{ref}|}{\max(S_d, S_{ref})} \tag{5.12}$$

$$m_3(Z, Z^*) = \frac{2}{\pi} \arctan\left|\frac{H_D}{W_D} - \frac{H_{ref}}{W_{ref}}\right| \tag{5.13}$$

Then a detection is said good if $m_1 \leq \epsilon_1$, $m_2 \leq \epsilon_2$ and $m_3 \leq \epsilon_3$ are simultaneously verified, where we experimentaly choose $\epsilon_1 = \epsilon_3 = 0.15$ and $\epsilon_2 = 0.5$. Finally, we can calculate, over the dataset, the (good) detection rate (DR), the false alarm rate (FAR) and then plot the corresponding ROC curve [2]. In [18], the authors propose two other interesting metrics which take care about another aspect of segmentation accuracy: the multiple trackers (MT) and multiple objects (MO). The first one represents the fact that multiple BBoxes are found on a unique target, the second one, the case of a unique BBox on multiple targets (see figure 5.13).

Figure 5.13: Multiple tracker (MT) and multiple objects (MO) definitions.

All these metrics permit to accurately evaluate the behavior and performances of any ATD algorithm.

Recognition

We consider two levels of classification: recognition and identification. The first one uses general classes (car, truck, tank, ...). The classes used by the second level correspond to detailed model of target (AMX30, Leclerc, T72, ...). In order to evaluate the performances of this kind of algorithms, we need to check if the class proposed by the ATR algorithm is or not the same as the reference class. The best way to summarize these results is to use confusion matrices [3].

5.2.3 Tracking

Here, we examine the tracking case performed by movement detection algorithms. Two points are needed to be evaluated: the target detection and the tracking itself. The detection case can be treated with the same metrics as described previously. Then we specifically add some metrics to deal with the performances of tracking. In [18], the authors address the behavior of a tracker, in the sense that the algorithm could assign successive trackers to a same target or a tracker initially assigned to one target could "jump" to another target. The first one is called the False Identified Tracker (FIT) and the second one the False Identified Object (FIO). For example, in figure 5.14, the first target is associated with Tracker1 but at a certain time the algorithm missed this tracker and create a new tracker: Tracker2, this is FIT. Tracker3 is assigned to the second target but jumps to the third target at a certain time, this corresponds to FIO.

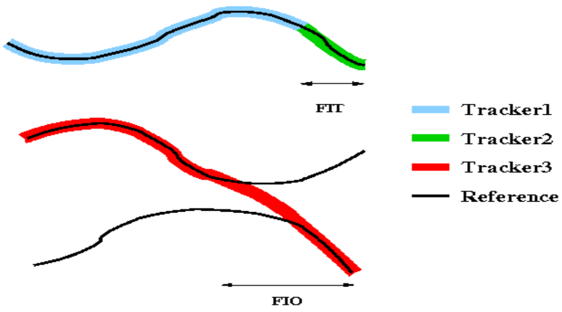


Figure 5.14: False Identified Trackers and Objects.

This complete methodology (called METRIC) to assess ATD/ATR/ATT algorithms on infrared images allowed us to do some evaluations in an exhaustive way. Since the publication [12], METRIC has become the "official"

DGA methodology to assess all ATD/ATR/ATT algorithms. It permits to easily compare the algorithms proposed to DGA and maintain an up to date "catalogue" of such algorithms for internal purposes.

Bibliography

[1] http://en.wikipedia.org/wiki/Jaccard_index.

[2] http://en.wikipedia.org/wiki/Receiver_operating_characteristic.

[3] http://en.wikipedia.org/wiki/Confusion_matrix.

[4] ROBIN competition. http://robin.inrialpes.fr/.

[5] A. Bartesaghi and G. Sapiro. Tracking of moving objects under severe and total occlusions. In International Conference on Image Processing (ICIP), Genoa, Italy, September 2005.

[6] L.M. Biberman, editor. Electro-Optical Imaging : system performance and modeling. SPIE Press, october 2000.

[7] J. Delon. Midway image equalization. Journal of Mathematical Imaging and Vision, 21(2):119–134, 2004.

[8] J. Delon. Movie and video scale-time equalization application to flicker reduction. IP, 15(1):241–248, January 2006.

[9] R.G. Driggers, P. Cox, and T. Edwards. Introduction to infrared and electro-optical systems. Artech House optoelectronics library. Artech House, 1999.

[10] J. Gilles, S. Landeau, T. Dagobert, P. Chevalier, and C. Bolut. Génération de bases de données images IR sous contraintes avec variabilité thermique intrinsèque des cibles. In GRETSI Symposium on Signal and Image Processing, Dijon, France, September 2009.

[11] J. Gilles, S. Landeau, T. Dagobert, P. Chevalier, and C. Bolut. IR image databases generation under target intrinsic thermal variability constraints. In International IR Target and Background Modeling & Simulation Workshop (ITBMS), Toulouse, France, June 2009.

[12] J. Gilles, S. Landeau, T. Dagobert, P. Chevalier, E. Stiée, D. Diaz, and J-L. Maillart. METRIC: a complete methodology for performances

evaluation of automatic target detection, recognition and tracking algorithms in infrared imagery. In OPTRO 2010, Paris, France, February 2010.

[13] J. Harris and Y-M. Chiang. Minimizing the "ghosting" artifact in scene-based nonuniformity correction. Behavioural and Brain Sciences, 16:48–9, 1993.

[14] J.G. Harris and Y.M. Chiang. Nonuniformity correction of infrared image sequences using the constant-statistics constraint. IEEE TIP, 8(8):1148–1151, August 1999.

[15] P. Jaccard. Distribution de la flore alpine dans le bassin des dranses et dans quelques régions voisines. Bulletin de la Société Vaudoise des Sciences Naturelles, 37:241–272, 1901.

[16] J.E. Pezoa, S.N. Torres, J.P. Córdova, and R.A. Reeves. An enhancement to the constant range method for nonuniformity correction of infrared image sequences. volume 3287 of Lecture Notes in Computer Science, pages 525–532. CIARP, Springer, 2004.

[17] G. Sapiro and V. Caselles. Histogram modification via differential equations. Journal of Differential Equations, 135(2):238–268, 1997.

[18] K. Smith, D. Gatica-Perez, J-M. Odobez, and S. Ba. Evaluating multi-object tracking. In CVPR'05: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Workshops, page 36. IEEE Computer Society, 2005.

[19] SPIE Security and Defense, Orlando. Image database generation using image metric constraints: an application within the CALADIOM project, US-Florida, 2006.

[20] Y. Tendero and J. Gilles. ADMIRE a locally adaptive single image non-uniformity correction and denoising algorithm. application to uncooled IR camera. In SPIE Defense, Security and Sensing conference, Baltimore, US, April 2012.

[21] Y. Tendero, J. Gilles, S. Landeau, and J-M. Morel. Efficient single image non-uniformity correction algorithm. In SPIE Security and Defence International Symposium, Toulouse, France, September 2010.

[22] S.N. Torres, E.M. Vera, R.A. Reeves, and S.K. Sobarzo. Scene-based non-uniformity correction method using constant range: Performance and analysis. pages 130–139. Proceedings of the 6th SCI, IX:224?229, 2002.

[23] Vollmerhausen, E.Jacobs, Hixon, and Friedman. NVTherm IP, the targetting task performance (TTP) metric. Night Vision and Electronic Sensors Directorate.

Chapter 6

Current work and perspectives

6.1 Texture modeling

In chapter 3, I presented my contributions concerning cartoon + textures models notably the theoretical results studying the behavior of such models with respect to the parameters. The theorem 2 is very useful because it tells us that we can use the decomposition as a preprocessing to characterize textures by bandpass filtering. I also have a lemma (not presented in this manuscript) which characterizes the decomposition in the specific case of noisy images. A forthcoming paper will present this approach in details.

One frustrating aspect of these decomposition models is their use for practical applications. Indeed, if many papers were published concerning the mathematical concerns of such models, no papers really use it in real applications. The above cited possibility to depict a new approach to analyse textures can be one of this application and, in a close future, I want to investigate this possibility for camouflaged target detection.

6.2 Image restoration

6.2.1 Atmospheric restoration

We saw in chapter 4 that considering the turbulence effects as the combination of two operators (deformation + blur) provides a good way to build efficient restoration algorithms. Today, we are studying the possibility to implement these algorithms in real time onto a small processing unit (like a DSP) for people from the D.O.D.

With Yohann Tendero, at the CMLA in ENS Cachan, we started to explore the ability to add some super-resolution capabilities to the stabilization algorithm and then reconstruct a more detailed image.

Beside these ideas, I'm currently mentoring an undergraduate student on a research project where we are doing a complete parametric analysis of the Fried kernel. Indeed, this kernel, as depicted in chapter 4, has two constant terms (X and $Y = V(Q, X)$) depending on four input parameters (D, λ, L and C_n^2). We aim to understand the behavior of this MTF with respect to these parameters in order to draw a more "mathematical" formulation which could be easier to handle in an optimization process. Preliminary results on the distribution of these constants seem to show that an "optimal" Y might exist regardless the conditions of operation and that X follows a probability law which depends only on λ (see Fig. 6.1 which shows the distribution of X for different λ). We are able to propose a kind of likelihood law $\Lambda(X)$ in the form $\Lambda_\lambda(X) = ae^{bX} + ce^{dX}$ where the parameters a, b, c, d can be evaluated by some regression method. We also found that the "optimal" X ($X_{opt}(\lambda)$) follow a law in powers of λ. The next step is now to use these results to build an optimal Fried deconvolution algorithm which will estimate both the reconstructed $\hat{u}$ and X_{opt}.
Another aspect concerning the Fried kernel and suggested by Michael Brenner from Harvard University is about the 5/3 exponent also called the Kolmogorov exponent. This exponent comes from Kolmogorov's turbulence theory which is well accepted by the turbulence community, but the point is that it comes from experimental measures and no mathematical proof of this value exists. A natural question arises connected to our deconvolution algorithm: is this 5/3 value the optimal one? Or asking differently, is Kolmogorov right or wrong in his theory? Then I started some investigations by changing the 5/3 value by a variable p in the Fried kernel and doing deconvolution on a large set of images by varying p. The first results seem to show that the best deconvolved image corresponds, for the whole test set, to an optimal value of p which is very close but not exactly equal to 5/3. These experiments open some new questions we are trying to solve.

6.2.2 Underwater restoration

I currently work on this topic in two collaborations: the first one with colleagues from the NATO group involved on modeling aspect of underwater imaging systems. The second is for EDF (the French electricity company) that wants to observe tagged numbers on uranium bars plunged in a pool in

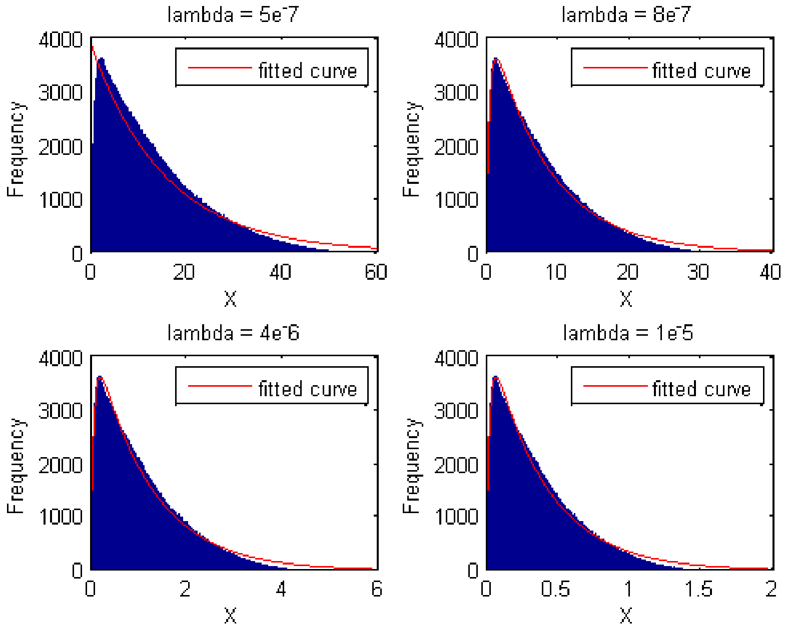


Figure 6.1: Histograms of X corresponding to different λ values

a nuclear plant. Like in the atmospheric case, underwater imaging is subject to turbulence. The water case is worse than hot air as the deformations can have larger magnitude. We will follow the same approach as in the atmospheric case by considering the combination of a distortion operator and a blur operator. If the stabilization algorithm proposed in chapter 4 must remain efficient, the Fried kernel for deblurring is no longer valid in the water. We recently were aware of the existence of the same kind of kernel but for underwater purposes. With Melissa Tong, a Ph.D. candidate at UCLA, we are exploring this new kernel and its behavior in order to adapt the deblurring process.

6.2.3 Digital holographic microscopy deconvolution

This collaboration has just started very recently with Thomas G. Dimiduk from the Department of Physics in Harvard University. Thomas is working on the development of a new equipment to do Digital Holographic Microscopy (DHM) to observe small cells directly in 3D. As the final goal is to capture the shape of the observed cell, the 3D image must have sharp edges. Nevertheless, the acquired 3D images are blurred because of the Point Spread Function (PSF) of the equipment itself. Thomas measured the 3D PSF of its system and it turns out that the measured PSF is extremely noisy and cannot be used directly in a deconvolution process. We are currently investigating two directions. Firstly, we are trying to denoise the PSF itself but classic denoising methods failed at this time and then we will try higher-order TV schemes which seems better suited for such functions. Secondly, different PSF analytical models exist in the litterature; we are exploring the opportunity to use such models.

6.3 QR-code restoration

Barcodes and in particular QR-codes (see Fig. 6.2) are widely used nowadays to distribute information on a product or a website. Smartphones are probably the most used devices to read such codes via many applications available on the different software markets. The main problem is that each smartphone uses its own camera to acquire the code and consequently, each camera has a specific PSF. This results in a phone dependent blur plus noise level. My collaboration with Yves Van Genip (UCLA) and Rustum Choksi (McGill University in Montreal) drive us to a method which, after a denoising of the image, uses the fixed patterns on the two top and left bottom corners (these squares are present in all QR-codes) to estimate the PSF of the equipment by an H^1 regularization; then we use it to do the deblurring and end with a binarization (Fig.6.3 shows a corrupted QR-code and its reconstructed version). In order to check the efficiency of our algorithm, we use QR-code reader software. The software is unable to read the corrupted code while, even if the reconstruction is not perfect, it succeeds to read the restored code. At this time, we plan to assess the overall performances of our method with respect to the influence of blur and noise, respectively. Otherwise, as today we just use a regularity assumption on the blur kernel, we also are interested to acquire some “real” kernels from different smartphones by, for example, taking many pictures of a Dirac function.

6.4 Empirical mode decomposition (EMD) theoretical study

I recently attented a talk about Empirical Mode Decomposition (EMD) which aimed to decompose a signal into different modulated waveforms accordingly to the signal itself. I was intrigued by this topic because today, in

Figure 6.2: Example of a QR-code.

Figure 6.3: A blurry and noisy input QR-code is shown on left and our reconstructed output on right.

my knowledge, there is only two proposed approaches to do this decomposition. The original one is based on some nonlinear filtering which prevents to the development of a “good” theoretical framework; the second one (which was the purpose of the talk I attented) proposes a functional based on sparsity and compressive sensing tools to obtain such decompositions. This last approach is interesting but remains difficult to interpret and is quite expensive in terms of computational time. After reading some papers about EMD, I have the feeling that we can built some kind of adaptive filtering approach to get this decomposition maybe by generalizing the demodulation process used in electronics to listen radio channels. I started to write some ideas and I will continue to explore this way in the future.

Chapter 7

Publications

This chapter gives the complete list of my publications. Let you note that the weak number of publications in journals is due to the fact that a large amount of my contributions were classified while I was at the DGA. Moreover, DGA's politics about publishing was to promote Defense conferences like SPIE.

Conferences with reviewing commitee

[1] J. Gilles. Choix d'un espace de représentation image adapté à la détection de réseaux routiers. Symposium Traitement et Analyse de l'Information: Méthodes et Applications (TAIMA), Hammamet (Tunisie), 2007.

[2] J. Gilles. Séparation en composantes structures, textures et bruit d'une image, apport de l'utilisation des contourlettes. In GRETSI Symposium on Signal and Image Processing, Dijon, France, September 2009.

[3] J. Gilles and B. Collin. Fast probabilist snake algorithm. International Conference on Image Processing (ICIP), Barcelona, 2003.

[4] J. Gilles, T. Dagobert, and C. De Franchis. Atmospheric turbulence restoration by diffeomorphic image registration and blind deconvolution. Advanced Concept for Intelligent Vision Systems Conference (ACIVS), Juan les Pins, France, 2008.

[5] J. Gilles, S. Landeau, T. Dagobert, P. Chevalier, and C. Bolut. Génération de bases de données images ir sous contraintes avec variabilité ther-

mique intrinsèque des cibles. In GRETSI Symposium on Signal and Image Processing, Dijon, France, September 2009.

[6] J. Gilles, S. Landeau, T. Dagobert, P. Chevalier, E. Stiée, D. Diaz, and J-L. Maillart. Metric: a complete methodology for performances evaluation of automatic target detection, recognition and tracking algorithms in infrared imagery. In OPTRO 2010, Paris, France, February 2010.

[7] J. Gilles, B.R. Sharma, W. Ferenc, H. Kastein, L. Lieu, R. Wilson, Y.R. Huang, A.L. Bertozzi, B. HomChaudhuri, S. Ramakrishnan, and M. Kumar. Robot swarming over the internet. In American Control Conference, Montreal, Canada, June 2012.

[8] N. Morizet and J. Gilles. A novel adaptive combination approach to score level fusion for multimodal biometrics combining wavelets ans statistical moments. International Symposium on Visual Computing (ISVC), Las Vegas (US), 2008.

Conferences based on abstract submission or invitation

[1] J. Gilles. Restoration algorithm and system performance evaluation for active imaging systems. SPIE Security and Defense, Florence (Italie), 2007.

[2] J. Gilles, S. Landeau, T. Dagobert, P. Chevalier, and C. Bolut. Ir image databases generation under target intrinsic thermal variability constraints. In International IR Target and Background Modeling & Simulation Workshop (ITBMS), Toulouse, France, June 2009.

[3] J. Gilles and S. Osher. Fried deconvolution. In SPIE Defense, Security and Sensing conference, Baltimore, US, April 2012.

[4] J. Gilles and S. Osher. Turbulence restoration: Atmospheric deblurring. In Mini-Symposium on Imaging Through Turbulence at the SIAM Imaging Sciences conference, Philadelphia, US, May 2012.

[5] J. Ma, Y. Yang, S. Osher, and J. Gilles. Image reconstruction in compressed remote sensing with low-rank and L1-norm regularization. In IGARSS Conference (IEEE International Geoscience and Remote Sensing Symposium), Munich, Germany, July 2012.

[6] Y. Mao and J. Gilles. Non rigid turbulence stabilization. In Mini-Symposium on Imaging Through Turbulence at the SIAM Imaging Sciences conference, Philadelphia, US, May 2012.

[7] Y. Mao and J. Gilles. Turbulence stabilization. In SPIE Defense, Security and Sensing conference, Baltimore, US, April 2012.

[8] Y. Tendero and J. Gilles. ADMIRE a locally adaptive single image non-uniformity correction and denoising algorithm. application to uncooled ir camera. In SPIE Defense, Security and Sensing conference, Baltimore, US, April 2012.

[9] Y. Tendero, J. Gilles, S. Landeau, and J-M. Morel. Efficient single image non-uniformity correction algorithm. In SPIE Security and Defence International Symposium, Toulouse, France, September 2010.

Journals with reviewing commitee

[1] J. Gilles. Noisy image decomposition: a new structure, texture and noise model based on local adaptivity. Journal of Mathematical Imaging and Vision (JMIV), 28(3):285–295, 2007.

[2] J. Gilles. Multiscale texture separation. Submited to SIAM Multiscale, Modeling and Simulation Journal, 2012.

[3] J. Gilles, Y. Mao, and S. Osher. Turbulent image restoration: from stabilization to deblurring. Submited in Journal of Optical Engineering for a special issue on imaging through atmospheric turbulence, 2012.

[4] J. Gilles and Y. Meyer. Properties of BV-G structures + textures decomposition models. application to road detection in satellite images. IEEE Transaction in Image Processing, 19(11):2793–2800, 2010.

[5] Y. Mao and J. Gilles. Non rigid geometric distortions correction - application to atmospheric turbulence stabilization. To appear in Inverse Problems and Imaging Journal, 2012.

Invited seminars, workshops

[1] J. Gilles. Apport de l'information géométrique en décomposition d'images. Workshop Contenu Informatif des Images Numériques (CIIN), ENS Cachan, 2004.

[2] J. Gilles. Décomposition des images en composantes géométriques, textures et bruit: application aux images de la défense. Workshop Mathématiques, Image et Gestaltisme (MIG), Ile de Berder, Golfe du Morbihan, 2004.

[3] J. Gilles. Décomposition d'images - applications. Séminaire du laboratoire Signaux et Systèmes (LSS), Supelec, 2004.

[4] J. Gilles. Décomposition des images: principes et application à la détection de réseaux routiers. Journée PRIDES, Laboratoire SIC - Université de Poitiers, 2007.

[5] J. Gilles. Décomposition d'images. Colloque Mathématiques et Images, Laboratoire de mathématiques et applications, Université de la Rochelle, 2007.

[6] J. Gilles and J.F. Aujol. Suppression de bruit multiplicatif par décomposition d'images - application aux images SAR. Journées imagerie polarimétrique - GdR ISIS - GdR Ondes, 2004.

[7] J. Gilles and J.F. Aujol. Modélisation des images par espaces de fonctions. Ecole analyse multirésolution des images, Université de Poitiers, 2008.

[8] J. Gilles and A.L. Bertozzi. Robot swarming over the internet. In 21st Southern California Nonlinear Control Workshop, California, US, October 2011. UCLA.

[9] J. Gilles, B.R. Sharma, W. Ferenc, H. Kastein, L. Lieu, R. Wilson, Y.R. Huang, A.L. Bertozzi, B. HomChaudhuri, S. Ramakrishnan, and M. Kumar. Robot swarming over the internet. In Frontiers of Real-World Multi-Robot Systems: Challenges and Opportunities Workshop, North Carolina, US, October 2011. Duke University.

[10] Y. Mao and J. Gilles. Turbulence image restoration by deformation field estimation and total variation regularization. In Turbulence Workshop,

Alexandria, Virginia US, August 2010. Night Vision and Electronic Sensors Directorate (NVESD).

[11] Y. Meyer and J. Gilles. Remarques sur un modèle de texture de L.Vese et S.Osher. Journées texture - GdR ISIS, 2003.

Unpublished

[1] J. Gilles and S. Osher. Bregman implementation of Meyer's G-norm for cartoon + textures decomposition. Technical Report CAM 11-73, UCLA - Department of Mathematics, November 2011.