Feature Extraction for Image Super-resolution using Finite Rate of Innovation Principles

by
Loïc Baboulaz

A Thesis submitted in fulfilment of requirements for the degree of
Doctor of Philosophy of Imperial College London

Communications and Signal Processing Group
Department of Electrical and Electronic Engineering
Imperial College London
2008

Statement of Originality

I certify that this thesis, and the research to which it refers, are the product of my own work, and that any ideas or quotations from the work of other people, published or otherwise, are fully acknowledged in accordance with the standard referencing practices of the discipline. I acknowledge the helpful guidance and support of my supervisor, Dr. Pier Luigi Dragotti. The material of this thesis has not been submitted for any degree at any other academic or professional institution.

Abstract

To understand a real-world scene from several multiview pictures, it is necessary to find the disparities existing between each pair of images so that they are correctly related to one another. This process, called image registration, requires the extraction of some specific information about the scene. This is achieved by taking features out of the acquired images. Thus, the quality of the registration depends largely on the accuracy of the extracted features.

Feature extraction can be formulated as a sampling problem for which perfect reconstruction of the desired features is wanted. The recent sampling theory for signals with finite rate of innovation (FRI) and the B-spline theory offer an appropriate new framework for the extraction of features in real images. This thesis first focuses on extending the sampling theory for FRI signals to a multichannel case and then presents exact sampling results for two different types of image features used for registration: moments and edges.

In the first part, it is shown that the geometric moments of an observed scene can be retrieved exactly from sampled images and used as global features for registration. The second part describes how edges can also be retrieved perfectly from sampled images for registration purposes. The proposed feature extraction schemes therefore allow in theory the exact registration of images. Indeed, various simulations show that the proposed extraction/registration methods overcome traditional ones, especially at low-resolution.

These characteristics make such feature extraction techniques very appropriate for applications like image super-resolution for which a very precise registration is needed. The quality of the super-resolved images obtained using the proposed feature extraction methods is improved by comparison with other approaches. Finally, the notion of polyphase components is used to adapt the image acquisition model to the characteristics of real digital cameras in order to run super-resolution experiments on real images.

Acknowledgments

I would like to thank the different persons that have all contributed to making this journey towards a PhD more enjoyable.

First of all, I would like to deeply thank my advisor Dr. Pier Luigi Dragotti. More than a thesis advisor, he gave me many sound advice years before I had even started my PhD at Imperial College. I am forever grateful to him for having given me the opportunity to undertake a PhD here in London. He always greeted me in a jovial mood at our weekly morning meetings and I thank him for his guidance and for having encouraged me all the way. His most precious advice was to enjoy my PhD as it is probably the only period in one's life where one can freely focus on a problem without worrying much about anything else. Thankfully, I enjoyed it. I would also like to thank the Department of Electrical and Electronic Engineering for the considerable financial help of their scholarship.

I would then like to thank my friends and colleagues from the C&SP research group. In particular, I am grateful to Beth, Flore, Jesse, Jon, Mark, Nicolas, Nikolay, Pancham, Vinesh and Yizhou (Eagle) for the good moments we had together during these years in London. I also wish to thank my old friends from the Lycée that I have not seen very often recently but who are always important to me: Ben, Emilien, Mathieu, Milena, Stephane as well as their relative partners. From my time in Lausanne, I would like to thank my friends who have each gone their own way and with whom I always enjoy to meet with: Catherine, David, Eva, Florence, Kathrin, Mathieu, Peng, Tiago and Wendy. And thank you to Byron for all the good times we have together wherever we meet.

Last but not least, I would like to thank my mum, my dad and my sister for their constant support and help. Finally, I am especially thankful to Natascha who supported me through the highs and lows of my PhD with patience and love.

Contents

List of Figures

Abbreviations and Notations

Abbreviations

A/D	Analog to digital;
AFBP	Annihilating filter based back projection;
CCD	Charge couple device (type of digital camera sensor);
CCTV	Closed-circuit television;
CG	Conjugate gradient;
CMOS	Complementary metal oxide semiconductor (type of digital camera sensor);
DSVD	Damped singular value decomposition;
ECG	Electrocardiogram;
EMLS	Expectation maximization least-square;
ESF	Edge spread function;
FRI	Finite rate of innovation;
GCV	Generalized cross validation;
HDTV	High definition television;
ISO	International standard organization;
LSF	Line spread function;
MRNSD	Modified residual norm steepest descent;
PSNR	Peak signal to noise ratio;
PSF	Point spread function;
RANSAC	Random sample consensus;
RC	Circuit composed resistors and capacitors;

S/H Sample and hold;

SLR Single lens reflex (camera);

SVD Singular value decomposition;

TSVD Truncated singular value decomposition;

Notations

$\otimes$ Kronecker product, or tensor product;

$*$ Convolution;

$\times$ Tensor product;

δ Dirac function;

ε Average geometric registration error;

$\varepsilon_{\max}$ Maximum geometric registration error;

λ Regularizing parameter;

φ Sampling Kernel;

$\hat{\varphi}$ Fourier transform of φ;

$\tilde{\varphi}$ Dual function of φ;

$\beta^P(t)$ 1-D B-spline function of order P ;

$\tilde{\beta}^P(t)$ 1-D Dual B-spline function of order P;

$B^P(f)$ Fourier transform of $\beta^P(t)$;

$b_T^P(k)$ Discrete B-spline of scale T and order P (in x and y directions);

$c_T^P(k)$ Discrete B-spline of scale T and order P and shifted by $1/2$;

$\beta^P(x,y)$ 2-D B-spline function of order P;

$\beta_{\underline{\alpha}}(t)$ 1-D E-spline function with parameters $\underline{\alpha}$;

$B_{\underline{\alpha}}(\omega)$ Fourier transform of $\beta_{\underline{\alpha}}(t)$;

L Support of the sampling kernel;

$c_n^{(p)}$ Coefficients for 1-D polynomial or exponential reproduction of order p;

$c_{m,n}^{(p,q)}$ Coefficients for 2-D polynomial or exponential reproduction of order p in x-direction and q in y-direction;

f	Continuous observed view;
g	Continuous blurred view;
$\mathbf{f}$	Discrete high resolution image of f used in simulations;
$\mathbf{M}$	Mapping matrix composed of blurring and downsampling operators;
$\mathbf{K}$	Blurring matrix;
$\underline{\mathbf{f}}$	Column-vector of the discrete high resolution image $\mathbf{f}$;
$\mathbf{g}$	Acquired discrete image (matrix);
$\underline{\mathbf{g}}$	Column-vector of the discrete acquired low-resolution image $\mathbf{g}$;
P	Order of approximation (polynomial or exponential);
ρ	Rate of innovation of a FRI signal;
L^2	Space of square integrable functions;
l^2	Space of square summable sequences;
T	Sampling period;
$V(\varphi)$	Shift-invariant space spanned by shifted versions of φ;
$r(t)$	Impulse response of an acquisition device for 1-D signals;
$r(x,y)$	Impulse response of an acquisition device for 2-D signals;
$h(x,y)$	Edge function;
$\mathbf{h}[m,n]$	Samples of $h(x,y)$;
$\vec{N}$	Vector normal to the edge;
$\mathbf{d}[m,n]$	Differentiated samples;
$\mathcal{T}$	Geometric transformation (translation, rotation, affine ...);
$\mathbf{A}$	2x2 matrix of an affine transformation (rotation, shear, scale);
$\mathbf{C}$	Covariance matrix;
$\mathbf{F}$	Cholesky decomposition of $\mathbf{C}$;
$\mathbf{R}$	Rotation matrix;
$\mathbf{t}$	2x1 vector of translation;
$x(t)$	1-D continuous time signal;
$\mathbf{y}[n]$	1-D sampled signal of $x(t)$;

$\Psi_{p,q}$ Continuous moments of order $(p+q)$;

$\underline{\Psi}_{p,q}$ Discrete moments of order $(p+q)$;

$m_{p,q}$ Continuous geometric moments of order $(p+q)$;

$\underline{m}_{p,q}$ Discrete geometric moments of order $(p+q)$;

$\mu_{p,q}$ Continuous central moments of order $(p+q)$;

$C_{p,q}$ Continuous complex moments of order $(p+q)$;

V_{in} Input voltage;

V_{out} Output voltage;

Chapter 1

Introduction

1.1 Background

OVER the last ten years, digital image acquisition devices have outnumbered traditional film cameras and have made inroads across many sectors of society. This recent shift in the way images are acquired has led to a large increase in the use of images in new areas and in the demand for new applications. Consumers have access to a variety of digital image acquisition devices ranging from web-cams and mobile phones to expensive digital single lens reflex (SLR) cameras (Figure 1.1(a)). Depending on their components, the quality of such devices can vary greatly. In the industry, intelligent automated systems increasingly rely on cameras and computer vision techniques to take decisions based on their environment in order to help or extend the users abilities, like for example in augmented cognition. In many cities, the use of large surveillance systems with CCTV cameras is widespread in places like streets, banks or supermarkets to thwart crime, analyze behavior and prevent accidents. Digital image processing plays also a major part in areas such as biomedical imagery, seismic imagery, satellite imagery or robot vision.

To be efficient, these technologies require good quality, detailed images of high resolution. However, in practice, there are many constraints which make the use of high-resolution cameras difficult, if not sometimes impossible. The quality and the resolution of a camera are mainly dictated by two elements: its lens and its sensor (CMOS or CCD). High quality lenses made of glass are fragile, heavy, bulky and very expensive to

Figure 1.1: (a) Digital image acquisition devices have made inroads across many sectors of society but their quality can vary greatly according to their use and cost. (b) A camera sensor network is made of independent sensors that communicate directly to the receiver; each sensor has its own observation of the phenomenon of interest.

manufacture whereas low quality lenses found on small-sized cameras are made of shock-resistant plastic and are relatively cheap to produce. Because perfect lenses do not exist, even the finest lenses act like a low-pass filter on the observed scene. This removes its high frequency content and adds blur to the image hitting the focal plane. Poor quality lenses blur the images more than fine ones. The camera sensor is described by its resolution *i.e.* the number of photosites, or pixels, it contains. Because of electronic limitations, it is not possible to both reduce the physical size of the camera sensor and keep the same resolution without inevitably introducing some noise (or errors) in the system and degrading the quality of the image. Thus cameras with high quality lenses and high resolution sensors have a large size, a heavy weight and are expensive. These characteristics are often not suitable in many situations. Furthermore, for autonomous cameras running on batteries, the compression of a high resolution image and/or its transmission between the camera and a receiver can drain most of its computing power that should be used instead to actually sense the environment.

Recently, advances in communications network technology and in signal processing have permitted a new model to emerge: the sensor network that analyzes the phenomenon of interest using several independent sensors disseminated inside or around it (Figure 1.1(b)). Because the acquisition is distributed among a set of sensors, the design constraints on each sensor can be relaxed. The framework of our research replaces the traditional single camera setup with a system composed of several cameras referred as distributed camera network. The same scene is observed by different sensors located at

different positions so that a multiple view of the scene is available. Each sensor is a small, independent, low power and low quality camera communicating only and directly to the receiver. These cameras are less obtrusive, lighter, more resistant to shocks and cheaper to produce. Most of the processing is done at the receiving end where the observed scene is analyzed by fusing the information coming from the set of cameras. Using several sensors to acquire a scene also makes a system that is naturally more robust to camera failure than a system where only a single camera is used.

Moreover having a multiple view of a scene allows to develop advanced techniques for scene analysis that are normally not possible with a single view. Such applications are for example super-resolution, occlusion removals, depth segmentation or photogrammetry. The very first task usually required when dealing with multiview images is their correct registration. This consists in finding the disparities existing between the different views so that the images can be overlaid as precisely as possible. Image registration is based on the extraction of features of interest in each image of the set. The more accurate their extraction, the more precise the registration. One example of application where a correct registration is critical is found in the challenging problem of image super-resolution.

1.2 Problem Statement and Motivations

Multiview camera systems are composed of several cameras positioned at different locations and focusing on the same scene of interest. The absolute and relative locations of the cameras are unknown. The i-th camera has its own (continuous) observation $f_i(x, y), x, y \in \mathbb{R}$, of the scene and is acquiring a sampled image $\mathbf{g}_i[m, n], m, n \in \mathbb{Z}$, at any given time. In many situations, the disparities between the images $\mathbf{g}_i, i = 0, 1, \ldots$, must be calculated as accurately as possible so that each image can be correctly overlaid on to another after appropriate warping. Thus the exact registration of two images consists in finding the transformation $\mathcal{T}_i$ between the view f_i of one camera and the view f_j of another camera. However, the observed continuous views are usually not available and image registration is based on the acquired sampled images $\mathbf{g}_i$ and $\mathbf{g}_j$.

By neglecting border effects during sampling, it is generally assumed that the dif-

ferent observations f_i can be related to a single observation of reference, say f_0, via a geometric transformation $\mathcal{T}_i$ of the coordinates (x, y):

$$f_i(x, y) = f_0(\mathcal{T}_i(x, y)), \qquad x, y \in \mathbb{R}, \tag{1.1}$$

where $\mathcal{T}_0 = \mathbf{1}$. The transformations $\mathcal{T}_i$ can be of various types depending on the level of complexity of the scene. These can range from simple translation to complex non-linear transformations (see Figure 1.2). In this research, we consider the case of linear transformations which are composed of translation, rotation, affine transformation and homography. This imaging situation corresponds either to a flat observed scene or a 3D scene that is observed from a distance much greater than the distance between cameras so that parallax effects are negligible [18, 31]. A homography is described by a 3x3 matrix with eight degrees of freedom (plus an arbitrary scaling factor). It represents the most complex of the linear transformations and can be retrieved from four pair of corresponding points. The affine transformation is simpler and is characterized by 6 degrees of freedom. In various situations, the homography transformation is well approximated by an affine transform. The goal of image registration is thus to find the different transformations $\mathcal{T}_i$ as accurately as possible given the available data which are the acquired images $\{\mathbf{g}_0, \mathbf{g}_1, \ldots\}$. As the resolution of these images decreases, less information is available and the correct estimation of the $\mathcal{T}_i$ gets more and more difficult.

For many applications, image registration is a preprocessing step that can jeopardize the entire success of the desired task. Since it is operating on the sampled images $\mathbf{g}_i$ instead of the original observed views f_i, image registration is an inverse problem and is dependent on the properties of the camera used for image acquisition: the sampling rate (or resolution) of the sensor, the amount of blur introduced by the lens, and the noise of the device. As the resolution of an image decreases, the local two-dimensional structure of the image degrades and a correct registration of two low-resolution images becomes increasingly difficult. In this respect, we observe that image registration and image sampling are intimately related. The first way to achieve exact registration of low-resolution images is to perfectly reconstruct the original views f_i and f_j from their sampled

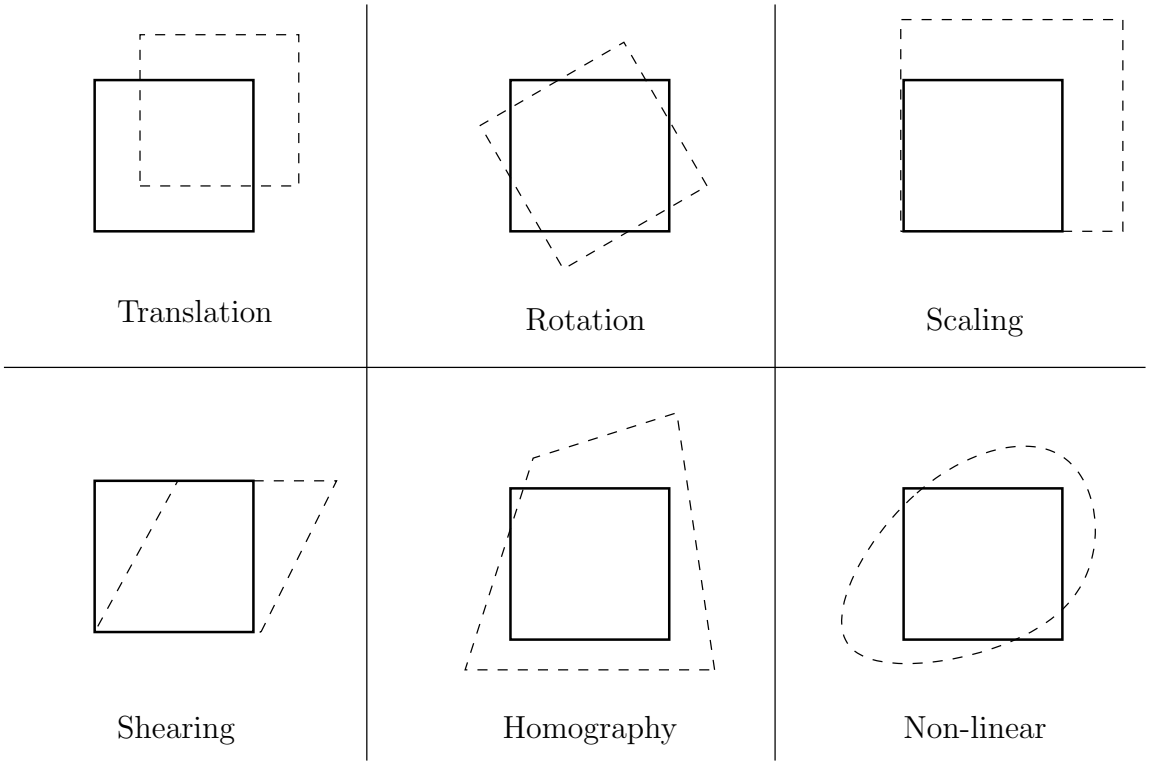


Figure 1.2: Different types of geometric transformations $\mathcal{T}$ considered in registration. Rigid transformations consist of translations and rotations only whereas affine transformations include also shearing and scaling. In certain imaging situations, the homography can be approximated by an affine transformation. Non-linear transformations are particularly important in biomedical images for instance.

versions $\mathbf{g}_i$ and $\mathbf{g}_j$ and run a registration algorithm on f_i and f_j. A second option would consist in extracting from $\mathbf{g}_i$ and $\mathbf{g}_j$ only the relevant image features of f_i and f_j using sampling theory.

The first option raises the fundamental question of signal reconstruction in sampling theory. For bandlimited signals, the well-known Nyquist-Shannon sampling theorem shows that exact reconstruction is possible from the samples only. Moreover, following the developments of wavelet theory (*e.g.* by focusing on the scaling function in a wavelet multiresolution framework rather the wavelet itself), new sampling schemes have recently emerged allowing perfect reconstruction of a certain class of non-bandlimited signals called signals with Finite Rate of Innovation (FRI) [126] [27]. This theory allows the perfect reconstruction from the samples of continuous signals which were previously impossible to recover exactly. The peculiarity of these signals resides in the fact that they can be completely described by a parametric expression having a finite number of degrees of freedom.

By taking into account known properties of the acquisition device, the parameters of the observed FRI signal can be exactly retrieved. Once these parameters are known, the original FRI signal is also known. The new sampling schemes for FRI signals proposed by Vetterli *et al* in [126] considers filters of infinite support like the sinc function. In [78], these results were extended to 2-D signals. In [27], it was shown that similar sampling schemes can also be obtained with physically realizable filters of compact support. In [110], Shukla *et al* presented exact methods for local reconstruction of multidimensional FRI signals from their samples using filters of compact support.

An alternative approach to signal reconstruction consists in extracting only relevant features in the set of images. This avoids the difficulty of reconstructing exactly the complete signal from its samples by looking only at interesting parts of the considered signals. For image registration, various features can be considered. A popular choice of features are corners or contours in the image which are points with a steep intensity gradient. However, other features can be considered. In particular, image moments are also extensively used in image registration either as global features when the moments of the whole image are considered, or as local features when moments of shapes within the image are considered. In all cases, the main challenge consists in recovering as accurately as possible the features occurring in f_i and f_j given the information from $\mathbf{g}_i$ and $\mathbf{g}_j$. Features of interest like corners or edges are basically non-bandlimited signals and are severely deteriorated during acquisition due to the low-pass effect of a real lens. When considering the feature extraction problem as a sampling problem, this suggests that the classical sampling theory for bandlimited signals is not suitable to retrieve exactly these features and that recent advances offered by the sampling theory for FRI signals make room for designing new feature extraction techniques that would strongly improve current registration methods.

The first goal in this work is to understand how the sampling theory for FRI signals fits in a multiview environment as this has not been studied yet. We thus first extend the scope of application of some selected new sampling schemes for FRI signals to a multichannel framework. Given these results, we then consider the case of real multiview images. In particular we want to develop new feature extraction techniques based on FRI

principles to obtain more accurate methods for registration of multiview images. Although the signals initially considered by the FRI theory are synthetic ones and are very different from real images, the main incentive in considering the theory for FRI signals is due to the fact that some features used for image registration can be modeled as FRI signals. Thus, if perfect reconstruction methods for FRI signals can lead to the exact extraction of features in an image, then in turn, an exact registration of multiview images becomes possible regardless of the resolution.

The main motivation of having a precise registration of low-resolution images is to be able to improve multiview applications like image super-resolution. Image super-resolution consists in generating at the receiver a single detailed high-resolution image by fusing the information of a set of different low-resolution multiview images acquired with a camera. One critical step in super-resolution is to obtain an accurate registration of the set of images especially when the resolution of the images decreases. This technique has recently received a lot of attention in the signal processing community and industry and the reasons for this enthusiasm are varied. Today, the transition between analog television (NTSC or PAL) to high definition television (HDTV) displays is fully engaged. The native display of HDTV must be able to accommodate a variety of video formats which often have a lower resolution than the display itself [63]. Simple upsampling methods like interpolation increase only the number of pixels but not the perceived resolution. Advanced techniques like image super-resolution allow the increase of both the pixel density and the perceived resolution. Besides, the drop in the price of low-quality image acquisition devices like webcams makes multiview image systems affordable to produce. Finally the miniaturization of cameras means that, in some situations, the hardware limitations have to be overcome with a software approach in order to produce images with the desired level of details.

Thus, overall, the research described in this thesis starts with a set of theoretical sampling results from which new approaches for accurate feature extraction in real images are developed. Registration methods using these extracted features are then proposed and present the advantage of being very precise for the registration of low resolution images. This characteristic makes them suitable for use in problems like image super-resolution

for which various experiments are also conducted. Throughout this research, two main assumptions are made: firstly, the transformation model between multiview images is a linear one, and secondly, the properties of the acquisition device, like the sampling kernel, are known exactly and follow the acquisition model used in the sampling theory for FRI signals [27].

1.3 Previous Related Work

In the literature, the problem of feature extraction is sometimes further split into two different problems: feature detection and feature location. Feature detection is merely concerned by the existence at pixel accuracy of a given feature; feature location, often written in parametric form, is concerned by the exact location of a feature, *i.e.* at subpixel accuracy. We consider throughout this thesis the problem of recovering the description of a feature as accurately as possible and we use the terms detection, location or extraction invariably.

Various surveys and books on image registration are available for an in-depth review of image registration techniques and feature extraction methods [13, 46, 132]. Registration techniques are either operating in the spatial domain or in the frequency domain. Frequency domain methods are computationally efficient but are generally limited to global rigid motions *i.e.* translations and rotations [35, 74, 124]. Registration techniques in the spatial domain are numerous. Earliest methods seek to maximize the cross-correlation between two images so that the two images are correctly overlaid [100]. Many variations of this model have been proposed like *e.g.* preprocessing the images with an edge detector in order to lower the computational load [3]. In this work, we are interested in the set of registration techniques occurring in the spatial domain and having the following sequential structure:

1. feature extraction in each image,

2. feature matching across images,

3. transformation estimation based on the matched features.

State-of-the-art feature extraction methods can be classified as contour-based, model-based or intensity-based. Contour-based algorithms first extract contours of objects and then locate features where the curvature on the contour is a local maximum [4,81,97]. Model-based methods aim at fitting a parametric model of the features of interest using minimization techniques [12, 23, 103]. In practice, most algorithms are intensity-based. Many detectors for corners or edges are traditionally implemented by filtering the signal with an operator that usually computes the first or second derivative. A measure of 'cornerness' is then calculated from the local maxima of the first derivatives or the zero-crossings of the second derivatives. Depending on its value and threshold, a corner is detected or not. Various automatic corner detectors have been proposed and they essentially differ in the way the local gradient and the measure of 'cornerness' are computed: we have in chronological order: Moravec [86], Harris-Plessey [54], Noble [91], Shi-Tomasi [108], Förstner [36], SUSAN [111]. In order to achieve subpixel accuracy, the position of the features is further improved by fitting a local paraboloid model estimated with a least-square method. A more detailed description and performance analysis of various corner detectors are available in [102, 105]. Experimental results showed that such corner detectors based on the gradient of the image can achieve an accuracy of approximately 1/3 to 1/4 of a pixel. The Harris-Plessey corner detector and its variants (with subpixel accuracy) are often considered as the current standard for detecting corners as this algorithm proved to be fairly robust to change in illumination, camera position and internal camera parameters [105]. The problem of feature detection is closely related to the problem of landmark detection often encountered in medical imaging [40,43,59]. The two problems are however different as automatic landmark detection usually consists in pruning a set of retrieved features (obtained *e.g.* with a Harris corner detector as in [59]) in order to keep only the features that correspond to the structures of the tissue of interest (*e.g.* a particular corner of the jaw's bone).

Similarly to corners, earliest edge detectors are obtained by convolving the image with a directional gradient operator like Roberts, Prewitt and Sobel, or with a Laplacian operator like Marr-Hildreth. A current standard in edge detection is the Canny edge detector [17] which simply looks at the gradient in four directions (horizontal, vertical, and

diagonals). Such a method does not give subpixel accuracy of edges. Subpixel accuracy can be achieved but this approach is generally computationally expensive. Using the local moments of an image calculated over a specific window together with least square fit allow to retrieve the parameters of various edge models with an accuracy up to one twentieth of a pixel [42, 55, 75, 99]. Steerable filters that can be rotated at any arbitrary angle by linear combination were also successfully used for detecting contours in noisy images [38, 61]. In [24], an overcomplete set of step functions called wedgelets was designed to find edges at subpixel accuracy. In addition to points of high curvature on a contour, centers of gravity (or barycenters) of closed contours can also be used as feature points for registration purposes [34, 47, 72].

Global features take into account the entire image to compute some global measurement of the image for registration and object recognition purposes. A desirable property of global features is that they provide a very concise representation of images as points in a multi-dimensional feature space. Nevertheless techniques based on global features may fail when objects in multiview images get occluded or merged. In all applications based on global features, it is either assumed that images contain just one object or that objects can be segmented beforehand from the background.

The prevailing global features used for image analysis are image moments which provide accurate descriptors for shapes and textures [117]. There exists a vast amount of literature available on moment theory and applications. Moments are used extensively in various computer vision and other applications like pattern recognition [57]. Moments have been used in a local fashion for image segmentation [71], image registration [34], interpolation [39], and optical flow estimation [115]. They have been successfully applied to the extraction of various features (*e.g.* DNA strand) in biomedical images [60, 115]. Finally, moments have also been used for the reconstruction of particular bilevel shapes [28, 82, 83, 110] or real images [87, 88, 131].

1.4 Outline of Thesis

In this thesis, we present new feature extraction techniques based on the sampling theory for FRI signals. We approach the problem of feature extraction from a sampling perspective and eventually present a complete image super-resolution algorithm implementing the developed methods.

We start in Chapter 2 by reviewing the sampling theory for FRI signals. We present different sampling results for both one and two dimensional FRI signals. We also present the important properties of the sampling kernel which play a fundamental role in the reconstruction procedure. We focus in particular on the family of B-spline kernels and E-spline kernels.

In Chapter 3, we extend the current sampling results to the multi-channel acquisition of FRI signals. We consider the case of 1-D FRI signals and the case of 2-D FRI signals. In 1-D, the set of multiview signals are different shifted versions of same observed signal. In 2-D, the multiview signals are related by various geometric transformations which depend both on the scene itself and on the positions of the sensors. The main goal is to have simpler acquisition devices and to be still able to reconstruct exactly each of the observed FRI signals. We show how in the case of bilevel polygons, the different views can be perfectly reconstructed using a method based purely on moments. Depending on the type of kernels used, different solutions having either a symmetric or an asymmetric distributed architecture are proposed. Finally, the case of analog-to-digital converters is considered. In particular, we show how the new sampling results for FRI signals can be used in a distributed fashion to improve problems associated with time interleaved analog-to-digital converters.

With this extension to multiview FRI signals, we then show in Chapter 4 how these results can be directly extended to low-resolution image acquisition devices and real-world scenes. We first describe in more detail our image acquisition model and how it compares with the FRI framework. In this chapter, two different feature extraction techniques are proposed. The first method follows directly from the results obtained in the case of multi-channel acquisition of FRI signals by retrieving the continuous moments of an image. The

second approach allows the local extraction of edges in real images. For both techniques, we show how these features can be used in order to obtain exact image registration results. We finally apply these algorithms in different experiments and analyze their accuracy.

In Chapter 5, we apply the proposed registration methods to image super-resolution as this application depends heavily on the quality of the registration step. We describe in more details the problem of image super-resolution and present in particular different methods for image reconstruction. We also compare the results obtained using our feature extraction methods with the results obtained from traditional feature extraction methods. We show that the quality of the super-resolved images is higher with the proposed feature extraction methods than with standard ones.

Finally, we conclude this thesis in Chapter 6 and present some ideas and remarks for future works.

1.5 Original Contribution

The main contribution of this thesis is in the development of exact feature extraction techniques for image registration using FRI principles. To this end, we first study the sampling theory for FRI signals in a multiview environment as detailed in chapter 3. Chapter 4 presents the novel approaches for feature extraction in real images and describes how to use those features for image registration purposes. Finally, chapter 5 presents various new results of image super-resolution based on the developed registration algorithms. To the best of our knowledge, chapters 3, 4 and 5 of this thesis contain the original research work which has led to the following publications:

- L. Baboulaz and P. L. Dragotti. Exact Feature Extraction using Finite Rate Of Innovation Principles with an Application to Image Super-resolution, *submitted to IEEE Transactions on Image Processing*, December 2007.

- P. Vandewalle, L. Baboulaz, P.L. Dragotti and M. Vetterli, Subspace-Based Methods for Image Registration and Super-Resolution, *to appear in Proc. of IEEE International Conference on Image Processing*, San Diego, USA, September 2008.

- L. Baboulaz and P. L. Dragotti. Local feature extraction for image super-resolution, *IEEE Int. Conf. on Image Processing*, pages 401–404, San Antonio, TX, USA, September 2007.

- L. Baboulaz and P. L. Dragotti. Image super-resolution with B-spline kernels, *7th IMA Int. Conf. on Mathematics in Signal Processing*, pages 114–117, Cirencester, UK, December 2006.

- L. Baboulaz and P. L. Dragotti. Distributed acquisition and image super-resolution based on continuous moments from samples, *IEEE Int. Conf. on Image Processing*, pages 3309–3312, Atlanta, GA, USA, October 2006.

- L. Baboulaz and P. L. Dragotti. Beyond vision with image super-resolution, *9th Great British Research Show for Early-Stage Researchers*, SET for Europe, House of Commons, Westminster, London, UK, March 2007[1].

- L. Baboulaz and P. L. Dragotti. Distributed Acquisition and Image Super-Resolution with Sampling Kernels reproducing Polynomials, *WavE 2006 Wavelets and Applications Conference*, EPFL, Lausanne, Switzerland, July 2006[2].

A particular effort has been made to ensure that the results presented in this thesis are also reproducible. Making research reproducible is an important aspect of science and is promoted in [123]. Matlab code and data are available on demand.

[1]Received GlaxoSmithKline/SET Commendation Award.
[2]Received EPFL/Bernoulli Center Travel Grant Award.

Chapter 2

Sampling Theory for Signals with Finite Rate of Innovation

2.1 Introduction

THE notion of signals with finite rate of innovation (FRI) was first introduced by Vetterli, Marziliano and Blu in 2002 [126]. They considered a set of 1-D signals that are neither bandlimited nor belong to a given subspace. Because such signals are not bandlimited, the classical sampling framework provided by Shannon's theory cannot provide perfect reconstruction strategies. They proposed new sampling schemes for perfect reconstruction of the original continuous signal out of the discrete representation provided by the samples. The particularity of the FRI signals resides in the fact that they possess a finite number of degrees of freedom per unit of time, also referred to as the rate of innovation. Examples of such signals are streams of Diracs or piecewise polynomial signals. In other terms, at any given time, the observed signal can be completely characterized by a finite number of parameters. The intuitive idea behind the sampling theory for FRI signals is that by recovering exactly those parameters using the available samples, the original signal can be reconstructed perfectly.

The sampling theory developed in [126] assumes an exact knowledge of the sampling kernel of the acquisition device which describes the pre-filtering module together with the sampling step. The kernels considered in [126] are the sinc and Gaussian kernels with

an infinite support and are therefore not physically realizable. The reconstruction of the FRI signals is based on a non-linear algorithm which involves the annihilating filter (*a.k.a.* Prony's method), a technique widely used in other fields like spectral estimation [113] and error correcting code [9]. In [26], Dragotti *et al* obtained the same results but with sampling kernels with finite support. Kernels with compact support facilitate the practical implementation of the exact sampling scheme for FRI signals which are also based on the annihilating filter method. Those sampling kernels have the particular property of reproducing polynomials or exponentials.

The sampling and the perfect reconstruction of 2-D FRI signals like sets of 2-D Diracs and bilevel polygons were also investigated by Maravić and Vetterli in the case of kernels with infinite support [77, 78]. More recently, Shukla and Dragotti proposed different new local and global reconstruction schemes for 2-D FRI signals like the directional derivatives approach, the complex moments approach or the tomographic approach [110]. In-depth treatments of sampling schemes for FRI signals can be found in the PhD theses of Marziliano [80], Maravić [76] and Shukla [109]. The performance of the proposed sampling schemes in the presence of noise has been analyzed for the case of 1-D signals in [27, 70, 79].

Sampling theory for FRI signals has already found applications in several areas such as resolution enhancement [27, 80], distributed compression [41], synchronization and channel estimation for ultra-wide band signals [70, 76], biomedical signals like ECG signals [52], A/D converters [65] and image super-resolution algorithms [5, 6].

This chapter first defines 1-D FRI signals, presents how to sample them and shows various reconstruction procedures. It then considers the case of sampling 2-D FRI signals. Finally more details are given on the functions that can be considered as valid sampling kernels for sampling FRI signals.

2.2 Sampling of 1-D FRI Signals

2.2.1 1-D Signals with Finite Rate of Innovation

Let us assume that the set of functions $\{\phi_r(t-t_k)\}_{r=0\ldots R-1}$ is known and define the signal with the following parametric form:

$$x(t) = \sum_{k\in\mathbb{Z}} \sum_{r=0}^{R-1} a_{k,r}\phi_r(t-t_k), \qquad t \in \mathbb{R}. \tag{2.1}$$

Since the $\{\phi_r(t-t_k)\}_{r=0\ldots R-1}$ are known, the only parameters that are unknown in the signal $x(t)$ are the coefficients $a_{k,r}$ and the time shifts t_k. We can introduce the function $C_x(t_a, t_b)$ that counts the number of unknown parameters in $x(t)$ over a given interval $[t_a, t_b]$. The rate of innovation of $x(t)$ is then defined as the average number of free parameters present in this signal:

$$\rho = \lim_{\ell\to\infty} \frac{1}{\ell} C_x\left(-\frac{\ell}{2}, \frac{\ell}{2}\right). \tag{2.2}$$

Thus, depending on the limit of Equation (2.2), one can characterize a signal on whether its rate of innovation is finite or not.

Definition 1 (Vetterli, Marziliano, and Blu, [126]). *A signal with a finite rate of innovation is a signal whose parametric representation is given in (2.1) and with a finite ρ as defined in (2.2)*

Interestingly, bandlimited signals satisfy Definition 1. For instance, let f_{max} be the maximum non-zero frequency in a bandlimited real signal $x_B(t)$. The well-known reconstruction formula for bandlimited signals with a sampling period $T = \frac{1}{2f_{max}}$ is:

$$x_B(t) = \sum_{n=-\infty}^{\infty} x_B(nT)\, sinc\left(\frac{t-nT}{T}\right). \tag{2.3}$$

This is a particular case of Equation (2.1) and the corresponding rate of innovation is $\rho = 1/T$. Therefore, a bandlimited signal is a signal with a finite rate of innovation since it is defined by a finite number of coefficients per unit of time.

The rate of innovation of a signal can be finite but is not necessarily constant with

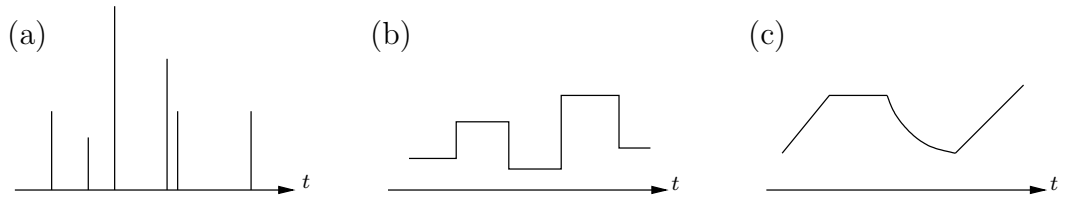


Figure 2.1: Examples of 1-D signals with finite rate of innovation; (a) streams of Diracs; (b) piecewise-constant signals; (c) piecewise-polynomial signals.

time. For this reason, one can define the local rate of innovation at time t over a moving window of size ℓ [126]:

$$\rho_\ell(t) = \frac{1}{\ell} C_x\left(t - \frac{\ell}{2}, t + \frac{\ell}{2}\right). \tag{2.4}$$

As $\ell \to \infty$, the local rate of innovation $\rho_\ell(t)$ tends to ρ. When designing a local reconstruction scheme for FRI signals, the local rate of innovation plays a particularly important role.

The first examples of FRI signals studied during the development of this sampling theory are streams of Diracs, streams of differentiated Diracs and piecewise polynomials (see Figure 2.1). Following Equation (2.1), a ν-periodic stream of Diracs with K Diracs per period can be expressed as:

$$x(t) = \sum_{n\in\mathbb{Z}} \sum_{k=0}^{K-1} a_k \delta(t - t_k - n\nu). \tag{2.5}$$

Since the k-th Dirac is described by its amplitude a_k and its location $t_k \in [0, \nu]$, the periodic stream of Diracs $x(t)$ has a rate of innovation equal to: $\rho = 2K/\nu$. Similarly, a ν-periodic stream of K differentiated Diracs per period can be written as:

$$x(t) = \sum_{n\in\mathbb{Z}} \sum_{k=0}^{K-1} \sum_{r=0}^{R_k-1} a_{k,r} \delta^{(r)}(t - t_k - n\nu), \tag{2.6}$$

where $\delta^{(r)}$ represents the r-th derivative of a Dirac which is defined as a distribution that satisfies the following property:

$$\int_{-\infty}^{\infty} f(t)\, \delta^{(r)}(t - t_0)\, dt = (-1)^r \frac{d^r f}{dt^r}(t_0).$$

Such a signal consists of K differentiated Diracs per period and $\sum_{k=0}^{K-1} R_k$ weights per period. It thus has a finite rate of innovation of $\rho = \left(K + \sum_{k=0}^{K-1} R_k\right)/\nu$. Finally, ν-periodic piecewise polynomials signals with K pieces per period with maximum polynomial degree $R-1 \geq 0$ have the following expression:

$$x(t) = \sum_{n\in\mathbb{Z}} \sum_{k=0}^{K-1} \sum_{r=0}^{R-1} a_{k,r} \left(t - t_k - n\nu\right)_+^r . \tag{2.7}$$

The R-th derivative of this signal leads to a stream of K differentiated Diracs, so the rate of innovation of a K-piecewise polynomial signal of order R is given by $\rho = \left(K + \sum_{k=0}^{K-1} R_k\right)/\nu$.

2.2.2 Sampling Setup

Sampling theory is concerned with the perfect reconstruction of a continuous signal from its discrete representation obtained by sampling. The sampling setup presented in Figure 2.2 is a typical sampling architecture composed of pre-filtering, sampling and reconstruction steps.

The impulse response $r(t)$ of the pre-filtering is the time-reversed version of the sampling kernel $\varphi(t)$. The signal $y(t)$ resulting from the convolution of $x(t)$ with $r(t)$ is then uniformly sampled at a period T to give the samples $\mathbf{y}[n], n \in \mathbb{Z}$. The samples can be equivalently expressed as:

$$\begin{aligned} \mathbf{y}[n] &= x(t) * r(t)\,|_{t=nT} \\ &= \int_{-\infty}^{\infty} x(t)\,\varphi(t/T - n)\,dt \\ &= \langle x(t), \varphi(t/T - n)\rangle . \end{aligned} \tag{2.8}$$

It is in this classical sampling setup that FRI signals are considered. We assume that we know perfectly the sampling kernel $\varphi(t)$ and the sampling period T. The sampling theory for FRI signals takes full advantage of this knowledge in the reconstruction procedures. There does not exist a single reconstruction algorithm for all FRI signals. Thus, given a particular type of FRI signal (*e.g.* stream of Diracs or piecewise polynomial

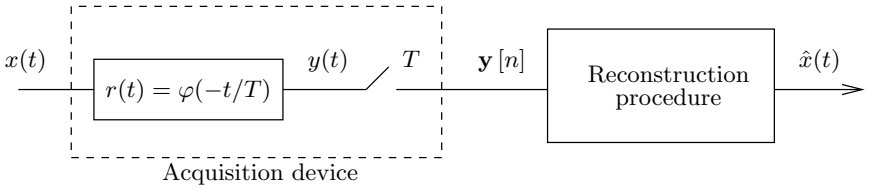


Figure 2.2: 1-D sampling setup. The input signal $x(t)$ is a 1-D continuous FRI signal; the impulse response of the acquisition device with a sampling period T is $r(t)$; the acquired samples are: $\mathbf{y}[n] = \langle x(t), \varphi(t/T - n)\rangle$. The reconstructed signal $\hat{x}(t)$ from the samples $\mathbf{y}[n]$ is equal to the original signal $x(t)$ in case of perfect reconstruction. The reconstructed procedure varies depending on the type of FRI signal.

signal), it is important to determine which sampling kernels can be employed and the corresponding reconstruction method. In-depth coverage and analysis of sampling kernels is provided in Section 2.4. At the moment we just mention that the sampling kernels we are considering have a compact support and are able to reproduce exactly polynomials or exponentials as introduced in [26]. This means that for a given kernel, there exists a set of coefficients $\left\{c_n^{(p)}\right\}_{p=0,1,\ldots,P}$ such that a linear combination of $\varphi(t)$ can either reproduce polynomials up to a certain degree P:

$$\sum_n c_n^{(p)} \varphi(t-n) = t^p, \tag{2.9}$$

or a finite set of exponentials $\{\alpha_p \in \mathbb{C}\}_{p=0,1,\ldots,P}$:

$$\sum_n c_n^{(p)} \varphi(t-n) = e^{\alpha_p t}, \tag{2.10}$$

where we consider that $\alpha_p = \alpha_0 + p\lambda$ and $\lambda \in \mathbb{C}$. It is also possible to define functions $\varphi(t)$ that are able to reproduce polynomials and exponentials [122].

2.2.3 Annihilating Filter Method and Reconstruction Procedure

The first part of this section is dedicated to the annihilating filter method as this technique is at the heart of the sampling theory for FRI signals. We then consider for the sake of clarity the case of reconstructing a stream of Diracs from its sampled version. This case constituted the driving motivation behind the emergence of the sampling theory for FRI signals and also provides much insight on the reconstruction procedures for other types of

FRI signals. Although the initial sampling kernels considered were the sinc and Gaussian kernels with infinite support [126], we consider here exclusively sampling kernels $\varphi(t)$ with compact support of length L as in [27].

Annihilating Filter Method

The annihilating filter method originates from the work of Baron de Prony on powersum series [22] and is thus sometimes referred as Prony's method in certain fields. A powersum series is defined as:

Definition 2. *A Powersum series of K components is given by:*

$$\tau_p = \sum_{k=0}^{K} a_k \left(u_k\right)^p, \quad p = 0, 1, \ldots, P, \tag{2.11}$$

where $\{a_k, u_k\} \in \mathbb{C}$ are K pairs of unknown parameters and $P+1$ is the number of available measurements.

Baron de Prony showed that the unknown parameters $\{a_k, u_k\}$ in such powersum series can be exactly calculated provided that the number P of available consecutive measurements τ_p is at least $2K-1$. The annihilating filter method is a widely used tool in high resolution spectral estimation or error correction coding. Vetterli *et al* introduced the use of the annihilating filter method in the case of sampling of FRI signals [126] as this method offers a powerful tool to isolate the innovative part of such signals.

Let τ_p be a powersum series as in Equation (2.11) where the parameters a_k and u_k are sought and K is known. The annihilating filter $B(z)$ is defined such that its roots are at the locations u_k:

$$B(z) = \sum_{p=0}^{K} b_p z^{-p} = \prod_{k=0}^{K-1} \left(1 - u_k z^{-1}\right).$$

This filter $B(z)$ has a finite impulse response b_p of length $K+1$. We then have:

$$b_p * \tau_p = \sum_{i=0}^{K} b_i \tau_{p-i} = \sum_{k=0}^{K-1} a_k u_k^p \underbrace{\sum_{i=0}^{K} b_i u_k^{-i}}_{=0} = 0. \tag{2.12}$$

The signal τ_p is thus annihilated by the filter $B(z)$. Assuming $b_0 = 1$, the K remaining

coefficients of $B(z)$ can be retrieved from Equation (2.12) if any $2K$ consecutive values of τ_p are known:

$$\begin{bmatrix} \tau_{K-1} & \tau_{K-2} & \cdots & \tau_0 \\ \tau_K & \tau_{K-1} & \cdots & \tau_1 \\ \vdots & \vdots & \ddots & \vdots \\ \tau_{2K-1} & \tau_{2K-2} & \cdots & \tau_{K-1} \end{bmatrix} \cdot \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_K \end{bmatrix} = - \begin{bmatrix} \tau_K \\ \tau_{K+1} \\ \vdots \\ \tau_{2K-1} \end{bmatrix}.$$

The filter coefficients b_i are obtained by solving the above Toeplitz system and the parameters u_k are obtained by finding the roots of $B(z)$. The a_k's are then determined by solving Equation (2.11) with K of the known values τ_p. This leads to solving the following Vandermonde system:

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ u_0 & u_1 & \cdots & u_{K-1} \\ \vdots & \vdots & \cdots & \vdots \\ u_0^{K-1} & u_1^{K-1} & \cdots & u_{K-1}^{K-1} \end{bmatrix} \cdot \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{K-1} \end{bmatrix} = \begin{bmatrix} \tau_0 \\ \tau_1 \\ \vdots \\ \tau_{K-1} \end{bmatrix}.$$

Such a system has always a unique solution if $u_k \neq u_l, \forall k \neq l$.

Reconstruction of a Stream of Diracs

Let $x(t)$ be a stream of Diracs defined by $\{a_k, t_k\}$ (respectively the amplitudes and locations of the Diracs) and which is sampled at a sampling period T by a kernel $\varphi(t)$. The sampling kernel has a compact support of size L and is able to reproduce polynomials or exponentials. The samples $\mathbf{y}[n]$ of $x(t)$ are given by:

$$\mathbf{y}[n] = \langle \varphi(t/T - n), x(t) \rangle.$$

In [27], it was shown that a stream of Diracs can be perfectly reconstructed from its samples $\mathbf{y}[n]$ as long as there are at most K Diracs in an interval of length $2KLT$.

The way to retrieve the parameters $\{a_k, t_k\}$ requires to compute the linear combinations of the samples $\mathbf{y}[n]$ with the coefficients $c_n^{(p)}$ used for polynomial reproduction as

in Equation (2.9) or for exponential reproduction as in Equation (2.10). In the case of coefficients $c_n^{(p)}$ used for polynomial reproduction, we have ($T = 1$):

$$\begin{aligned}
\tau_p &= \sum_n c_n^{(p)} \mathbf{y}[n] \\
&= \left\langle x(t), \sum_n c_n^{(p)} \varphi(t-n) \right\rangle && \text{from Equation (2.8),} \\
&= \int_{-\infty}^{\infty} x(t)\, t^p dt && \text{from Equation (2.9),} \\
&= \sum_{k=0}^{K-1} a_k t_k^p, \qquad p = 0, 1, \ldots, P,
\end{aligned}$$

whereas in the case of coefficients $c_n^{(p)}$ used for exponential reproduction, we have:

$$\begin{aligned}
\tau_p &= \sum_n c_n^{(p)} \mathbf{y}[n] \\
&= \left\langle x(t), \sum_n c_n^{(p)} \varphi(t-n) \right\rangle && \text{from Equation (2.8),} \\
&= \int_{-\infty}^{\infty} x(t)\, e^{\alpha_p t} dt && \text{from Equation (2.10),} \\
&= \sum_{k=0}^{K-1} a_k e^{\alpha_p t_k}, \qquad p = 0, 1, \ldots, P.
\end{aligned}$$

Notice from the third line of both expressions that, in the polynomial case, the measurements τ_p are actually the moments of the FRI signal $x(t)$ and that, in the exponential case, the τ_p are the Fourier coefficients of $x(t)$ when α_p is a purely imaginary number.

We also observe that, in both cases, the linear combination leads to measurements τ_p that can be expressed as a powersum series (see Definition 2). Thus, provided that the sampling kernel can reproduce up to $P \geq 2K - 1$ polynomials or exponentials, it is possible to find $2K$ consecutive measurements τ_p and find exactly the parameters of the K Diracs using the annihilating filter method.

2.2.4 Reconstruction of noisy FRI signals

The introduction of these new sampling schemes for FRI signals is fairly new and only two studies have recently considered the presence and the effect of noise on these reconstruction methods [79] [10].

In any real acquisition system, the presence of noise is almost inevitable. Noise can arise from many different sources and can in some cases substantially degrade the desired performances if it is not appropriately considered during processing. The perturbations

on the system have usually three main origins: channel mismatch (*e.g.* use of the wrong sampling kernel), analog noise on the input signals and digital noise on the samples. In order to compensate for the noise, the sampling rate has to be increased so that redundancy in the available data is introduced.

It is in this context that Blu *et al* analyzed the reconstruction of a ν-periodic stream of Diracs $x(t)$ in a noisy environment and with a sinc window of a given bandwidth as sampling kernel [10]. The proposed method to solve the noisy case combines a total least-squares approximation based on a singular value decomposition (see [98]) and an iterative signal enhancement algorithm based on Cadzow's work [15].

In the case of a single Dirac ($K = 1$), Blu *et al* derived the theoretical lower Cramér-Rao bounds for the minimal uncertainties on the retrieved location and amplitude of the Dirac from noisy samples using an unbiased algorithm. For $K \geq 2$, the Cramér-Rao bounds derived for one Dirac are still giving a good estimate as long as the Diracs are sufficiently far apart. They showed empirically in the case of two Diracs that the proposed reconstruction algorithm is almost optimal for noise level above 5dB of SNR power for which the computed uncertainties reach the theoretical lower Cramér-Rao bounds.

2.3 Sampling of 2-D FRI Signals

2.3.1 2-D Signals with Finite Rate of Innovation

The extension of the definition of FRI signals from 1-D to 2-D is possible and quite straightforward. Given a set of known 2-D functions $\{\phi_r(x - x_i, y - y_j)\}_{r=0\ldots R-1}$, we can define a 2-D signal with the following parametric form:

$$f(x,y) = \sum_{i,j\in\mathbb{Z}} \sum_{r=0}^{R-1} a_{i,j,r}\phi_r(x - x_i, y - y_j). \tag{2.13}$$

The only unknown parameters in the signal $f(x,y)$ are the coefficients $a_{i,j,r} \in \mathbb{R}^2$ and the spatial shifts $(x_i, y_j) \in \mathbb{R}$. Let $C_f(x_c, y_c, \ell_x, \ell_y)$ be the function that counts the number of unknown parameters in $f(x,y)$ over the window of size $\ell_x \times \ell_y$ and centered in (x_c, y_c).

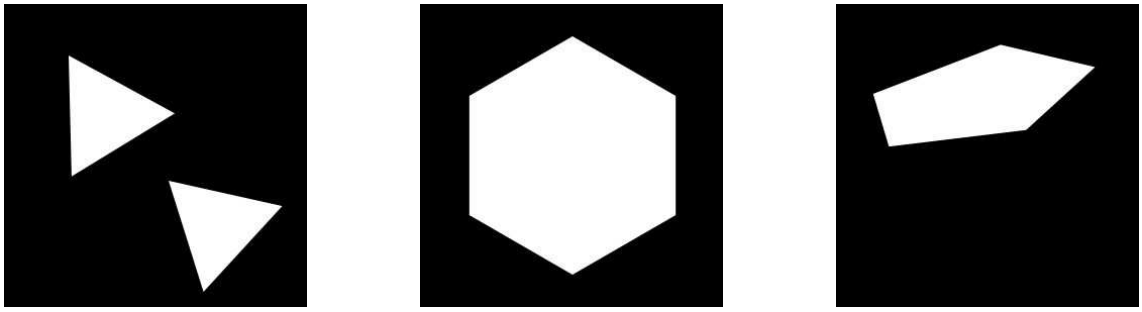


Figure 2.3: Examples of 2-D FRI signals: bilevel polygonal images. Because the polygons are convex, they are completely characterized by the location of their corners.

The local rate of innovation of $f(x,y)$ can then be defined as follows:

$$\rho_{\ell_x \ell_y} = \frac{1}{\ell_x \ell_y} C_x(x_c, y_c, \ell_x, \ell_y). \tag{2.14}$$

If $\phi_r(x,y) = \delta_r(x,y)$, the resulting 2-D FRI signal is a set of 2-D Diracs which can be considered as a particular realization of a 2-D Poisson process. Various geometrical shapes that have a finite degree of freedom also fall into the definition of FRI signals. For example, 2-D lines, convex bilevel polygons and algebraic curves like ellipsis can all be characterized by a finite number of parameters. Figure 2.3 presents examples of convex bilevel polygonal images which are completely characterized by the location of the polygon's vertices.

2.3.2 2-D Sampling Setup

The sampling setup for 2-D FRI signals is shown in Figure 2.4. The input signal $f(x,y)$ is a 2-D FRI signal and is convolved with a 2-D sampling kernel $\varphi(-x/T_x, -y/T_y)$ to give rise to a smoothed version of f denoted by $g(x,y)$. The signal $g(x,y)$ is then uniformly sampled at periods $T_x \in \mathbb{R}^+$ and $T_y \in \mathbb{R}^+$ along the x axis and y directions respectively. We obtain the set of samples $\mathbf{g}[m,n]$:

$$\begin{aligned} \mathbf{g}[m,n] &= f(x,y) * r(x,y)\,|_{(x,y)=(mT_x, nT_y)} \\ &= \int_{-\infty}^{\infty} f(x,y)\,\varphi(x/T_x - m, y/T_y - n)\,dt \\ &= \langle f(x,y), \varphi(x/T_x - m, y/T_y - n) \rangle. \end{aligned} \tag{2.15}$$

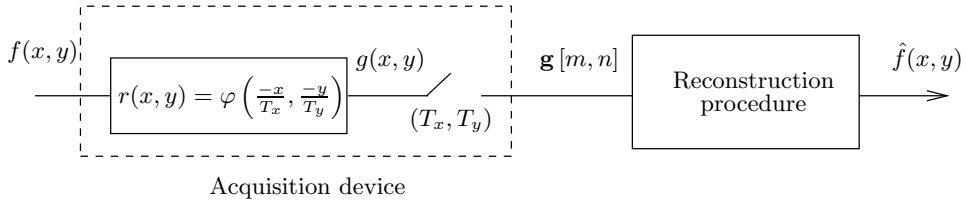


Figure 2.4: 2-D sampling setup. The input signal $f(x,y)$ is a 2-D continuous FRI signal; the impulse response of the acquisition device with a sampling period (T_x,T_y) is $r(x,y)$; the acquired samples are $\mathbf{g}[m,n]=\langle f(x,y),\varphi(x/T_x-m,y/T_y-n)\rangle$. The reconstructed signal $\hat{f}(t)$ from the samples $\mathbf{g}[m,n]$ is equal to the original signal $f(x,y)$ in case of perfect reconstruction. The reconstructed procedure varies depending on the type of FRI signal.

Unless explicitly specified, we assume $T_x=T_y=T$. As in 1-D, we assume that we have a perfect knowledge of the sampling kernel $\varphi(x,y)$ and the sampling period in the reconstruction procedures. We assume at present that the considered sampling kernels have a compact support and that they are able to reproduce 2-D polynomials exactly. Thus there exists a set of coefficients $\left\{c_{m,n}^{(p,q)}\right\}$, $p=0,1,...,P$ and $q=0,1,...,Q$ such that a linear combination of $\varphi(x,y)$ can reproduce polynomials up to a certain degree P in x direction and Q in y direction:

$$\sum_n c_{m,n}^{(p,q)}\varphi(x-m,y-n)=x^p y^q. \tag{2.16}$$

Again, unless explicitly specified, we assume $P=Q$.

2.3.3 Reconstruction Procedures

The first schemes for the sampling and perfect reconstruction of 2-D FRI signals are attributed to Maravić *et al* in [76, 78]. In the proposed methods, the Fourier coefficients are computed from the samples of a periodic set of 2-D Diracs or bilevel polygons obtained with a sinc kernel. The Fourier coefficients are then used in the annihilating filter method to calculate the innovation of the FRI signal. A second approach uses the Radon transform to project a 2-D FRI signal along several directions. Each of the projections is a 1-D FRI signal that can be perfectly recovered from its Fourier coefficients using the sampling theory for 1-D signals. Provided that enough projections are available, the 2-D FRI signal can be reconstructed using back-projection (see Figure 2.5). Thus a set of K 2-D Diracs

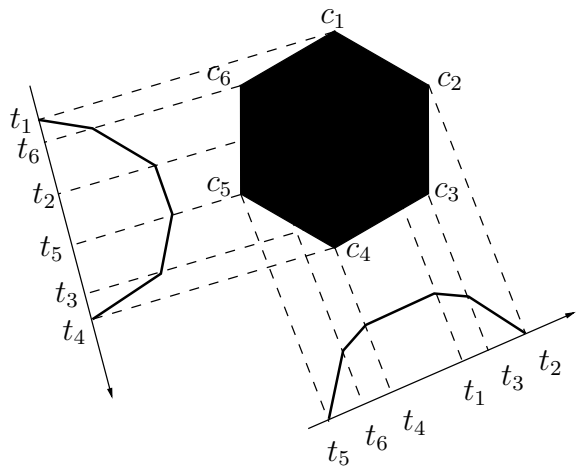


Figure 2.5: The Radon projections along different directions of a 2-D FRI signal are 1-D FRI signals. Here the projections of a bilevel convex polygon result in piecewise-linear signals. Each projection can be reconstructed independently (using Fourier coefficients or moments). By back-projecting the reconstructed FRI projections, it is possible to reconstruct the 2-D convex polygon.

requires $K+1$ projections, each one being a stream of K 1-D Diracs. Similarly, a bilevel polygon with K vertices can be reconstructed perfectly from $K+1$ projections if $2K$ samples for each projection are available. In that case, each projection consists of 1-D piecewise linear signals.

In [109, 110], Shukla *et al* considered using finite support sampling kernels that reproduce polynomials. This type of kernel is physically realizable and makes it particularly appealing. While Maravić used the Fourier coefficients, Shukla considered the moments of FRI signals and proposed three different approaches to reconstruct perfectly different 2-D FRI signals from their samples.

The first approach examines the problem of recovering planar polygons using directional derivatives. Consider one particular corner of a polygon and the two sides joined by that corner; by applying two successive finite differences on the samples in the directions of the two sides, it is shown that the location of the considered corner can be exactly retrieved from its samples.

The second approach, called the Annihilating Filter based Back-Projection algorithm (AFBP), is a Radon-based approach that makes use of the annihilating filter method. It allows the reconstruction of 2-D polynomials with convex polygonal boundaries from

the samples. It is shown that the exact moments of the Radon projection for a given orientation can be retrieved from the moments of the samples differentiated along the same direction. Using the annihilating filter method, the Radon projections can be reconstructed one by one until there are enough Radon projections to allow the reconstruction of the original 2-D polynomials with convex polygonal boundaries.

Finally, the third approach also makes use of the annihilating filter method and reconstructs bilevel polygons from the samples by determining directly the exact locations of the vertices from the complex moments. After recalling the theory of moments, this method is described in more details in the next chapter as it is used in an extension to multiview 2-D FRI signals.

2.4 Sampling Kernels

2.4.1 Kernel Properties

We now give more detail on the sampling kernels considered in the sampling theory of FRI signals. In most acquisition devices, the observed signal is usually first filtered before sampling. Very often, this filter corresponds to intrinsic physical properties of the acquisition device and it is therefore very difficult, if not impossible, to design it freely. The sampling kernel is the time reversed version of the filter's impulse response of the acquisition device. It is therefore important to consider a class of kernels as general as possible in order to accommodate many types of devices. In [27], a wide range of kernels $\varphi(t)$ are proposed to study FRI signals and are divided into three families:

1. Polynomial reproducing kernels: the set of shifted versions of $\varphi(t)$ can reproduce polynomials up to a certain degree P:

$$\sum_n c_n^{(p)} \varphi(t-n) = t^p, \quad p = 0, \dots P, \tag{2.17}$$

for given coefficients $c_m^{(p)}$.

2. Exponential reproducing kernels: the set of shifted versions of $\varphi(t)$ can reproduce

up to $P+1$ complex exponentials $\alpha_p \in \mathbb{C}$, $p = 0, 1, \ldots, P$:

$$\sum_n c_n^{(p)} \varphi(t-n) = e^{\alpha_p t}, \quad \alpha_p = \alpha_0 + p\lambda, \quad \lambda \in \mathbb{C}, \tag{2.18}$$

for given coefficients $c_m^{(p)}$.

3. Rational kernels whose Fourier transform can be written with the following rational form:

$$\hat{\varphi}(\omega) = \frac{\prod_{i=0}^{I}(j\omega - b_i)}{\prod_{p=0}^{P}(j\omega - \alpha_p)}, \quad I < P, \quad \alpha_p = \alpha_0 + p\lambda, \quad p = 0, 1, \ldots, P, \tag{2.19}$$

where $\hat{\varphi}(\omega)$ is the Fourier transform of $\varphi(t)$

The parameter P here will depend on the rate of innovation of the signal under observation. In this thesis, we only consider the first two families and we focus in particular on polynomial reproducing kernels.

In those two cases, the set of shifted functions $\varphi(t-n)$ spans the shift-invariant function space $V(\varphi)$ defined as:

$$V(\varphi) : \left\{ f_V(t) = \sum_{n=-\infty}^{\infty} c_n \varphi(t/T - n) : c_n \in l^2 \right\}.$$

Any function $f_V(t)$ in $V(\varphi)$ is completely characterized by a unique discrete time sequence c_n which can be considered as its coordinates in the basis formed by the integer shifts of φ. The space $V(\varphi)$ is a closed subspace of L^2 if $\varphi(t-n)$ forms a Riesz basis [2]:

$$A\|c\|_{l^2}^2 \leq \left\| \sum_{n=-\infty}^{\infty} c_n \varphi(t-n) \right\|^2 \leq B\|c\|_{l^2}^2, \qquad B > A > 0,$$

where A and B are two constants.

Let $f(t)$ be a function in L^2 not necessarily in $V(\varphi)$, then its least-square representation $f_V(t) \in V(\varphi)$ is the orthogonal projection of $f(t)$ onto $V(\varphi)$. In this case, $f_V(t)$ is identical to $f(t)$ if and only if $f(t)$ is already in $V(\varphi)$. The basis $\varphi(t-n)$ is generally not orthogonal so the coordinates c_n of $f_V(t)$ are obtained by inner product with the dual

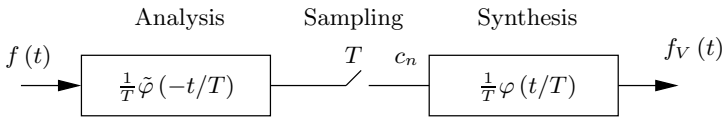


Figure 2.6: Approximation theory principles: the analysis filter $\tilde{\varphi}$ projects orthogonally the function $f(t)$ on the shift invariant space V_φ spanned by $\varphi(t/T-n)$. The projected function $f_V(t)$ has for coordinates c_n in V_φ and is only equal to $f(t)$ if $f(t) \in V_\varphi$ initially. The functions φ and $\tilde{\varphi}$ are dual functions of each other.

(biorthogonal) function $\tilde{\varphi}(t)$ (see Figure 2.6) [11]:

$$c_n = \int_{-\infty}^{\infty} f(x)\,\tilde{\varphi}(x/T-n)\,\frac{dx}{T}. \tag{2.20}$$

If the basis $\varphi(t-n)$ is orthogonal, then $\tilde{\varphi}(t) = \varphi(t)$. If not, the dual function $\tilde{\varphi}(t)$ still belongs to the same space $V(\varphi)$ and can be expressed as a linear combination of the synthesis function $\varphi(t)$ [2]:

$$\tilde{\varphi}(t) = \left((a)^{-1} * \varphi\right)(t) \quad \text{where} \quad a_n = \int_{-\infty}^{\infty} \varphi(x)\,\varphi(x-n)\,dx, \tag{2.21}$$

or equivalently in the Fourier domain:

$$\widehat{\tilde{\varphi}}(\omega) = \frac{\hat{\varphi}(\omega)}{\widehat{a}(\omega)}, \quad \widehat{a}(\omega) = \sum_n \left|\hat{\varphi}(\omega+2n\pi)\right|^2, \tag{2.22}$$

where $\widehat{\tilde{\varphi}}$ and $\widehat{\varphi}$ are the Fourier transforms of $\tilde{\varphi}$ and φ respectively.

In [114], Strang and Fix enounced the necessary and sufficient conditions for a function $\varphi(t)$ to be able to reproduce a polynomial up to a degree P:

$$\begin{cases} \widehat{\varphi}(0) \neq 0 \quad \text{and} \\ \frac{d^{(p)}\widehat{\varphi}(2i\pi)}{d\omega^p} = 0 \quad \text{for} \quad i \in \mathbb{Z}^*, \quad p = 0, \ldots, P, \end{cases} \tag{2.23}$$

where $\widehat{\varphi}(\omega)$ is the Fourier transform of $\varphi(t)$. These conditions are referred as the Strang-Fix conditions of order $P+1$ and are well-known in the wavelet theory in which the approximation order of a scaling function is directly related to the number of vanishing moments of a wavelet. There exists a variety of functions satisfying Strang-Fix conditions. On Figure 2.7, a Daubechies scaling function of order 6 (orthogonal function) is used

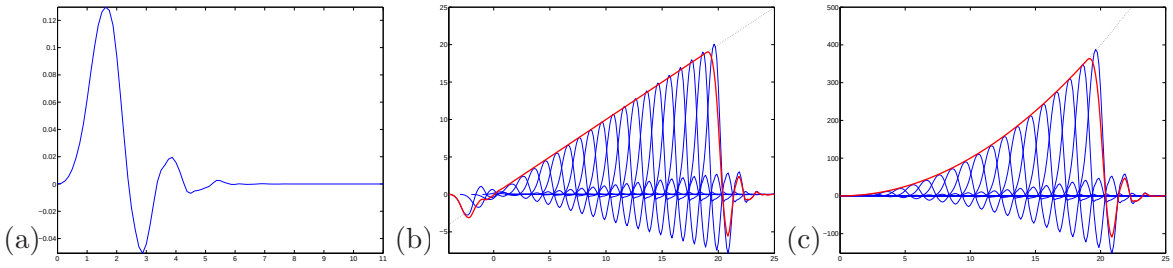


Figure 2.7: (a) Daubechies scaling function of order 6; (b) The summation of several scaled and shifted versions of the Daubechies function (in blue) can reproduce locally the monomial t (reproduction is the red curve and the actual monomial is the dotted black curve); (c)The summation of several scaled and shifted versions (in blue) of the Daubechies function can reproduce locally (in red) the monomial t^2 (in black).

to reproduce the polynomials t and t^2. The space V_φ is the shift invariant polynomial subspace of degree P and Equation (2.20) becomes:

$$c_n^{(p)} = \frac{1}{T}\int_{-\infty}^{\infty} t^p \tilde{\varphi}\left(t/T - n\right) dt \qquad p = 0, \ldots, P. \tag{2.24}$$

We notice that the coefficients $c_n^{(p)}$ are the moments of order p of the dual function $\tilde{\varphi}\left(t - n\right)$. Let $\tilde{m}_p$ be the moments of the centered dual function $\tilde{\varphi}\left(t\right)$:

$$\tilde{m}_p = \int_{-\infty}^{\infty} t^p \tilde{\varphi}\left(t\right) dt, \quad p \in \mathbb{Z}.$$

The knowledge of the first $P+1$ moments $\tilde{m}_p$ is in theory sufficient to compute all the coefficients $c_n^{(p)}$ since Equation (2.24) can be rewritten by using the binomial theorem as follows:

$$c_n^{(p)} = T^p \sum_{k=0}^{p} \binom{p}{k} n^{p-k} \tilde{m}_k, \qquad p = 0, \ldots, P. \tag{2.25}$$

Thus, in addition to the biorthogonal case, it is possible to define a less constraining basis $\tilde{\varphi}_q\left(t\right)$ called quasi-biorthogonal basis. In this scheme, the analysis function $\tilde{\varphi}_q\left(t\right)$ can be any function as long as its first $P+1$ moments are equal to the moments $\tilde{m}_p$ of $\tilde{\varphi}\left(t\right)$ for $p = 0, \ldots, P$:

$$\int_{-\infty}^{\infty} t^p \tilde{\varphi}_q\left(t\right) dt = \tilde{m}_p.$$

Similar results can be derived for two dimensional kernels $\varphi(x,y)$. In particular, we consider 2-D kernels $\varphi(x,y)$ that are variable separable. Such kernels are obtained by tensor product of two 1-D functions $\varphi_1(x)$ and $\varphi_2(y)$:

$$\varphi(x,y) = \varphi_1(x) \times \varphi_2(y).$$

Computing the coefficients $c_{m,n}^{(p,q)}$ for reproducing polynomials $x^p y^q$ with $\varphi(x,y)$ is in this case simply obtained by multiplying the coefficients $c_m^{(p)}$ and $c_n^{(q)}$ for $\varphi_1(x)$ and $\varphi_2(y)$ as follows:

$$\sum_{m,n\in\mathbb{Z}} c_{m,n}^{(p,q)} \varphi_1(x/T - m)\,\varphi_2(y/T - n) = x^p y^q \quad p = 0,\ldots,P, \quad q = 0,\ldots,Q, \tag{2.26}$$

where

$$c_{m,n}^{(p,q)} = c_m^{(p)} c_n^{(q)}.$$

The theory related to exponential reproduction is quite recent and relies at the moment on a single family of functions called exponential spline, or E-splines [122]. These functions are treated in more detail in Section 2.4.3.

2.4.2 B-spline kernels

An important family of functions that reproduce polynomials is the family of polynomials spline functions [119]. A polynomial spline is a continuous piecewise polynomial curve and is characterized by the maximum degree of the polynomials constituting its segments. There is some confusion on this matter in the literature: mathematicians usually define the order of a spline in the sense of its approximation order, *i.e.* a spline of order $P+1$ can reproduce polynomials from degree 0 to P and each segment of the spline is therefore made of a polynomial of degree P [106]; researchers in the signal processing community define the order of a spline as the maximum degree of the polynomial of each segment so that a spline of order P can reproduce polynomials from degree 0 to P [120]. To avoid confusion, this research has adopted this last convention and we refer indifferently to the order or the degree of a spline as the maximum degree of the polynomials constituting its

pieces.

The segments of a spline of order P are connected at points called knots (or nodes) and are such that the $(P-1)$-th derivatives of the spline are continuous (*i.e.* a function of class C^{P-1}). Although the space between knots can be nonuniform, in this work we consider only splines with knots distant by one unit. In this framework, the splines of odd orders have knots located at integers $\{i\}, i \in \mathbb{Z}$, whereas for spline of even orders, they are located at half of integers $\left\{i + \frac{1}{2}\right\}, i \in \mathbb{Z}$. For each order P, there exists a unique polynomial spline with a compact support of shortest size $P+1$ and centered in 0. This function $\beta^P(t)$ is symmetrical and is called centered B-spline of order P. This function is originally defined as the evaluation of a certain integral [106]:

$$\beta^P(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \left(\frac{2 \sin u/2}{u} \right)^{P+1} e^{iut} du \quad P = 0, 1, \ldots, \tag{2.27}$$

which is the inverse Fourier transform of the sinc function $\mathrm{sinc}(x) = \sin(x)/x$ elevated to the power $P+1$. This function can also be written in the time domain as:

$$\beta^P(t) = \sum_{j=0}^{P+1} \frac{(-1)^j}{P!} \binom{P+1}{j} \left(t + \frac{P+1}{2} - j \right)^P \cdot u\left(t + \frac{P+1}{2} - j \right), \quad t \in \mathbb{R},$$

where $u(t)$ is the unit step function:

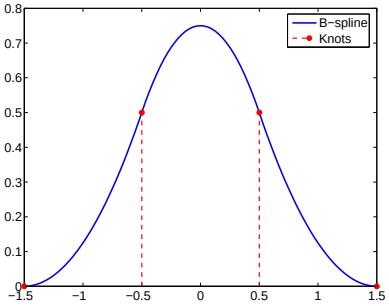
$$u(t) = \begin{cases} 0, & t < 0 \\ 1, & t \geq 0 \end{cases}$$

The B-splines of order 2 and 3 are plotted in Figure 2.8 together with the location of the knots. From its definition in Equation (2.27), the Fourier transform $B^P(\omega)$ of $\beta^P(t)$ is given by:

$$\beta^P(t) \quad \Leftrightarrow \quad B^P(f) = [\mathrm{sinc}(f)]^{P+1}.$$

It can be verified that the B-splines satisfy Strang-Fix conditions (Equations (2.23)):

$$\begin{cases} B^P(0) = 1 \\ \frac{d^p B^P(2\pi k)}{df^p} = 0, \quad k \in \mathbb{Z}, \quad p = 0, 1, \ldots, P. \end{cases}$$

(a) 

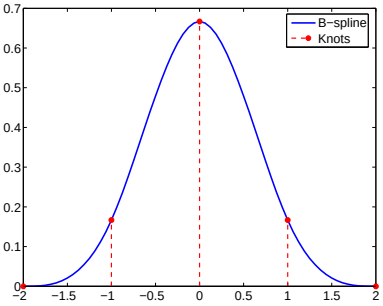
(b) 

Figure 2.8: (a) B-spline of order 2: this function has a support of size 3 and has knots at locations $t_{\mathbf{knots}} = [-1.5, -0.5, 0.5, 1.5]$; (c) B-spline of order 3: this function has a support of size 4 and has knots at locations $t_{\mathbf{knots}} = [-2, -1, 0, 1, 2]$. Each segment between two knots is a polynomial of same degree as the order of the considered B-spline.

The B-splines also satisfy Riesz conditions so the shifted versions of B-splines form a basis of a closed vector space $V\left(\beta^P\right)$ which comprises all spline functions $s\left(t\right)$ of degree P with unit spacing between knots:

$$V\left(\beta^P\right) : \left\{ s(t) = \sum_{n=-\infty}^{\infty} c_n \beta^P (t-n) : c \in l^2 \right\}.$$

By scaling, or expanding the B-spline by a factor T, it is possible to obtain a multiscale representation and to design spline subspaces at different scales [121]:

$$V\left(\beta_T^P\right) : \left\{ s_T(t) = \sum_{n=-\infty}^{\infty} c_n \beta^P (t/T-n) : c \in l^2 \right\}.$$

When scaling by a sampling period T, the spacing between the knots of $\beta_T^P\left(t\right) = \beta^P(t/T)$ is equal to T. As observed from its Fourier transform, the B-splines can be recursively obtained by convolution of the box spline $\beta^0\left(t\right)$:

$$\beta^P(t) = \underbrace{\beta^0(t) * \ldots * \beta^0(t)}_{P+1 \text{ times}} \quad , \quad \beta^0(t) = \begin{cases} 1, & |t| < \frac{1}{2} \\ \frac{1}{2}, & |t| = \frac{1}{2} \\ 0, & \text{otherwise} \end{cases} \tag{2.28}$$

Discrete B-spline kernels $b_T^P\left(k\right)$ are obtained by sampling the continuous B-spline functions $\beta_T^P(t)$:

$$b_T^P\left(k\right) = \beta^P\left(t/T\right))\big|_{t=k}, \quad k \in \mathbb{Z}.$$

It is possible to generate discrete B-splines by using a convolution procedure similar to Equation (2.28) but one has to be careful about the parity of the order and scale of the desired B-spline to implement it properly [120]. Thus we have the following:

- T odd > 1:

$$b_T^P(k) = \frac{1}{T^P}(\underbrace{b_T^0 * b_T^0 * \ldots * b_T^0}_{\text{P+1 times}}) * b_1^P(k)$$

- P odd and T even > 1:

$$b_T^P(k) = \frac{1}{T^P} \cdot \delta_{\frac{P+1}{2}} * (\underbrace{b_T^0 * b_T^0 * \ldots * b_T^0}_{\text{P+1 times}}) * b_1^P(k)$$

- P even and T even > 1:

$$b_T^P(k) = \frac{1}{T^P} \cdot \delta_{\frac{P+2}{2}} * (\underbrace{b_T^0 * b_T^0 * \ldots * b_T^0}_{\text{P+1 times}}) * c_1^P(k)$$

where $c_T^P(k)$ is the discrete B-spline shifted by 1/2:

$$c_T^P(k) = \beta^P(k/T + \frac{1}{2}),$$

and $\delta_i(k)$ is the shift operator such that $\delta_i * s(k) = s(k-i)$. The starting conditions are given by:

$$b_T^0(k) = \begin{cases} 1 & \text{for } T/2 \leq k \leq T/2 \\ 0 & \text{otherwise} \end{cases},$$

$$c_T^0(k) = \begin{cases} 1 & \text{for } 1-T \leq k \leq 0 \\ 0 & \text{otherwise} \end{cases},$$

and the following recursive equations are used to evaluate the values of $b_1^P(k)$ and $c_1^P(k)$ at higher orders:

$$b_T^P(k) = \frac{(\frac{k}{T} + \frac{P+1}{2})c_T^{P-1}(k) + (\frac{P+1}{2} - \frac{k}{T})c_T^{P-1}(k-T)}{P},$$

$$c_T^P(k) = \frac{(\frac{k}{T} + \frac{P+2}{2})b_T^{P-1}(k+T) + (\frac{P}{2} - \frac{k}{T})b_T^{P-1}(k)}{P}.$$

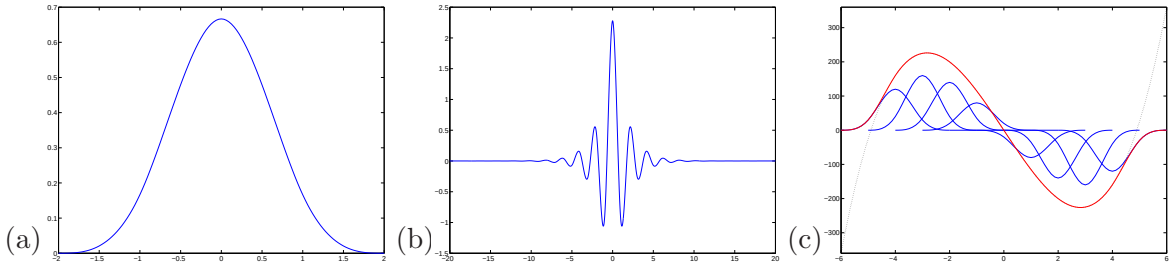


Figure 2.9: (a) cubic B-spline; (b) cubic dual B-spline; (c) The summation of several scaled and shifted versions (in blue) of the cubic B-spline can reproduce locally (in red) the polynomial $t^3 - 120t$ (in black).

The set of shifted versions of the B-spline is not an orthogonal basis. It is therefore necessary to define the dual B-spline function $\tilde{\beta}^P(t)$ in order to have a set of biorthogonal basis functions for the subspace $V\left(\beta^P\right)$. Using Equations (2.21), it is possible to show that for any spline $s\left(t\right) = \sum_n d_n \beta^P\left(t-n\right)$, there exists a dual spline function $\tilde{s}\left(t\right)$ expressed as [121]:

$$\tilde{s}\left(t\right) = \sum_{n=-\infty}^{\infty} \left(d * b_1^{2P+1}\right)^{-1}\left(n\right) \cdot \beta^P\left(t-n\right),$$

where:

$$\left(d * b_1^{2P+1}\right)^{-1}\left(n\right) \leftrightarrow \left(D\left(z\right) B^{2\rho+1}\left(z\right)\right)^{-1}. \tag{2.29}$$

Here, $D\left(z\right)$ and $B^{2\rho+1}\left(z\right)$ are respectively the z-transforms of $d\left[n\right]$ and $b_1^{2\rho+1}\left[n\right]$. The dual B-spline $\tilde{\beta}^P(t)$ is thus defined as

$$\tilde{\beta}^P(t) = \sum_{n=-\infty}^{\infty} \left(b_1^{2P+1}\right)^{-1}\left(n\right) \cdot \beta^P\left(t-n\right).$$

Figure 2.9(a) and (b) show respectively a cubic B-spline and its dual. Figure 2.9(c) presents the reproduction of the polynomial $x^3 - 120x$ over the interval $[-4, 4]$ with shifted cubic B-splines. 2-D B-spline kernels are obtained by tensor product of two 1-D B-spline functions. Figure 2.10(a) presents a 2-D B-spline kernel obtained by tensor product of two cubic B-splines. Such a kernel is rotationally symmetric as both underlying functions are symmetric. Finally Figure 2.10(b) shows the reproduction of the polynomial plane x^3 using this kernel.

The B-spline kernels are interesting functions in image acquisition. When an ac-

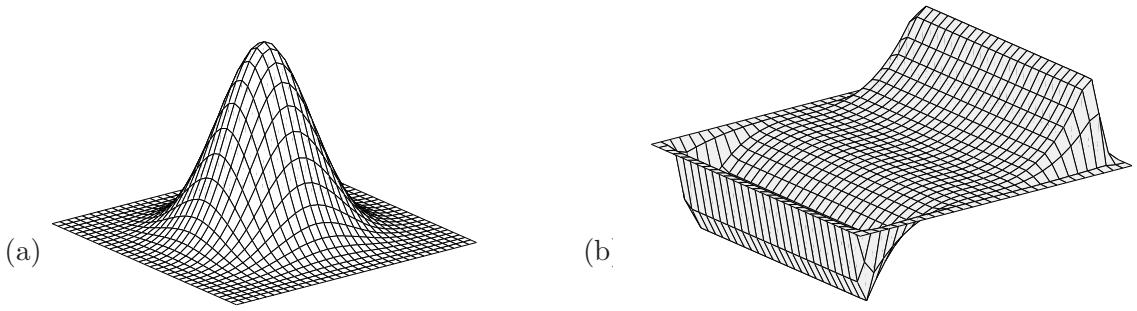


Figure 2.10: (a) 2-D cubic B-spline; (b) reproduction of the plane x^3 (Different grid scale).

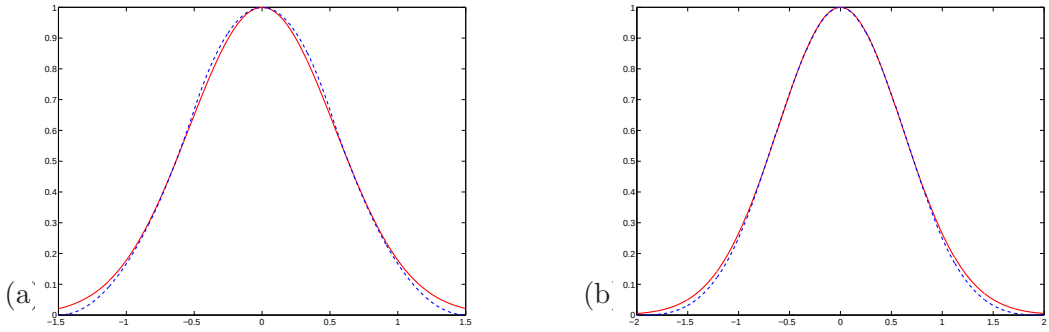


Figure 2.11: Comparison of B-splines (dashed line) and Gaussians (solid line). (a) quadratic B-spline ($\rho = 2$); (b) cubic B-spline ($\rho = 3$).

quisition device is modeled, the point-spread function of the lens is very often assumed to be a Gaussian pulse. As mentioned in [121], B-splines of order $P \geq 2$ are increasingly similar to Gaussian functions. It can also be demonstrated that for $P \to \infty$, the B-spline becomes a Gaussian curve. This similarity is shown in the two plots of Figure 2.11 where a scaled B-spline of order 2 and a scaled B-spline of order 3 are compared to a Gaussian pulse. The variance of each Gaussian was calculated so that the curve has the same full width at half maximum as the B-spline they are compared with.

2.4.3 E-spline kernels

The theory of functions reproducing exponentials is rather recent and is based on the theory of exponential splines (E-splines) developed by Unser *et al* [122]. An exponential

spline is a piecewise exponential curve connected at knots in such a way that it is continuous with its $(P-1)$ derivatives. As with a B-spline, an E-spline is defined by its order P but it is also characterized by $(P+1)$ complex parameters $\underline{\alpha} = (\alpha_0, \cdots, \alpha_P) \in \mathbb{C}^p$. These parameters correspond to the exponentials that an E-spline $\beta_{\underline{\alpha}}(t)$ can actually reproduce. For a given order P, the E-spline is the shortest possible exponential spline with support equal to $(P+1)$. The set of shifted E-splines $\{\beta_{\underline{\alpha}}(t-k)\}_{k\in\mathbb{Z}}$ constitutes a biorthogonal (Riesz) basis to the exponential spline vector space $V_{\underline{\alpha}}$ defined as:

$$V_{\underline{\alpha}} = \left\{ s\left(t\right) = \sum_{n\in\mathbb{Z}} c_n \beta_{\underline{\alpha}}\left(t-k\right) | c_n \in l_2 \right\}.$$

The E-spline $\beta_{\alpha_0}(t)$ of order 0 with a single parameter α and its Fourier transform $B_\alpha\left(\omega\right)$ are defined as:

$$\beta_\alpha(t) = \begin{cases} e^{\alpha t}, & 0 \le t < 1, \\ 0, & \text{otherwise.} \end{cases} \quad \longleftrightarrow \quad B_\alpha\left(\omega\right) = \frac{1-e^{\alpha - j\omega}}{j\omega - \alpha}.$$

The construction of E-splines of a higher order with parameter $\underline{\alpha} = (\alpha_0, \cdots, \alpha_P)$ is very similar to the construction of B-splines, as it is obtained by successive convolutions of the zero-th order E-splines with the desired parameters:

$$\beta_{\underline{\alpha}}(t) = \beta_{\alpha_0} * \beta_{\alpha_1} * \ldots * \beta_{\alpha_P}(t).$$

The Fourier transform of an E-spline of order P is thus given by:

$$B_{\underline{\alpha}}\left(\omega\right) = \prod_{p=0}^{P} \frac{1-e^{\alpha_p - j\omega}}{j\omega - \alpha_p}.$$

We now see that the E-splines are natural extensions of the B-splines. If the $P+1$ desired parameters are all zeros, *i.e.* $\underline{\alpha} = \mathbf{0}$, then the obtained function is a B-spline of order P. Interestingly, if the E-spline of order P has a single parameter α of multiplicity $P+1$, then the E-spline reproduces exponential polynomials of the type $t^p e^{\alpha t}$, $p = 0, \ldots, P$ which is a generalized version of the polynomial B-spline. The E-spline of order 0 with real parameter $\alpha_0 = -0.1$ and the E-spline of order 1 with parameters $\underline{\alpha} = (-0.05 - 0.5j, -0.5j)$ are

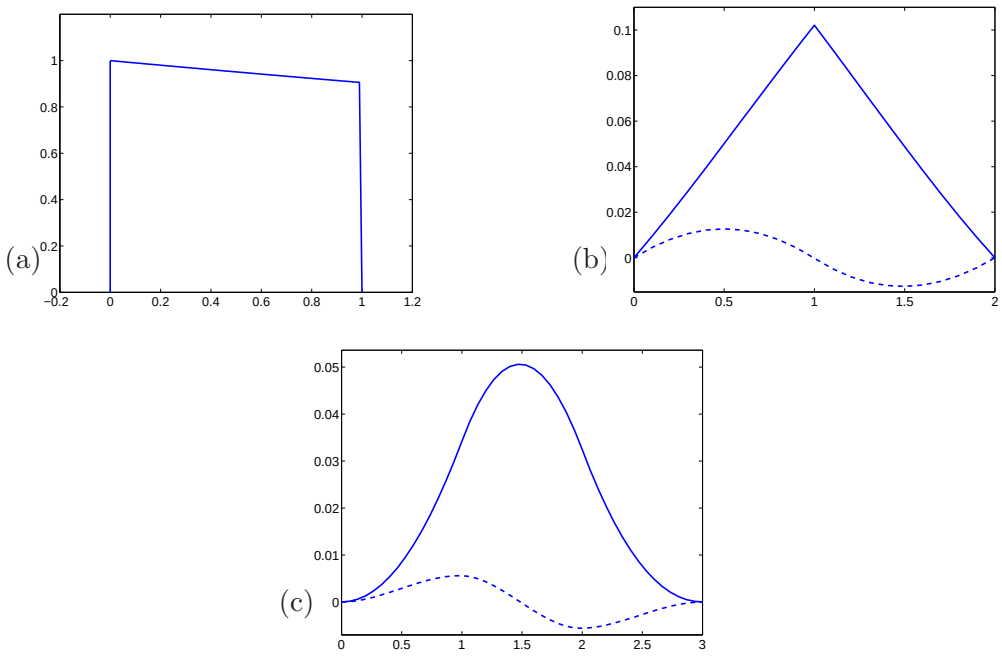


Figure 2.12: (a) Real E-spline of order 0 and parameter $\alpha_0 = -0.1$; (b) Complex E-spline of order 1 with parameters $\alpha_0 = -0.05 - 0.5j$, $\alpha_1 = -0.5j$; (c) Complex E-spline of order 2 obtained by convolution of the two previous E-splines ($\alpha_0 = -0.05 - 0.5j$, $\alpha_1 = -0.5j$, $\alpha_2 = -0.1$). The solid line corresponds to the real part and the dashed line to the imaginary part.

sketched in Figure 2.12 (a) and (b) respectively. In Figure 2.12 (c), the E-spline of order 2 obtained by convolution of the two previous E-splines is shown. This complex E-spline has three parameters $\underline{\alpha} = (-0.05 - 0.5j, -0.5j, -0.1)$.

The reconstruction of FRI signals like streams of Diracs with an E-spline sampling kernel requires the necessary coefficients $c_n^{(q)}$ to be found in order to be able to reproduce exponentials (see Equation (2.10)). The method to find the coefficients $c_n^{(q)}$ is however more challenging than in the B-spline case since the dual E-spline is not well defined in the literature. One way to overcome this difficulty is to solve the set of equations in order to find the right coefficients. Suppose we want to reproduce the exponential e^{α_p} with an E-spline of order P:

$$e^{\alpha_p t} = \sum_n c_n^{(i)} \beta_{\underline{\alpha}}(t - n),$$

then for $t = 1$:

$$\begin{aligned} e^{\alpha_p} &= \sum_n c_n^{(p)} \beta_{\underline{\alpha}}(1-n) \\ &= \sum_{n=-\lfloor (P+1)/2 \rfloor}^{\lfloor (P+1)/2 \rfloor - 1} c_n^{(p)} \beta_{\underline{\alpha}}(1-n) \end{aligned} \tag{2.30}$$

since the size support of the E-spline $\beta_{\underline{\alpha}}(t)$ is $P+1$ and it is equal to 0 at 0 and $P+1$. We then observed that :

$$c_n^{(p)} = e^{n\alpha_p} \cdot c_0^{(p)}. \tag{2.31}$$

By using Equation (2.31) in Equation Equation (2.30) and by solving for $c_0^{(p)}$, we obtain finally:

$$c_n^{(p)} = \begin{cases} e^{n\alpha_p} \cdot c_0^{(p)}, & n \neq 0, \\ e^{\alpha_p} \Big/ \sum_{k=-\lfloor (P+1)/2 \rfloor}^{\lfloor (P+1)/2 \rfloor - 1} e^{k\alpha_p} \beta_{\underline{\alpha}}(1-k), & n = 0. \end{cases} \tag{2.32}$$

Note that this equation requires the value of $\beta_{\underline{\alpha}}(t)$ at integers $k = 1, \ldots, P$ to be known. This can either be obtained by adequately convolving discrete E-splines of order 0, or by using the analytical expression of the desired E-spline as given below for the order 0, 1, and 2:

- $\beta_{\alpha_0}(t) = e^{\alpha_0 t}, \qquad t \in [0, 1[.$

- $\beta_{(\alpha_0,\alpha_1)}(t) = \begin{cases} \frac{e^{\alpha_0 t} - e^{\alpha_1 t}}{\alpha_0 - \alpha_1}, & t \in [0, 1[, \\ \frac{e^{\alpha_1 t + \alpha_0 - \alpha_1} - e^{\alpha_0 t + \alpha_1 - \alpha_0}}{\alpha_0 - \alpha_1}, & t \in [1, 2[. \end{cases}$

$$\bullet \quad \beta_{(\alpha_0,\alpha_1,\alpha_2)}(t) = \begin{cases} \frac{e^{\alpha_0 t}-e^{\alpha_2 t}}{(\alpha_0-\alpha_1)(\alpha_0-\alpha_2)} - \frac{e^{\alpha_1 t}-e^{\alpha_2 t}}{(\alpha_0-\alpha_1)(\alpha_1-\alpha_2)}, & t \in [0,1[, \\ \frac{e^{\alpha_2 t}}{\alpha_0-\alpha_1}\left[\frac{e^{\alpha_0-\alpha_2}-e^{(t-1)(\alpha_0-\alpha_2)}-e^{t(\alpha_0-\alpha_2)+\alpha_1-\alpha_0}+e^{\alpha_1-\alpha_2}}{\alpha_0-\alpha_2}\right. & \\ \quad \left. -\frac{e^{\alpha_1-\alpha_2}-e^{(t-1)(\alpha_1-\alpha_2)}-e^{t(\alpha_1-\alpha_2)+\alpha_0-\alpha_1}+e^{\alpha_0-\alpha_2}}{\alpha_1-\alpha_2}\right], & t \in [1,2[, \\ \frac{e^{\alpha_2 t}}{\alpha_0-\alpha_1}\left[\frac{e^{(t-1)(\alpha_0-\alpha_2)+\alpha_1-\alpha_0}-e^{\alpha_0+\alpha_1-2\alpha_2}}{\alpha_0-\alpha_2}\right. & \\ \quad \left. -\frac{e^{(t-1)(\alpha_1-\alpha_2)+\alpha_0-\alpha_1}-e^{\alpha_0+\alpha_1-2\alpha_2}}{\alpha_1-\alpha_2}\right], & t \in [2,3[. \end{cases}$$

As it can be observed, the complexity of the analytical expression grows very rapidly with the order of the E-spline. An implementation to find the analytical expressions of E-splines is available in Mathematica (see [122]). Using the coefficients from Equation (2.32) and the E-spline of order 2 as shown in Figure 2.12 (c), we show the reproduction of exponentials $e^{-0.1t}$, $e^{-0.5jt}$ and $e^{(-0.05-0.5j)t}$ in Figure 2.13, 2.14 and 2.15 respectively.

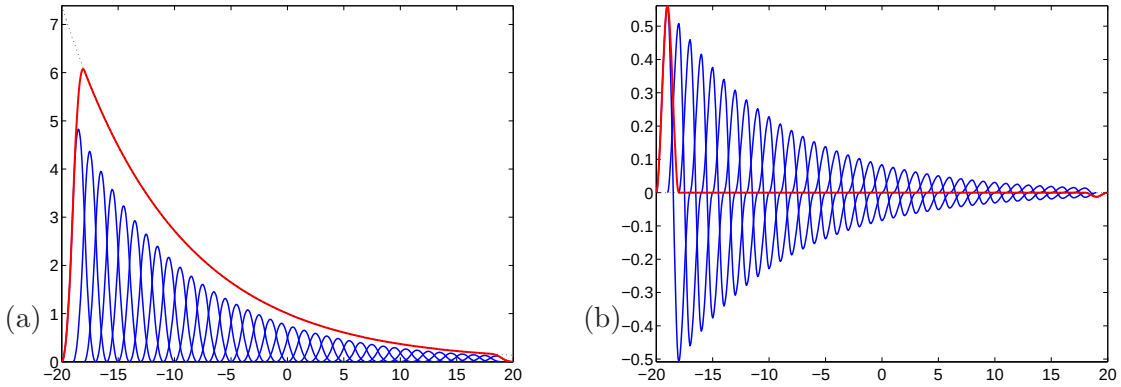


Figure 2.13: The summation of several scaled and shifted versions (in blue) of a complex E-spline of order 2 with parameters $\alpha_0 = -0.05 - 0.5j$, $\alpha_1 = -0.5j$, $\alpha_2 = -0.1$ can reproduce locally (in red) the exponential $e^{-0.1t}$ (in black). (a) real part; (b) imaginary part.

2.5 Conclusion and further considerations

The sampling theory for FRI signals thus uses a completely deterministic framework in order to recover the innovations in the sampled signal. The strength of this theory is in providing a fast algorithmic method that provides a unique solution to the sampling problem. Moreover, this theory also provides the minimum number of measurements that

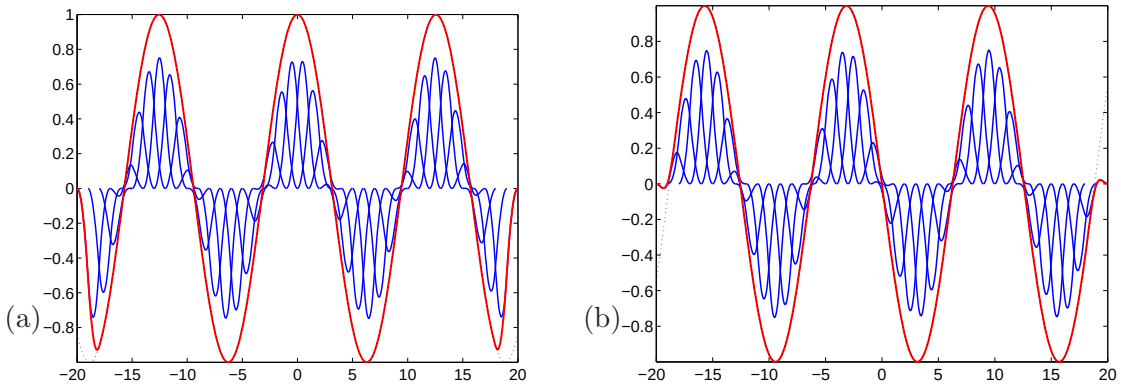


Figure 2.14: The summation of several scaled and shifted versions (in blue) of a complex E-spline of order 2 with parameters $\alpha_0 = -0.05 - 0.5j$, $\alpha_1 = -0.5j$, $\alpha_2 = -0.1$ can reproduce locally (in red) the exponential $e^{-0.5jt}$ (in black). (a) real part; (b) imaginary part.

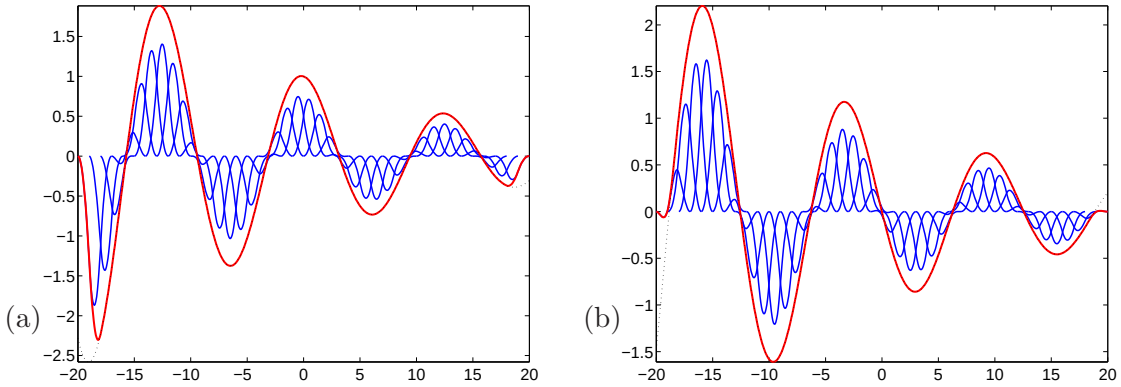


Figure 2.15: The summation of several scaled and shifted versions (in blue) of a complex E-spline of order 2 with parameters $\alpha_0 = -0.05 - 0.5j$, $\alpha_1 = -0.5j$, $\alpha_2 = -0.1$ can reproduce locally (in red) the exponential $e^{(-0.05-0.5j)t}$ (in black). (a) real part; (b) imaginary part.

is required to reconstruct a given FRI signal in the noiseless scenario. With regard to noise issue, optimal theoretical bounds (Cramér-Rao in particular) can be derived to evaluate the robustness of reconstruction algorithms.

The problem of retrieving a stream of Diracs could also be formulated in various ways. For example, a statistical formulation of the parameter determination can be expressed as:

$$p(a_1, a_2, ...a_K, t_1, t_2, ...t_K | x(t)) = \frac{p(x(t)|a_1, a_2, ...a_K, t_1, t_2..t_K) p(a_1, a_2, ..., a_K) p(t_1, t_2, ...t_K)}{p(x(t))},$$

where $x(t)$ is a stream of K Diracs with real parameters $\{a_k, t_k\}, k = 1, \ldots, K$. The

set of parameters maximizing the left hand side above are the parameters of the Diracs occurring in $x(t)$. However, the search of the optimal parameters can be difficult as it operates in a multidimensional space and requires the knowledge of the joint distribution $p(x(t)|a_1, a_2, ...a_K, t_1, t_2..t_K)$. Such a search over a possibly large multidimensional space poses the problems of developing suitable algorithms that are fast enough and which ensure the unicity of solution. Such difficulties are not encountered using the sampling theory for FRI signal. This probabilistic framework however can easily be accommodated to deal with the presence of noise.

The approach taken by compressed sensing in [25] and [16] can be related to the sampling of FRI signals because of the sparsity of such signals. The main difference though resides in the fact that the retrieved parameters can take real values (*e.g.* the locations of the Diracs) in the FRI framework whereas the compressed sensing approach is designed on a discrete grid. As described in [10], the problem of recovering a stream of Diracs can be considered using a compressed sensing framework if the locations of the Diracs are assumed to be on a known discrete grid $\{\theta_{n'}\}_{n'=0,1,...,(N'-1)}$:

$$\min_{x'_0,x'_1,...,x'_{N'-1}} \sum_{n'=0}^{N'-1} \left|x'_{n'}\right| \quad \text{under the constaint} \quad \sum_{n=1}^{N} \left|\mathbf{y}[n] - \sum_{n'=0}^{N'-1} x'_{n'} \varphi(nT - \theta_{n'})\right|^2 \leq N\sigma^2, \tag{2.33}$$

where σ^2 represents the noise power. The minimization in ℓ^1 of the above equation can provide the parameters of the discrete Diracs with high probabilities [16], but requires iterative methods and significantly more measurements ($O(K \log N')$) than by using the direct annihilating filter method ($2K + 1$ samples). Various optimization algorithms using gradient projections can be considered to solve Equation (2.33) (see *e.g.* [32] for the state-of-the-art algorithms used in compressed sensing). The compressed sensing approach does not put the same tight requirements on the sampling kernel (*e.g.* reproduction of polynomial, annihilation property) and can thus be more adaptable than the FRI framework. In the presence of noise, the advantage of using the ℓ^1 norm, as it is in compressed sensing, becomes less evident and no comparison with the Cramér-Rao bounds is available for validating the optimality of this technique.

Chapter 3

Distributed acquisition of FRI Signals

3.1 Introduction

IN various areas of computer science and electrical engineering, a similar trend has been observed over the past decade and consists in the distribution to several independent devices a task that is normally completed by only one single instrument. The motivation behind distributing the workload is the possibility to design smaller, simpler and lighter devices that require less power and that are also cheaper to produce. Using a swarm of devices instead of a unique one also brings robustness to the whole system: if one device fails, the system can still often operate normally. This is clearly illustrated for example in computer networks which saw the emergence of peer-to-peer networks: the weaknesses of a network relying on a single centralized server with a fixed allocated bandwidth can be overcome by considering a distributed client/server network architecture which adapts to the number of users and does not have a single point of failure.

The same developments can also be seen in signal processing with for example the concept of distributed sensor networks. The acquisition and recording of a physical phenomenon into a digital or analog format that can be handled for further processing represents a critical stage. In a traditional setup with a unique sensor, acquiring an accurate discrete representation $\mathbf{g}$ of an observed scene f usually requires a relatively

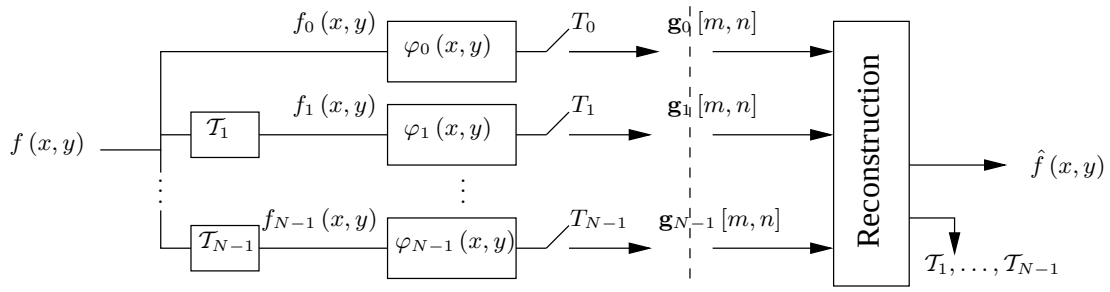


Figure 3.1: Model of a distributed acquisition system with N sensors. Each sensor $i = 1 \ldots N-1$ observes a modified version $f_i(x,y)$ of the scene of reference $f_0(x,y)$ observed by the 0-th sensor. The transformation between the i-th view and the view of reference is denoted $\mathcal{T}_i$. Each sensor outputs a set of samples $\mathbf{g}_i[m,n]$. The reconstruction procedure takes into account all the sets of samples $\mathbf{g}_i[m,n], i = 0,\ldots,N-1$ to find the various transformations $\mathcal{T}_i$ and the observed signal $\hat{f}(x,y)$. In case of perfect reconstruction, we have $\hat{f}(x,y) = f(x,y)$.

high sampling rate, accurate electronic components and a high power consumption. By distributing the acquisition to several sensors instead of one, each sensor can work at a lower sampling rate, use less power and be part of an architecture that is more robust to failure. Finally, new applications can be designed as the phenomenon of interest is observed from different perspectives (in time or space).

A model of a distributed acquisition system is presented in Figure 3.1. Each sensor observes a signal f_i, $i = 0,\ldots,N-1$, which is a transformed version of the signal of reference, $f = f_0$. The transformation between f and f_i, $i = 1,\ldots,N-1$ is denoted by $\mathcal{T}_i$. Since the sampling rate at each sensor is lower, less information is actually available from the sampled signals $\mathbf{g}_i$ so that it is usually not possible to reconstruct f_i given only $\mathbf{g}_i$. The aim instead is to be able to reconstruct the signal f_i, $i = 0,\ldots,N-1$, from the joint observation of the sampled signals $\mathbf{g}_i$, $i = 0,\ldots,N-1$. One way to achieve that is to reconstruct f and find exactly the transformations $\mathcal{T}_i$, $i = 1,\ldots,N-1$ so that all the different views are found.

The main part of this chapter is dedicated to the distributed acquisition of FRI signals. The results of the sampling theory for FRI signals are extended to the distributed scenario for which there is no known study. The framework considered here is when a FRI signal is acquired from different and unknown locations in such a way that current reconstruction methods cannot be applied at each sensor independently. For example, if

K close Diracs are acquired by a device whose kernel φ cannot reproduce at least $2K$ polynomials, then the set of Diracs cannot be recovered. This chapter investigates ways to use the samples from different sensors jointly so that the different observed FRI signals can be exactly reconstructed. This problem raises the feasibility of sampling FRI signals in a truly distributed fashion. With such an architecture, none of the sensors can recover the original signal independently. As shown in the next sections, the nature of the sampling kernel is of considerable importance.

Two cases are treated in this chapter: in the first part, 2-D FRI signals like convex bilevel polygons are considered together with polynomial reproducing sampling kernels. In the second part, the case of streams of Diracs sampled with exponential reproducing sampling kernels is considered.

The reconstruction approach considered for the case of convex bilevel polygons is based on complex moments. Hence, the chapter starts by recalling the basics of the moment theory before detailing the sampling result that allows the reconstruction of polygons from its samples as described in [110]. We then present a distributed architecture where one sensor of reference can reconstruct its view f directly from its own samples and where the other sensors need to calculate the transformations $\mathcal{T}_i$ relative to the sensor of reference in order to find their own views. This chapter does not focus on how to retrieve the transformations $\mathcal{T}_i$. Instead, we are merely interested in determining the characteristics needed by the acquisition devices in a distributed environment in order to achieve perfect reconstruction of FRI signals.

In the 1-D case, we consider the distributed acquisition of a stream of Diracs with kernels reproducing exponentials and this allows us to exhibit some noticeable differences with polynomial kernels. In 1-D, the set of observed signals are delayed versions in time. We show how the observed streams of Diracs and the different time delays can be retrieved exactly in a truly distributed architecture. We finally broaden the scope of this chapter by presenting an application of the use of exponential kernels to time-interleaved analog-to-digital converters where the framework of FRI signals is relaxed and sampled signals do not need to be FRI signals.

3.2 Theory of Moments

3.2.1 Definitions

Moments are features of signals and are characterized by a type and an order p. A general mathematical definition of the moments of a function $f(t)$ can be expressed as:

$$\Psi_p = \int_{-\infty}^{\infty} f(t)\, \Psi_p(t)\, dt, \quad p \in \mathbb{N}, \tag{3.1}$$

where the function $\Psi_p(t)$ is referred as the basis function (or moment weighting kernel). Depending on the choice of the basis function, many different types of moment can defined: geometric, central, complex, Legendre, Zernike, Hermite or Chebyshev (Tchebichef) ... (see *e.g.* [107]). Orthogonal polynomials like Zernike, Legendre, Hermite or Chebyshev are used as basis functions to generate the corresponding type of moments. The geometric moments m_p of order $p \geq 0$ of a function $f(t)$ are obtained by using the monomial t^p as basis function:

$$m_p = \int_{-\infty}^{\infty} f(t)\, t^p dt, \quad p \in \mathbb{N}. \tag{3.2}$$

A special point t_b called barycenter of the function $f(t)$ is defined at:

$$t_b = \frac{m_1}{m_0}.$$

Let $m_p^{(1)}$ and $m_p^{(2)}$ be the moments of the function $f(t)$ and the translated function $f(t - \Delta t)$ respectively. It is quite straightforward to show the relation between those two sets of moments:

$$\begin{aligned} m_p^{(2)} &= \int_{-\infty}^{\infty} f(t - \Delta t)\, t^p dt \\ &= \int_{-\infty}^{\infty} f(x)\, (x + \Delta t)^p\, dx \\ &= \int_{-\infty}^{\infty} f(x) \sum_{k=0}^{p} \binom{p}{k} x^k \Delta t^{p-k} dx \\ &= \sum_{k=0}^{p} \binom{p}{k} \Delta t^{p-k} m_k^{(1)}. \end{aligned} \tag{3.3}$$

In fact, the translation Δt can easily be retrieved from the barycenters $t_b^{(1)}$ and $t_b^{(2)}$ of $f(t)$ and $f(t-\Delta t)$ respectively as follows:

$$\Delta t = t_b^{(2)} - t_b^{(1)}. \tag{3.4}$$

Given the value of t_b, it is interesting to calculate the moments of the function $f(t+t_b)$ centered in t_b. These moments, called the central moments $\mu_{p,q}$ of $f(t)$, are expressed as:

$$\begin{aligned} \mu_p &= \int_{-\infty}^{\infty} f(t)\,(t-t_b)^p\,dt \\ &= \sum_{k=0}^{p} \binom{p}{k} (-t_b)^{p-k}\, m_k. \end{aligned} \tag{3.5}$$

The weighted local moments are defined on a portion only of the function $f(t)$ by using a symmetrical positive window (or weight) function $w(t)$ of support $\mathcal{O}$. The local weighted geometric moments centered at t_0 are thus expressed as:

$$m_p(t_0) = \int_{\mathcal{O}} w(t-t_0)\, f(t)\, t^p dx, \quad p \in \mathbb{N}. \tag{3.6}$$

The use of a simple box function as the window $w(t)$ is usually not used in practice as it is often more desirable to increase the weights around the center of the window and to gradually decrease the weights to zero as we move away from it. Therefore, various local weighted moments can be defined depending on the choice of $w(t)$. In [7], the choice of a Lorentzian function was considered to computed the weighted moments for pattern recognition purposes whereas in [115], a B-spline of a certain degree is chosen for the detection of DNA strands and the estimation of the motion field. In [130], the notion of cross-weighted moments between two functions is defined for registration and matching applications.

3.2.2 Image moments

Equation (3.1) can be applied to multiple dimensions in a straightforward manner. For two dimensional signals like images, we have:

$$\Psi_{p,q} = \iint_{-\infty}^{\infty} f(x,y)\, \Psi_{p,q}(x,y)\, dxdy, \quad p,q \in \mathbb{N}, \tag{3.7}$$

which thus gives for geometric and central moments of order $p+q$, $p,q \in \mathbb{N}$:

$$m_{p,q} = \iint_{-\infty}^{\infty} f(x,y)\, x^p y^q dxdy, \tag{3.8}$$

$$\mu_{p,q} = \iint_{-\infty}^{\infty} f(x,y)\, (x-x_b)^p (y-y_b)^q\, dxdy \tag{3.9}$$

$$= \sum_{k=0}^{p}\sum_{l=0}^{q} \binom{p}{k}\binom{q}{l} (-x_b)^{p-k} (-y_b)^{q-l}\, m_{k,l}, \tag{3.10}$$

where the barycenter (x_b, y_b) of $f(x,y)$ is given by:

$$x_b = \frac{m_{1,0}}{m_{0,0}} \qquad \text{and} \qquad y_b = \frac{m_{0,1}}{m_{0,0}}.$$

Interestingly, the moments can provide some basic information about the geometry of the image $f(x,y)$. Let $\mathbf{C}$ be the covariance matrix of $f(x,y)$ given by:

$$\mathbf{C} = \begin{bmatrix} \mu_{2,0} & \mu_{1,1} \\ \mu_{1,1} & \mu_{0,2} \end{bmatrix}.$$

The orientation of an image is given by the orientation of the eigenvector associated with the largest eigenvalue of $\mathbf{C}$. This direction is called the principal axis and is the line going through the center of mass and around which the first order moments are minimized [62]. It can be shown that the orientation angle $-\pi/4 \leq \theta \leq \pi/4$ is given by the following relation :

$$\theta = \frac{1}{2} \arctan\left(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}}\right). \tag{3.11}$$

The complex moments $C_{p,q}$ of order $p+q$ are another type of moment that can be

defined for 2-D functions on the complex image plane $z = x + jy$, $j = \sqrt{-1}$ as follows (often attributed to Abu-Mostafa *et al* in [1] but already defined by Davis *et al* in [21]):

$$C_{p,q} = \iint_{-\infty}^{\infty} f(x,y)\,(x+jy)^p\,(x-jy)^q\,dxdy \tag{3.12}$$

$$= \sum_{k=0}^{p}\sum_{l=0}^{q}\binom{p}{k}\binom{q}{l} j^{p-k+q+l}\,(-1)^{q-l}\,m_{k+l,p-k+q-l}. \tag{3.13}$$

Interestingly, the complex moments $C_{p,q}^{(1)}$ and $C_{p,q}^{(2)}$ of $f(x,y)$ and their rotated version by an angle α are related by the following expression:

$$C_{p,q}^{(2)} = e^{-j(p-q)\alpha} C_{p,q}^{(1)}. \tag{3.14}$$

As it can be observed from Equations (3.10) and (3.13), the knowledge of the geometric moments of a function f is sufficient to calculate the values of other types of moments with a linear combination. Besides, it is also possible to derive similar relations for other basis functions like Legendre polynomials. In this respect, the geometric moments are considered as the fundamental set of moments in the theory of moments from which other type of moments can be derived.

Moments have been extensively studied and used in signal processing [57, 117]. Moment theory has been most successful in areas like pattern recognition, object matching or registration. One of the earliest works in image processing is provided by Hu on visual pattern recognition which gives the following fundamental theorem [57]:

Theorem 1 (Hu's Uniqueness Theorem, [57].). *Let $f(x,y)$ be a piecewise continuous signal with a compact support $\mathcal{O}$ of the xy-plane where it is non zero. Then the set of geometric moments $\{m_{p,q}\}$ is uniquely determined by $f(x,y)$, and conversely $f(x,y)$ is uniquely determined by $\{m_{p,q}\}$.*

Therefore, piecewise continuous images can be uniquely represented by an infinite set of moments $\{m_{p,q}\}$ and can be reconstructed exactly using the following expression [62]:

$$f(x,y) = \iint_{-\infty}^{\infty} e^{-j2\pi(x\xi_1+y\xi_2)} \left[\sum_{p=0}^{\infty}\sum_{q=0}^{\infty} m_{p,q}\frac{(j2\pi)^{p+q}}{p!q!}\xi_1^p\xi_2^q\right] d\xi_1 d\xi_2. \tag{3.15}$$

However, if only a finite set of moments $\{m_{p,q}\}$ is available, it is in general not possible to reconstruct exactly $f(x,y)$. This is essentially due to the fact that the summation and the integration signs in (3.15) are not interchangeable because the Fourier transform of the term $(j2\pi\xi_1)^p$ is not bounded [62]. However, some signals like bilevel polygons can actually be perfectly represented by a finite set of moments.

3.2.3 On the Definitions of Continuous Moments and Discrete Moments

As seen from Equation (3.7), the basis function $\Psi_{p,q}(x,y)$ and the observed signal $f(x,y)$ are both continuously-defined functions. However, in practice, the signal $f(x,y)$ is not available as-is and is approximated by its sampled version $\mathbf{g}[m,n]$ during acquisition. Therefore, the integrals in Equation (3.7) are usually approximated by discrete summations which lead to numerical errors in the calculated moments [88]. Therefore, strictly speaking, these moments are different. However, in the majority of the works on moments, they are often assumed which is not the case here.

In this thesis, it is fundamental to remember that a distinction is made between the exact moments (from Equation (3.7)) called the *continuous moments* and their approximated calculations with discrete summations which are referred to as *discrete moments*:

$$\text{CONTINUOUS MOMENTS:} \qquad \Psi_{p,q} = \iint_{-\infty}^{\infty} f(x,y)\,\Psi_{p,q}(x,y)\,dxdy,$$

$$\text{DISCRETE MOMENTS:} \qquad \underline{\Psi}_{p,q} = \sum_{m,n} \mathbf{g}[m,n]\,\Psi_{p,q}(m,n).$$

Discrete and continuous moments have been defined in a different way in some works as in [87, 88, 131] where the distinction between continuous and discrete moments was made upon on whether the basis functions $\Psi_{p,q}$ are continuous orthogonal polynomials (*e.g.* Legendre) or discrete orthogonal polynomials (*e.g.* Chebyshev). A set of discrete orthogonal polynomials $t_m(x)$ are satisfying the orthogonality condition are for discrete values of x only [116].

3.3 Reconstruction of Bilevel Polygons with Complex Moments

The method of perfectly reconstructing convex bilevel polygons from samples as described in [110] are given in this section. Such signals belong to the class of FRI signals since convex polygons are completely and uniquely characterized by the location of their vertices. The reconstruction approach focuses on retrieving the exact locations of the corners of the polygon which is done by looking at the complex moments.

Let $f(x,y)$ be the function describing a convex bilevel polygon with K vertices $\{(x_k,y_k)\,|k=1,\ldots,K\}$. We now consider the image complex plane $z=x+jy, j=\sqrt{-1}$ where the vertices of the polygon are given by $\{z_k=x_k+jy_k|k=1,\ldots,K\}$. In [20], Davis proved that any complex analytic function $h(z)$ in the closure Γ (see [104] for more details on analytic functions) satisfies:

$$\iint_{\Gamma} f(x,y)\,h''(z)\,dxdy=\sum_{k=1}^{K}a_k h(z_k)\,, \tag{3.16}$$

where $h''(z)$ denotes the second derivative of $h(z)$ with respect to z. The complex coefficients $\{a_k\in\mathbb{C}|k=1,\ldots,K\}$ are independent of the function $h(z)$. What the Equation (3.16) essentially tells us is that the evaluation of the double integral is equal to a linear combination of the function h evaluated at the location of the vertices z_k of the polygon.

In [83], Milanfar *et al* considered Equation (3.16) by setting $h(z)=z^p$ which yields to:

$$\begin{aligned}\iint_{\Gamma} f(x,y)\,h''(z)\,dxdy &= p(p-1)\underbrace{\iint_{\Gamma} f(x,y)\,z^{p-2}dxdy}_{=C_{p-2,0}},\\ &= p(p-1)\,C_{p-2,0},\\ &= \sum_{k=1}^{K}a_k z_k^p.\end{aligned} \tag{3.17}$$

Here $C_{p-2,0}$ are the complex moments of the polygon function $f(x,y)$. In this case, the complex moments $C_{p-2,0}$ weighted by $p(p-1)$ are equal to a powersum series with unknown parameters $\{a_k,z_k\}$. By defining the weighted complex moments as

$\tau_p = p\,(p-1)\,C_{p-2,0}$ with $\tau_0 = \tau_1 = 0$, Equation (3.17) can be rewritten in the following more compact form:

$$\tau_p = \sum_{k=1}^{K} a_k z_k^p. \tag{3.18}$$

Thus, by utilizing the annihilating filter method (Section 2.2.3), it is possible to find the unknown parameters $\{a_k, z_k\}$ provided that at least $2K$ consecutive weighted complex moments τ_p are available. Since $\tau_0 = \tau_1 = 0$, this requires us to know $2K-2$ consecutive complex moments like the set $\{C_{0,0}, C_{1,0}, \ldots, C_{2K-3,0}\}$.

Finally, in [110], Shukla *et al* considered the problem of reconstructing perfectly a sampled polygon in the light of the theory of sampling for FRI signals. Let $\mathbf{g}_{m,n}$ be the samples of the polygon $f\,(x,y)$ obtained with a sampling kernel $\varphi\,(x,y)$ reproducing polynomials:

$$\mathbf{g}\,[m,n] = \langle f\,(x,y)\,, \varphi\,(x/T_x - m, y/T_y - n)\rangle\,. \tag{3.19}$$

Now similarly to Equation (2.13), we have for $p,q = 0,1,\ldots,P$:

$$\begin{aligned}
\tau_{p,q} &= \sum_{m,n} c_{m,n}^{(p,q)} \mathbf{g}\,[m,n] \\
&= \left\langle f\,(x,y)\,, \sum_{m,n} c_{m,n}^{(p,q)} \varphi\,(x-m, y-n) \right\rangle \\
&= \iint_{-\infty}^{\infty} f\,(x,y)\,x^p y^q dxdy, \\
&= m_{p,q},
\end{aligned} \tag{3.20}$$

where the coefficients $c_{m,n}^{(p,q)}$ are for example given in Equation (2.26). The above equation shows that it is possible to retrieve the *exact geometric moments* of the bilevel polygon $f\,(x,y)$ from its sampled version by properly combining the samples with some coefficients related to the polynomial reproduction formula of the sampling kernel. When the moments are calculated with Equation (3.20), these moments are referred to as the *continuous moments* since they are the exact moments of the continuous function f and not the moments of the discrete version $\mathbf{g}$. Now, given the continuous geometric moments, it is possible to find the exact complex moments of the polygon using Equation (3.13) and therefore to find the vertices of the polygon. The following result can now be enounced:

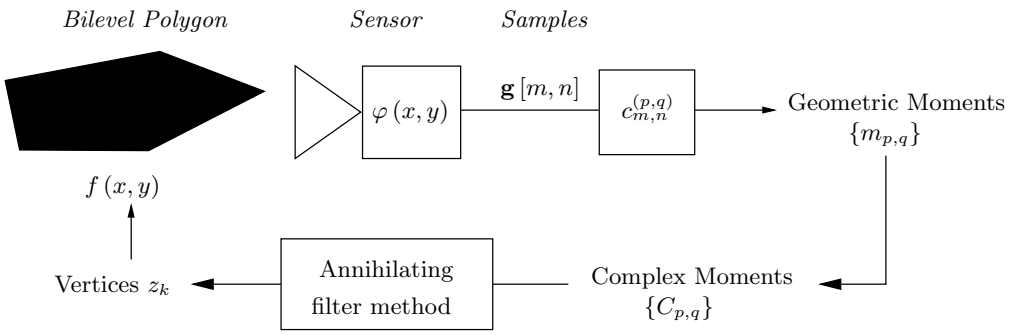


Figure 3.2: From the acquisition to the perfect reconstruction of a bilevel polygon. The polygon $f(x,y)$ is acquired by a sensor with a sampling kernel $\varphi(x,y)$ to give the samples $\mathbf{g}[m,n]$. The samples are combined with the coefficients $c_{p,q}^{(m,n)}$ related to $\varphi(x,y)$ to obtain the exact continuous geometric moments $m_{p,q}$ of $f(x,y)$. From these moments, the continuous complex moments $C_{p,q}$ are computed and used in the annihilating filter method to retrieve the exact locations of the vertices of the original polygon $f(x,y)$.

Proposition 2 (Shukla and Dragotti, [110].). *A bilevel and convex polygon $f(x,y)$ with K corner points is uniquely determined from the samples $\mathbf{g}[m,n] = \langle f(x,y), \varphi(x/T_x - m, y/T_y - n)\rangle$ provided that the sampling kernel $\varphi(x,y)$ is able to reproduce polynomials up to a degree $2K-3$ along the Cartesian axes x and y.*

The power of this sampling result resides in the fact that the vertices of the polygon can be located exactly given the samples. The whole approach of sampling and perfectly reconstructing a bilevel polygon using the complex moments is summarized in the block diagram of Figure 3.2. A simulation showing the reconstruction of a pentagon ($K = 5$ corners) from its samples is presented in Figure 3.3. Figure 3.3 (a) shows the original bilevel pentagon (512x512 px). This image is then artificially sampled with a B-spline of order $P = 2K-3 = 7$ and a scale $T = 16$ so that the resulting sampled image has a size of 32x32 px. The retrieved corners are plotted against the original image in Figure 3.3 (c).

The experiment shown in Figure 3.4 is based on a real acquired image. An image of a scene containing a rectangle was acquired with a Nikon D70s digital camera and a lens set out of focus. This image of size 2014x3039 px is then decimated by a factor 6 to give a low-resolution image of size 335x507 px (Figure 3.4 (a)). An area of size 22x22 px is then selected around the samples of the rectangle ($K = 4$) as shown on Figure 3.4 (b). In this experiment, the sampling kernel is modeled by a B-spline of order $P = 5$ and a scale $T = 6$

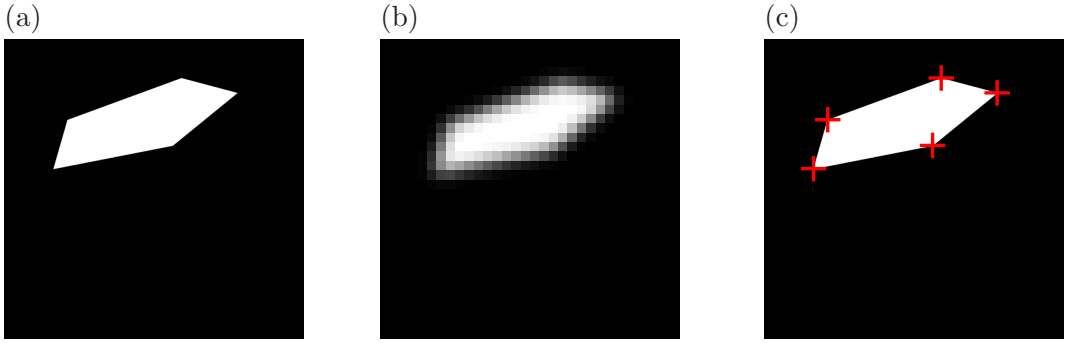


Figure 3.3: Simulation of sampling and reconstructing 2-D bilevel convex polygons. (a) Original bilevel polygon with 5 corners $f(x, y)$ (512x512 px); (b) Artificially sampled image $\mathbf{g}[m, n]$ of size 32x32 px. The sampling kernel is a B-spline with $P = 7$ and $T = 16$; (c) Original polygon and extracted corners (+).

(decimation rate). The four corners of the rectangle retrieved from the complex moments are shown in Figure 3.4 (c) and Figure 3.4 (d) at a different zoom level. As opposed to the simulation presented in Figure 3.3, it is important to notice that the samples in this experiment are not artificially generated with a known sampling kernel but are the actual samples from the camera. More details on how the camera is modeled are given in the next chapter where the image acquisition model is described. It is the first time that such experiments of bilevel polygons reconstruction from real samples have been conducted.

3.4 Distributed Acquisition with Kernels Reproducing Polynomials

3.4.1 Distributed Acquisition of Bilevel Polygons

Problem statement

From the previous analysis, the perfect reconstruction of a bilevel polygon with K corners from its samples requires $2K - 2$ consecutive moments or equivalently a sampling kernel that can reproduce polynomials up to degree $P = 2K - 3$. For the sake of clarity in this section, we call the order P of a sensor when we want to refer to the maximum polynomial reproduction degree P of the sampling kernel that constitutes that sensor. We now consider the case of several sensors observing the same polygon with K corners from different locations. We investigate here how the order of the sensors can be lowered such

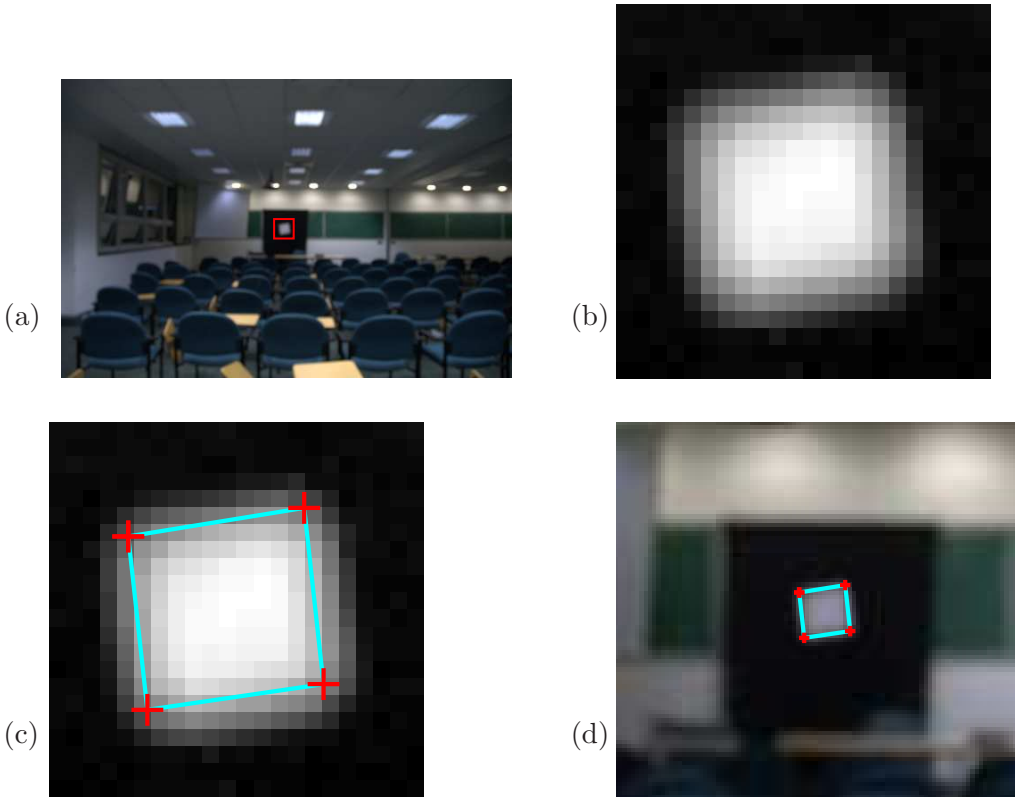


Figure 3.4: Experiment of reconstruction of a rectangle from real data. (a) Considered image of size 335x506 px resulting from a decimation by $T = 6$ of a blurred image taken out of focus. The square framed area is the region of interest of size 22x22 px; (b) Region of interest containing the samples of the rectangle; (c) Retrieved corners (+) and boundaries of the reconstructed rectangle. The sampling kernel is modeled by a B-spline with $P = 5$ and $T = 6$; (d) Retrieved rectangle in the considered image.

that $P < 2K - 3$ while still preserving perfect reconstruction of the observed polygon at each sensor. Since $P < 2K - 3$, independent reconstruction of the polygon at each sensor is not conceivable and this suggests a joint reconstruction method that fuses the information available from several sensors.

Using a sensor of lower order means that the sampling kernel has a shorter support and that less samples are necessary. Take for example a B-spline kernel $\beta_P(t)$ of order P and suppose that a single Dirac is located somewhere in $[0, 1]$. Because a B-spline of order P has a support of length $P + 1$, the number of samples affected by the Dirac is exactly $P + 1$. Thus, the higher the order of a sensor, the larger the support of the kernel and the more samples that are required.

We first look at the case where all the sensors have the same order $P < 2K - 3$

and call this setup the symmetric distributed acquisition architecture. We then consider the case where one main sensor has an order $P_{main} = 2K - 3$ and all the other sensors have the same order $P_{aux} < 2K - 3$. This configuration is called asymmetric distributed acquisition architecture. In each case, we first want to see whether perfect reconstruction is possible at each sensor and what the minimum order is.

Symmetric Distributed Acquisition Architecture

Each sensor has an order $P < 2K - 3$ and observes a different transformed version of the same polygon with K corners. The idea here is to examine whether each sensor can generate only a subset of the necessary moments that are required to reconstruct the polygon. As we shall see, this is not possible.

Suppose first that each sensor can extract the same subset of moments from order 0 up to order $P < 2K - 3$. Suppose the simplest case where two observed polygons f_i and f_j are related by a translation (t_1, t_2). If an oracle gives the exact translation (t_1, t_2) between f_i and f_j, then one can relate the moments $m_{p,q}^{(i)}$ of f_i and the moments $m_{p,q}^{(j)}$ of f_j as follows:

$$m_{p,q}^{(j)} = \sum_{k=0}^{p} \sum_{l=0}^{q} \binom{p}{k} \binom{q}{l} t_1^{p-k} t_2^{q-l} m_{k,l}^{(i)}. \tag{3.21}$$

This is a direct extension to 2-D of Equation (3.3). This equation naturally shows that the knowledge of the moments of a particular order of different view of a scene cannot give the moments of higher order for another view. In other words, there is no new information about a scene contained in the moments of the same order of two different views of the same scene.

One could also perhaps design sensors that can output a different subset of moment each so that by union of all the subsets, all the necessary moments are obtained. For example, one sensor provides only the moments 0, 1 and 2 and another sensor provides only the moments 3, 4 and 5. However, this is also not possible in the proposed framework because, if a sampling kernel satisfies the Strang-Fix conditions and can reproduce a polynomial of degree 3, it must also, by construction, reproduce all the polynomials of degree 0 to 2. It is therefore impossible to distribute equally among several sensors the

computation of moments.

Asymmetric Distributed Acquisition Architecture

Since it is not possible to have all the sensors with an order $P < 2k - 3$, we now look at the case where one main sensor has an order $P_{main} = 2K - 3$ and all the other sensors have the same order $P_{aux} < 2K - 3$ so that $P_{main} > P_{aux}$.

Since the main sensor can output all the necessary moments, it can reconstruct the observed polygon independently. However, the other sensors cannot. The solution for these auxiliary sensors is to calculate exactly the transformations $\mathcal{T}_i$ between them and the main sensor so the reconstructed scene at the main sensor can be properly transformed to give the view as observed by each sensor.

In this framework, the multiview reconstruction problem becomes a registration problem and the minimum order necessary at each auxiliary sensor is the order required to find the transformations $\mathcal{T}_i$. It turns out that:

- if the $\mathcal{T}_i$'s are translations, they can be retrieved from the moments of order 0 and 1: $P_{aux} = 1$;

- if the $\mathcal{T}_i$'s are rigid transformations (rotation and translation), they can be retrieved from the moments of order 0 to 2: $P_{aux} = 2$.

- if the $\mathcal{T}_i$'s are affine transformations, they can be retrieved from the moments of order 0 to 3 (see Section 4.3.2): $P_{aux} = 3$.

Thus, as presented in Figure 3.5, perfect reconstruction at each sensor can be achieved with an asymmetric distributed acquisition architecture whose main sensor has order $P_{main} = 2K - 3$ and auxiliary sensors have an order P_{aux} =1,2 or 3 depending on the assumed transformation between the sensors. The only possible case where P_{aux} and P_{main} are equal occurs when the observed polygon is a triangle ($K = 3$) and the assumed transformations are affine transformations.

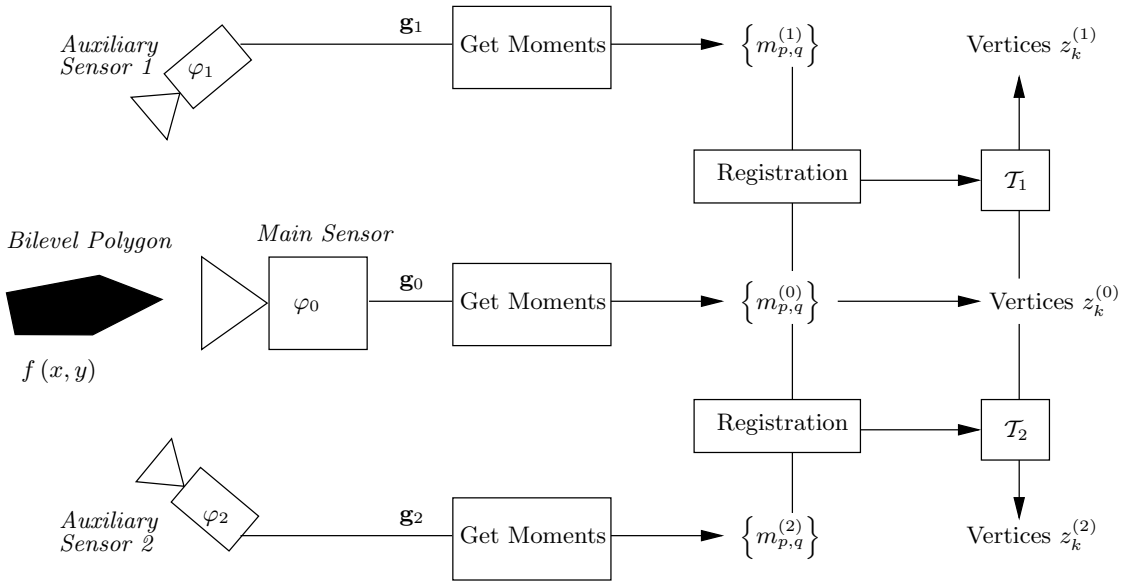


Figure 3.5: Asymmetric architecture for distributed acquisition of bilevel polygons. The main camera retrieves independently the vertices $z_k^{(0)}$ of the observed polygon from its moments whereas auxiliary cameras cannot. The relative transformation $\mathcal{T}_i$ between the auxiliary sensors and the main sensor are obtained with a moment-based registration method. The vertices $z_k^{(i)}$ of the polygon observed by the i-th sensor are then retrieved after transforming the vertices $z_k^{(0)}$ with $\mathcal{T}_i$.

3.4.2 Simulations and Experiments

We first show in Figure 3.6 a simulation of the distributed acquisition of a synthetic pentagon with the architecture in Figure 3.5. The original pentagon observed by the main sensor is shown in the middle column of Figure 3.6(a). The scene observed by the first auxiliary sensor is shown on the left column of Figure 3.6(a). It is related to the pentagon of the main sensor by the following affine transformation $\mathbf{x}' = \mathcal{T}_1(\mathbf{x})$:

$$\mathcal{T}_1 : \left\{ \mathbf{x}' = \mathbf{A}_1\mathbf{x} + \mathbf{t}_1, \quad \mathbf{A}_1 = \begin{bmatrix} 1.3 & -0.5 \\ -0.1 & 1.4 \end{bmatrix} \quad \text{and} \quad \mathbf{t}_1 = \begin{bmatrix} -100 \\ -80 \end{bmatrix} \right\}.$$

The second sensor observes the scene depicted in the right column of Figure 3.6(a). It is also related to the view of the main sensor by the following affine transformation $\mathbf{x}' = \mathcal{T}_2(\mathbf{x})$:

$$\mathcal{T}_2 : \left\{ \mathbf{x}' = \mathbf{A}_2\mathbf{x} + \mathbf{t}_2, \quad \mathbf{A}_2 = \begin{bmatrix} 1.2 & 0.4 \\ 0.7 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{t}_2 = \begin{bmatrix} 13 \\ 50 \end{bmatrix} \right\}.$$

Each sensor samples independently its own view. Since the main sensor must be able to reconstruct the pentagon ($K = 5$), it has a sampling kernel of order $2K - 3 = 7$. The sampling kernel used here is a B-spline of order $P = 7$ and sampling period $T = 24$. The auxiliary sensors only need to be able to estimate the affine transformation between themselves and the main sensor. Thus, the sampling kernels are a B-spline of order $P = 3$ with scale $T = 64$. The resulting sampled images are respectively displayed on the middle, left and right columns of Figure 3.6(b). The resolution of the image at the main sensor is 32x32 px whereas the resolution of the images at both auxiliary sensors is 8x8 px. The continuous moments are computed at each sensor, then:

1. the corners of the polygon $z_k = x_k + jy_k, k = 1, \ldots, 5$ are retrieved at the main sensor;

2. the affine transformations $\mathcal{T}_1$ and $\mathcal{T}_2$ are calculated to register both auxiliary sensors; and

3. the corners of the polygon observed by the auxiliary sensors are inferred by applying $\mathcal{T}_i$ on $(x_k, y_k), k = 1, \ldots, 5$.

The positions of the estimated corners are shown for each sensor in Figure 3.6(c) where the corners are plotted against the original scenes. The estimated affine transformations $\widetilde{\mathcal{T}}_1$ and $\widetilde{\mathcal{T}}_2$ computed from the continuous moments are:

$$\widetilde{\mathbf{A}}_1 = \begin{bmatrix} 1.3012 & -0.5283 \\ -0.0781 & 1.3935 \end{bmatrix} \quad \text{and} \quad \widetilde{\mathbf{t}}_1 = \begin{bmatrix} -99.3889 \\ -79.6741 \end{bmatrix},$$

and

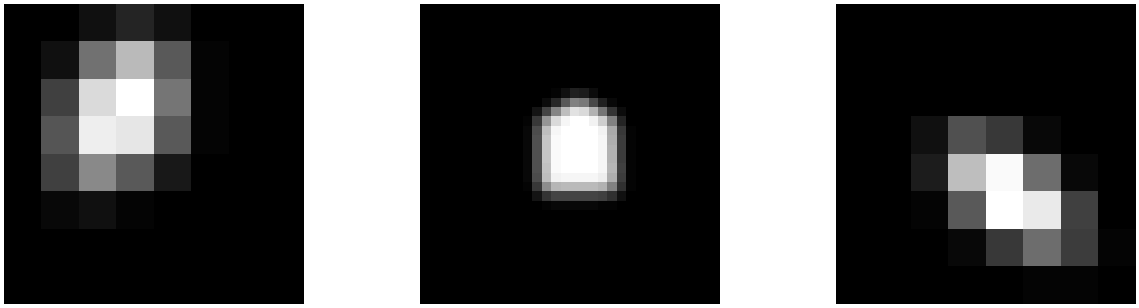
$$\widetilde{\mathbf{A}}_2 = \begin{bmatrix} 1.2023 & 0.4027 \\ 0.6975 & 0.9989 \end{bmatrix} \quad \text{and} \quad \widetilde{\mathbf{t}}_2 = \begin{bmatrix} 13.3637 \\ 50.0388 \end{bmatrix}.$$

In a second experiment presented in Figure 3.7, the samples are not generated artificially but are the real samples acquired with a digital camera. Two out-of-focus pictures of a rectangular polygon ($K = 4$) are taken from two different positions with a SLR Nikon D70s digital camera, giving two blurred images of size 2014x3039 px. The sampling kernel at the main sensor is modeled with a B-spline of order $P_{main} = 5$ and

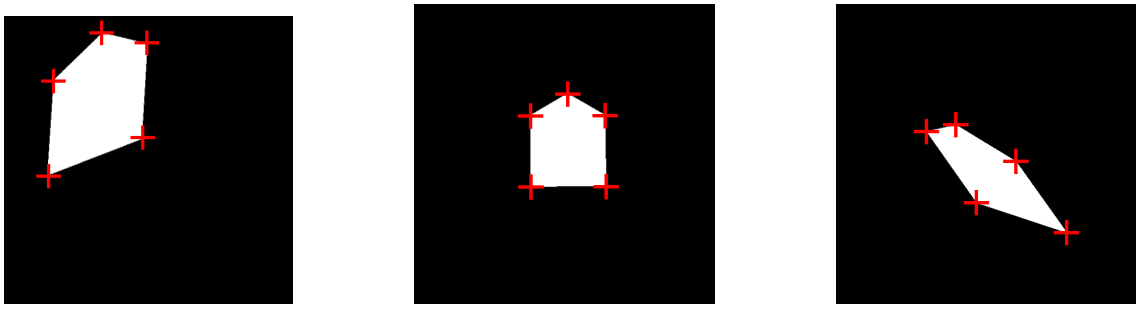
scale $T_{main} = 6$ whereas the sampling kernel at the auxiliary sensors is modeled with cubic B-spline ($P_{aux} = 3$) and a scale $T_{aux} = 9$. Notice that in both cases, the support of the kernel has a length of 36px. The two pictures are then decimated by T_{main} and T_{aux} respectively to give an image of size 335x506 px at the main sensor and 224x338 px at the auxiliary sensors (Figure 3.7(a)).

A set of samples containing the polygon is then selected on each image (red box in

(a) Original scenes (512x512 px each). Left: auxiliary sensor #1; middle: main sensor; right: auxiliary sensor #2.



(b) Artificially sampled images. Left: auxiliary sensor #1: 8x8 px; middle: main sensor: 32x32 px; right: auxiliary sensor #2: 8x8 px.



(c) Retrieved corners (+) plotted on original scenes (512x512 px). Left: auxiliary sensor #1; middle: main sensor; right: auxiliary sensor #2.

Figure 3.6: Distributed acquisition of a pentagon. The sampling kernel of the main sensor is a B-spline of order $P_{main} = 7$ and scale $T_{main} = 16$, and the sampling kernels of the auxiliary sensors are a B-spline of order $P_{aux} = 3$ and scale $T_{aux} = 64$.

3.7(a)). The regions of interest have a size of 22x22 px for the main sensor and 15x15 px for the auxiliary sensor. As in the previous simulation, the corners of the polygon and the affine transformation are calculated using the continuous moments. The locations of the corners and the edges of the polygon recovered for each sensor is presented in Figure 3.7 (b) and (c) at a different zoom level. For information, the retrieved affine transformation between the auxiliary and the main sensors is:

$$\widetilde{\mathbf{A}}_1 = \begin{bmatrix} 0.821 & 0.0498 \\ -0.0461 & 0.9157 \end{bmatrix} \quad \text{and} \quad \widetilde{\mathbf{t}}_1 = \begin{bmatrix} -8.1361 \\ 1.1065 \end{bmatrix},$$

The images presented in Figure 3.7 (d) are the "ideal", focused and high resolution images to help a visual comparison of the retrieved polygons and their locations in the scenes.

3.5 Distributed Acquisition with Kernels Reproducing Exponentials

3.5.1 Distributed Acquisition of a Stream of Diracs with E-spline Kernels

In this section, the acquisition of 1-D signals is considered with sensors whose sampling kernels can reproduce exponentials. Let us consider an acquisition system composed of N sensors and let $x_i(t), i = 0, \ldots, N-1$, be the 1-D continuous signals observed by the corresponding sensors. By convention, each sensor $i > 0$ observes a delayed version of $x_0(t)$ such that:

$$x_i(t) = x_0(t - \Delta t_i),$$

where the time delays $\Delta t_i \in \mathbb{R}$ are unknown. The sampling kernel $\varphi_i(t)$ at each sensor is modeled by an E-spline of order P with a support of length $L = P + 1$:

$$\varphi_i(t) = \beta_{\underline{\alpha}^{(i)}}(t), \quad \text{with} \quad \underline{\alpha}^{(i)} = \left\{\alpha_0^{(i)}, \ldots, \alpha_P^{(i)}\right\},$$

(a) Low-resolution images from two locations. Left: main sensor 335x506 px; right: auxiliary sensor 224x338 px. Red boxes show the regions of interest.

(b) Retrieved corners (+). Left: main sensor, 22x22 px; right: auxiliary sensor, 15x15px.

(c) Retrieved rectangle on each low-resolution images. Left/right: main/auxiliary sensor.

(d) High resolution and focused images of the same scenes for visual comparison.

Figure 3.7: Experiment of distributed acquisition of bilevel polygons. The sampling kernel of the main sensor on the left is modeled by a B-spline with $P = 5$ and $T = 6$. The auxiliary sensor on the right has a sampling kernel modeled by a B-spline with $P = 3$ and $T = 9$.

such that:

$$\alpha_p^{(i)} = \alpha_0^{(i)} + p\lambda^{(i)}, \quad p = 0, 1, \ldots, P \quad \text{for given } \lambda^{(i)} \in \mathbb{C} \text{ and } \alpha_0^{(i)} \in \mathbb{C}.$$

We assume for simplicity that $\lambda^{(i)} = \lambda$. The samples $\mathbf{y}_i[n]$ acquired at the i-th sensor are thus expressed as:

$$\mathbf{y}_i[n] = \left\langle x_i(t), \beta_{\underline{\alpha}^{(i)}}(t/T - m) \right\rangle.$$

Let us now consider the case where $x_0(t)$ is an infinite stream of Diracs with the k-th Dirac described by the parameters (a_k, t_k) corresponding respectively to its amplitude and its location. In the single sensor scenario, if the kernel $\varphi_0(t)$ can reproduce up to $2K$ exponentials (*i.e.* $P = 2K - 1$) as:

$$\sum_n c_n^{(0,p)} \varphi_0(t - n) = e^{\alpha_p^{(0)} t},$$

and provided that there are at most K Diracs in any interval of length $2KLT$, then the following measurements $\tau_p^{(0)}$:

$$\begin{aligned} \tau_p^{(0)} &= \sum_n c_n^{(0,p)} \mathbf{y}_0[n] \\ &= \int_{-\infty}^{\infty} x_0(t)\, e^{\alpha_p^{(0)} t} dt \\ &= \sum_{k=0}^{K-1} a_k e^{\alpha_p^{(0)} t_k}, \qquad p = 0, 1, \ldots, 2K-1. \end{aligned}$$

provide enough information to retrieve exactly each Diracs with the annihilating filter method [27]. Notice from the second line of the expression above that if $\underline{\alpha}^{(0)}$ is a set of purely imaginary numbers, then $\tau_p^{(0)}$ are the Fourier coefficients of $x_0(t)$.

As mentioned in the previous section, distributed acquisition aims to have sensors of lower order P. As opposed to the polynomial case where only an asymmetric architecture is feasible, it is possible to truly distribute the acquisition of FRI signals with kernels reproducing exponentials. Thus, the $2K$ parameters can be freely distributed among the N available sensors. The reason why this is possible with kernels reproducing exponentials is because such kernels offer a more general framework than kernels repro-

ducing polynomials. With splines for example, B-splines kernels are a particular case of E-splines with parameters $\alpha_p = 0, \forall p$. It is now possible to design different kernels $\beta_{\underline{\alpha}^{(i)}}(t)$ and $\beta_{\underline{\alpha}^{(j)}}(t), j \neq i$ having the same order P but which are reproducing different sets of exponentials $\underline{\alpha}^{(i)} \neq \underline{\alpha}^{(j)}$. Thus the proposed architecture is symmetric in the sense that each sensor now has the same order. However the reproduced exponentials are different at each kernel.

In a distributed architecture, it is first necessary to register precisely the different acquired signals so that they can be adequately combined. Assume that one parameter is common to the sets $\underline{\alpha}^{(0)}$ and $\underline{\alpha}^{(1)}$ of the sensors observing $x_0(t)$ and $x_1(t)$ respectively. For example, this parameter can be the first and the last parameter such that:

$$\alpha_P^{(0)} = \alpha_0^{(1)} = \alpha.$$

The corresponding measurements $\tau_P^{(0)}$ and $\tau_0^{(1)}$ are then expressed as:

$$\begin{aligned} \tau_P^{(0)} &= \int_{-\infty}^{\infty} x_0(t)\, e^{\alpha t} dt, \\ \tau_0^{(1)} &= \int_{-\infty}^{\infty} x_1(t)\, e^{\alpha t} dt. \end{aligned}$$

Since $x_1(t) = x_0(t - \Delta t_1)$, we can rewrite $\tau_0^{(1)}$ as:

$$\tau_0^{(1)} = e^{\alpha \Delta t_1} \int_{-\infty}^{\infty} x_0(t)\, e^{\alpha t} dt,$$

which finally yields to the following relation:

$$\tau_0^{(1)} = e^{\alpha \Delta t_1} \cdot \tau_P^{(0)}. \tag{3.22}$$

This relation is a general expression of the well-known Fourier Shift Theorem for $\alpha \in \mathbb{C}$ not necessarily a pure imaginary number. Thus, if one common parameter is shared between two sensors, the time delay Δt_1 between the two sensors can be retrieved exactly using

Equation (3.22) as:

$$\Delta t_1 = \frac{\ln \frac{|\tau_0^{(1)}|}{|\tau_P^{(0)}|} + j\left(\arg\left(\tau_0^{(1)}\right) - \arg\left(\tau_P^{(0)}\right)\right)}{\alpha}. \tag{3.23}$$

Once Δt_1 is known, Equation (3.22) can be used directly to register the measurements $\tau_p^{(1)}, p > 0$ with respect to the sensor of reference so that the annihilating filter method can be run to retrieve the complete FRI signal.

As an illustration, consider that the observed signal is an infinite stream of Diracs with at most $K = 3$ Diracs per interval $\Omega = 2KLT$. With a single sensor, the required sampling kernel must be able to reproduce $2K$ exponentials $\underline{\alpha} = \{\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5\}$. Since $L = 7$ in this case, $\Omega = 42T$. With two sensors of order P_0 and P_1 observing respectively $x_0(t)$ and $x_1(t) = x_0(t - \Delta t)$, the parameters can be distributed in the two following sets where α_3 is repeated: $\underline{\alpha}_0 = \{\alpha_0, \alpha_1, \alpha_2, \alpha_3\}$ and $\underline{\alpha}_1 = \{\alpha_3, \alpha_4, \alpha_5\}$. Since α_3 is shared between the two sensors, the time delay Δt can be calculated with Equation (3.23). If the first sensor is chosen as reference, 3 Diracs can be perfectly retrieved by applying the annihilating filter method with $\{\tau_0, \tau_1, \tau_2, \tau_3\}$ from the first sensor and the registered data $\left\{e^{-\alpha_4 \Delta t} \cdot \tau_4, e^{-\alpha_5 \Delta t} \cdot \tau_5\right\}$ from the second sensor. Since the order at each sensor is decreased, the support of the corresponding sampling kernel is also reduced. Consequently, less samples are affected by a given Dirac at each sensor and $\Omega = \max(2KL_0T, 2KL_1T) = 30T$ as $L_0 = 5$ and $L_1 = 4$. The streams of Diracs that can now be reconstructed exhibit a higher rate of innovation ρ as the interval Ω where 3 Diracs can occur is reduced.

If K and N are given, the minimum order P required at each sensor to reconstruct the stream of Diracs and find the $N - 1$ time delays of the sensors is:

$$P = \left\lceil \frac{2K}{N} \right\rceil + 1.$$

The extreme case occurs when there are $N = 2K - 1$ sensors of order $P = 2$: each sensor has one parameter that tracks the novelty of the observed signal and one parameter that is shared to find the delay with other sensors.

3.5.2 Time-interleaved Analog-to-Digital Converters

In this section, previous results are used in a more general framework than FRI signals. Analog-to-digital converters (A/D converters) are necessary in the majority of signal processing and communications systems. They are often decisive as far as the performance of an electronic component is concerned. There exists a variety of different technologies for A/D converters such as flash, pipelined, successive-approximation or sigma-delta converters [127].

The conversion time of an analog signal into its digital representation is however not zero so if the input signal varies more than one least significant bit during that time, the output will be incorrect. One way to handle this problem is to add a sample-and-hold circuit (S/H) which holds the value of the signal so that the A/D converter has a stable signal and some time to perform an accurate conversion. A S/H circuit is an analog circuit with analog input/output and a digital control input [64]. A simplified model of a S/H circuit is presented in Figure 3.8 [96]. It is composed of a first order RC circuit and a switch driven by a clock. When the switch is closed ("Sample" mode), the input voltage $V_{in}(t)$ is connected to the capacitor C and the output signal $V_{out}(t)$ follows the input signal. When the switch is open ("Hold" mode), the capacitor is disconnected from the input and maintains the last value the input had when the switch opened. We consider that the time required by the A/D converter is short enough to assume that the hold value is constant with time.

The accuracy of a S/H circuit is mainly dictated by how quickly the input value is held according to the clock signal. Indeed, delays in the various components of a S/H circuit can weaken its accuracy. The most important delays in such circuits are the following [67]:

- the analog propagation delay Δt_b in the buffer amplifier;
- the switch driver digital delay Δt_d;and
- the aperture time Δt_s of the switch, which is the time required to physically disconnect the hold capacitor from the input voltage.

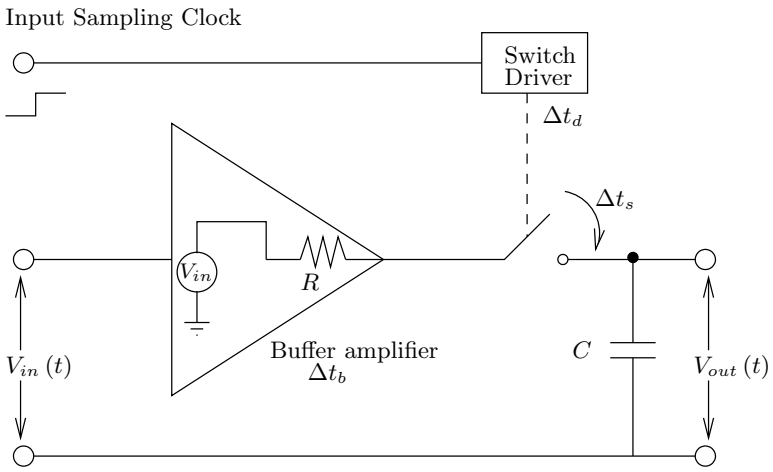


Figure 3.8: Simplified Sample-And-Hold circuit modeled as a first order RC circuit. When the switch is closed ("Sample" mode), the input signal $V_{in}(t)$ is connected to the capacitor C and the output signal $V_{out}(t)$ follows the input signal. When the switch is open ("Hold" mode), the capacitor C is disconnected from the input and maintains the last value the input had when the switch opened. Various time delays can weaken the accuracy of the circuit: the delay Δt_b in the buffer, the delay Δt_d in the switch driver and the aperture time Δt_s in the switch itself.

The overall delay Δt called effective aperture delay time, aperture delay or time skew, is given by [67]:

$$\Delta t = \Delta t_d - \Delta t_b + \frac{\Delta t_s}{2}.$$

In addition to aperture delays, the A/D converters suffer also from other errors like amplitude offset errors or gain errors which are not considered here. In the same conditions, all A/D converters with S/H circuit experience a different static aperture delay Δt. The aperture delay of a given A/D converter can however vary if the temperature of the chip changes, for example at start-ups or at different workloads.

A/D converter technologies are limited by the difficulty to achieve both high sampling rate and high resolution. The limitation of the resolution of an A/D converter is mainly determined by fundamental and inevitable imperfections of the physical device like sampling jitter or sample-and-hold non-linearity. To overcome the speed limitation of A/D conversion, an array of time-interleaved A/D converters can be considered [8]. This technique uses a number of A/D converters in parallel whose sampling times are interleaved in order to achieve a high sampling rate. A model of time-interleaved A/D converter array is presented in Figure 3.9. Each one of the N converters operates at a

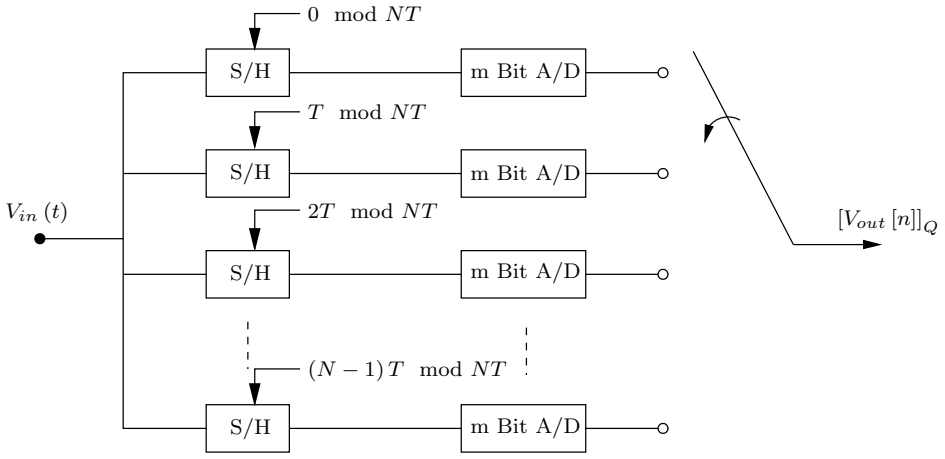


Figure 3.9: Time-interleaved converter arrays can convert at high sampling rate by using N A/D converters in parallel whose sampling times are interleaved. Each converter operates at a period NT but the output signal of the array is sampled at a period T.

sampling period NT but the overall conversion of the device is done at a sampling period T. One particularly interesting advantage of this technique is that the die-size and process requirements are reduced compared to other high-speed techniques. However, in addition to problems common to all A/D converters, time-interleaved A/D converters also suffers from new errors coming from the parallelization. Because in practice all manufactured A/D converters are different, each channel of the time-interleaved A/D converter contains errors called channel mismatches which results in a degradation of the resolution. Among the possible causes of channel mismatch errors, the aperture delay different for each channel is probably the most harmful error as the resulting sampling becomes periodically non-uniform [73] [69]. Indeed, since the A/D converters outputs data alternately, the time difference between two consecutive samples is not T anymore. The corresponding error is a function of the aperture delay and the slew rate (defined as dV/dt) of the input signal [85]. The error shows a periodic pattern with the same period as each individual converter. To compensate for different aperture delays and improve the resolution of the time-interleaved A/D converter array, it is necessary to estimate as precisely as possible the relative aperture delays Δt_i of each channel [29].

We now suppose that each S/H circuit has a sampling period $NT = 1$ and ignore quantization. Therefore, the time-interleaved A/D converter has an overall sampling pe-

riod $T = 1/N$. The physical RC circuit of each S/H device (Figure 3.8) can be described by the following differential equation when the switch is closed (*i.e.* in "Sample mode"):

$$\frac{V_{in}(t) - V_{out}(t)}{R} = C\frac{dV_{out}(t)}{dt}.$$

The solution of this differential equation is:

$$V_{out}(t) = \int_{-\infty}^{t} V_{in}(\lambda)\, e^{\frac{\lambda - t}{RC}} d\lambda. \tag{3.24}$$

By defining the function:

$$\varphi(t) = \begin{cases} e^{\frac{t}{RC}}, & t \leq 0, \\ 0, & t > 0, \end{cases}$$

Equation (3.24) can be rewritten at time $t = n \in \mathbb{Z}$ as:

$$\begin{aligned} V_{out}[n] &= \int_{-\infty}^{\infty} V_{in}(t)\, \varphi(t - n)\, dt \\ &= \langle V_{in}(t), \varphi(t - n) \rangle. \end{aligned}$$

Since each S/H circuit in the time-interleaved array has a different static aperture delay Δt_i, the analog value hold at the output of S/H circuit for the i-th channel of the time-interleaved A/D converter is:

$$V_{out}^{(i)}[n] = \langle V_{in}(t - \Delta t_i), \varphi(t - n) \rangle.$$

Similarly to [27], the filtering of $V_{out}^{(i)}[n]$ with $\frac{1}{\alpha}(e^{\alpha} z - 1)$ with $\alpha = 1/RC$ yields:

$$\begin{aligned} \mathbf{y}_i[n] &= \frac{1}{\alpha}\left(e^{\alpha} V_{out}^{(i)}[n+1] - V_{out}^{(i)}[n]\right) \\ &= \frac{1}{\alpha}\langle V_{in}(t - \Delta t_i), e^{\alpha}\varphi(t - n - 1) - \varphi(t - n)\rangle \\ &= \langle V_{in}(t - \Delta t_i), \beta_{\alpha}(t - n)\rangle, \end{aligned}$$

where $\beta_\alpha(t)$ is an E-spline of order 0 with parameter α. By computing:

$$\tau^{(i)} = \sum_n c_n^{(i)} \mathbf{y}_i[n],$$

where $c_n^{(i)}$ are the coefficients necessary to reproduce e^α with $\beta_\alpha(t-n)$, the different aperture delays Δt_i of the channels relatively to the first channel can be effectively retrieved with Equation (3.23):

$$\Delta t_i = \frac{1}{\alpha} \ln \frac{|\tau^{(i)}|}{|\tau^{(0)}|}. \tag{3.25}$$

Once the information on the aperture delays is known, post processing correction methods relocate the samples at the correct positions. The effect of the aperture delay mismatch is removed and the different A/D converters at each channel behave identically. Finally, the estimations of Equation (3.25) can be run appropriately to find the changing aperture delays at different working temperatures of the chip.

Chapter 4

Image Feature Extraction and Registration

4.1 Introduction

IN this chapter, we do not consider exclusively the problem of multiview acquisition of FRI signals but we look at the case of real images that are not necessarily FRI. In particular, we focus on the case of multiview images and on the problem of their accurate registration. However, the approach taken to tackle this problem is very similar to the framework employed during the FRI analysis.

The registration of two images f_1 and f_2 is the problem of finding the transformation $\mathcal{T}$ that relates them. This is often modeled as:

$$f_2(x, y) = f_1(\mathcal{T}(x, y)).$$

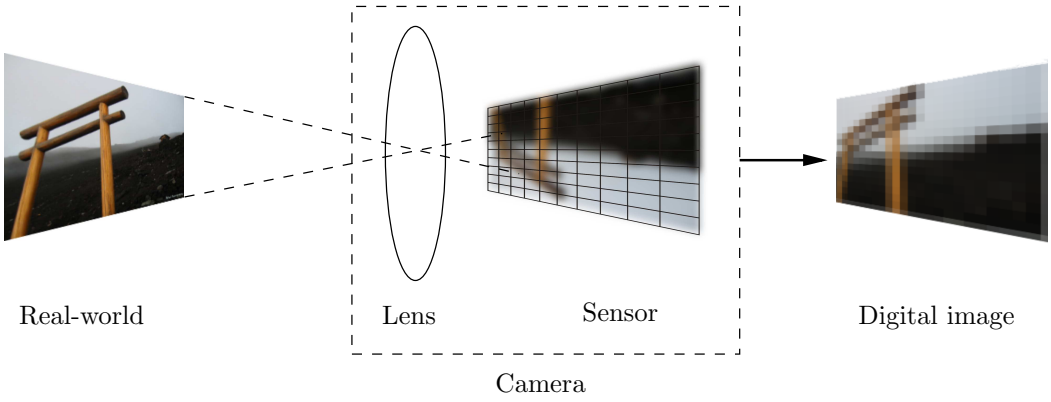
The difficulty of this problem comes from the fact that f_1 and f_2 are not available and the registration has to be performed using the corresponding discrete images $\mathbf{g}_1$ and $\mathbf{g}_2$ that have been captured with an acquisition device like a digital camera. In addition to sampling the observed scenes, the acquisition device also introduces non-linear modifications to the input signals. Thus, the estimation of $\mathcal{T}$ given $\mathbf{g}_1$ and $\mathbf{g}_2$ should take into account as much as possible the various characteristics of the acquisition device and, when possible, take advantage of them.

The approach taken in this study is to consider the set of feature-based registration methods whose accuracy depends on the way features are extracted from $\mathbf{g}_1$ and $\mathbf{g}_2$. This chapter introduces two novel approaches for feature extraction that can be used for registration purposes. Those two methods are based on FRI principles and achieve an exact extraction of the desired features in real images.

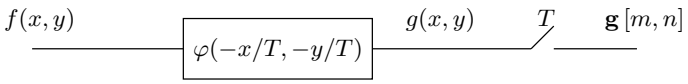
The chapter starts by introducing the model of a digital camera that is considered for image acquisition. In light of the previous chapter, it introduces an extraction method for global features based on the continuous moments and show how they can be used to retrieve up to an affine transform between images. A new local feature extraction technique is then presented. It allows the exact location of multiple parallel step edges in images from which corners can be inferred for registration. For each method, experiments are conducted to evaluate the registration accuracies and results are compared with other existing methods.

4.2 Image Acquisition Model

This section reviews the idealized image formation model considered in this thesis and describes how the samples are related to the observed view via the Point-Spread Function (PSF) of the camera lens. The diagram in Figure 4.1(a) presents the main components of a camera which lead to a digital image given the observed view. The light rays are first focused by the lens of the camera but, because a lens is never perfect, they are inevitably blurred before hitting the image sensor. As the quality of the material and the size of the lens decrease, the amount of blur introduced increases. Various other elements such as motion blur or atmospheric blur can also deteriorate the observed view. However they are not considered in this research. The blur introduced by the lens is characterized by the Point-Spread Function (PSF). In the literature, the PSF of a lens is very often modeled by a Gaussian pulse. In this research however, the PSF is modeled with B-spline functions for mainly two reasons. First, the shape of B-spline functions is very similar to a Gaussian pulse [121]. Second, B-splines possess properties such as polynomial reproduction that we want take to advantage of. B-splines have already been used as a PSF model ([95])



(a) From the real observed view to its digital representation



(b) Model of a camera in terms of filter and analog to digital converter

Figure 4.1: Camera model; (a) the incoming irradiance light field is blurred by the lens and sampled at the image sensor; (b) Equivalent model: f is the irradiance light field, φ is the point-spread function of the lens, g is the blurred irradiance light field, T is the sampling period and g is the sampled image.

but their polynomial reproduction capabilities have not been exploited yet. The last main component of a camera is the image sensor, either a CCD (Charge Coupled Device) or CMOS (Complementary Metal Oxide Semiconductor) array. It measures the amount of light received and outputs a sampled image, the term digital image usually referring to the quantized sampled image.

Figure 4.1(b) presents the equivalent idealized model to Figure 4.1(a) in terms of filter and analog-to-digital converter. The incoming continuous irradiance light-field $f(x,y)$ is first filtered with the function $\varphi(x,y)$. This two-dimensional function is the PSF that characterizes the lens and is assumed as known. The blurred observation $g(x,y) = f(x,y) * \varphi(-x/T,-y/T)$ is then uniformly sampled so that the discrete representation of

the observed view is given by the following equivalent expressions:

$$\begin{aligned}\mathbf{g}[m,n] &= g(m,n) \\ &= {}_{T}\downarrow(\varphi(-x/T,-y/T) * f(x,y)) \\ &= \iint_{\Omega} f(x,y)\,\varphi(x/T-m,y/T-n)\,dxdy \\ &= \langle f(x,y), \varphi(x/T-m,y/T-n)\rangle \end{aligned} \tag{4.1}$$

where $x, y \in \mathbb{R}$, $m, n \in \mathbb{Z}$. We assume throughout this research that the sampling period is equal to T in both dimensions in order to simplify notations. As seen on Equation (4.1), the impulse response of the filter representing the lens is expanded by a factor T corresponding to the sampling period. As in the sampling theory of FRI signals, the PSF is thereafter referred to as the sampling kernel $\varphi(x,y)$ of the acquisition device. The sampling kernel is the time-reversed version of the impulse response of the filter in Figure 4.1(b).

The whole sampling process has noticeable properties and various assumptions can be made to simplify the problem. The sampling process described by Equation (4.1) is part of a more general framework represented by the so-called two-dimensional Fredholm integral equation of the first kind defined as [37]:

$$\mathbf{g}[m,n] = \iint f(x,y)\,\varphi(x/T,y/T,m,n)\,dxdy. \tag{4.2}$$

Any problem written in this form and which tries to find f given φ and $\mathbf{g}$, is an *ill-posed inverse problem* in the Hadamard sense [48]: a small perturbation on $\mathbf{g}$ can cause large variations on the estimated solution. If we assume that the blur introduced by the lens does not depend on the position in the image, then the sampling kernel φ in Equation (4.2) is *spatially invariant*:

$$\varphi(x/T,y/T,m,n) = \varphi(x/T-m,y/T-n).$$

Plugging this expression in Equation (4.2) gives the sampling Equation (4.1). The kernel

$\varphi(x,y)$ can also have a *compact support* $\Omega = [-L,L] \times [-L,L]$:

$$\varphi(x,y) = 0 \quad \forall (x,y) \notin \Omega.$$

Besides 2-D sampling kernels are often *variable separable kernels* which implies that the two-dimensional function $\varphi(x,y)$ is a tensor product of two one-dimensional functions:

$$\varphi(x,y) = \varphi_1(x) \otimes \varphi_2(y).$$

Under this assumption, Equation (4.1) becomes:

$$\mathbf{g}[m,n] = \iint f(x,y)\, \varphi_1(x/T - m)\, \varphi_2(y/T - n)\, dxdy. \tag{4.3}$$

Two dimensional B-splines obtained by tensor product of the same 1-D function with variables x and y are an example of 2-D kernels that are both variable separable and rotationally symmetric.

4.3 Global Feature Extraction

4.3.1 Continuous moments of an image

Since the first work of Hu [57] on image moments, functions of moments have been extensively used in pattern recognition to build features that are invariant to a given transformation [56]. Moments have also been used in various ways to perform image registration [34] [72] [130]. Traditionally, the observed view $f(x,y)$ is not available so the true moments $m_{p,q}$ of the continuous function $f(x,y)$ cannot be directly computed. Instead, they are approximated from the acquired image $\mathbf{g}$ using the discretized version of Equation (3.8):

$$\underline{m}_{p,q} = \sum_{m,n} \mathbf{g}[m,n]\, (mT)^p\, (nT)^q. \tag{4.4}$$

When the resolution of $\mathbf{g}$ gets low, the discrete moments $\underline{m}_{p,q}$ do not provide a good approximation anymore and the discrepancy between the true and the discrete moments degrades the performance of any moment-based techniques.

As shown in [27] and [110], it is possible to compute the exact moments of a FRI signal from its sampled version, provided that the sampling kernel is known and satisfies the Strang-Fix conditions. In fact, this result is correct not only in the case of FRI signals but also for any signals. We thus propose to use these results on real images in order to extract the true *continuous moments* of a real object f from its samples $\mathbf{g}$. Following Equation (2.13) and using Fubini's theorem, the continuous moments are obtained by linear combinations of the samples with the coefficients $c_{m,n}^{(p,q)}$ for polynomial reproduction (Equation (2.16)) as follows:

$$\begin{aligned} m_{p,q} &= \iint f(x,y)\, x^p y^q dxdy \\ &\overset{(a)}{=} \iint f(x,y) \sum_m \sum_n c_{m,n}^{(p,q)} \varphi(x/T - m, y/T - n)\, dxdy \\ &= \sum_m \sum_n c_{m,n}^{(p,q)} \iint f(x,y)\, \varphi(x/T - m, y/T - n)\, dxdy \\ &\overset{(b)}{=} \sum_m \sum_n c_{m,n}^{(p,q)} \mathbf{g}[m,n], \end{aligned} \tag{4.5}$$

where (a) and (b) come respectively from Equations (2.16) and (4.1). Once the continuous geometric moments are obtained, other types of continuous moments (*e.g.* central or complex moments) can be calculated as well with Equations (3.9) and (3.12).

Figure 4.2 shows how the estimation of the moments using Equation (4.4) or (4.5) changes when the resolution decreases. For this experiment, 20 standard images (*e.g.* Lena, Goldhill, Peppers, Mandril) of size 512x512 are each artificially blurred and down-sampled to generate different square images with resolutions 256, 128, 64, 32, 16 and 8 pixels. Given these low resolution images, the estimated moments $\hat{m}_{p,q}$ using either Equations (4.4) or (4.5) are compared to the true moments $m_{p,q}$ of the original image before sampling by calculating the normalized distance between them:

$$d_k = \sum_{i+j=k} \frac{(\hat{m}_{i,j} - m_{i,j})^2}{m_{i,j}^2},$$

where k defines the order of the moments considered. Figure 4.2 shows the variation of the average normalized distance d_2 with respect to the resolution of the sampled images for moments of order 2. When the sampling kernel is known and reproduces polynomials,

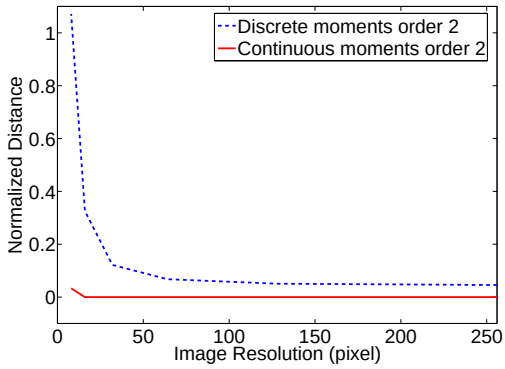


Figure 4.2: Effects of the change of resolution in the estimation of image moments of order 2. Dashed lines represent the normalized distances between the real moments and the discrete moments obtained from Equation (4.4). Solid lines represent the normalized distances between the real moments and the continuous moments obtained from Equation (4.5). As the resolution decreases, the discrete moments diverge whereas the continuous moments remain very accurate, even in the presence of noise. Similar observations are made for higher order moments.

the moments obtained with Equation (4.5) provide much more accurate results than those obtained with Equation (4.4).

4.3.2 Affine Registration of Signals using Moments

Moments of an image can be used as features for registration. Since they are obtained from all the samples of the considered image, moments convey global information on the image. Thus, in order to use the moments for registration, the observed views should not have new objects appearing or disappearing. We consider the case of objects which are always visible on a uniform background.

Let $\mathbf{g}_1$ and $\mathbf{g}_2$ be two acquired images of the views f_1 and f_2 obtained as in Equation (4.1). Using the continuous moments, we want to find the transformation $\mathcal{T}_2$ which relates the coordinates of f_2 to the coordinates of f_1 (see Equation (1.1)). We assume that the transformation $\mathcal{T}_2$ is an affine transformation represented by a translation $\mathbf{t}$ in x and y directions and by a 2x2 matrix $\mathbf{A}$ composed of a rotation θ, a scaling $(\mathrm{X}_{scale}, \mathrm{Y}_{scale})$ and a shear $(\mathrm{X}_{shear}, \mathrm{Y}_{shear})$:

$$\mathcal{T}_2\left(x, y\right) = \mathbf{A}\left[x \quad y\right]^T + \mathbf{t} \tag{4.6}$$

with

$$\mathbf{A} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \cdot \begin{bmatrix} \mathrm{X}_{scale} & \mathrm{X}_{shear} \\ \mathrm{Y}_{shear} & \mathrm{Y}_{scale} \end{bmatrix}.$$

Let $\mathbf{C}_i, i = 1, 2$ be the covariance matrix of f_1 and f_2 respectively:

$$\mathbf{C}_i = \begin{bmatrix} \mu_{2,0}^{(i)} & \mu_{1,1}^{(i)} \\ \mu_{1,1}^{(i)} & \mu_{0,2}^{(i)} \end{bmatrix},$$

where $\mu_{p,q}^{(i)}$ are the central moments of order $(p+q)$ of f_i. Sprinzak and Werman showed in [112] that the problem of finding the affine transformation matrix $\mathbf{A}$ between two point sets is equivalent to finding a 2x2 orthonormal matrix $\mathbf{R}$ between the point sets in their canonical forms. The canonical form occurs when the covariance matrices $\mathbf{C}_i$ are equal to the identity matrix. In [56], Heikkilä uses the following transform to convert the views into their canonical forms. Let $\mathbf{F}_i$ be the Cholesky decomposition of $\mathbf{C}_i$:

$$\mathbf{C}_i = \mathbf{F}_i \mathbf{F}_i^T \quad \text{where} \quad \mathbf{F}_i = \begin{bmatrix} \sqrt{\mu_{2,0}^{(i)}} & 0 \\ \frac{\mu_{1,1}^{(i)}}{\sqrt{\mu_{2,0}^{(i)}}} & \sqrt{\mu_{0,2}^{(i)} - \frac{\mu_{1,1}^{(i)2}}{\mu_{2,0}^{(i)}}} \end{bmatrix}.$$

The affine matrix $\mathbf{A}$ can now be expressed as:

$$\mathbf{A} = \mathbf{F}_2 \mathbf{R} \mathbf{F}_1^{-1}. \tag{4.7}$$

The previous transform on f_1 and f_2 gives rise to $\bar{f}_1$ and $\bar{f}_2$ respectively:

$$\bar{f}_i(x, y) = f_i\left(\mathbf{F}_i^{-1}(x - m_{1,0}, y - m_{0,1})\right). \tag{4.8}$$

These views are now related only by a rotation with a single unknown parameter α instead of 4 parameters in the case of an affine transform:

$$\bar{f}_2(x, y) = \bar{f}_1(\mathbf{R}(x, y)), \tag{4.9}$$

where $\mathbf{R}$ is the orthonormal matrix representing either a rotation or a mirror reflection

and a rotation. By considering the complex moments $\bar{C}_{p,q}^{(i)}$ of $\bar{f}_i$, it is possible to calculate the rotation parameter. Indeed, we have:

$$\bar{C}_{p,q}^{(2)} = e^{j(p-q)\alpha} \cdot \bar{C}_{p,q}^{(1)} \quad \Rightarrow \quad \alpha = \frac{arg\left\{\bar{C}_{p,q}^{(2)}\right\} - arg\left\{\bar{C}_{p,q}^{(1)}\right\}}{p-q} \mod \frac{2\pi}{p-q},$$

where $arg\{\cdot\}$ is the phase angle of a complex number. Moreover, the complex moments $\bar{C}_{p,q}, p+q=3$, are related to the central moments $\mu_{p,q}, p+q \leq 3$ as follows [56]:

$$\begin{aligned}
Re\left\{\bar{C}_{2,1}\right\} &= \left(\mu_{3,0}\mu_{0,2} - 2\mu_{2,1}\mu_{1,1} + \mu_{2,0}\mu_{1,2}\right)\mu_{2,0}^{-1/2}\kappa^{-1}, \\
Im\left\{\bar{C}_{2,1}\right\} &= \left(-\mu_{3,0}\mu_{1,1}\mu_{0,2} + \mu_{2,1}\mu_{2,0}\mu_{0,2} + 2\mu_{1,1}^2\mu_{2,1} - 3\mu_{1,1}\mu_{2,0}\mu_{1,2} + \mu_{2,0}^2\mu_{0,3}\right)\mu_{2,0}^{-1/2}\kappa^{-3/2}, \\
Re\left\{\bar{C}_{3,0}\right\} &= \left(\mu_{3,0}\mu_{0,2}\mu_{2,0} - 4\mu_{3,0}\mu_{1,1}^2 + 6\mu_{2,1}\mu_{1,1}\mu_{2,0} - 3\mu_{2,0}^2\mu_{1,2}\right)\mu_{2,0}^{-3/2}\kappa^{-1} \\
Im\left\{\bar{C}_{3,0}\right\} &= (-3\mu_{3,0}\mu_{1,1}\mu_{0,2}\mu_{2,0} + 4\mu_{1,1}^3\mu_{3,0} + 3\mu_{2,1}\mu_{2,0}^2\mu_{0,2} - 6\mu_{1,1}^2\mu_{2,1}\mu_{2,0}, \\
&\quad +3\mu_{1,1}\mu_{2,0}^2\mu_{1,2} - \mu_{2,0}^3\mu_{0,3})\mu_{2,0}^{-3/2}\kappa^{-3/2},
\end{aligned} \tag{4.10}$$

where $\kappa = \mu_{2,0}\mu_{0,2} - \mu_{1,1}^2$, and $Re\{\cdot\}$ and $Im\{\cdot\}$ refer respectively to the real and imaginary part of a complex number.

Thus the whitening transform does not need to be applied explicitly as only the continuous central moments of the original views are necessary and can be obtained using Equation (4.5) and Equation (3.8). In this way, the matrices $\mathbf{R}$ and $\mathbf{F}_i$ can be computed to estimate the matrix $\mathbf{A}$. The translation $\mathbf{t}$ is calculated using the barycenters of each image and Equation (4.6):

$$\mathbf{t} = \begin{bmatrix} x_b^{(2)} & y_b^{(2)} \end{bmatrix}^T - \mathbf{A}\begin{bmatrix} x_b^{(1)} & y_b^{(1)} \end{bmatrix}^T.$$

Since in the absence of noise we can retrieve the exact continuous moments, it is thus possible in theory to register low-resolution images exactly.

4.3.3 Registration Experiments with Continuous Moments

To measure the accuracy of the estimated transformation, the average and maximum geometric registration error, ε and $\varepsilon_{\max}$ respectively, are calculated as in [129]:

$$\varepsilon = \frac{1}{N^2} \sum_x \sum_y \| \mathcal{T}(x,y) - \tilde{\mathcal{T}}(x,y) \|_2, \tag{4.11}$$

$$\varepsilon_{\max} = \max_{x,y} \| \mathcal{T}(x,y) - \tilde{\mathcal{T}}(x,y) \|_2, \tag{4.12}$$

where $\tilde{\mathcal{T}}$ is the estimated affine transformation, $\mathcal{T}$ is the exact affine transformation and N is the size of the considered images f_1 and f_2. These error measures compute the distance between the calculated locations of a point (x, y) after transformation with $\mathcal{T}$ and after $\tilde{\mathcal{T}}$. If the estimated transformation is perfectly calculated, then the distance between the two new points is zero. In the case of translation only, ε and $\varepsilon_{\max}$ are the same since the registration error is the same over the whole image plane. However, if rotation occurs, $\varepsilon_{\max}$ and ε are different as an error in the estimation of the angle means that the registration error will increase with the distance from the center of rotation.

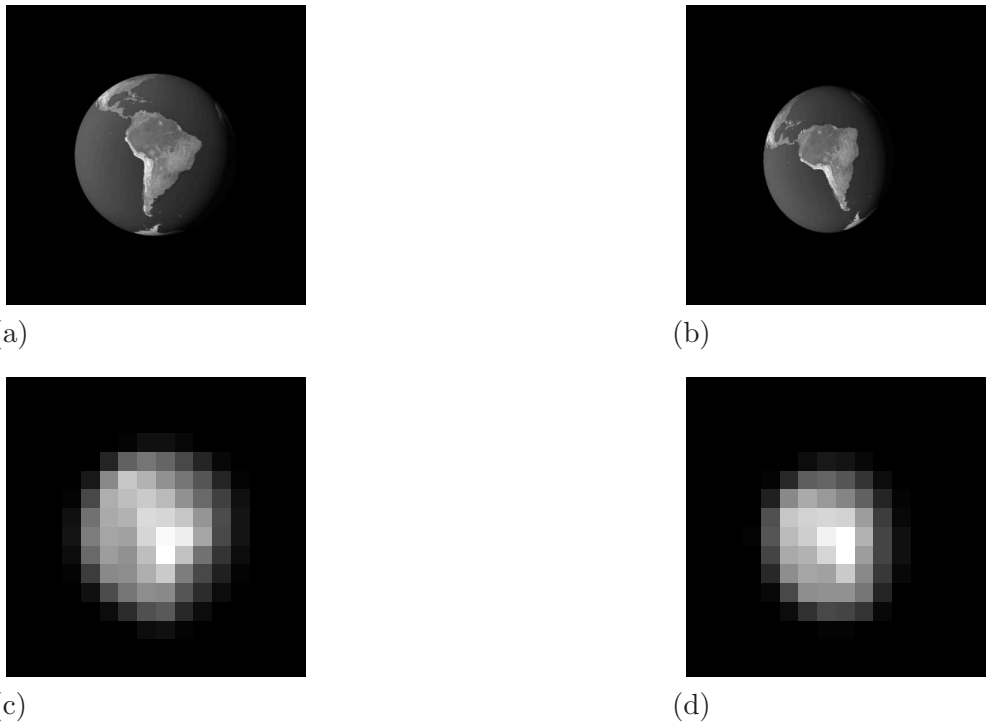
In Figure 4.3 (a)-(b), two high resolution images of size 512x512 pixels are considered as the two different views f_1 and f_2 of the same scene. The affine transformation between f_1 and f_2 consists of a rotation of an angle $\theta = 36$ degrees, a scaling factor of $(X_{scale}, Y_{scale}) = (0.8, 0.9)$, a shear factor of $(X_{shear}, Y_{shear}) = (0.1, -0.1)$ and a translation of -12 pixels and 7 pixels in X and Y direction respectively. We therefore have:

$$\mathcal{T}_2(x,y) = \mathbf{A} \begin{bmatrix} x & y \end{bmatrix}^T + \mathbf{t}$$

with

$$\mathbf{A} = \begin{bmatrix} 0.706 & -0.4481 \\ 0.3893 & 0.7869 \end{bmatrix} \quad \text{and} \quad \mathbf{t} = \begin{bmatrix} -12 \\ 7 \end{bmatrix}.$$

These two views are sampled with a cubic B-spline to generate two low-resolution images $\mathbf{g}_1$ and $\mathbf{g}_2$ of size 16x16 pixels (decimation factor of 32) as shown in Figure 4.3 (c)-(d). Given these two low-resolution images, we apply the registration method with the



(a) (b)

(c) (d)

Figure 4.3: (a)-(b) The two original views f_1 and f_2: 512x512 pixels each; (c)-(d) The two corresponding acquired low-resolution images g_1 and g_2: 16x16 pixels each. Original image "Blue Marble" by NASA's Earth Observatory.

continuous moments to estimate $\mathbf{A}$ and $\mathbf{t}$. The calculated affine transform is:

$$\tilde{\mathbf{A}} = \begin{bmatrix} 0.7064 & -0.4471 \\ 0.3885 & 0.7872 \end{bmatrix} \quad \text{and} \quad \tilde{\mathbf{t}} = \begin{bmatrix} -12.004 \\ 7.007 \end{bmatrix}.$$

The average and maximum geometric registration errors are:

$$\begin{aligned} \varepsilon &= 0.0413 \quad \text{pixel}, \\ \varepsilon_{\max} &= 0.148 \quad \text{pixel}. \end{aligned}$$

When calculating the affine transform from the true moments computed from f_1 and f_2, the same registration errors are obtained. For comparison, the same simulation is run with the discrete moments $\underline{m}_{p,q}$. As expected, they do not perform as well at this resolution.

The retrieved transformation is in this case:

$$\tilde{\mathbf{A}} = \begin{bmatrix} 0.7244 & -0.4475 \\ 0.3941 & 0.7953 \end{bmatrix} \quad \text{and} \quad \tilde{\mathbf{t}} = \begin{bmatrix} -12.49 \\ 8.23 \end{bmatrix}$$

for average and maximum errors of:

$$\begin{aligned} \varepsilon &= 11.2 \quad \text{pixel}, \\ \varepsilon_{\max} &= 40.4 \quad \text{pixel}. \end{aligned}$$

The improvement of the average registration error is by a factor of 280 in this simulation.

4.4 Local Feature Extraction

4.4.1 Step Edge as an FRI Signal

In this section, the considered features are local. When working at low-resolution, features are usually more difficult to find and locate accurately as each sample integrates a larger part of the original scene. Thus the properties of very localized features such as corners can be lost when images are acquired at low resolution. One dimensional features like edges are however more resilient to downsampling as they usually span a larger part of the scene. We now focus on the extraction of straight step edges.

A straight step edge is described by three parameters, namely its amplitude α, its orientation θ and its offset γ with respect to a given axis. This model of step edges is presented in Figure 4.4. We demonstrate how to retrieve the exact parameters from the samples. Let $\vec{N}$ be the vector normal to the edge and $\vec{d}$ the vector of any point (x, y) in $\mathbb{R}^2$:

$$\vec{N} = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix} \quad , \quad \vec{d} = \begin{pmatrix} x - \gamma \\ y \end{pmatrix}.$$

Given $\vec{N}$ and $\vec{d}$, a step edge function $h(x, y)$ can be expressed as:

$$h(x, y) = \alpha H\left(\left\langle \vec{d}, \vec{N} \right\rangle\right), \tag{4.13}$$

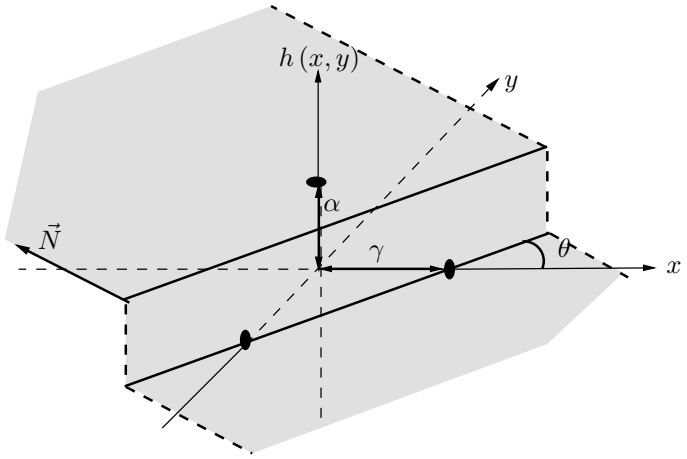


Figure 4.4: Step edge model. A straight step edge is described by 3 parameters: its amplitude α, its orientation θ and its offset γ.

where H is the unit step function defined as:

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

4.4.2 Step Edge Extraction

We now assume that the signal $h(x, y)$ is sampled using a 2-D B-spline sampling kernel $\varphi(x, y)$. Therefore, following Equation (4.1), the obtained samples $\mathbf{h}[m, n]$ are simply:

$$\mathbf{h}[m, n] = \langle h(x, y), \varphi(x/T - m, y/T - n) \rangle .$$

The set of samples is filtered with a finite difference operator to give the differentiated samples $\mathbf{d}[m, n]$:

$$\mathbf{d}[m, n] = \mathbf{h}[m + 1, n] - \mathbf{h}[m, n] . \tag{4.14}$$

It is shown in [27] that the samples $d[m, n]$ are the samples that would have been directly obtained from the inner product of the derivative of $h(x, y)$ along x and the modified kernel $\varphi(x, y) * \beta_0(x)$. The differentiation of a discrete sampled signal is thus related to

the derivation of its continuous counterpart as follows:

$$\begin{aligned}\mathbf{d}[m,n] &= \left\langle \frac{dh(x,y)}{dx}, \varphi(x/T-m, y/T-n) * \beta_0(x/T-m) \right\rangle \\ &= \left\langle \frac{dh(x,y)}{dx}, (\varphi_1(x/T-m) * \beta_0(x/T-m)) \otimes \varphi_2(y/T-n) \right\rangle,\end{aligned}$$

where the kernel φ has been assumed to be separable. Compared to $\varphi(x/T, y/T)$, the new kernel $\varphi(x/T, y/T) * \beta_0(x/T)$ can reproduce polynomials with one degree higher than φ along the x direction and has a support increased by one unit on the x axis. When the sampling kernel is $\varphi(x,y) = \beta_{\mathrm{P}}(x) \otimes \beta_{\mathrm{P}}(y)$, the modified kernel is a 2-D B-spline kernel of degree $\mathrm{P}+1$ along x and P along y:

$$\mathbf{d}[m,n] = \left\langle \frac{dh(x,y)}{dx}, \beta_{\mathrm{P}+1}(x/T-m) \otimes \beta_{\mathrm{P}}(y/T-n) \right\rangle.$$

Moreover, the first derivative of the unit step function is given by:

$$\frac{dh(x,y)}{dx} = -\alpha \sin\theta \cdot \delta\left(\vec{d} \cdot \vec{N}\right),$$

which finally yields to the following relation:

$$\mathbf{d}[m,n] = -\alpha \sin\theta \cdot \left\langle \delta\left(\vec{d} \cdot \vec{N}\right), \varphi(x/T-m, y/T-n) * \beta_0(x/T-m) \right\rangle.$$

We now compute the weighted sum of the differentiated samples affected by the edge in row n with the coefficients $c_m^{(p)}$ used for reproduction of polynomial x^p with the modified kernel $\varphi_1(x/T-m) * \beta_0(x/T-m)$:

$$\tau_{p,n} = \sum_{m \in \mathcal{S}_n} c_m^{(p)} \mathbf{d}[m,n], \tag{4.15}$$

where $\mathcal{S}_n$ is the set of column indices of the samples affected by the edge in row n. It can be shown that (see Appendix A):

$$\tau_{p,n} = -\alpha \sum_{j=0}^{p} \binom{p}{j} \frac{m_{p-j}}{(\tan\theta)^{p-j}} \left(\gamma + \frac{n}{\tan\theta}\right)^j, \tag{4.16}$$

where m_j are the moments of the sampling kernel:

$$m_j = \int_{\infty}^{\infty} t^j \varphi_2(t)\, dt. \tag{4.17}$$

Since $\varphi_2(t)$ is known, its moments m_j can be computed numerically once and stored. Besides, we have $m_0 = 1$ because $\varphi_2(t)$ satisfies partition of unity. Also, for symmetric functions like B-splines, the odd order moments are equal to zero: $m_{2j+1} = 0$. In the case of B-spline, it turns out that the even order moments m_{2j} can be calculated analytically. To simplify notations, we now write the quantity $u_n = \gamma + \frac{n}{\tan\theta}$. For $p = 0, 1, 2, 3$, formula (4.16) becomes:

$$\begin{cases} \tau_{0,n} = -\alpha m_0 \\ \tau_{1,n} = -\alpha \left[m_0 u_n + \frac{m_1}{\tan\theta} \right] \\ \tau_{2,n} = -\alpha \left[m_0 u_n^2 + 2\frac{m_1}{\tan\theta} u_n + \frac{m_2}{(\tan\theta)^2} \right] \\ \tau_{3,n} = -\alpha \left[m_0 u_n^3 + 3\frac{m_1}{\tan\theta} u_n^2 + 3\frac{m_2}{(\tan\theta)^2} u_n + \frac{m_3}{(\tan\theta)^3} \right]. \end{cases}$$

Solving directly this system of equations for α, γ and θ leads to an ambiguity about the sign of the angle θ of the edge. To overcome this and find the angle θ, we consider instead two consecutive rows, *i.e.* n and $n+1$ and compute $\tau_{k,n}$ and $\tau_{k,n+1}$. It turns out that this approach gives a simple relation for θ:

$$\tan\theta = \frac{\tau_{0,n}}{\tau_{1,n+1} - \tau_{1,n}}. \tag{4.18}$$

The complete solution for a single step edge is then given by:

$$\begin{cases} \alpha &= -\tau_{0,n}, \\ \tan\theta &= \frac{\tau_{0,n}}{\tau_{1,n+1} - \tau_{1,n}}, \\ \gamma &= \frac{(n+1)\tau_{1,n} - n\tau_{1,n+1}}{\tau_{0,n}}, \end{cases} \tag{4.19}$$

where it has been assumed that $m_0 = 1$ and $m_1 = 0$. Thus Equations (4.19) allow the calculation of the exact parameters of a step edge from a sampled version using only two consecutive rows.

It is possible to extend this analysis to any number K of parallel step edges. Such

signal is specified by $2K+1$ parameters, namely (α_k, γ_k) for $k = 1, \ldots, K$ and the angle θ. It is expressed as:

$$h(x,y) = \sum_{k=1}^{K} \alpha_k H\left(\left\langle \vec{d_k}, \vec{N} \right\rangle\right), \tag{4.20}$$

with $\vec{d_k} = (x - \gamma_k, y)^T$. We prove in Appendix B that the angle θ can be retrieved using the same formula as in Equation (4.18). Moreover, by defining the quantity $\widehat{\tau}_{p,n}$ as follows:

$$\widehat{\tau}_{p,n} = \begin{cases} \tau_{0,n}, & p = 0, \\ \tau_{p,n} - \sum_{j=0}^{p-1} \binom{p}{j} \frac{m_{p-j}}{m_0 (\tan\theta)^{p-j}} \widehat{\tau}_{j,n}, & p > 0, \end{cases} \tag{4.21}$$

we show in Appendix C that this quantity can be written in the form of a powersum series:

$$\widehat{\tau}_{p,n} = \sum_{k=1}^{K} \lambda_k \cdot (u_{k,n})^p \qquad p = 0, 1, \ldots, M-1, \tag{4.22}$$

where $u_{k,n} = \gamma_k + \frac{n}{\tan\theta}$ and $\lambda_k = -\alpha_k m_0$. The K pairs of unknowns $\{\lambda_k, u_{k,n}\}$ can then be retrieved by applying the annihilating filter method (*a.k.a.* Prony's method) provided that $M \geq 2K$ [27,113,126]. Finally, these results can be extended to polynomial edges and higher order derivatives by considering the generalized annihilating filter method. Figure 4.5 shows the case of linear edges and is treated in details in Appendix D. A linear edge (or roof edge) can be modeled as:

$$h(x,y) = H\left(\left\langle \vec{d}, \vec{N} \right\rangle\right) \left[\lambda\left(-x\sin\theta + y\cos\theta\right) + \alpha\right]. \tag{4.23}$$

Interestingly, Equations (4.18), (4.21) and (4.22) are valid in this case too.

To determine the correct set of samples affected by a given edge, we first run a simple edge detector, *e.g.* a Canny edge detector. We then retrieve the samples on the row surrounding each position labeled as an edge. Since the kernel has a compact support, the number of samples affected by an edge is finite. The samples are then used to compute the parameters of potential step edges from Equations (4.19). Edges having the same parameters are fused together by averaging the parameters together and by increasing the weight of this edge by one. Thus, a step edge that has been extracted k times has a weight equal to k after fusion. Finally, edges with a weight below a given threshold are discarded

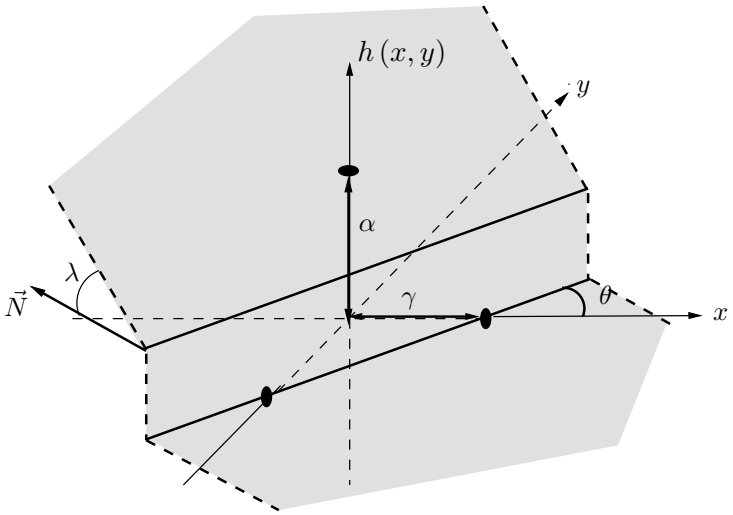


Figure 4.5: Linear edge model. A straight linear edge is described by 4 parameters: the amplitude α of the plane at the origin, the slope of the plane λ, the edge orientation θ and the edge offset γ.

in order to keep only edges with sufficiently large weights. This procedure is described in the pseudo-code of Algorithm 1. By considering only horizontal differentiations as in Equation (4.14), horizontal edges cannot be extracted. In practice, Algorithm 1 is run on the sampled image and on the transposed image to find all possible step edges, and both results are fused again.

Figure 4.6 shows how edge extraction is performed with algorithm 1 on a synthetic image presenting different step edges with various orientations, amplitudes and offsets. Figure 4.6(a) shows the original scene before acquisition. Figure 4.6(b) is the acquired image of size 64x64 pixels obtained with a quadratic B-spline sampling kernel. Figure 4.6(c) shows the differentiated samples $d_{m,n}$. Figure 4.6(d) shows the position of potential step edges using the Canny Edge detector. The retrieved edges are presented in Figure 4.6(e) and are also plotted against the original scene in Figure 4.6(f).

4.4.3 Registration Experiment from Extracted Edges

The step edge extraction technique described in Algorithm 1 is used to find possible step edges in low-resolution images. Then, edge intersections are used as local features for registration. To assess the accuracy of the proposed feature extraction method in the context of image registration, we compare it to the Harris corner detector in the following

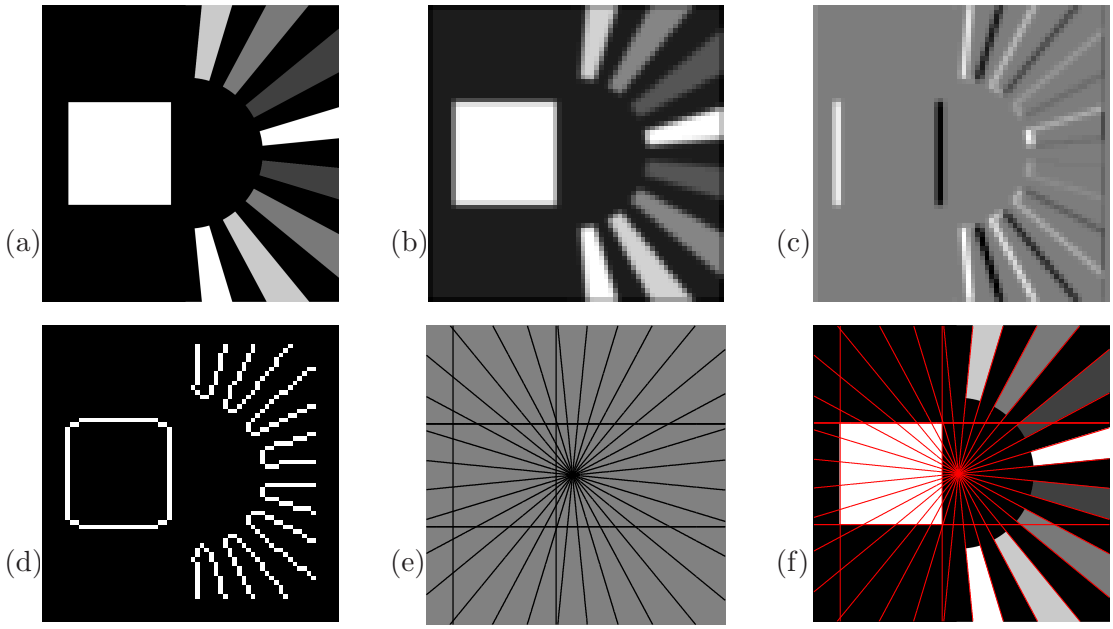


Figure 4.6: (a) Original image (1024x1024pixels); (b) Sampled image with a B-spline sampling kernel (64x64pixels); (c) Horizontally differentiated samples; (d) Canny edge detection; (e) Retrieved edges with Algorithm 1; (f) Retrieved edges plotted against the original image.

experiment.

Setup

A high resolution image of a simple scene favorable to both our method and the Harris corner detector is first acquired. The scene contains step edges, sharp corners, letters and textures. The acquired picture is then cropped at two different known locations to create two pictures f_1 and f_2 of size 512x512 pixels each with a slightly different field of view (see Figure 4.7 (a) and (b)). The transformation, here a translation, between each picture is therefore known exactly and is given by:

$$\mathbf{t} = \begin{bmatrix} 0 \\ 28 \end{bmatrix}.$$

Each image is then artificially downsampled with a quadratic B-spline of scale 8 giving two images of size 64x64 pixels each (see Figure 4.7 (c) and (d)). Features are then extracted from $\mathbf{g}_1$ and $\mathbf{g}_2$ using either the Harris corner detector or our step edge detector.

Algorithm 1 Subpixel Edge Location

1: Define $weight = 1$,
2: Run a Canny-like edge detector on sampled image (*e.g.* `edge(I,'canny')` in Matlab),
3: Compute the differentiated samples using Equation (4.14),
4: **for all** position (i, j) detected as an edge **do**
5: find the differentiated samples in the neighborhood of (i, j),
6: calculate $\tau_{0,j}$, $\tau_{1,j}$ and $\tau_{1,j+1}$ using Equation (4.15),
7: calculate α, $\tan\theta$ and γ using Equations (4.19),
8: store $[\alpha \quad \tan\theta \quad \gamma \quad weight]$ as a candidate edge
9: **end for**
10: **while** there exists similar edges **do**
11: Merge similar edges *i.e.* average $[\alpha \quad \tan\theta \quad \gamma]$ and add *weight* together,
12: **end while**
13: Discard edges having a too small weight.

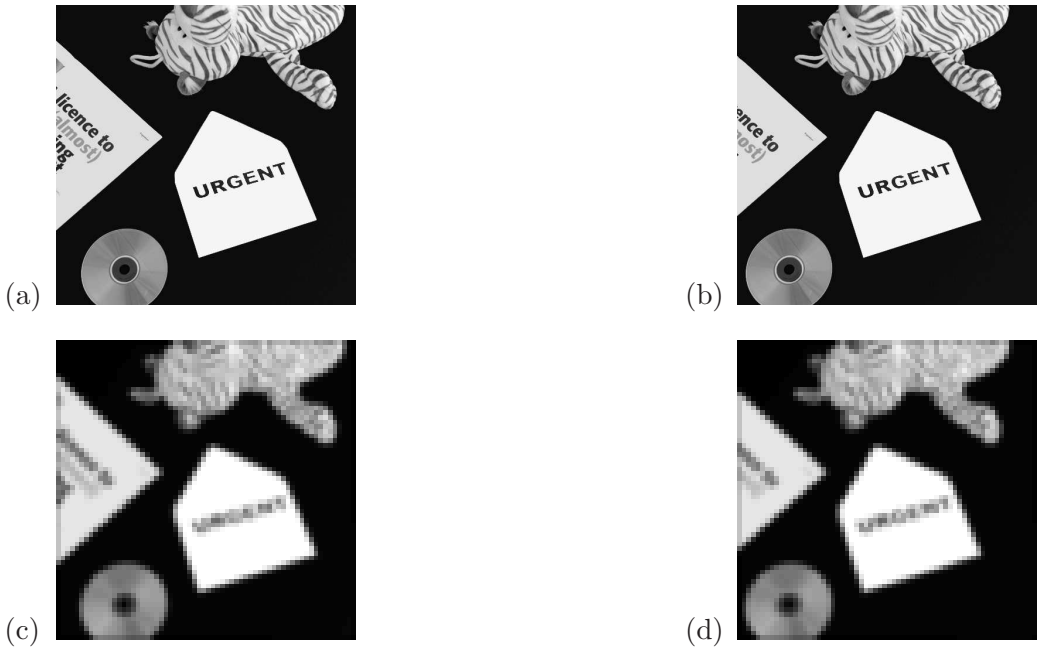


Figure 4.7: (a)-(b) Two high-resolution images of size 512x512 pixels with different field of view (translation $= (0, 28)$); (c) - (d) The same two images after sampling with a quadratic B-spline of order 8. Each image has a size 64x64 pixels.

Harris Corner Detector

The Harris corner detector was first presented by Harris and Stephens in [54]. This detector is based on the image intensities and models a corner as a point where the intensity changes greatly in two different directions.

For different shifts $[m, n]$, the intensity change is first calculated over a small win-

dow:

$$E[m,n] = \sum_{u,v} w[u,v] \left(\mathbf{g}[u+m, v+n] - \mathbf{g}[u,v]\right)^2,$$

where $w[u,v]$ is the integration window function *e.g.* a circular Gaussian window. In practice, if the shifts $[m,n]$ are small, the intensity change is approximated by:

$$E[m,n] = [u \quad v] \cdot M \cdot [u \quad v]^T,$$

where

$$M = \sum_{u,v} w[u,v] \begin{bmatrix} \mathbf{g}_x^2 & \mathbf{g}_x\mathbf{g}_y \\ \mathbf{g}_x\mathbf{g}_y & \mathbf{g}_y^2 \end{bmatrix}.$$

Here, $\mathbf{g}_x$ and $\mathbf{g}_y$ are the image derivatives of $\mathbf{g}$ after applying a finite difference operator along x and y directions respectively. Based on the eigenvalue analysis of the matrix M, a measure of corner response R is calculated as:

$$\begin{aligned} R &= \lambda_1\lambda_2 - k\left(\lambda_1 + \lambda_2\right)^2 \\ &= \det M - k\left(\mathrm{trace}M\right)^2, \end{aligned}$$

where λ_1 and λ_2 are the two eigenvalues of M and k is an empirical parameter usually set between 0.04 and 0.06. Now depending on the value of R, we have:

- $|R|$ small: the region considered is flat;
- $R < 0$: the region considered presents an edge; and
- R large: the region considered contains a corner.

The Harris corner detector used in our experiment is the subpixel implementation by Kovesi that can be found in [68]. Figure 4.8 shows the extracted features using both feature extraction techniques. Figure 4.8(a) and (b) show the Harris corners in two low-resolution images while Figure 4.8(c) and (d) show the features extracted using the proposed approach. Note that in this last case, no features are extracted on the contour of the tiger plush as model considered if the step edge, thus reducing the number of available features for registration.

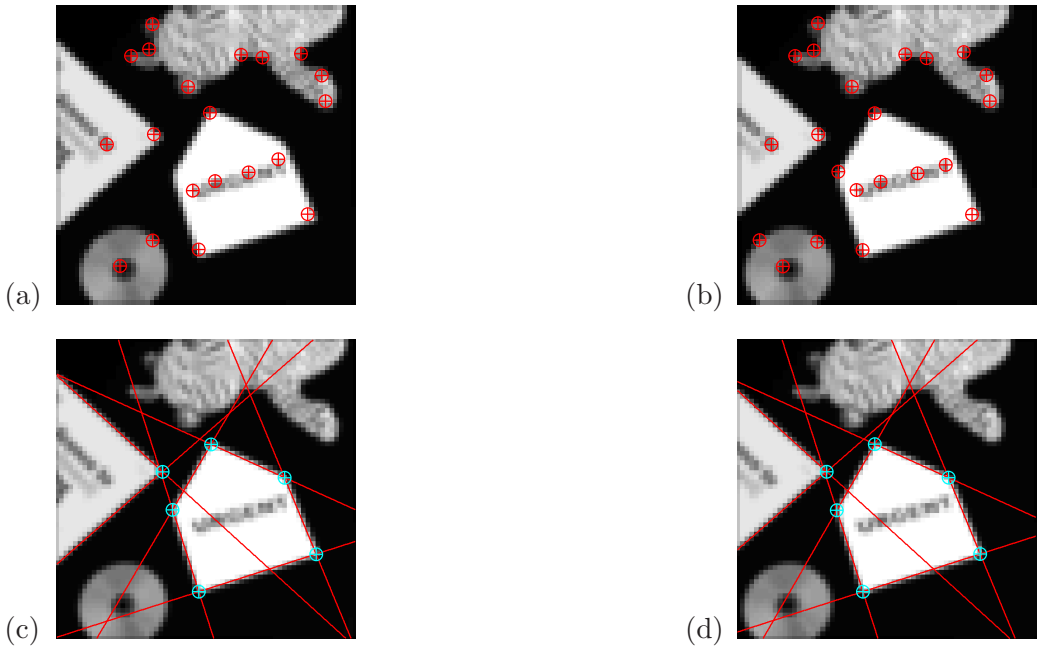


Figure 4.8: (a)-(b) Harris features detected in the low-resolution images; (c) - (d) Features detected using our step edge detector in the low-resolution images.

Feature-based Registration Procedure

We follow a similar registration procedure described by Capel and Zisserman in [18]. The local features considered in their work are corners extracted using the Harris corner detector. Although this detector does not achieve a high sub-pixel accuracy, precise registration can be obtained provided that a large number of corresponding features (several hundreds) are available on each image. As the number of extracted features increases, the inaccuracy of the Harris corner detector is counterbalanced by the large number of features involved in the estimation of the transformation. However on low-resolution images, only a small number of features, say between 10 to 20, can sometimes be extracted in each image and matched. In this case, the registration error can be large.

The same registration procedure as given in Algorithm 2 is used for the case of Harris features and our features. The functions implementing the correlation matching and the RANSAC robust estimation are available from [68]. The RANSAC (RANdom SAmple Consensus) algorithm first introduced by Fischler and Bolles in [33] is used to discard bad extracted features so that they are not taken into account for the estimation of the transformation. Figure 4.9 shows the feature correspondences after RANSAC estimation

Algorithm 2 Registration of two images with local features [18]

1: Extract local features at sub-pixel accuracy in both images,
2: Find putative feature correspondences using *e.g.* a correlation-based method on the feature neighborhoods,
3: n=1, {Start RANSAC}
4: **repeat**
5: Select randomly 3 putative correspondences to calculate the affine transformation $\mathcal{T}$,
6: Compute geometric image distance error for each putative correspondence,
7: Calculate number of inliers consistent with $\mathcal{T}$ as the number of feature correspondences whose distance error is less than a threshold,
8: n=n+1,
9: **until** n=N
10: Select the transformation $\mathcal{T}$ with the highest number of inliers {End RANSAC},
11: Re-calculate the transformation $\mathcal{T}$ using all the features considered as inliers.

with both types of features. With the Harris detector, eighteen corners have been successfully matched across images and are plotted against f_1 and f_2 on Figure 4.9(a). With our features, the six corners have been successfully matched across the two images and are plotted on Figure 4.9(a) against the high resolution images for a visual appreciation of the subpixel accuracy.

Registration Accuracy

The translation $\mathbf{t}$ between the two images is then calculated in each case by averaging the space differences between each feature correspondence. The estimated translation using the proposed features is:

$$\tilde{\mathbf{t}} = \begin{bmatrix} -0.15 \\ 28.13 \end{bmatrix}.$$

Because we only have a translation, the average and maximum registration errors as defined by Equations (4.11) and (4.12) are equal:

$$\varepsilon = \varepsilon_{\max} = 0.039 \quad \text{pixel.}$$

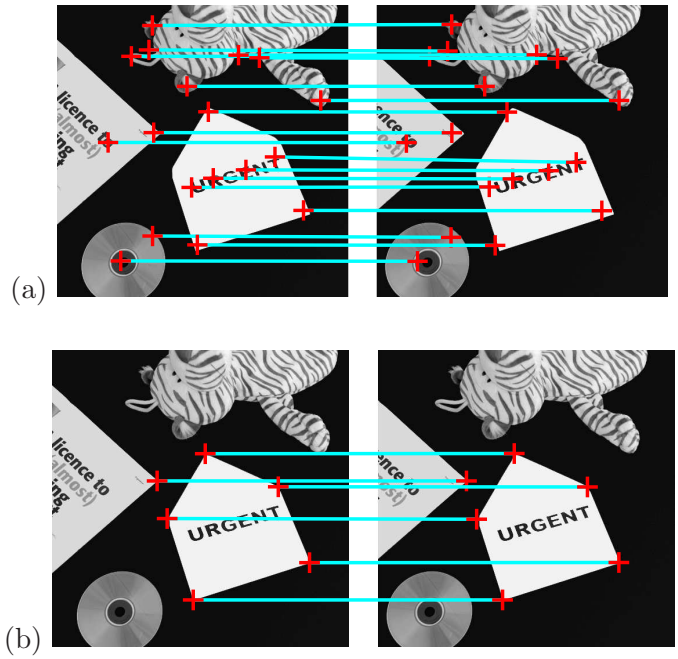


Figure 4.9: (a) Feature correspondences plotted against the high resolution images after RANSAC estimation with Harris features: 18 corners are matched; (b) Feature correspondences plotted against the high resolution images after RANSAC estimation with the features from our step edge detector: 6 corners are matched.

As for the Harris feature extraction case, the estimated translation is:

$$\tilde{\mathbf{t}} = \begin{bmatrix} -0.94 \\ 26.93 \end{bmatrix},$$

and the corresponding registration errors are:

$$\varepsilon = \varepsilon_{\max} = 2.04 \quad \text{pixel}.$$

Thus, although only half the number of corner points have been extracted with the proposed approach by comparison with the Harris corner detector, the registration accuracy is improved by a factor of 50 using the step edge extractor.

Chapter 5

Application to Image Super-resolution

5.1 Introduction

IMAGE super-resolution techniques aim at constructing a single detailed high-resolution image from a set of several low-resolution images of the scene taken from different positions. Because each low-resolution image captures a different view of the scene, it is possible to reconstruct an image in which details that cannot be seen on any of the acquired images become visible. Image super-resolution is therefore a solution to overcome the physical limitations of hardware capabilities.

The problem of image super-resolution can be conceptually divided into two sub-problems known as image registration and image reconstruction (see Figure 5.1). Image registration is necessary to find the disparity between the low-resolution images and image reconstruction consists of fusing the set of images into a single image and restoring the fused image by removing any blur and noise introduced by the acquisition device.

Since its first formulation by Tsai and Huang [58], the super-resolution problem has received much attention in the signal processing community. Most earlier works focused on the restoration stage assuming that traditional registration methods provided a solution sufficiently accurate and robust. However the quality of the restoration depends heavily on the accuracy of the registration. Indeed, the restoration process is a typical inverse

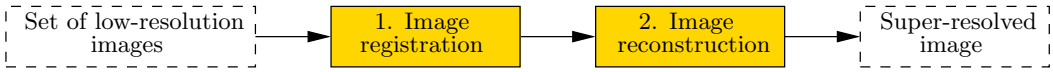


Figure 5.1: The two main stages of image super-resolution: image registration and image reconstruction. Input low-resolution images need first to be registered before a super-resolved image can be generated. The image reconstruction step involves image restoration whose output can be greatly improved with a correct registration.

ill-posed problem for which the presence of noise makes the search for a good solution very difficult. As it can be seen in Figure 5.1, the registration precedes the restoration step. Therefore the noise at the input of the restoration stage resulting from misregistration can be greatly reduced if registration is accurately performed. This would result in producing a super-resolved image of higher quality.

There has therefore been a recent shift in research on super-resolution towards the problem of achieving an accurate registration of low-resolution signals [128] [101]. Since the same motivation is also behind this research, it is therefore quite natural to test our registration algorithms in the context of image super-resolution.

5.2 Discrete Problem Formulation

Let $\mathbf{f}_i$ be the desired, ideal, high-resolution discrete image of size NxN representing the continuous scene f_i observed by the i-th camera and let $\mathbf{g}_i$ be the corresponding acquired image of size MxM, with $M < N$. Equation (4.1) can be discretized using quadrature methods into a linear system relating $\mathbf{g}_i$ and $\mathbf{f}_i$ with sampling period $T = N/M$ as:

$$\underline{\mathbf{g}}_i = \mathbf{M} \cdot \underline{\mathbf{f}}_i, \quad \text{where} \quad \underline{\mathbf{f}}_i = \mathrm{vec}\,(\mathbf{f}_i) \quad \text{and} \quad \underline{\mathbf{g}}_i = \mathrm{vec}\,(\mathbf{g}_i)\,. \tag{5.1}$$

The 'vec' operator stacks the columns of a matrix X into a column-vector $\underline{\mathbf{x}}$. The dimensions of $\underline{\mathbf{f}}_i$ and $\underline{\mathbf{g}}_i$ are respectively N^2x1 and M^2x1. Usually dense and large (M^2xN^2), the mapping matrix $\mathbf{M}$, obtained by discretization of the sampling kernel φ, represents the blurring and the downsampling process of the acquisition device. The i-th row of $\mathbf{M}$ consists of the weights used in the linear combination of all the elements of $\underline{\mathbf{f}}$ which give rise to the pixel in the i-th row of $\underline{\mathbf{g}}$. If we assume spatial invariance and compactness of the sampling kernel, $\mathbf{M}$ becomes sparse and exhibits some particular structures (*e.g.*

circulant-like, Toeplitz-like or Hankel-like) which depend on the boundary conditions considered outside the field of view. For example, by assuming periodic boundary conditions, the matrix $\mathbf{M}$ becomes a block-circulant matrix with cyclic shift value equal to T [89].

We now focus on the case where the sampling kernel is separable. Equation (5.1) becomes in this case:

$$\underline{\mathbf{g}}_i = (\mathbf{M}_1 \otimes \mathbf{M}_2) \cdot \underline{\mathbf{f}}_i \quad \Leftrightarrow \quad \mathbf{g}_i = \mathbf{M}_2 \cdot \mathbf{f}_i \cdot \mathbf{M}_1^T. \tag{5.2}$$

Here, $\mathbf{M}_1$ and $\mathbf{M}_2$ are both of size $M\text{x}N$. We therefore have $\mathbf{M} = \mathbf{M}_1 \otimes \mathbf{M}_2$. If the sampling kernel is considered rotationally symmetric, then $\mathbf{M}_1 = \mathbf{M}_2$. The main advantage of considering a space-invariant separable kernel is computational since only two $M\text{x}N$ matrices are necessary to describe the large $M^2\text{x}N^2$ matrix $\mathbf{M}$ due to the properties of the Kronecker product. It is also much more efficient as far as matrix-vector operations and singular values decompositions are concerned. The matrix $\mathbf{M}_2$ blurs and downsamples the columns of $\mathbf{f}$ whereas $\mathbf{M}_1$ operates on the rows of $\mathbf{f}$. We define $\mathrm{k}_1[m]$ as the discrete version of the 1-D sampling kernel $\varphi_1(x)$ by expanding it by T and then sampling it:

$$\mathrm{k}_1[m] = \varphi_1\left(\frac{m}{T} + \ell\right), \quad m \in \mathbb{Z}, \tag{5.3}$$

where ℓ is a desired offset. The sampling period therefore determines the resolution of the discrete approximation of $\varphi_1(x)$. Assuming periodic boundary conditions, the $[m, n]$ entry of $\mathbf{M}_1$ is given by:

$$\mathbf{M}_1[m, n] = \mathrm{k}_1[n - 1 - T(m - 1) \mod N].$$

It can be seen as a circulant matrix to which only one out of T rows have been kept, *i.e.*

the product of a downsampling matrix and blurring matrix. Thus for $T = 2$, we have:

$$\mathbf{M}_1 = \begin{pmatrix} \mathrm{k}_1[0] & \mathrm{k}_1[1] & \mathrm{k}_1[2] & \dots & \mathrm{k}_1[N-1] \\ \mathrm{k}_1[N-2] & \mathrm{k}_1[N-1] & \mathrm{k}_1[0] & \dots & \mathrm{k}_1[N-3] \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \mathrm{k}_1[3] & \cdots & \mathrm{k}_1[0] & \mathrm{k}_1[1] & \mathrm{k}_1[2] \\ \mathrm{k}_1[1] & \cdots & \mathrm{k}_1[N-2] & \mathrm{k}_1[N-1] & \mathrm{k}_1[0] \end{pmatrix}_{M\mathrm{x}N}.$$

$\mathbf{M}_2$ is obtained similarly after discretization of $\varphi_2(y)$. Unless stated otherwise, we consider in this research the case of a spatially invariant and separable sampling kernel.

5.3 Image Restoration

In image super-resolution, the image reconstruction step consists in interpolating the missing samples on the desired high resolution grid given the available registered samples. The resulting image $\tilde{\mathbf{g}}$ of size $N\mathrm{x}N$ is blurred since the effect of the lens has not yet been removed. Because of the interpolation step and possible misregistration, the image $\tilde{\mathbf{g}}$ is also perturbed by noise. The restoration problem consisting in removing the blur is expressed as:

$$\underline{\tilde{\mathbf{g}}} = \mathbf{K} \cdot \underline{\mathbf{f}}, \tag{5.4}$$

with

$$\underline{\tilde{\mathbf{g}}} = \underline{\mathbf{g}} + \underline{\mathbf{e}}, \tag{5.5}$$

where:

- $\underline{\mathbf{g}} = \mathrm{vec}(\mathbf{g})$ is the noiseless blurred image,
- $\underline{\mathbf{f}} = \mathrm{vec}(\mathbf{f})$ is the desired image,
- $\underline{\mathbf{e}}$ is the perturbation,
- $\mathbf{K}$ is the $N^2\mathrm{x}N^2$ matrix representing the pure blurring process.

We notice that the restoration problem is a particular case of the sampling problem (see Equation (5.1)) where $\mathbf{K} = \mathbf{M}$ for $T = 1$ (and $M = N$).

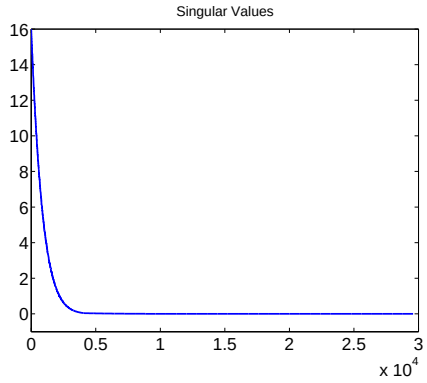


Figure 5.2: Singular values of the blurring matrix K obtained by discretization of the cubic B-spline with $N = 512$. The condition number of this matrix is equal to 7.8×10^{33}.

5.3.1 Discrete Ill-posed Problems

Because the sampling kernel φ has a smoothing effect, the matrix $\mathbf{K}$, of rank r, is inherently ill-conditioned: the singular values σ_i of $\mathbf{K}$ decay progressively to zero. For example, let φ be a 2-D cubic B-spline with scale 16 and let $N = 512$. The corresponding matrix $\mathbf{K}$ of size 512^2x512^2 is the Kronecker product of two matrices $\mathbf{K}_1$ and $\mathbf{K}_2$ so that: $\mathbf{K} = \mathbf{K}_2 \otimes \mathbf{K}_1$. The singular values of $\mathbf{K}$ are shown in Figure 5.2 for which the condition number, defined as the ratio between the largest and the smallest singular values, is equal to 7.8×10^{33}.

The problem of finding $\underline{\mathbf{f}}$ given $\underline{\tilde{\mathbf{g}}}$ and $\mathbf{K}$ is a linear discrete ill-posed inverse problem for which a square summable solution with finite l_2-norm is searched. In this case, standard inverse methods do not work since the perturbations within $\tilde{\mathbf{g}}$ are amplified by the very large singular values of $\mathbf{K}^{-1}$. Indeed, the singular value decomposition (SVD) of $\mathbf{K}$ is given by:

$$\mathbf{K} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T,$$

where $\mathbf{U}$ and $\mathbf{V}$ are matrices with orthonormal columns $\underline{\mathbf{u}}_i$ and $\underline{\mathbf{v}}_i$, and $\mathbf{\Sigma}$ is the diagonal matrix of singular values arranged in descending order. The direct inverse solution to

Equation (5.4) is:

$$\begin{aligned}
\underline{\mathbf{f}} &= \mathbf{K}^{-1}\underline{\tilde{\mathbf{g}}} \\
&= \sum_{i=1}^{r} \frac{\underline{\mathbf{u}}_i^T \underline{\tilde{\mathbf{g}}}}{\sigma_i} \underline{\mathbf{v}}_i \\
&= \sum_{i=1}^{r} \frac{\underline{\mathbf{u}}_i^T \left(\underline{\mathbf{g}} + \underline{\mathbf{e}}\right)}{\sigma_i} \underline{\mathbf{v}}_i \\
&= \sum_{i=1}^{r} \frac{\underline{\mathbf{u}}_i^T \underline{\mathbf{g}}}{\sigma_i} \underline{\mathbf{v}}_i + \underbrace{\sum_{i=1}^{r} \frac{\underline{\mathbf{u}}_i^T \underline{\mathbf{e}}}{\sigma_i} \underline{\mathbf{v}}_i}_{\text{error}} .
\end{aligned}$$

Thus, as the contributions of the small singular values of $\mathbf{K}$ are taken into account, the noise is amplified and the solution starts to be dominated by the noise.

5.3.2 Regularization Methods

Direct Regularization Methods

Discrete ill-posed problems are usually solved using regularization methods which transform an ill-posed problem into a well-posed one by dampening or filtering out the part of the solution corrupted by noise (*i.e.* due to small singular values) and by adding a constraint on a priori knowledge about the solution called the regularization term. One standard way to solve the restoration problem:

$$\underline{\tilde{\mathbf{g}}} = \mathbf{K} \cdot \underline{\mathbf{f}}$$

is to minimize the residual norm that measures the quality of the solution:

$$\min_{\underline{\mathbf{f}}} \left\| \underline{\tilde{\mathbf{g}}} - \mathbf{K} \cdot \underline{\mathbf{f}} \right\|_2 \tag{5.6}$$

Various constraints can be added to the above minimization problem. Direct regularization methods offer a framework where the estimated solution can be computed in one step. A common regularization term $\Omega\left(\underline{\mathbf{f}}\right)$ is to require that the norm of the solution, or of its derivatives, is also minimized. Probably the most popular method, the Tikhonov

regularization is expressed as [118]:

$$\underline{\mathbf{f}}_\lambda = \text{argmin} \left\{ \left\| \underline{\tilde{\mathbf{g}}} - \mathbf{K}\underline{\mathbf{f}} \right\|_2^2 + \lambda^2 \left\| L\left(\underline{\mathbf{f}} - \underline{\mathbf{f}}^*\right) \right\|_2^2 \right\}, \tag{5.7}$$

or as:

$$\underline{\mathbf{f}}_\lambda = \text{argmin} \left\| \begin{bmatrix} \mathbf{K} \\ \lambda L \end{bmatrix} \underline{\mathbf{f}} - \begin{bmatrix} \underline{\tilde{\mathbf{g}}} \\ \lambda L \underline{\mathbf{f}}^* \end{bmatrix} \right\|_2^2$$

where:

- $\underline{\mathbf{f}}^*$ is an optional initial estimate;
- λ is the regularization parameter that tunes the importance of the constraint; and
- L is an operator like the identity matrix if the norm of $\underline{\mathbf{f}}$ is minimized or a banded matrix of finite differences if the norm of the derivative is minimized.

The optimal value of λ needs to be estimated separately. It can be shown that for $L = I$ and $\underline{\mathbf{f}}^* = 0$, the Tikhonov solution of Equation (5.7) is [49]:

$$\begin{aligned} \underline{\mathbf{f}}_\lambda &= \left(\mathbf{K}^\mathbf{T}\mathbf{K} + \lambda^2 \mathbf{I}\right)^{-1} \mathbf{K}^T \underline{\tilde{\mathbf{g}}} \\ &= \sum_{i=1}^{r} \Phi_i \frac{\underline{\mathbf{u}}_i^T \underline{\tilde{\mathbf{g}}}}{\sigma_i} \underline{\mathbf{v}}_i, \end{aligned}$$

where

$$\Phi_i = \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2}. \tag{5.8}$$

Initially close to 1 for $\sigma_i \gg \lambda$, the filter factors Φ_i decays smoothly down to zero when the singular values $\sigma_i \ll \lambda$. Thus, the regularization method tends to gradually filter out the contribution of $\frac{\underline{\mathbf{u}}_i^T \underline{\tilde{\mathbf{g}}}}{\sigma_i} \underline{\mathbf{v}}_i$ in the solution as σ_i gets smaller. The Tikhonov factors Φ_i of Equation (5.8) are in this case identical to the Wiener filter which operates in the frequency domain [94].

A large set of regularization methods differ depending on how the decay of the filter factors is specified. The general expression of the solution is written as:

$$\underline{\mathbf{f}}_\lambda = \sum_{i=1}^{r} \Phi_i \frac{\underline{\mathbf{u}}_i^T \underline{\tilde{\mathbf{g}}}}{\sigma_i} \underline{\mathbf{v}}_i, \tag{5.9}$$

For example, the filter factors Φ_i for different methods are can be the following [49]:

- Least-square estimation:

$$\Phi_i = 1 \quad \forall i.$$

  This is the least square estimate when there is no regularization term ($\lambda = 0$). It usually gives a poor quality solution for discrete ill-posed problems.

- Truncated Singular Value Decomposition (TSVD):

$$\Phi_i = \begin{cases} 1, & \text{if} \quad i \leq \lambda \\ 0, & \text{if} \quad i > \lambda \end{cases}$$

  The filter factors implement a step filter where the first λ biggest singular values are taken into account and the remaining ones are discarded.

- Damped Singular Value Decomposition (DSVD):

$$\Phi_i = \frac{\sigma_i}{\sigma_i + \lambda}. \tag{5.10}$$

  The DSVD offers a smoother cut-off than the TSVD and the filter decays more slowly than the Tikhonov regularization of Equation (5.8).

Iterative Regularization Methods

The iterative Conjugate Gradient (CG) method is popular to solve linear problem with a sparse symmetric positive definite matrix:

$$\underline{\tilde{\mathbf{g}}} = \mathbf{K} \cdot \underline{\mathbf{f}}, .$$

If the system is not symmetric positive definite, the CG is run on the normal equations: $\mathbf{K}^T \underline{\tilde{\mathbf{g}}} = \mathbf{K}^T \mathbf{K} \cdot \underline{\mathbf{f}}$. The implementation of the CG however does not compute explicitly the product $\mathbf{K}^T \mathbf{K}$. The CG has inherent regularizing properties [51]: as the algorithm iterates, the low frequency components of the solution converge faster than the high frequency ones. Consequently, there is an optimal number of iteration after which the CG starts

to introduce errors in the estimated solution. The number of iteration constitutes the regularization parameter λ.

An alternative algorithm to CG is LSQR [93] which avoids the need to write the problem in terms of normal equations as for the CG algorithm. The LSQR is mathematically equivalent to the CG but it allows the use of the Lanczos bidiagonalization algorithm to build the so-called Lanczos vectors [44]. The reorthogonalization at each iteration of the Lanczos vectors accelerates in theory the convergence of LSQR, but this is not usually done in practice as it is too computationally demanding [51].

A different approach to restoration is to consider iterative minimization algorithms for solving linear systems that constrain the solution to be nonnegative. Enforcing nonnegativity is particularly relevant for images and often leads to more accurate results although it is also more computationally demanding. The algorithm introduced in [66] as EM-LS (Expectation Maximization Least Square) and in [90] as MRNSD (Modified Residual Norm Steepest Descent) solves the following problem:

$$\min_{\underline{\mathbf{f}}} \left\| \underline{\tilde{\mathbf{g}}} - \mathbf{K} \cdot \underline{\mathbf{f}} \right\|_2 \qquad \text{subject to} \quad \underline{\mathbf{f}} \geq 0.$$

At each iteration, the direction of the negative gradient is taken and a line search is applied to minimize the residual norm $\left\| \underline{\tilde{\mathbf{g}}} - \mathbf{K} \cdot \underline{\mathbf{f}} \right\|$ in such a way that the nonnegativity of the solution is maintained. The MRNSD approach, as all the steepest descent algorithms, has a slow convergence which can be accelerated using preconditioning. Finally, there exists also with this approach an optimal number of iterations that gives the best solution possible.

Because the exact moments of $\underline{\mathbf{f}}$ can be retrieved with the proposed feature extraction approach, different restoration methods which constrain the moments up to order 3 of the solution to be equal to the moments of $\underline{\mathbf{f}}$ were tried out. However, no significant improvements were obtained as far as the quality of the restored image is concerned. As also suggested by theorem 1, this is due to the fact that a large number of moments is required to describe the content of an image so that the restoration process is helped to converge towards the correct solution.

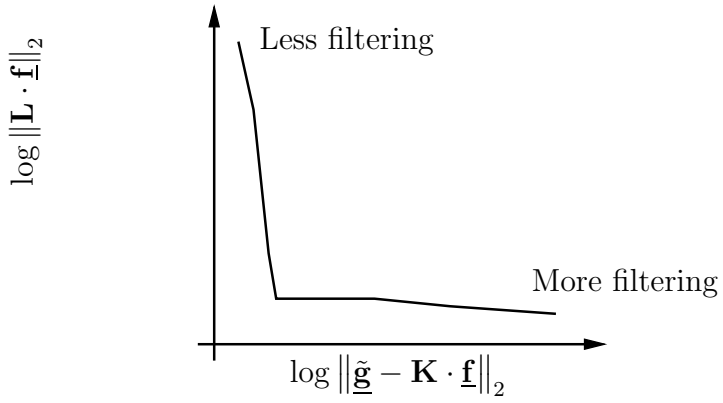


Figure 5.3: The log-log curve of the residual norm versus the solution norm often leads to an L-shaped curve as above. The optimal value for the regularizing parameter is found at the corner of the curve where both the residual norm and the solution norm are minimized. Problems can occur when the corner is not sharp enough.

Regularization Parameter Estimation

Direct and iterative methods both require the estimation of the optimal value of the regularization parameter λ. The regularization parameter can have different meanings: it can correspond either to a particular singular value (*e.g.* Tikhonov, DSVD or TSVD methods) or to the number of iteration (*e.g.* with the CG or MRNSD methods). The two most popular methods are the L-curve method and the Generalized Cross-Validation method.

The L-curve is a method that seeks to minimize the residual norm $\left\|\underline{\tilde{\mathbf{g}}} - \mathbf{K} \cdot \underline{\mathbf{f}}\right\|_2$ and the norm of the solution $\|\mathbf{L} \cdot \underline{\mathbf{f}}\|_2$ as shown on Figure 5.3 [84] [50]. The L-curve is a log-log plot of the two norms for different λ and has the shape of an "L". The optimal regularization parameter is found at the corner of the "L" where both norms are minimized. Thus, after computations of $\left\|\underline{\tilde{\mathbf{g}}} - \mathbf{K} \cdot \underline{\mathbf{f}}\right\|_2$ and $\|\mathbf{L} \cdot \underline{\mathbf{f}}\|_2$ for several values of λ, the optimal regularization parameter can be estimated by finding the maximum of the curvature of the L-curve. One necessary assumption to ensure that the L-curve has an "L" shape is that the Discrete Picard Condition is satisfied for the noiseless problem [50]: the exact SVD coefficients $|\underline{\mathbf{u}}_i^T \underline{\mathbf{g}}|$ decay faster than the singular values σ_i of $\mathbf{K}$. This condition ensures that, in the ideal noiseless case, the least-square solution does not have a large norm and that it can be approximated by the regularized solution [125].

The Generalized Cross-Validation (GCV) is a function of the regularizing parameter

λ whose minimum is achieved at the optimal λ. The GCV function can be expressed as [51]:

$$G(\lambda) = \frac{\left\|\mathbf{K}\underline{\mathbf{f}}_{\lambda} - \underline{\mathbf{g}}\right\|_2^2}{(trace\,(\mathbf{I} - \mathbf{K}\mathbf{K}_{\lambda}))}\quad '$$

where $\mathbf{K}_{\lambda}$ is such that:

$$\underline{\mathbf{f}}_{\lambda} = \mathbf{K}_{\lambda}\underline{\mathbf{g}}$$

The exact expression of the GCV function varies depending on the regularization method used and the type of regularization parameter looked for (see [45]).

5.4 Image Super-resolution: Simulations

5.4.1 Comparison of Restoration Methods

The set of restoration methods described in the previous section are now used in the restoration step of a super-resolution algorithm. The assessment of the quality of the super-resolved image is based on the PSNR value and the visual quality of the reconstructed image. The original image has size 172x172 pixels and is used to generate a set of 15 different translated images and sampled by a cubic B-spline with scale $T = 4$. Each low-resolution image has a size of 43x43 pixels.

The set of low-resolution images is registered using the continuous moments and fused on the same high resolution grid where missing samples are evaluated by cubic interpolation. The size of the resulting blurred image is the same as the original image: 172x172 pixels. The original image, one low-resolution image and the reconstructed blurred image are shown in Figure 5.4.

The image presented in Figure 5.4(c) is then restored using various algorithms. The resulting restored images are shown in the following figures:

- Figure 5.5: Truncated SVD with GCV method,
- Figure 5.6: Damped SVD with GCV method,
- Figure 5.7: Tikhonov regularization with L-curve method,
- Figure 5.8: Wiener filter (`deconvwnr` in Matlab),

(a) (b) (c)

Figure 5.4: (a) Original image (172x172 px); (b) Sampled image (43x43 px); (c) Reconstructed super-resolved image (172x172 px) before restoration using 15 translated images. Original image "Blue Marble" by NASA's Earth Observatory.

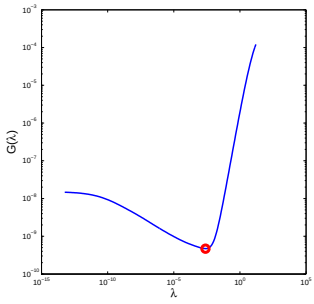
- Figure 5.9: Conjugate Gradient with L-curve method,
- Figure 5.10: Modified Residual Norm Steepest Descent.

From the obtained results, the Wiener filter and the MRNSD algorithm provide the images with the best quality both visually and in terms of PSNR (above 33 dB). Therefore these two restoration methods are selected for further super-resolution experiments. The Wiener filter has the advantage of being a direct method compared to the MRNSD algorithm which requires one to find the optimal number of iteration.

(a) (b)

Figure 5.5: (a) GCV function; (b) Super-resolved image with Truncated SVD and GCV method: PSNR = 7.9 dB. The minimum of the GCV function is achieved at k= 10873 **largest singular values.**

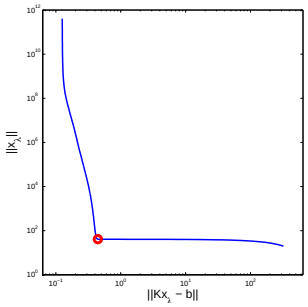
In the next two experiments, larger problems are considered with super-resolved images of size 512x512 pixels obtained from sets of low-resolution images of size 64x64 pixels ($T = 8$). In the first case shown in Figure 5.11, the assumed transformation between low-resolution images is a translation. A set of 40 low-resolution images is used



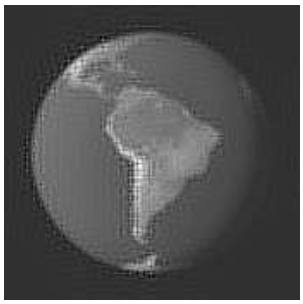
(a)

(b)

Figure 5.6: (a) GCV function; (b) Super-resolved image with Damped SVD and GCV method: PSNR = 13.5 dB. The minimum of the GCV function is achieved at $\lambda = 2.5 \cdot 10^{-3}$.



(a)



(b)

Figure 5.7: (a) L-curve; (b) Super-resolved image with Tikhonov regularization and L-curve method: PSNR = 15.7 dB. The corner of the L-curve is achieved at $\lambda = 1.3 \cdot 10^{-2}$.

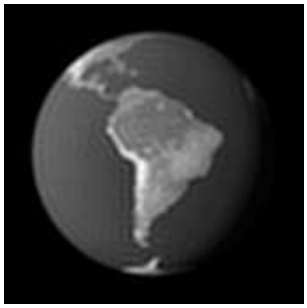
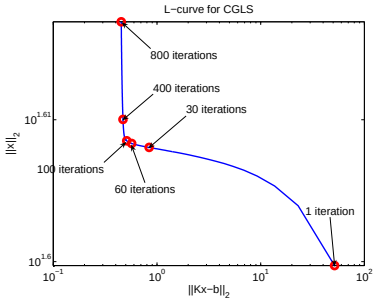


Figure 5.8: Super-resolved image with Wiener filter: PSNR = 33.6 dB.

to construct the super-resolved image. Figure 5.11(a) shows the ground-truth image and Figure 5.11(b) shows one the 40 low-resolution images. The registration is based on the continuous moments and both the Wiener filter and the MRNSD are considered in the restoration step (Figure 5.11(c) and (d)). Both super-resolved images are visually good and the image restored with the Wiener filter has a higher PSNR (24.2 dB) than the image

(a)



(b)

Figure 5.9: (a) L-curve; (b) Super-resolved image with Conjugate Gradient algorithm and L-curve method: PSNR = 23.8 dB. The corner of the L-curve is achieved at 100 iterations approximately.

Figure 5.10: Super-resolved image with Modified Residual Norm Steepest Descent algorithm after 400 iterations (experimentally estimated): PSNR = 33 dB.

restored with the MRNSD approach (23 dB). However this is not the case with the second experiment shown in Figure 5.12 where the low-resolution images are now related by affine transformations. As previously, the registration is based on the continuous moments and both the Wiener filter and the MRNSD algorithms are used for restoration. In this more challenging situation, the super-resolved image obtained with the MRNSD algorithm (Figure 5.12(d)) has much better visual quality and a higher PSNR (30.5 dB) than the image restored with Wiener filter (17.1 dB). Because affine transformations are more difficult to estimate exactly, the image to restore is noisier than in the case of low-resolution images related by translations. We thus observe that the MRNSD is more robust to noise than the Wiener filter but has the disadvantage of being slow to converge to the solution. The straightforward approach offered by the Wiener filter and its variants has recently been used as a restoration method in several super-resolution techniques [101] [53].

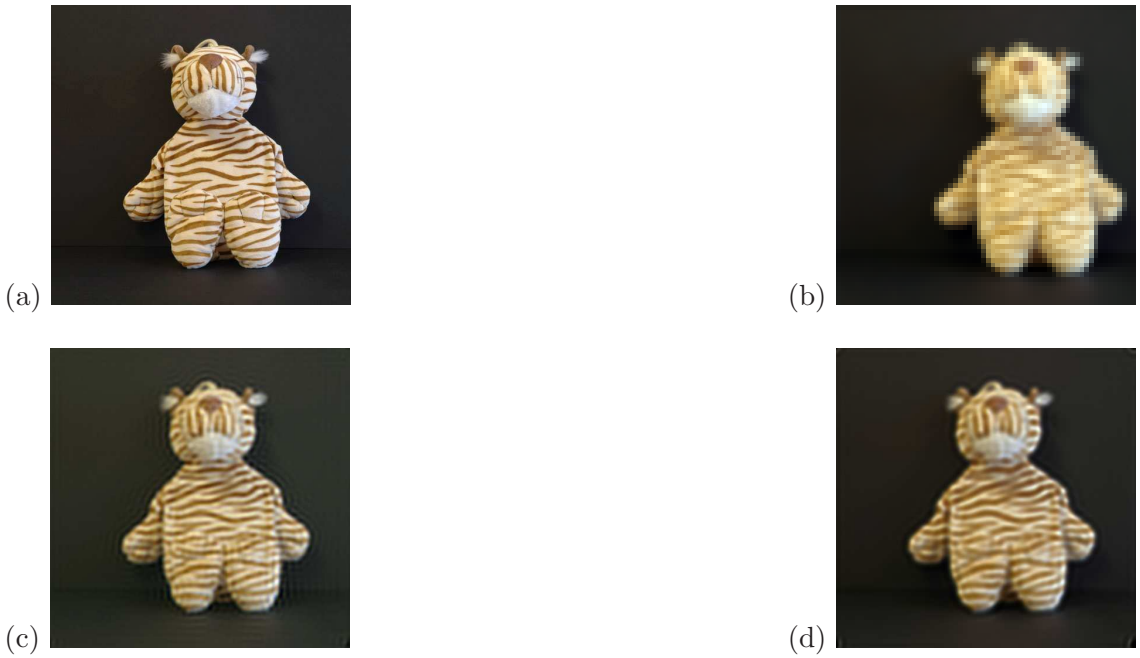


Figure 5.11: Image super-resolution of translated images with registration based on continuous moments; (a) Original high resolution image (512x512pixels); (b) One of the 40 translated low-resolution images (64x64 pixels) used in the super-resolution simulation; (c) Super-resolved image obtained with Wiener filter, 512x512 pixels, PSNR = 24.2 dB; (d) Super-resolved image obtained with the Modified Residual Norm Steepest Descent method (100 iterations), 512x512 pixels, PSNR = 23 dB.

5.4.2 Moment-based Registration: Discrete vs Continuous Moments

The following simulation illustrates that the quality of the super-resolved image is increased when accurate features for registration are considered. The discrete and continuous moments are used here for comparison and the experiment is shown in Figure 5.13. As with the previous experiments, we use a single high resolution image (Satellite image, 512x512 pixels) shown in Figure 5.13(a) to generate 24 other images related by translations. Each of these images is blurred and downsampled with a cubic B-spline to give a low-resolution image of size 64x64 pixels (see *e.g.* Figure 5.13(b)). This set of low-resolution images is then used as input for super-resolution. The translations are retrieved either from the discrete or the continuous moments of each image for comparison. In the case of the discrete moments, the registration error averaged over the 24 frames is $\bar{\epsilon} = 0.11$ pixels with a maximum registration error of $\bar{\epsilon}_{\max} = 2.5$ pixels. In the case of the continuous moments, the registration error averaged over the 24 frames is $\bar{\epsilon} = 3.5 \cdot 10^{-29}$ pixels with a maximum

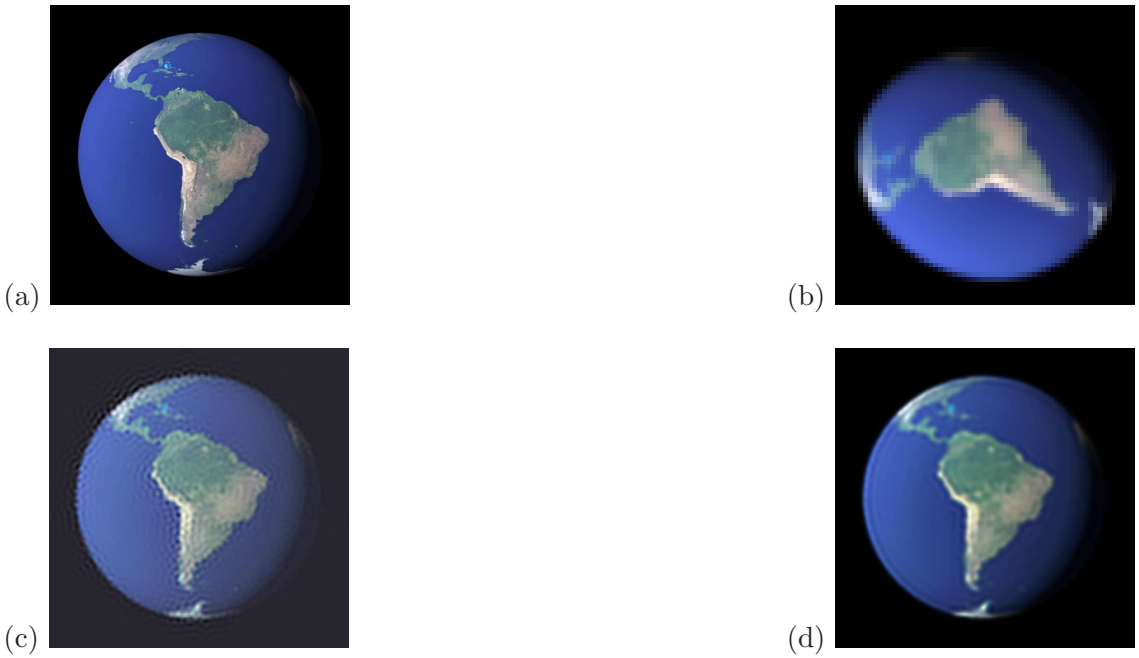


Figure 5.12: Image super-resolution with affine registration based on continuous moments; (a) Original high resolution image (512x512pixels); (b) One of the 16 affine low-resolution images (64x64 pixels) used in the super-resolution simulation; (c) Super-resolved image obtained with Wiener filter, 512x512 pixels, PSNR = 17.1 dB; (d) Super-resolved image obtained with the Modified Residual Norm Steepest Descent method (80 iterations), 512x512 pixels, PSNR = 30.5 dB. Original image "Blue Marble" by NASA's Earth Observatory.

registration error of $\bar{\epsilon}_{\max} = 8 \cdot 10^{-28}$ pixels. For a fair comparison, only the Wiener filter is used for restoration as it is fast and does not involve iterations. The super-resolved images are shown in Figure 5.13(c) and (d). The image in Figure 5.13(c) is obtained after registration with the discrete moments. Visual artefact are clearly visible and the PSNR is equal to 16.8dB. The super-resolved image shown in Figure 5.13(d) results from the utilization of the continuous moments for registration. Ringing effects are less visible and the image has a higher PSNR = 19.6 dB. Thus, by considering more accurate features like the continuous moments instead of the discrete moments, the registration is improved and can lead to super-resolved images of higher quality.

5.4.3 Edge-based Registration: Edge Extraction vs Harris features

In this second experiment, we consider the registration based on the extraction of step edges. As in the previous section, we generate 20 images of a scene (Figure 5.14(a))

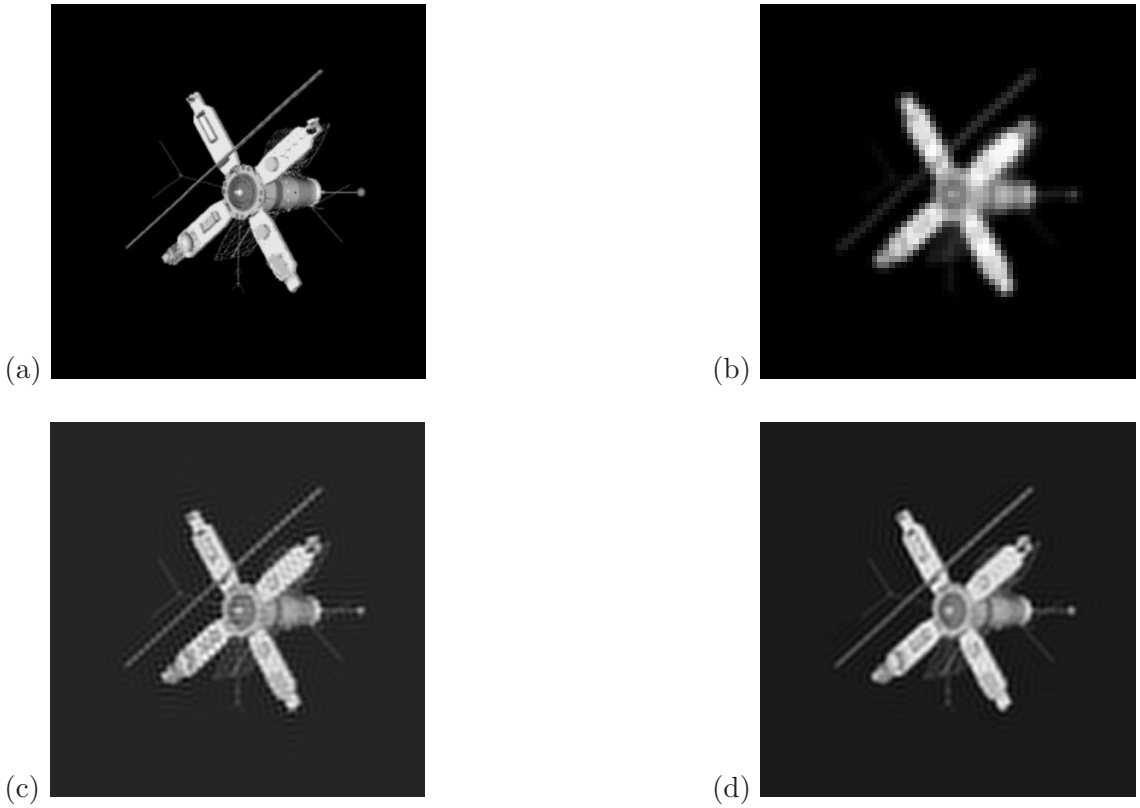


Figure 5.13: Image super-resolution from translated images with registration based on moments; (a) Original high resolution image (512x512pixels); (b) One of the 24 low-resolution images (64x64 pixels) used in the super-resolution simulation; (c) Super-resolved image obtained from the discrete moments and the Wiener filter, 512x512 pixels, PSNR = 16.8 dB; (d) Super-resolved image obtained from the continuous moments and the Wiener filter, 512x512 pixels, PSNR = 19.6 dB.

by cropping a high resolution image at different locations. The set of images are thus related by translations and each image is then blurred and downsampled with a quadratic B-spline to generate 20 low-resolution images of size 64x64 pixels (Figure 5.14(b)). In this simulation, two feature extraction methods are considered: the proposed step edge extractor and the Harris corner detector. As previously, in order to do a fair comparison, we do not use an iterative method and consider the Wiener filter as the restoration method in both cases.

With Harris features, the registration error averaged over the 20 frames is $\bar{\epsilon} = 0.44$ pixels and the maximum registration error observed in the 20 images is $\bar{\epsilon}_{\max} = 2.04$ pixels. The number of matched features varies between 13 and 21 corners. With our extracted features, the average registration error averaged over the 20 frames is $\bar{\epsilon} = 0.044$ pixels and

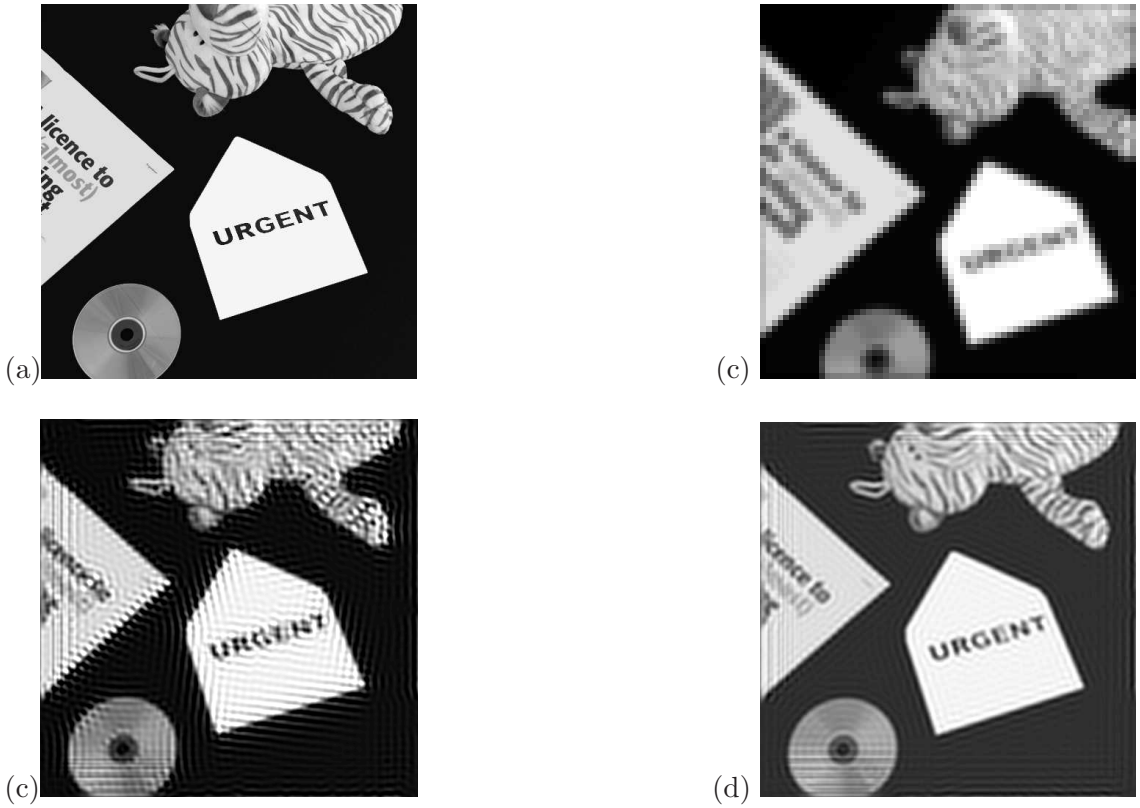


Figure 5.14: Image super-resolution from translated images with registration from the extracted edges and detected corners; (a) Original high resolution image (512x512pixels); (b) One of the 20 low-resolution images (64x64 pixels) used in the super-resolution simulation; (c) Super-resolved image obtained from features of the Harris corner detector, 512x512 pixels, PSNR = 14.1 dB; (d) Super-resolved image obtained with features from the proposed edge detector, 512x512 pixels, PSNR = 15.1 dB.

the maximum registration error observed in the 20 images is $\bar{\epsilon}_{\max} = 0.14$ pixels. Although the number of matched features with our approach is only 6 throughout the whole set of images, the registration accuracy is improved by a factor 10. Figure 5.14(c) presents the super-resolved image obtained with Harris features. This image has strong artefact and a PSNR = 14.1 dB. On the other hand, the super-resolved image obtained with the proposed registration technique shows a much better visual quality and a PSNR = 15.1 dB. The most visible artefact on both images occurs on the borders due to boundary effects and to the lack of available samples for interpolation. This simulation, thus, better highlights the fact that more accurate registration leads to better super-resolution.

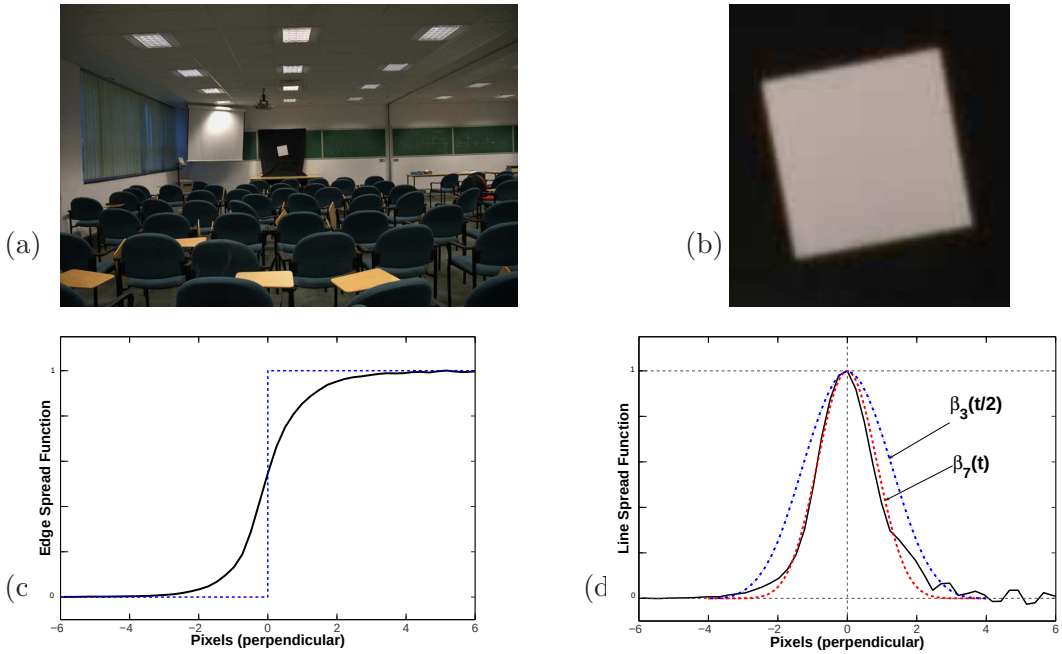


Figure 5.15: Estimation of the PSF with the slanted Edge Method; (a) Image of a slanted white square with step edges for Point Spread Function estimation (acquired with a Nikon D70s digital camera); (b) Zoom on the target; (c) Measured Edge Spread Function (solid line) and ideal step edge (dashed line); (d) Measured Line Spread Function (solid line). Its support has length 8 ranging from -4 to 4. The sampling kernel $\varphi(t)$ can be modeled for example by either a centered B-spline of degree 7, $\beta_7(t)$ (dashed line) or a centered B-spline of degree 3 scaled by 2, $\beta_3(t/2)$ (dash-dot line). In both case, the sampling kernel has support 8.

5.5 Image Super-resolution: Real-case Scenario

5.5.1 Estimation of the Sampling Kernel

In this section, we consider the case of image super-resolution from real images using the samples as they are acquired by a digital camera. The registration approach considered here is based only on the continuous moments. Since the registration approach takes a sampling point of view, we want our image samples to be modified as little as possible by internal post-processing occurring in a digital camera after acquisition. The set of images is thus acquired by a SLR digital camera (Nikon D70s) in RAW format with no edge sharpening. In a first experiment, pictures are taken in a classroom. The focal length is set at 18mm (35mm equivalent: 27mm) and other settings are: F16, 1/60s and ISO 200.

To estimate the support and the shape of the PSF, the slanted edge method is used. The PSF is the response of an imaging system to an infinitely small point light source.

Measuring it directly can thus be very difficult. The slanted edge method estimates the PSF indirectly by measuring the Edge Spread Function (ESF). It requires a picture of a slanted step edge to be taken so that the sampling rate of the step edge is increased. By differentiating along the edge's normal direction, the Line Spread Function (LSF) can be obtained (the response of the camera to a single line). The LSF represents the cross-section of the PSF in a given direction. The PSF is assumed to be circularly symmetric and spatially invariant in the considered region, so that only one LSF is necessary to characterize the PSF. An implementation following ISO standards for Matlab of the slanted edge method is freely available in [14].

Results of the estimation of the PSF are presented in Figure 5.15. The acquired image is presented in Figure 5.15(a) and the target for the PSF estimation is shown in Figure 5.15(b). The measured ESF is the solid line in Figure 5.15(c) while the estimated LSF is the solid line in Figure 5.15(d). From both functions, it can be observed that the support of the PSF is approximately 8 pixels, ranging from -4 to 4. The bell shape of the PSF is fitted with a B-spline of support 8 as well. Two different fits are presented in Figure 5.15(d): a B-spline of degree 7, $\beta_7(t)$ (dashed line), and a B-spline of degree 3 scaled by 2, $\beta_3(t/2)$ (dash-dot line). Although $\beta_7(t)$ seems to fit better the measured LSF, both functions are considered.

5.5.2 Super-resolution results

Keeping the same camera settings, 40 images of a scene of interest are taken from random positions by slightly moving the camera horizontally and vertically between each acquisition. Figure 5.16 (a) presents one of the acquired images with the scene to super-resolve at its center. This region of interest of size 128x128 pixels is selected in each image (*e.g.* Figure 5.16 (b)) and only this region is used for registration and restoration.

The sampling kernel is first modeled by a B-spline of degree 7. We register the images using the continuous moments of the regions of interest and the fused image is restored with the MRNSD method. The obtained super-resolved image is shown in Figure 5.16 (c). In the second case, the sampling kernel is modeled by a B-spline of degree 3 and scale 2. Because the kernel is scaled by two, the device is oversampling by a factor two

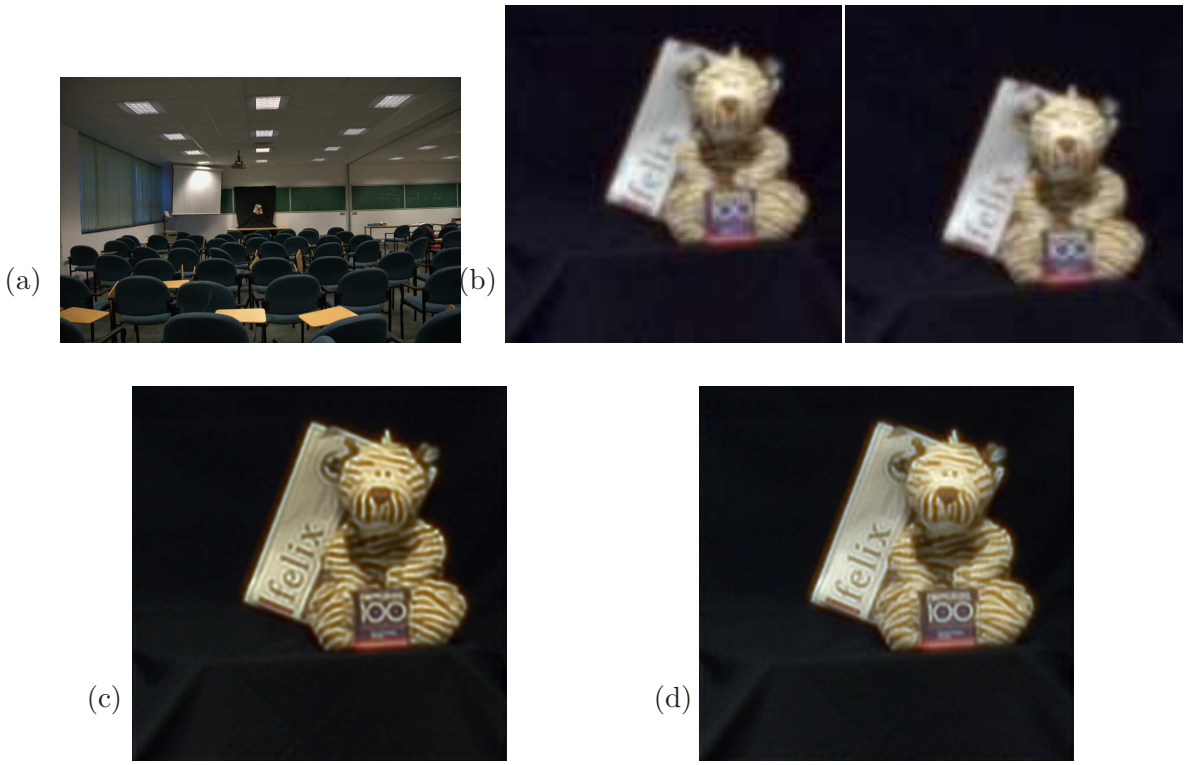


Figure 5.16: Real-case scenario of image super-resolution from 40 images acquired with a Nikon D70s SLR camera and a lens at a focal length of 18mm; (a) Acquired image of reference (3039x2014 pixels); the region of interest for super-resolution is visible at the center of the image; (b) Two out of 40 regions of interest (128x128 pixels) used for super-resolution; (c) Super-resolved image of size 1024x0124 pixels (SR factor = 8). The PSF in this case is modeled by a B-spline of order 7 (scale 1); (d) Super-resolved image of size 1024x0124 pixels (SR factor = 8). The PSF in this case is modeled by a B-spline of order 3 (scale 2). Both images have been restored after 60 iterations of the MRNSD method.

with respect to our model. The sampling period is consequently reduced by a factor two and the samples can be written as:

$$\mathbf{g}\left[m,n\right] = \left\langle f\left(x,y\right), \varphi\left(x/T - m/2, y/T - n/2\right)\right\rangle.$$

Two consecutive samples are now distant by $T/2$ instead of T and even and odd samples must be treated independently so that polynomial reproduction is satisfied. We thus

decompose the observed samples into their four polyphase components [27]:

$$\begin{cases} \mathbf{g}[2m, 2n] &= \langle f(x,y), \varphi(x/T - m, y/T - n) \rangle \\ \mathbf{g}[2m, 2n+1] &= \langle f(x,y), \varphi(x/T - m, y/T - n - 1/2) \rangle \\ \mathbf{g}[2m+1, 2n] &= \langle f(x,y), \varphi(x/T - m - 1/2, y/T - n) \rangle \\ \mathbf{g}[2m+1, 2n+1] &= \langle f(x,y), \varphi(x/T - m - 1/2, y/T - n - 1/2) \rangle. \end{cases}$$

For each region of size 128x128 pixels (Figure 5.16(b)), four sub-images of size 64x64 pixels are considered separately, each one corresponding to a polyphase component. The continuous moments of each polyphase component are computed and used for registration of the polyphase component of each view. For each image pairs, four estimations of the registration are obtained and then averaged. The super-resolved image achieved after registration from the polyphase components is shown in Figure 5.16 (d). As in the previous case, we restored the image using 60 iterations of the MRNSD algorithm. The super-resolved image presents a good level of detail and is comparable to the image obtained with B-spline of level 7. Moreover, this image looks less saturated than the previous one.

Another experiment is presented in Figure 5.17. Sixty pictures of the moon are taken with a digital SLR camera and a lens with a focal length of 38mm (35mm equivalent: 57mm) and settings: F16, 1/60s, ISO 200. The PSF in this case is not estimated as previously and is empirically chosen to be a cubic B-spline at scale 1. This choice was decided in this case empirically by selecting the PSF that lead to the visually most pleasing super-resolved image. The final super-resolved image is obtained after 60 iterations of the MRNSD restoration method. Figure 5.17(a) shows the moon as acquired by the camera. Figure 5.17(b) presents the obtained super-resolved image where details of the moon can be observed.

(a)

(b)

Figure 5.17: Real super-resolution of the moon from 60 images acquired with a Nikon D70s SLR camera and a lens (18-70mm, F3.5-4.5) set at a focal length of 38mm (35mmm equiv.: 57mm). (a) The moon as acquired by the camera (60x60 px); (b) Super-resolved image of the moon (600x600 px) with MRNSD restoration method.

Chapter 6

Conclusions

6.1 Thesis Summary

WE have presented in this thesis new feature extraction methods for the registration of multiview images. The proposed techniques allow the exact extraction of the considered features and are very efficient for image features that are acquired at low resolution. The problem of feature extraction was formulated as a multichannel sampling problem for which results from the sampling theory for signals with finite rate of innovation were considered. The main requirement necessary to use this theory is that the characteristics of the acquisition device, like the sampling kernel, must be known in advance or through a preliminary calibration procedure. The sampling kernel assumed in the major part of this thesis are B-spline functions which exhibit appealing properties (compact support, reproduction of polynomial, dual basis) as well as being suitable for modeling real camera devices.

The new sampling schemes for FRI signals have been first extended to a multichannel acquisition setup and, depending on the assumed sampling kernel (polynomial- or exponential-reproducing one), their performances in this new framework have been analyzed. It appears that kernels reproducing polynomials are not suitable for a truly distributed and symmetric multichannel setup and requires an asymmetric setup. Experiments of distributed acquisition of bilevel polygons have been carried out and, although they are somewhat artificial and mostly for illustration purposes, some applications can be

found such as in vision-based localization used in robotics. As opposed to kernels reproducing polynomials, it is possible to design a truly symmetric and distributed acquisition system using exponential reproducing kernels. We then proposed an application of the exponential case to time-interleaved analog-to-digital converters for which unwanted delays introduced by the electronic components can be calculated precisely from the samples in order to improve the accuracy.

We then considered the case of feature extraction for image registration. Exact extraction methods have been proposed for two different features: image moments and step edges. The geometric moments are global features and do not explicitly require any feature correspondence between images. The registration method used in this thesis and based on various types of moments (central and complex moments) allows the retrieval of up to an affine transformation occurring between two different views. This assumes that the observed scene is flat or that the perspective effect is weak so that the homography can be approximated by an affine transformation and there is no known method for more complex transformations. Extraction of step edges, on the contrary, provides local features for registration from the contours, intersection and corners information. This approach is more difficult to implement and, as with any local technique, it requires the correct correspondence of features for registration. The corresponding pairs of corners allow the estimation of more complicated transformations like homographies. The performance of the proposed feature extraction methods is presented and simulations of image registration based on these features demonstrated their superiority when compared to similar registration methods using standard features.

In a traditional framework, sampling theory is concerned with the problem of reconstructing a signal from its samples. However perfect reconstruction schemes are rarely available for highly complex signals like a real-world scene: it is not possible to acquire a low-resolution image and have it displayed on a high resolution screen without inevitably introducing undesirable visual artefact. One solution to overcome this problem is provided by the technique of super-resolution. Developing new accurate registration techniques that can improve image super-resolution algorithms was the initial and main motivation throughout this research. We thus implemented an image super-resolution algorithm that

incorporates the proposed feature extraction techniques in the registration part. With the local and the global features, the super-resolved images obtained from artificial data showed an improvement of the quality, both visually and in terms of PSNR, when compared to the super-resolved images obtained with the same implementation but using standard feature extraction methods. These results emphasize the need for very precise registration methods for super-resolution.

Given these favorable results, we then ran real experiments of image super-resolution from images acquired by a real camera. To satisfy our model, simple calibration measurements of intrinsic parameters of the camera, like the point-spread function, were carried out and the notion of polyphase components of an image was used for registration. The obtained super-resolved images from real data set show a real improvement in their resolving power compared to the acquired images. Besides, this also underlines the fact that the proposed image acquisition framework is appropriate for modeling existing cameras. Finally, the last super-resolution experiment was conducted without prior calibration and the settings were chosen in order to obtain the best super-resolved image possible.

Along the same line as the sampling theory for FRI signals, the work presented in this thesis on new feature extraction techniques takes into account the inherent defects and imperfections of an acquisition device like a cameras and turns these shortcomings to an advantage. Similarly to [124] where aliasing was exploited to provide more information, it is here the knowledge of the sampling kernel that allows to develop exact methods for extracting features almost irrespectively to original resolution of the acquired signal. Although the knowledge of the sampling kernel can be seen as another constraint in the design of a device, the potential benefits that may be obtained from it can be far greater than the inconvenience caused.

6.2 Future Research

The feature extraction techniques proposed in this work lead to rather different registration procedures. The main limitations of using the moments as image features are linked to the impossibility of dealing with occlusion and disappearance of objects in the scene, and with

non-uniform backgrounds. In these cases, accurate segmentation techniques like level sets methods [92] should be used to separate the objects of interest from the still background in order to register different images with moments. It would be interesting to apply the approach of continuous moments to another popular application for moments that is pattern recognition. One of the main advantage of the proposed method with moments is that it works very well with very low-resolution images which thus allows to operate with very compact representations of the patterns of interest. The step edges were used as features because it was possible to derive a parametric model (FRI signal). In addition to edges, other types of feature can be envisaged and modeled as FRI signals (lines, circles...) and used in man-made scenes like urban areas or in more specific applications (*e.g.* bar codes reader).

Since it was not the main focus on this thesis, the restoration techniques employed in the super-resolution experiments are fairly standard ones (with the exception of the MRNSD algorithm). More advanced restoration/regularization techniques based, for example, on the minimization of the ℓ_1-norm of the solution ([32]) would probably lead to super-resolved images of better quality. In particular, the exact positions of the step edges in the images extracted for registration were not taken into account although this could be useful information for restoration. Various algorithms are likely to give even sharper images like for example Total Variation methods [19] [30]. Besides, the interpolation step does not consider the fact that the sampling kernel is known. Different interpolation schemes can be developed so that the blurred image contains less errors and is ultimately better restored.

A thorough analysis of the impact of noise on the proposed algorithms similarly to the work of Blu *et al* [10]. In particular, deriving performance bounds in the case of the edge features, such as the Cramér-Rao bounds, would provide much insights as to how the noise affects the feature extraction schemes and which signal enhancement methods (*e.g.* Cadzow [15]) are most suitable.

In addition to standard digital cameras, it would be of interest to apply these techniques to different acquisition devices with various sampling kernels like microscopes, satellite imaging, or low quality cameras found on webcam or mobile phones. The B-splines

offer a flexible framework to model a variety of sampling kernels as well as reproducing polynomials. Although less studied, the E-splines can offer a similar framework and lead to new sampling results. Since the registration takes into account the parameters of the acquisition device, auto-calibration techniques for cameras can be developed through a possibly recursive process by choosing the parameters that lead to the best registration of multiview images.

Appendix A

Proof of Equation (4.16): Single Step Edge Case

Let $h(x,y)$ be the single step edge function with parameters (α,θ,γ) as in Equation (4.14) and let $\mathbf{h}[m,n] = \langle h(x,y), \varphi(x/T-m, y/T-n)\rangle$ be its samples. The sampling kernel $\varphi(x,y)$ can reproduce polynomials and is variable separable so that $\varphi(x,y) = \varphi_1(x)\otimes\varphi_2(y)$. We denote by $\mathbf{d}[m,n] = \mathbf{h}[m+1,n] - \mathbf{h}[m,n]$ the differentiated samples of $h(x,y)$ and by $c_m^{(p)}$ the coefficients used for polynomial reproduction with the modified kernel $\varphi_1(x/T-m) * \beta_0(x/T-m)$. We now prove the following result:

Proposition 3. *The linear combination of the differentiated samples* $\mathbf{d}[m,n]$ *with the coefficients* $c_m^{(p)}$*:*

$$\tau_{p,n} = \sum_m c_m^{(p)} \mathbf{d}[m,n],$$

is equal to:

$$\tau_{p,n} = -\alpha \sum_{i=0}^{p} \binom{p}{i} \frac{m_i}{(\tan\theta)^i} \left(\gamma + \frac{n}{\tan\theta}\right)^{p-i}$$

where $m_i = \int_\infty^\infty t^i \varphi_2(t)\,dt$ *are the moments of* $\varphi_2(y)$*.*

Proof. We first recall the following relation between the moments m_p of function $f(t)$ and the moments m'_p of the translated function $f(t-\Delta t)$:

$$m'_p = \int_{-\infty}^{\infty} f(t-\Delta t)\, t^p dt = \sum_{k=0}^{p} \binom{p}{k} \Delta t^k m_{p-k} \tag{A.1}$$

Let $\Omega = [-L, L]$ be the support of the sampling kernel $\varphi_2(t)$. For simplicity, we assume $T = 1$:

$$
\begin{aligned}
\tau_{p,n} &= \sum_{m \in \mathcal{S}_n} c_m^{(p)} \mathbf{d}[m,n] \\
&= -\alpha \sin\theta \sum_{m \in \mathcal{S}_n} c_m^{(p)} \left\langle \delta\left(\vec{d}\cdot\vec{N}\right), \varphi(x-m, y-n) * \beta_0(x-m) \right\rangle \\
&= -\alpha \sin\theta \left\langle \delta\left(\vec{d}\cdot\vec{N}\right), \left(\sum_{m \in \mathcal{S}_n} c_m^{(p)} \varphi_1(x-m) * \beta_0(x-m) \right) \otimes \varphi_2(y-n) \right\rangle \\
&= -\alpha \sin\theta \left\langle \delta\left(\vec{d}\cdot\vec{N}\right), x^p \otimes \varphi_2(y-n) \right\rangle \\
&= -\alpha \sin\theta \int_{n-L}^{n+L} \varphi_2(y-n) \int_{-\infty}^{\infty} \delta(-x\sin\theta + y\cos\theta + \gamma\sin\theta)\, x^p dx dy \\
&= -\alpha \sin\theta \int_{n-L}^{n+L} \varphi_2(y-n) \int_{-\infty}^{\infty} \delta(-t + y\cos\theta + \gamma\sin\theta) \left(\frac{t}{\sin\theta}\right)^p \frac{dt}{\sin\theta} dy \\
&= -\alpha \int_{n-L}^{n+L} \varphi_2(y-n) \left(\frac{y\cos\theta + \gamma\sin\theta}{\sin\theta}\right)^p dy \\
&= -\alpha \int_{n-L}^{n+L} \varphi_2(y-n) \left(\frac{y}{\tan\theta} + \gamma\right)^p dy \\
&= -\alpha \int_{-L}^{L} \varphi_2(t) \left(\frac{t+n}{\tan\theta} + \gamma\right)^p dt \\
&= \frac{-\alpha}{(\tan\theta)^p} \int_{-L}^{L} \varphi_2(t) (t + n + \gamma\tan\theta)^p\, dt, \qquad \text{then applying Equation (A.1),} \\
&= \frac{-\alpha}{(\tan\theta)^p} \sum_{i=0}^{p} \binom{p}{i} m_{p-i} (n + \gamma\tan\theta)^i, \qquad \text{where} \quad m_i = \int_{-\infty}^{\infty} t^i \varphi_2(t)\, dt, \\
&= -\alpha \sum_{i=0}^{p} \binom{p}{i} \frac{m_{p-i}}{(\tan\theta)^{p-i}} \left(\gamma + \frac{n}{\tan\theta}\right)^i.
\end{aligned}
$$

For information, we give below the expression of the even order moments for the B-splines of degree 0 to 3 (odd order moments are null due to the symmetry of B-splines):

$$
\begin{aligned}
\varphi_2(t) = \beta_0(t) &\implies m_{2j} = \frac{1}{2^{2j}(2j+1)} \\
\varphi_2(t) = \beta_1(t) &\implies m_{2j} = \frac{2}{(2j+2)(2j+1)} \\
\varphi_2(t) = \beta_2(t) &\implies m_{2j} = \frac{3^{2j+3} - 3}{2^{2j+2}(2j+3)(2j+2)(2j+1)} \\
\varphi_2(t) = \beta_3(t) &\implies m_{2j} = \frac{2^{2j+5} - 8}{(2j+4)(2j+3)(2j+2)(2j+1)}.
\end{aligned}
$$

□

Appendix B

Proof of Equation (4.18): K Parallel Step Edges Case

Let $h(x,y)$ be the function representing K parallel step edges with parameters $\{\alpha_k, \gamma_k, \theta\}$, $k = 1, \ldots, K$, as in Equation (4.20) and let $\mathbf{h}[m,n] = \langle h(x,y), \varphi(x/T - m, y/T - n)\rangle$ be its samples. The sampling kernel $\varphi(x,y)$ can reproduce polynomials and is variable separable so that $\varphi(x,y) = \varphi_1(x) \otimes \varphi_2(y)$. We denote by $\mathbf{d}[m,n] = \mathbf{h}[m+1,n] - \mathbf{h}[m,n]$ the differentiated samples of $h(x,y)$ and by $c_m^{(p)}$ the coefficients used for polynomial reproduction with the modified kernel $\varphi_1(x/T - m) * \beta_0(x/T - m)$. We now prove the following result:

Proposition 4. *By defining $\tau_{p,n}$ as the linear combination of the differentiated samples with the coefficients $c_m^{(p)}$:*

$$\tau_{p,n} = \sum_m c_m^{(p)} \mathbf{d}[m,n],$$

the angle θ parameter of the K parallel step edges can be exactly retrieved as:

$$\tan\theta = \frac{\tau_{0,n}}{\tau_{1,n+1} - \tau_{1,n}}$$

Proof. From Appendix A, it is straightforward to show that in the case of K parallel step

edges, we have:

$$\begin{aligned}
\tau_{p,n} &\equiv \sum_{m\in\mathcal{S}_n} c_m^{(p)}\mathbf{d}\left[m,n\right] \\
&= \sum_{k=1}^{K} -\alpha_k \sin\theta \sum_{m\in\mathcal{S}_n} c_m^{(p)} \left\langle \delta\left(\vec{d_k}\cdot\vec{N}\right), \varphi\left(x-m,y-n\right) * \beta_0\left(x-m\right)\right\rangle \\
&= \sum_{k=1}^{K} -\alpha_k \sum_{i=0}^{p} \binom{p}{i} \frac{m_{p-i}}{(\tan\theta)^{p-i}} \left(\gamma_k + \frac{n}{\tan\theta}\right)^i \\
&= \sum_{i=0}^{p} \binom{p}{i} \frac{m_{p-i}}{(\tan\theta)^{p-i}} \left(\sum_{k=1}^{K} -\alpha_k \left(u_{k,n}\right)^i\right), \qquad \text{(B.1)}
\end{aligned}$$

where $u_{k,n} = \gamma_k + \frac{n}{\tan\theta}$. We then have:

$$\begin{aligned}
\tau_{1,n+1} - \tau_{1,n} &= \left(\frac{m_1}{\tan\theta}\sum_{k=1}^{K} -\alpha_k + m_0 \sum_{k=1}^{K} -\alpha_k u_{k,n+1}\right) - \left(\frac{m_1}{\tan\theta}\sum_{k=1}^{K} -\alpha_k + m_0 \sum_{k=1}^{K} -\alpha_k u_{k,n}\right) \\
&= m_0 \sum_{k=1}^{K} -\alpha_k \left(u_{k,n+1} - u_{k,n}\right) \\
&= m_0 \sum_{k=1}^{K} \frac{-\alpha_k}{\tan\theta} \\
&= \frac{\tau_{0,n}}{\tan\theta},
\end{aligned}$$

which finally leads to the desired result. □

Appendix C

Proof of Equation (4.22): K Parallel Step Edges Case

Let $h(x,y)$ be the function representing K parallel step edges with parameters $\{\alpha_k, \gamma_k, \theta\}$, $k = 1, \ldots, K$, as in Equation (4.20) and let $\mathbf{h}[m,n] = \langle h(x,y), \varphi(x/T - m, y/T - n)\rangle$ be its samples. The sampling kernel $\varphi(x,y)$ can reproduce polynomials and is variable separable so that $\varphi(x,y) = \varphi_1(x) \otimes \varphi_2(y)$. We denote by $\mathbf{d}[m,n] = \mathbf{h}[m+1,n] - \mathbf{h}[m,n]$ the differentiated samples of $h(x,y)$ and by $c_m^{(p)}$ the coefficients used for polynomial reproduction with the modified kernel $\varphi_1(x/T - m) * \beta_0(x/T - m)$. We now prove the following result:

Proposition 5. *Let the quantity $\widehat{\tau}_{p,n}$ be recursively defined as:*

$$\widehat{\tau}_{p,n} = \begin{cases} \tau_{0,n}, & p = 0, \\ \tau_{p,n} - \sum_{j=0}^{p-1} \binom{p}{j} \frac{m_{p-j}}{m_0 (\tan\theta)^{p-j}} \widehat{\tau}_{j,n}. & p > 0, \end{cases}$$

where $\tau_{p,n}$ is defined in Equation (B.1), then $\widehat{\tau}_{p,n}$ can be expressed as a powersum series as:

$$\widehat{\tau}_{p,n} = \sum_{k=0}^{K} \lambda_k (u_{k,n})^p,$$

with $\lambda_k = -m_0 \alpha_k$ and $u_{k,n} = \gamma_k + \frac{n}{\tan\theta}$.

Proof. First, the case for $p = 0$ is straightforward since by definition, we have:

$$\widehat{\tau}_{0,n} = \tau_{0,n} = m_0 \sum_{k=0}^{K} -\alpha_k = \sum_{k=0}^{K} \lambda_k$$

For $p > 0$, we have:

$$\widehat{\tau}_{p,n} = \tau_{p,n} - \sum_{j=0}^{p-1} \binom{p}{j} \frac{m_{p-j}}{m_0 (\tan\theta)^{p-j}} \widehat{\tau}_{j,n}.$$

Moving the summation term to the left side of the expression gives:

$$\widehat{\tau}_{p,n} + \sum_{j=0}^{p-1} \binom{p}{j} \frac{m_{p-j}}{m_0 (\tan\theta)^{p-j}} \widehat{\tau}_{j,n} = \tau_{p,n}.$$

After assimilating $\widehat{\tau}_{p,n}$ in the left summation and recalling the definition of $\tau_{p,n}$ in Equation (B.1), we obtain:

$$\sum_{j=0}^{p} \binom{p}{j} \frac{m_{p-j}}{m_0 (\tan\theta)^{p-j}} \widehat{\tau}_{j,n} = \sum_{i=0}^{p} \binom{p}{i} \frac{m_{p-i}}{(\tan\theta)^{p-i}} \left(\sum_{k=1}^{K} -\alpha_k (u_{k,n})^i \right).$$

Then by identifying each term of the summation on each side of the expression above, we have:

$$\frac{1}{m_0} \widehat{\tau}_{i,n} = \sum_{k=1}^{K} -\alpha_k (u_{k,n})^i .$$

We can then conclude with the desired result:

$$\begin{aligned} \widehat{\tau}_{i,n} &= \sum_{k=1}^{K} -m_0 \alpha_k (u_{k,n})^i \\ &= \sum_{k=1}^{K} \lambda_k (u_{k,n})^i \end{aligned}$$

□

Appendix D

Proof of Equations (4.18) and (4.22) in the Linear Edge Case

Let $h(x,y)$ be the function representing a linear edge with parameters $\{\alpha, \lambda, \gamma, \theta\}$ (see Figure 4.5):

$$h(x,y) = H\left(\left\langle \vec{d}, \vec{N} \right\rangle\right)\left[\lambda\left(-x\sin\theta + y\cos\theta\right) + \alpha\right]. \tag{D.1}$$

The function $H(t)$ is the unit step function and:

- α is the amplitude of the plane at the origin,
- λ is the slope of the plane,
- γ is the offset of the straight edge,
- θ is the angle of the straight edge.

The first and second derivative of $h(x,y)$ are respectively ($\lambda \neq 0$):

$$\begin{aligned}
\frac{dh(x,y)}{dx} &= -\sin\theta\delta\left(\left\langle \vec{d}, \vec{N} \right\rangle\right)\left[\lambda\left(-x\sin\theta + y\cos\theta\right) + \alpha\right] - \lambda\sin\theta H\left(\left\langle \vec{d}, \vec{N} \right\rangle\right), \\
\frac{d^2h(x,y)}{dx^2} &= \sin^2\theta\delta^{(1)}\left(\left\langle \vec{d}, \vec{N} \right\rangle\right)\left[\lambda\left(-x\sin\theta + y\cos\theta\right) + \alpha\right] + 2\lambda\sin^2\theta\delta\left(\left\langle \vec{d}, \vec{N} \right\rangle\right), \\
&= \lambda\sin^2\theta\left[\delta^{(1)}\left(\left\langle \vec{d}, \vec{N} \right\rangle\right)\left(\left(-x\sin\theta + y\cos\theta\right) + \frac{\alpha}{\lambda}\right) + 2\delta\left(\left\langle \vec{d}, \vec{N} \right\rangle\right)\right],
\end{aligned} \tag{D.2}$$

where $\delta^{(1)}(t)$ is the differentiated Dirac function. Similarly to the step edge case, the function $h(x,y)$ is sampled with a polynomial reproducing kernel $\varphi(x,y) = \varphi_1(x) \otimes \varphi_2(y)$ so that we observe the following samples:

$$\mathbf{h}[m,n] = \langle h(x,y), \varphi(x/T - m, y/T - n) \rangle \tag{D.3}$$

We consider the second order differentiation of the samples obtained by filtering and giving the following differentiated samples $\mathbf{d}[m,n]$:

$$\mathbf{d}[m,n] = \mathbf{h}[m+2,n] - 2\mathbf{h}[m+2,n] + \mathbf{h}[m,n] \tag{D.4}$$

Using the result of [27], the differentiated samples are equal to:

$$\mathbf{d}[m,n] = \left\langle \frac{d^2 h(x,y)}{dx^2}, \varphi_1 * \beta_0 * \beta_0 (x-m) \otimes \varphi_2 (y-n) \right\rangle \tag{D.5}$$

Let $c_m^{(p)}$ be the coefficients used for polynomial reproduction of degree p with the modified kernel $\varphi_1 \beta_0 * * \beta_0 (x/T - m)$ and let $\tau_{p,n}$ be the linear combination of these coefficients and the differentiated samples. We therefore have:

$$\begin{aligned}
\tau_{p,n} &= \sum_m c_m^{(p)} \mathbf{d}[m,n] \\
&= \sum_m c_m^{(p)} \left\langle \frac{d^2 h(x,y)}{dx^2}, \varphi_1 * \beta_0 * \beta_0 (x-m) \otimes \varphi_2 (y-n) \right\rangle \\
&= \left\langle \frac{d^2 h(x,y)}{dx^2}, x^p \otimes \varphi_2 (y-n) \right\rangle.
\end{aligned}$$

Assuming that φ_2 has a compact support $\Omega = [-L, L]$, then:

$$\begin{aligned}
\tau_{p,n} &= \lambda \sin^2\theta \int_{n-L}^{n+L} \varphi_2(y-n) \int_{-\infty}^{\infty} \Big[\delta^{(1)} \left(\left\langle \vec{d}, \vec{N} \right\rangle \right) \left((-x\sin\theta + y\cos\theta) + \frac{\alpha}{\lambda} \right) + \\
&\qquad 2\delta \left(\left\langle \vec{d}, \vec{N} \right\rangle \right) \Big] x^p dx dy, \\
&= \lambda \sin^2\theta \int_{n-L}^{n+L} \varphi_2(y-n) \int_{-\infty}^{\infty} \Big[\delta^{(1)} (-x\sin\theta + y\cos\theta + \gamma\sin\theta) \times \\
&\qquad \left((-x\sin\theta + y\cos\theta) + \frac{\alpha}{\lambda} \right) + 2\delta (-x\sin\theta + y\cos\theta + \gamma\sin\theta) \Big] x^p dx dy,
\end{aligned}$$

Let $t = x \sin\theta$, we have:

$$\begin{aligned}\tau_{p,n} &= \lambda \sin^2\theta \int_{n-L}^{n+L} \varphi_2(y-n) \int_{-\infty}^{\infty} \Big[\delta^{(1)}(-t + y\cos\theta + \gamma\sin\theta) \times \\ &\quad \Big((-t + y\cos\theta) + \frac{\alpha}{\lambda}\Big) + 2\delta(-t + y\cos\theta + \gamma\sin\theta)\Big] \left(\frac{t}{\sin\theta}\right)^p \frac{dt}{\sin\theta} dy,\end{aligned}$$

Now, since $\int_{-\infty}^{\infty} f(t)\,\delta^{(r)}(t - t_0)\,dt = (-1)^r f^{(r)}(t_0)$:

$$\begin{aligned}\tau_{p,n} &= \lambda \sin\theta \int_{n-L}^{n+L} \varphi_2(y-n) \left[(-1)\left(-\left(\frac{t}{\sin\theta}\right)^p + p\left(-t + y\cos\theta + \frac{\alpha}{\lambda}\right)\left(\frac{t}{\sin\theta}\right)^{p-1}\right) + \right. \\ &\qquad \left. 2\left(\frac{t}{\sin\theta}\right)^p\right]_{t = y\cos\theta + \gamma\sin\theta} dy \\ &= \lambda \sin\theta \int_{n-L}^{n+L} \varphi_2(y-n) \left[3\left(\frac{t}{\sin\theta}\right)^p - p\left(-t + y\cos\theta + \frac{\alpha}{\lambda}\right) \times \right. \\ &\qquad \left. \left(\frac{t}{\sin\theta}\right)^{p-1}\right]_{t = y\cos\theta + \gamma\sin\theta} dy \\ &= \lambda \sin\theta \int_{n-L}^{n+L} \varphi_2(y-n) \left[3\left(\gamma + \frac{n}{\tan\theta}\right)^p - p\left(\frac{\alpha}{\lambda} - \gamma\sin\theta\right)\left(\gamma + \frac{n}{\tan\theta}\right)^{p-1}\right] dy.\end{aligned}$$

Introducing the notation $u_n = \gamma + \frac{n}{\tan\theta}$ yields:

$$\begin{aligned}\tau_{p,n} &= 3\lambda \sin\theta \int_{n-L}^{n+L} \varphi_2(y-n)\,(u_n)^p\,dy - \\ &\qquad p\lambda \sin\theta \left(\frac{\alpha}{\lambda} - \gamma\sin\theta\right) \int_{n-L}^{n+L} \varphi_2(y-n)\,(u_n)^{p-1}\,dy.\end{aligned}$$

Using the results from Appendix A, we finally obtain:

$$\begin{aligned}\tau_{p,n} &= 3\lambda \sin\theta \sum_{i=0}^{p} \binom{p}{i} \frac{m_{p-i}}{(\tan\theta)^{p-i}} (u_n)^i - \\ &\qquad p\lambda \sin\theta \left(\frac{\alpha}{\lambda} - \gamma\sin\theta\right) \sum_{i=0}^{p-1} \binom{p-1}{i} \frac{m_{p-1-i}}{(\tan\theta)^{p-1-i}} (u_n)^i,\end{aligned}$$

where m_i are the moments of order i of φ_2. Let A and B be defined as:

$$A = 3\lambda \sin\theta \tag{D.6}$$

$$B = -\lambda \sin\theta \left(\frac{\alpha}{\lambda} - \gamma \sin\theta\right), \tag{D.7}$$

so that:

$$\tau_{p,n} = A \sum_{i=0}^{p} \binom{p}{i} \frac{m_{p-i}}{(\tan\theta)^{p-i}} (u_n)^i + p \cdot B \sum_{i=0}^{p-1} \binom{p-1}{i} \frac{m_{p-1-i}}{(\tan\theta)^{p-1-i}} (u_n)^i . \tag{D.8}$$

Proposition 6. *The angle θ and slope λ can be retrieved using $\tau_{0,n}$, $\tau_{1,n}$ and $\tau_{1,n+1}$ as:*

$$\tan\theta = \frac{\tau_{0,n}}{\tau_{1,n+1} - \tau_{1,n}}, \tag{D.9}$$

and

$$\lambda = \frac{\tau_{0,n}}{3 \sin\theta m_0}. \tag{D.10}$$

Proof. We have:

$$\begin{aligned} \tau_{0,n} &= A m_0 \\ \tau_{1,n} &= A \left[m_0 (u_n) + \frac{m_1}{\tan\theta} \right] - B m_0 \\ \tau_{1,n+1} &= A \left[m_0 (u_{n+1}) + \frac{m_1}{\tan\theta} \right] - B m_0, \end{aligned} \tag{D.11}$$

and:

$$\tau_{1,n+1} - \tau_{1,n} = A m_0 \frac{1}{\tan\theta} = \frac{\tau_{0,n}}{\tan\theta}.$$

Therefore:

$$\tan\theta = \frac{\tau_{0,n}}{\tau_{1,n+1} - \tau_{1,n}}.$$

Using Equations (D.11) and (D.6), we also directly conclude that:

$$\lambda = \frac{\tau_{0,n}}{3 \sin\theta m_0}.$$

□

Proposition 7. *Let the quantity $\widehat{\tau}_{p,n}$ be recursively defined as:*

$$\widehat{\tau}_{p,n} = \begin{cases} \tau_{0,n}, & p = 0, \\ \tau_{p,n} - \sum_{j=0}^{p-1} \binom{p}{j} \frac{m_{p-j}}{m_0 (\tan\theta)^{p-j}} \widehat{\tau}_{j,n}. & p > 0, \end{cases}$$

then $\widehat{\tau}_{p,n}$ can be expressed as:

$$\widehat{\tau}_{p,n} = A m_0 (u_n)^p + pB m_0 (u_n)^{p-1}. \tag{D.12}$$

Proof. For $p = 0$, we directly have:$\widehat{\tau}_{0,n} = \tau_{0,n} = Am_0$. For $p > 0$, we have:

$$\widehat{\tau}_{p,n} = \tau_{p,n} - \sum_{j=0}^{p-1} \binom{p}{j} \frac{m_{p-j}}{m_0 (\tan\theta)^{p-j}} \widehat{\tau}_{j,n},$$

Therefore:

$$\begin{aligned} \widehat{\tau}_{p,n} + \sum_{k=0}^{p-1} \binom{p}{k} \frac{m_{p-k}}{m_0 (\tan\theta)^{p-k}} \widehat{\tau}_{k,n} &= \tau_{p,n}, \\ \sum_{k=0}^{p} \binom{p}{k} \frac{m_{p-k}}{m_0 (\tan\theta)^{p-k}} \widehat{\tau}_{k,n} &= A \sum_{k=0}^{p} \binom{p}{k} \frac{m_{p-k}}{(\tan\theta)^{p-k}} (u_n)^k + \\ &\qquad p \cdot B \sum_{k=0}^{p-1} \binom{p-1}{k} \frac{m_{p-1-k}}{(\tan\theta)^{p-1-k}} (u_n)^k. \\ \sum_{k=0}^{p} \binom{p}{k} \frac{m_{p-k}}{m_0 (\tan\theta)^{p-k}} \widehat{\tau}_{k,n} &= A \sum_{k=0}^{p} \binom{p}{k} \frac{m_{p-k}}{(\tan\theta)^{p-k}} (u_n)^k + \\ &\qquad p \cdot B \sum_{k=1}^{p} \binom{p-1}{k-1} \frac{m_{p-k}}{(\tan\theta)^{p-k}} (u_n)^{k-1}. \end{aligned}$$

Now, since $p\binom{p-1}{k-1} = k\binom{p}{k}$, we have:

$$\begin{aligned} \sum_{k=0}^{p} \binom{p}{k} \frac{m_{p-k}}{m_0 (\tan\theta)^{p-k}} \widehat{\tau}_{k,n} &= A \sum_{k=0}^{p} \binom{p}{k} \frac{m_{p-k}}{(\tan\theta)^{p-k}} (u_n)^k + \sum_{k=1}^{p} k \binom{p}{k} \frac{m_{p-k}}{(\tan\theta)^{p-k}} B (u_n)^{k-1}. \\ \sum_{k=0}^{p} \binom{p}{k} \frac{m_{p-k}}{m_0 (\tan\theta)^{p-k}} \widehat{\tau}_{k,n} &= \sum_{k=0}^{p} \binom{p}{k} \frac{m_{p-k}}{(\tan\theta)^{p-k}} \left[A (u_n)^k + kB (u_n)^{k-1} \right]. \end{aligned}$$

Then by identifying each term of the summation on each side of the expression above, we

finally obtain:

$$\begin{aligned} \frac{1}{m_0}\widehat{\tau}_{k,n} &= A\left(u_n\right)^k + kB\left(u_n\right)^{k-1} \\ \widehat{\tau}_{k,n} &= Am_0\left(u_n\right)^k + kBm_0\left(u_n\right)^{k-1}. \end{aligned}$$

□

Provided that enough consecutive values of $\widehat{\tau}_{p,n}, p = i-1, i, i+1\ldots$ are known, the values of u_n, A and B can be retrieved using a generalized version of the annihilating filter method. Once u_n is known, the parameter γ can be retrieved as:

$$\gamma = u_n - \frac{n}{\tan\theta}. \tag{D.13}$$

Then, using the value of B, the remaining parameter α can be also retrieved as:

$$\alpha = \lambda\gamma\sin\theta - \frac{B}{\sin\theta}. \tag{D.14}$$

Bibliography

[1] Y. Abu-Mostafa and D. Psaltis. Recognitive aspects of moments invariants. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, PAMI-6(6):698–706, November 1984.

[2] A. Aldroubi and M. Unser. Sampling procedures in function spaces and asymptotic equivalence with Shannon's sampling theory. *Numerical Functional Analysis and Optimization*, 15:1–21, 1994.

[3] P. E. Anuta. Spatial registration of multispectral and multitemporal digital imagery using FFT techniques. *IEEE Trans. on Geoscience Electronics*, GE-8:353–368, October 1970.

[4] H. Asada and M. Brady. The curvature primal sketch. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 8(1):2–14, January 1986.

[5] L. Baboulaz and P. L. Dragotti. Distributed acquisition and image super-resolution based on continuous moments from samples. *IEEE Int. Conf. on Image Processing*, pages 3309–3312, October 2006.

[6] L. Baboulaz and P. L. Dragotti. Local feature extraction for image super-resolution. *IEEE Int. Conf. on Image Processing*, pages 401–404, September 2007.

[7] I. Balslev, K. Døring, and R. D. Eriksen. Weighted central moments in pattern recognition. *Pattern Recogn. Lett.*, 21(5):381–384, 2000.

[8] W. C. J. Black and D. A. Hodges. Time interleaved converter arrays. *IEEE Journal of Solid-state circuits*, 15(6):1022–1029, December 1980.

[9] R. E. Blahut. *Theory and Practice of Error Control Codes*. Addisson-Wesley, 1983.

[10] T. Blu, P.-L. Dragotti, M. Vetterli, P. Marziliano, and L. Coulot. Sparse sampling of signal innovations. *IEEE Signal Processing Magazine*, 25(2):31–40, March 2008.

[11] T. Blu and M. Unser. Approximation error for quasi-interpolators and (multi-) wavelet expansions. *Applied and Computational Harmonic Analysis*, 6(2):219–251, March 1999.

[12] P. Brand and R. Mohr. Accuracy in image measure. In S. El-Hakim, editor, *SPIE Conf. on Videometrics III*, volume 2350, pages 218–228, November 1994.

[13] L. G. Brown. A survey of image registration techniques. Survey, Dept of Computer Science, Columbia University, January 1992.

[14] P. Burns. Slant edge analysis tool SFRMAT 2.0. *Available at http : //www.i3a.org/downloads_iso_tools.html*, December 2007.

[15] J. A. Cadzow. Signal enhancement: a useful signal processing tool. In *Fourth Annual ASSP Workshop on Spectrum Estimation and Modeling*, pages 162–167, August 1988.

[16] E. Candes and T. Tao. Decoding by linear programming. *IEEE Transactions on Information Theory*, 51(12):4203–4215, Dec. 2005.

[17] J. F. Canny. A computational approach to edge detection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 8(6):679–698, November 1986.

[18] D. Capel and A. Zisserman. Computer vision applied to super-resolution. *IEEE Signal Processing Magazine*, pages 75–86, May 2003.

[19] T. Chan, S. Esedoglu, F. Park, and A. Yip. *Handbook of Mathematical Models in Computer Vision*, chapter 2: Recent developments in total variation image restoration. Springer, 2004.

[20] P. Davis. Triangle formulas in the complex plane. *Mathematics of Computation*, 18(88):569–577, October 1964.

[21] P. Davis and H. Pollak. On the analytic continuation of mapping functions. *Trans. Amer. Math. Soc.*, 87(1):198–225, January 1958.

[22] B. G. R. de Prony. Essai expérimental et analytique: sur les lois de la dilatabilité de fluides élastiques et sur celles de la force expansive de la vapeur de l'alcool, à diffrentes températures. *Journal de l'École Polytechnique*, 1(22):24–76, 1795.

[23] R. Deriche and G. Giraudon. A computational approach for corner and vertex detection. *Int. Journal Computer Vision*, 10(2):101–124, April 1993.

[24] D. Donoho. Wedgelets: Nearly minimax estimation of edges. *Annals of Statistics*, 27(3):859–897, June 1999.

[25] D. Donoho. Compressed sensing. *IEEE Transactions on Information Theory,*, 52(4):1289–1306, April 2006.

[26] P. L. Dragotti, M. Vetterli, and T. Blu. Exact sampling results for signals with finite rate of innovation using Strang-Fix and local kernels. In *Proc. IEEE ICASSP*, Philadelphia, USA, March 2005.

[27] P. L. Dragotti, M. Vetterli, and T. Blu. Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets Strang-Fix. *IEEE Trans. on Signal Processing*, 55(5):1741–1757, May 2007.

[28] M. Elad, P. Milanfar, and G. H. Golub. Shape from moments – an estimation theory perspective. *IEEE Trans. on Signal Processing*, 52(7):1814–1829, July 2004.

[29] J. Elbornsson, F. Gustafsson, and J.-E. Eklund. Blind adaptive equalization of mismatch errors in a time-interleaved A/D converter system. *IEEE Transaction on Circuits and Systems -I: Regulare papers*, 51(1):151–158, January 2004.

[30] S. Farsiu, M. D. Robinson, M. Elad, and P. Milanfar. Fast and robust multiframe super resolution. *IEEE Trans. on Image Processing*, 13(10):1327–1344, October 2004.

[31] O. Faugeras. *Three-Dimensional Computer Vision – A Geometric Viewpoint*. The MIT Press, Cambridge, 1993.

[32] M. Figueiredo, R. Nowak, and S. Wright. Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems. *Selected Topics in Signal Processing, IEEE Journal of*, 1(4):586–597, Dec. 2007.

[33] M. A. Fischler and R. C. Bolles. Random sample consensus: a paradigm for model fitting with application to image analysis and automated cartography. *Communications of the ACM*, 24, June 1981.

[34] J. Flusser and T. Suk. A moment-based approach to registration of images with affine geometric distortion. *IEEE Trans. on Geoscience and Remote Sensing*, 32(2):382–387, March 1994.

[35] H. Foroosh, J. B. Zerubia, and M. Berthod. Extension of phase correlation to subpixel registration. *IEEE Trans. on Image Processing*, 11(3):188–200, March 2002.

[36] W. Förstner. A framework for low level feature extraction. *European Conf. on Computer Vision*, pages 383–394, 1994.

[37] E. Fredholm. Sur une classe d'equations fonctionnelles. *Acta Mathematica*, 27(1):365–390, December 1903.

[38] W. T. Freeman and E. H. Adelson. The design and use of steerable filters. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 13(9):891–906, September 1991.

[39] Q. Gao and F. Yin. Two-dimensional direction-based interpolation with local centered moments. *Graph. Models Image Process.*, 61(6):323–339, 1999.

[40] J. D. García-Arteaga and J. Kybic. Automatic landmark detection for cervical image registration validation. *Proceedings of the SPIE Medical Imaging 2007: Computer-Aided Diagnosis*, 6514, 2007.

[41] N. Gehrig and P. L. Dragotti. Distributed sampling and compression of scenes with finite rate of innovation in camera sensor networks. *Data Communication Conference (DCC)*, March 2006.

[42] S. Ghosal and R. Mehrotra. A moment-based unified approach to image feature detection. *IEEE Trans. on Image Processing*, 6(6):781–793, June 1997.

[43] J. Glaunès, M. Vaillant, and M. I. Miller. Landmark matching via large deformation diffeomorphisms on the sphere. *J. Math. Imaging Vis.*, 20(1-2):179–200, 2004.

[44] G. H. Golub and C. F. van Loan. *Matrix Computations*. Johns Hopkins University Press, 3rd edition, 1996.

[45] G. H. Golub and U. von Matt. Generalized cross-validation for large-scale problems. *Journal of Computational and Graphical Statistics*, 6(1):1–34, March 1997.

[46] A. Goshtasby. *2-D and 3-D Image Registration,*. John Wiley and Sons, New York, Feb. 2005.

[47] A. Goshtasby, G. Stockman, and C. Page. A region-based approach to digital image registration with subpixel accuracy. *IEEE Trans. on Geoscience and Remote Sensing*, 24(3):390–399, May 1986.

[48] J. Hadamard. *Lectures on Cauchy's Problem in Linear Partial Differential Equations*. Yale University Press, New Haven, 1923.

[49] P. C. Hansen. *Rank-Deficient and Discrete Ill-posed Problems: Numerical Aspect of Linear Inversion*. Society for Industrial and Applied Mathematics, Philadelphia, 1999.

[50] P. C. Hansen. *The L-Curve and its Use in the Numerical Treatment of Inverse Problems*, pages 119–142. Number 5 in Advances in Computational Bioengineering. WIT Press, Southampton, 2001.

[51] P. C. Hansen. Regularization tools version 4.0 for matlab 7.3. *Numerical Algorithms*, 46:189–194, 2007.

[52] Y. Hao, P. Marziliano, M. Vetterli, and T. Blu. Compression of ECG as a signal with finite rate of innovation. In *IEEE Int. Conf. of Engineering in Medicine and Biology Society*, pages 7564–7567, Shanghai, China, September 1-4 2005.

[53] R. C. Hardie. A fast image super-resolution algorithm using adaptive Wiener filter. *IEEE Trans. on Image Processing*, 16(12):2953–2964, December 2007.

[54] C. Harris and M. Stephens. A combined corner and edge detector. In *4th Alvey Vision Conference*, pages 147–151, 1988.

[55] J. Heikkila. Moment and curvature preserving technique for accurate ellipse boundary detection. In *Int. Conf. on Pattern Recognition*, volume 1, page 734. IEEE Computer Society, August 1998.

[56] J. Heikkilä. Pattern matching with affine moment descriptors. *Pattern Recognition*, 37(9):1825–1834, March 2004.

[57] M. K. Hu. Visual pattern recognition by moments invariants. *IRE Trans. on Information Theory*, 8:179–187, 1962.

[58] T. S. Huang and R. Y. Tsai. *Multiple frame image reconstruction and registration*, chapter 7. Advances in Computer Vision and Image Processing. Greenwich, 1984.

[59] A. Innes, V. Ciesielski, J. Mamutil, and S. John. Landmark detection for cephalometric radiology images using pulse coupled neural networks. *Proceedings of the International Conference on Artificial Intelligence*, pages 511–517, 2002.

[60] M. Jacob, T. Blu, and M. Unser. An exact method for computing the area moments of wavelet and spline curves. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23(6):633–642, Jun 2001.

[61] M. Jacob and M. Unser. Design of steerable filters for feature detection using canny-like criteria. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 26(8):1007–1019, August 2004.

[62] A. K. Jain. *Fundamentals of Digital Image Processing*. Prentice Hall Information and Systems Sciences. Prentice Hall International, 1989.

[63] Y. Jia, Q. D. Le, L. Pearlstein, and P. Swan. Video processing in hdtv receivers for recovery of missing picture information: De-interlacing, frame-rate conversion, and super-resolution. *Information Display Magazine*, 23(11):24–29, November 2007.

[64] U. Jorges, G. Jummel, and M. Reinhold. A macromodel of sample-and-hold circuits. *Int. Journal of Circuit Theory and Applications*, 25(6):438–502, 1997.

[65] I. Jovanovíc and B. Beferull-Lozano. Oversampled a/d conversion and error-rate dependence of nonbandlimited signals with finite rate of innovation. *IEEE Trans. on Image Processing*, 54(6):2140–2154, June 2006.

[66] L. Kaufman. Maximum likelihood, least squares, and penalized least squares for PET. *IEEE Trans. on Medical Imaging*, 12(2):200–214, June 1993.

[67] W. Kester, editor. *The Data Conversion Handbook*, chapter 2.3, page 2.68. Newnes, 2005.

[68] P. D. Kovesi. MATLAB and Octave functions for computer vision and image processing. The University of Western Australia. Available from: <http://www.csse.uwa.edu.au/∼pk/research/matlabfns/>.

[69] N. Kurosawa, H. Kobayashi, K. Maruyama, H. Sugawara, and K. Kobayashi. Explicit analysis of channel mismatch effects in time-interleaved ADC systems. *IEEE Trans. on Circuits and Systems I - Fundamental Theory and Applications*, 48(3):261–271, March 2001.

[70] J. Kusuma. *Economic sampling of parametric signals*. PhD thesis, Massachusetts Intitute of Technology, Cambridge, MA 02139, USA, 2006.

[71] B.-C. Li and J. Shen. Fast calculation of local moments and application to range image segmentation. *Proc. International Conference on Pattern Recognition*, C:298–301, 1992.

[72] H. Li, B. S. Manjunath, and S. K. Mitra. A contour-based approach to multisensor image registration. *IEEE Trans. on Image Processing*, 4(3):320–334, March 1995.

[73] P. Löwenborg, H. Johansson, and L. Wanhammar. Analysis of gain and time-skew errors in filter bank based a/d converters. *IEEE Midwest Symposium on Circuits and Systems*, 1:263–266, 2001.

[74] L. Lucchese and G. M. Cortelazzo. A noise-robust frequency domain technique for estimating planar roto-translations. *IEEE Transaction on Signal Processing*, 48(6):1769–1786, June 2000.

[75] E. P. Lyvers, O. R. Mitchell, M. L. Akey, and A. P. Reeves. Subpixel measurements using a moment-based edge operator. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 11(12):1293–1309, December 1989.

[76] I. Maravić. *Sampling methods for parametric non-bandlimited signals : extensions and applications*. PhD thesis, Audio-Visual Communication Laboratory, Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland, 2004.

[77] I. Maravić and M. Vetterli. A sampling theorem for the radon transform of finite complexity objects. *IEEE Conf. on Acoustics, Speech and Signal Processing*, 2:1197–1200, April 2002.

[78] I. Maravić and M. Vetterli. Exact sampling results for some classes of parametric nonbandlimited 2-d signals. *IEEE Trans. on Signal Processing*, 52(1):175–189, January 2004.

[79] I. Maravić and M. Vetterli. Sampling and reconstruction of signal with finite rate of innovation in the presence of noise. *IEEE Trans. on Signal Processing*, 53(8):2788–2805, August 2005.

[80] P. Marziliano. *Sampling Innovations*. PhD thesis, Audio-Visual Communication Laboratory, Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland, 2001.

[81] G. Medioni and Y. Yasumoto. Corner detection and curve representation using cubic B-spline. *Computer Vision, Graphics, and Image Processing*, 39(3):267–278, September 1987.

[82] P. Milanfar, W. C. Karl, and A. S. Willsky. Reconstructing binary polygonal objects from projections: a statistical view. *CVGIP: Graph. Models Image Process.*, 56(5):371–391, 1994.

[83] P. Milanfar, G. C. Verghese, W. Clem Karl, and A. S. Willsky. Reconstructing polygons from moments with connections to array processing. *IEEE Trans. on Signal Processing*, 43(2):432–443, February 1995.

[84] K. Miller. Least squares methods for ill-posed problems with a prescribed bound. *SIAM Journal on Mathematical Analysis*, 1(1):52–74, 1970.

[85] A. Montijo and K. Rush. Accuracy in interleaved ADC systems. *Hewlett-Packard Journal*, pages 38–46, October 1993.

[86] H. Moravec. Toward automatic visual obstacle avoidance. *Int. Joint Conf. Artificial Intelligence*, page 584, 1977.

[87] R. Mukundan. Some computational aspects of discrete orthonormal moments. *Image Processing, IEEE Transactions on*, 13(8):1055–1059, Aug. 2004.

[88] R. Mukundan, S. Ong, and P. Lee. Image analysis by tchebichef moments. *Image Processing, IEEE Transactions on*, 10(9):1357–1364, Sep 2001.

[89] J. Nagy and M. Kilmer. Kronecker product approximation for three-dimensional imaging applications. *IEEE Trans. on Image Processing*, 15:604–613, 2006.

[90] J. Nagy and Z. Strakos. Enforcing nonnegativity in image reconstruction algorithms. *Mathematical Modeling, Estimation, and Imaging, David C. Wilson, et.al., Eds.*, pages 182–190, 2000.

[91] A. Noble. *Descriptions of image surfaces*. PhD thesis, Oxford University, Department of Engineering Science, 1989.

[92] S. J. Osher and R. P. Fedkiw. *Level Set Methods and Dynamic Implicit Surfaces*. Springer, October 2002.

[93] C. C. Paige and M. A. Saunders. LSQR: An algorithm for sparse linear equations and sparse least squares. *ACM Trans. on Mathematical Software*, 8(1):43–71, March 1982.

[94] A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw Hill, 3rd edition, 1991.

[95] M. Peyrovian and A. Sawchuk. Image restoration by spline functions. *Applied Optics*, 17:660, 1978.

[96] H. Pichler, F. Pavuza, and T. Sommer. S&H circuits for fast ad converters. *EUROMICRO, System Architecture and Integration*, pages 603–609, September 1994.

[97] A. Pikaz and I. Disntein. Using simple decomposition for smoothing and feature point detection of noisy digital curves. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 16(8):808–813, August 1994.

[98] V. Pisarenko. The retrieval of harmonics from a covariance function. *Geophysical Journal International*, 33:347–366, September 1973.

[99] I. Popovici and D. Withers. Locating thin lines and roof edges by custom-built moments. *Proc. IEEE Int. Conf. on Image Processing*, pages 753–756, October 2006.

[100] W. K. Pratt. Correlation techniques of image registration. *IEEE Trans. on Aerospace and Electronic Systems*, AES-10(3):353–358, May 1974.

[101] D. Robinson, S. Farsiu, and P. Milanfar. Optimal registration of aliased images using variable projection with applications to superresolution. *The Computer Journal*, April/May 2007.

[102] V. Rodehorst and A. Koschan. Comparison and evaluation of feature point detectors. *International Symposium Turkish-German Joint Geodetic Days*, March 2006.

[103] K. Rohr. Recognizing corners by fitting parametric models. *Int. Journal on Computer Vision*, 9(3):213–230, December 1992.

[104] W. Rudin. *Real and complex analysis, 3rd ed.* McGraw-Hill, Inc., 1987.

[105] C. Schmid, R. Mohr, and B. C. Evaluation of interest point detectors. *Int. Journal on Computer Vision*, 37(2):151–172, June 2000.

[106] I. J. Schoenberg. Contribution to the problem of approximation of equidistant data by analytic functions. *Quarterly of Applied Mathematics*, 4:45–99, 112–141, 1946.

[107] J. Shen, W. Shen, and D. Shen. On geometric and orthogonal moments. *Int. J. of Pattern Recognition and Artificial Intelligence*, 14(7):875–894, 2000.

[108] J. Shi and C. Tomasi. Good features to track. In *IEEE Conf. on Computer Vision and Pattern Recognition*, Seattle, June 1994.

[109] P. Shukla. *Sampling Schemes for Multidimensional Nonbandlimited Signals*. PhD thesis, Imperial College London, United Kingdom, 2007.

[110] P. Shukla and P. L. Dragotti. Sampling schemes for multidimensional signals with finite rate of innovation. *IEEE Trans. on Signal Processing*, pages 3670–3686, July 2007.

[111] S. Smith and J. Brady. SUSAN – a new approach to low level image processing. *Int. Journal on Computer Vision*, 23:45–78, 1997.

[112] J. Sprinzak and M. Werman. Affine point matching. *Pattern Recognition Letters*, 15:337–339, April 1994.

[113] P. Stoica and R. Moses. *Introduction to spectral analysis*. Prentice-Hall, 2000.

[114] G. Strang and G. Fix. A Fourier analysis of the finite element variational method. In *Constructive Aspect of Functional Analysis*, pages 796–830, Rome, Italy, 1971.

[115] M. Suhling, M. Arigovindan, P. Hunziker, and M. Unser. Multiresolution moment filters. *International Conference on Image Processing*, 1:393– 396, 2002.

[116] G. Szegö. Orthogonal polynomials. *4th ed. New York: Amer. Math. Soc*, 23, 1975.

[117] C. Teh and R. T. Chin. On image analysis by the methods of moments. *IEEE Trans. on Pattern Analysis and Machine Iintelligence*, 10(4):496–513, July 1988.

[118] A. Tikhonov. Solution of incorrectly formulated problems and the regularization method. *Soviet Math. Dokl.*, 4:1035–1038, 1963.

[119] M. Unser. Splines: A perfect fit for signal and image processing. *IEEE Signal Processing Magazine*, 16(6):22–38, November 1999.

[120] M. Unser, A. Aldroubi, and M. Eden. B-Spline signal processing: Part I—Theory and Part II—Efficient design and applications. *IEEE Trans. on Signal Processing*, 41(2):821–848, February 1993.

[121] M. Unser, A. Aldroubi, and M. Eden. The L_2 polynomial spline pyramid. *IEEE Trans. on Pattern Analysis and Machine Iintelligence*, 15(4):364–379, April 1993.

[122] M. Unser and T. Blu. Cardinal Exponential splines: Part I–Theory and Filtering Algorithms and Part II–Think Analog, Act digital. *IEEE Trans. on Signal Processing*, 53(4):1425–1449, April 2005.

[123] P. Vandewalle, G. Barrenetxea, I. Jovanovic, A. Ridolfi, and M. Vetterli. Experiences with Reproducible Research in Various Facets of Signal Processing Research. In *IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, volume 4, pages 1253–1256, 2007.

[124] P. Vandewalle, L. Sbaiz, J. Vandewalle, and M. Vetterli. How to take advantage of aliasing in bandlimited signals. *IEEE Conf. on Acoustics, Speech and Signal Processing*, 3:948–951, May 2004.

[125] J. Varah. Pitfalls in the numerical solution of linear ill-posed problems. *SIAM Journal Sci. Stat. Comput.*, 4(2):164–176, 1983.

[126] M. Vetterli, P. Marziliano, and T. Blu. Sampling signals with finite rate of innovation. *IEEE Trans. on Signal Processing*, 50(6):1417–1428, June 2002.

[127] R. Walden. Analog-to-digital converter survey and analysis. *IEEE Journal on Selected Areas in Communications*, 17(4):539–550, April 1999.

[128] N. Woods, N. Galatsanos, and A. Katsaggelos. Stochastic methods for joint registration, restoration, and interpolation of multiple undersampled images. *IEEE Trans. on Image Processing*, 15(1):201–213, January 2006.

[129] M. Xia and B. Liu. Image registration by super-curves. *IEEE Trans. on Image Processing*, 13(5):720–732, May 2004.

[130] Z. Yang and F. Cohen. Cross-weighted moments and affine invariants for image registration and matching. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 21(8):804–814, August 1999.

[131] P. Yap, P. Raveendran, and S. Ong. Chebyshev moments as a new set of moments for image reconstruction. *Neural Networks, 2001. Proceedings. IJCNN '01. International Joint Conference on*, 4:2856–2860, 2001.

[132] B. Zitova and J. Flusser. Image registration methods: a survey. *Image and Vision Computing*, 21:977–1000, October 2003.