Exact Feature Extraction using Finite Rate of Innovation Principles with an Application to Image Super-resolution

Loïc Baboulaz*, *Student Member, IEEE,* and Pier Luigi Dragotti, *Member, IEEE,*

EDICS: MRP-WAVL (Wavelets), ISR-RGST (Registration / alignment) and ISR-SUPR (Super-resolution)

Abstract

The accurate registration of multiview images is of central importance in many advanced image processing applications. Image super-resolution, for example, is a typical application where the quality of the super-resolved image is degrading as registration errors increase. Popular registration methods are often based on features extracted from the acquired images. The accuracy of the registration is in this case directly related to the number of extracted features and to the precision at which the features are located: images are best registered when many features are found with a good precision. However, in low-resolution images, only a few features can be extracted and often with a poor precision. By taking a sampling perspective, we propose in this paper new methods for extracting features in low resolution images in order to develop efficient registration techniques. We consider in particular the sampling theory of signals with finite rate of innovation [10] and show that some features of interest for registration can be retrieved perfectly in this framework, thus allowing an exact registration. We also demonstrate through simulations that the sampling model which enables the use of finite rate of innovation principles is well-suited for modeling the acquisition of images by a camera. Simulations of image registration and image super-resolution of artificially sampled images are first presented, analyzed and compared to traditional techniques. We finally present favorable experimental results of super-resolution of real images acquired by a digital camera available on the market.

I. Introduction

Multiview camera systems are composed of a set of cameras positioned at different locations and focusing on the same scene of interest. Thus, at any given time, the i-th camera acquires a sampled image $\mathbf{g}_i[m, n]$ of

The authors are with the Communications and Signal Processing Group, Imperial College London. E-mails: {L.Baboulaz, P.Dragotti}@imperial.ac.uk. Address: Communications and Signal Processing Group, Electrical and Electronic Engineering, Imperial College London, Exhibition Road, London SW7 2AZ, England; Tel: +44 (0)20 759-46192 ; Fax: +44 (0)20 759-46234.

This work was in part presented at ICIP'06 and ICIP'07 [2][3].

the particular view $f_i(x, y)$ it has of the scene. To take advantage of such systems and use advanced image processing techniques like super-resolution, motion estimation or occlusion removals, the first processing task of utmost importance is the accurate registration of the acquired images. The exact registration of two images consists in finding the geometric transformation that exists between the view f_i of one camera and the view f_j of a second camera with a different location. However, the continuous views are usually not available and the registration is done on the acquired digital sampled images $\mathbf{g}_i$ and $\mathbf{g}_j$.

Various surveys and books on image registration are available for an in-depth review of recent and more classic image registration techniques [5], [49], [19]. These techniques operate either in the spatial domain or the frequency domain of the image. Frequency domain methods can be computationally efficient but are often limited to global rigid motion as the translations and rotation are estimated from the aliased spectra [28], [18], [45]. There are a wide variety of registration techniques in the spatial domain. Earliest methods involved the use of the cross-correlation between images as it is maximized when the two images are correctly overlaid [35]. Many variations of this model have been proposed, *e.g.* by preprocessing the images with an edge detector [1]. Registration based on mutual information is also a standard method used in medical imaging [34]. In [23], the Taylor expansion of the first order is used to find the parameters of the registration. A large set of methods consist first in extracting features in images and then matching them across images in order to calculate the existing transformations. In that case, the registration is based only on the retrieved features. Local features are points of interest in the image like the center of gravity of closed boundary regions [17], [27] or corners. Various automatic corner detectors have been proposed, like the famous Harris-Plessey detector [20] or the SUSAN detector [39]. More details on different corner detectors can be found in [37]. In [7], correspondence between features is efficiently carried out by first computing putative correspondences with a correlation matching algorithm and then refining them with a RANSAC algorithm [16]. Another types of features are global features which take into account the whole image. They do not require any feature correspondence step but a single transformation occurring between any pair of images must be assumed. Image moments are the prevailing global features used for estimating image disparity and allow the retrieval of affine transformations [40], [21], [2].

Image registration is an inverse problem as it tries to estimate from the sampled images $\mathbf{g}_i, i = 1, 2, \ldots, N$ the transformation that occurred between the views $f_i, i = 1, 2, \ldots, N$. It is also dependent on the properties of the camera used for image acquisition like the sampling rate (or resolution) of the sensor, the imperfection of the lens that adds blur, and the noise of the device. As the resolution decreases, the local two-dimensional structure of an image degrades and an exact registration of two low-resolution images becomes increasingly difficult. In this respect, we observe that registration and sampling are intimately related. In order to achieve perfect registration, one "brute-force" solution would consist in reconstructing perfectly the continuous signals f_i from their sampled versions $\mathbf{g}_i$ and run a registration algorithm on the reconstructed continuous signals. Another

solution is to carry out a perfect estimation of relevant image features in f_i from $\mathbf{g}_i$ and run a registration algorithm based on such features alone.

Perfect reconstruction of a signal from its sampled version is the fundamental problem in sampling theory. For bandlimited signals, the well-known Nyquist-Shannon sampling theorem shows that exact reconstruction is possible from the samples only. Moreover, 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) [46]. The peculiarity of these signals resides in the fact that they can be completely described by a parametric expression with a finite number of degrees of freedom. By taking into account known properties of the sampling setup, the parameters of the observed FRI signal can be exactly retrieved. The new sampling schemes for FRI signals were first proposed by Vetterli *et al* in [46] and then extended in [10] to the case of kernels with compact support. The multidimensional scenario has been considered in [29] and [38].

In this paper, we provide exact results for feature extraction in low-resolution images for registration purposes. We are primarily interested in finding the exact localization of features. In that respect, feature extraction (or localization) algorithms differ from feature detection techniques which are merely concerned with the existence or not of feature in an image. We assume that the sampling kernel is known and consider the latest theoretical results from the sampling theory of FRI signals to derive new feature extraction techniques. The underlying motivation in using the FRI sampling theory is due to the fact that some features used for registration can be modelled as FRI signals for which perfect reconstruction methods can be established. In particular, we present two novel feature extraction techniques that allow to retrieve exactly global features like moments or local features like step edges in low-resolution images. From the extracted features,standard feature-based registration algorithms are used to perform image super-resolution as this application requires a registration of high quality. We first show super-resolution results on synthetic images obtained by simulating an acquisition device with known characteristics. We then apply the proposed algorithms to image super-resolution of real images captured with a camera available on the market. The quality of the super-resolved images in both cases gives evidence of the accuracy of the registration and also highlights the validity of our acquisition model.

The organization of the paper is as follows. Section II presents the principles of sampling FRI signals and describes the sampling kernels and their properties. It also introduces the image acquisition model considered and details the problem of image registration. In Section III, we present an exact method that retrieves global features which are then used for registration purposes. In Section IV, we demonstrate a local approach of feature extraction. We focus on the extraction of step edges from which corners are inferred. In Section V, both registration methods are used in the context of image super-resolution and various experimental results are provided. We then conclude in Section VI.

II. PROBLEM SET-UP

A. Image Acquisition Model

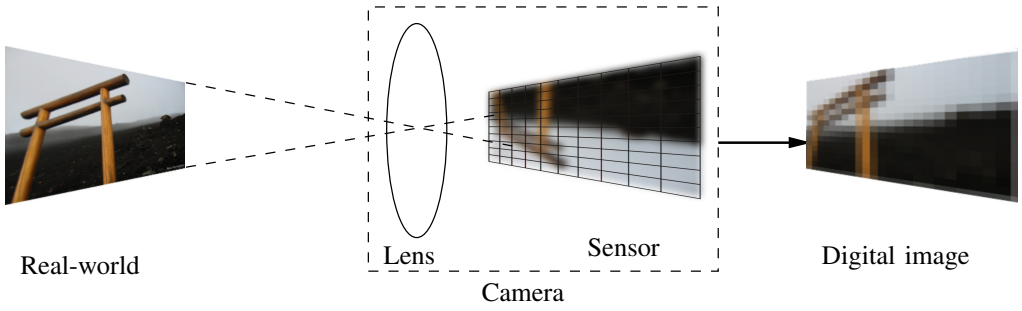
This section reviews the idealized image formation model considered in this paper and describes how the samples are related to the observed view via the Point-Spread Function (PSF) of the camera. The diagram in Figure 1(a) presents the main components of a camera which lead to a digital image of a given 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 sources of blur also participate to the overall filtering of the ray lights. Thus, a diffraction phenomenon occurs when the size of the camera aperture is of the same order as the wavelength of the light rays. This introduces interferences on the camera plane which are often referred as the "airy disk" or "blob". Also, some amount of blur caused by motion or atmospheric conditions can be introduced and deteriorate the observed view.

Another 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. Since each pixel value results from the integration of the incoming light over a finite spatial region defined by the characteristics of the sensor array (physical size, number of pixel, technology), the sensor array also contributes in blurring the image. In this research, we consider that the overall blur introduced by the camera is characterized by the PSF. The PSF is often modeled by a Gaussian pulse in the literature but we model here the PSF with B-spline functions for mainly two reasons. First, B-splines are very similar to a Gaussian pulse [44]. Second, B-splines possess properties like polynomial reproduction that we want to take advantage of. B-splines have already been used as a PSF model in [33] but their polynomial reproduction capabilities have not yet been exploited.

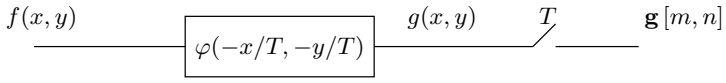
Figure 1(b) presents the equivalent idealized model to Figure 1(a) in terms of filter and analog-to-digital converter. The incoming continuous irradiance light-field $f(x,y)$ is first filtered with the function $\varphi(x,y)$. This two-dimensional function is the PSF that characterizes the camera and is assumed 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:

$$\begin{aligned} \mathbf{g}[m,n] = g(m,n) &= \iint f(x,y)\,\varphi(x/T - m, y/T - n)\,dxdy \\ &= \langle f(x,y), \varphi(x/T - m, y/T - n) \rangle, \end{aligned} \tag{1}$$

where $x, y \in \mathbb{R}$, $m, n \in \mathbb{Z}$. We assume throughout this paper that the sampling period is equal to T in both dimensions in order to simplify notations. As seen in Equation (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



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



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

Fig. 1. Camera model: (a) the incoming irradiance light field is blurred by the lens and sampled by 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 $\mathbf{g}$ is the sampled image.

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 1(b). When writing Equation (1), it is implicitly assumed that the sampling kernel is spatially invariant. Finally if the kernel is separable, it can be written as a tensor product of two 1D functions: $\varphi(x, y) = \varphi_1(x) \otimes \varphi_2(y)$ so that: $\mathbf{g}[m, n] = \iint f(x, y)\, \varphi_1(x/T - m)\, \varphi_2(y/T - n)\, dxdy$.

B. Sampling Kernels and Finite Rate of Innovation

The recent sampling theory for FRI signals is concerned with the problem of perfect reconstruction of non-bandlimited signals. Examples of FRI signals are streams of Diracs or piecewise polynomial functions. Such signals are non-bandlimited but follow a parametric expression with a finite number of degrees of freedom. For example, a stream of Diracs is completely defined by the location and the amplitude of each Dirac.

In an acquisition device, the incoming signal $x(t)$ is first filtered and then sampled. The obtained samples are given by $y_n = \langle x(t), \varphi(t/T - n) \rangle$ where the function $\varphi(t)$ is the sampling kernel. In this study, we consider the set of sampling kernels known as *polynomial reproducing kernels* as described in [10]. The particular property of these kernels is their ability to reproduce polynomials up to a certain degree by linearly combining several shifted versions such as:

$$\sum_{n \in \mathbb{Z}} c_n^{(p)} \varphi(t/T - n) = t^p \quad p = 0, \ldots, P, \tag{2}$$

where $c_n^{(p)}$ are known coefficients and P depends on the kernel itself. Such functions are said to have an approximation order equal to $(P + 1)$. In [42], Strang and Fix proved the necessary and sufficient conditions

for a function $\varphi(x)$ to possess polynomial reproduction qualities:

$$\widehat{\varphi}(0) \neq 0 \qquad \text{and} \qquad \frac{d^{(k)}\widehat{\varphi}(2i\pi)}{d\omega^k} = 0 \quad \text{for} \quad i \in \mathbb{Z} - \{0\}, \quad k = 0, \ldots, P, \tag{3}$$

where $\widehat{\varphi}(\omega)$ is the Fourier transform of $\varphi(t)$. These conditions are called the Strang-Fix conditions of order $(P+1)$. The coefficients $c_n^{(p)}$ used for the reproduction of the monomial t^p are computed as follows [4]:

$$c_n^{(p)} = \langle t^p, \tilde{\varphi}(t/T - n) \rangle = \frac{1}{T} \int t^p \tilde{\varphi}(t/T - n)\, dt, \qquad p = 0, \ldots, P, \tag{4}$$

where $\tilde{\varphi}(t)$ is the dual of $\varphi(t)$. The computation of the coefficients $c_n^{(p)}$ is straightforward when $\varphi(t)$ is orthogonal, that is, when $\langle \varphi(t), \varphi(t-n) \rangle = \delta_n$. In this case, $\tilde{\varphi}(t) = \varphi(t)$ and $c_n^{(p)} = \langle t^p, \varphi(t/T - n) \rangle$. It is more involved when $\varphi(t)$ is not orthogonal. In this second situation, one has to evaluate the dual function first and then the coefficients. The computation of the dual of a B-spline is discussed later in this section.

There exists a variety of functions satisfying Strang-Fix conditions. We consider in this paper the family of B-spline [43]. This function called centered B-spline of degree P is denoted by $\beta_{\mathrm{P}}(x)$ and can be obtained recursively by successive convolution of the box B-spline $\beta_0(x)$:

$$\beta_{\mathrm{P}}(x) = \underbrace{\beta_0(x) * \ldots * \beta_0(x)}_{(\mathrm{P}+1) \text{ times}} \quad , \quad \beta_0(x) = \begin{cases} 1, & |x| < \frac{1}{2} \\ \frac{1}{2}, & |x| = \frac{1}{2} \\ 0, & \text{otherwise.} \end{cases}$$

The Fourier transform of $\beta_{\mathrm{P}}(x)$ is given by:

$$\widehat{\beta}_{\mathrm{P}}(\omega) = [\mathrm{sinc}(\omega)]^{\mathrm{P}+1} \tag{5}$$

and satisfies Strang-Fix conditions of order $(\mathrm{P}+1)$. Now assume that $\varphi(x)$ is a B-spline $\beta_{\mathrm{P}}(x)$. The dual B-spline $\tilde{\varphi}(x)$ that forms a biorthonormal basis with $\varphi(x)$ so that $\langle \varphi(x-m), \tilde{\varphi}(x) \rangle = \delta_m$ can be computed as [44]:

$$\tilde{\varphi}(x) = \sum_{n=-\infty}^{\infty} (b_{2\mathrm{P}+1})^{-1}[n] \cdot \beta_{\mathrm{P}}(x-n),$$

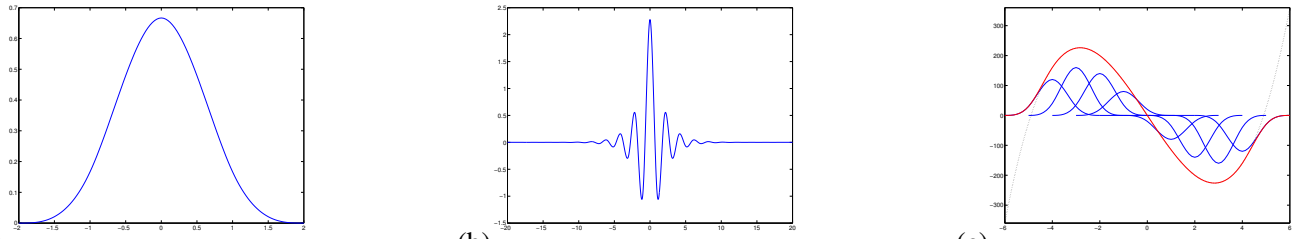
where:

$$(b_{2\mathrm{P}+1})^{-1}[n] \leftrightarrow (B_{2\mathrm{P}+1}(z))^{-1}.$$

Here, $B_{2\mathrm{P}+1}(z)$ is the z-transform of $b_{2\mathrm{P}+1}[n] = \beta_{2\mathrm{P}+1}(n)$, $n \in \mathbb{Z}$. For further details on the computation of the dual we refer to [44]. Figure 2(a) and (b) show, as example, the cubic B-spline and its dual. Figure 2(c) presents the reproduction with B-splines of the polynomial $5x^3 - 120x$ over the interval $[-4, 4]$.

In two dimensions, similar results can be derived. In particular, when $\varphi(x, y)$ is a separable kernel, each of its components is reproducing polynomials so that we have:

$$\sum_{m,n \in \mathbb{Z}} c_{m,n}^{(p,q)} \varphi_1(x/T - m)\, \varphi_2(y/T - n) = x^p y^q, \quad p = 0, \ldots, P, \quad q = 0, \ldots, Q, \text{ with } \quad c_{m,n}^{(p,q)} = c_m^{(p)} \cdot c_n^{(q)}.$$



(a) (b) (c)

Fig. 2. (a) Cubic B-spline; (b) Cubic Dual spline used to calculate the coefficient $c_m^{(p)}$; (c) Reproduction (in red) of the polynomial $5x^3 - 120x$ (dotted line) with a linear combination of scaled and shifted cubic B-splines (in blue) over the support $[-4, 4]$.

A closer look at Equation (1) also reveals that the sampling period T is also used to rescale the sampling kernel. The reason of this coupling is to maintain a unit spacing (relative to φ) between two consecutive samples in order to satisfy the polynomial reproduction property as reflected in Equation (2). However, it can happen that, in a real scenario, the shift between two samples is smaller than T. For example, image acquisition devices like mobile phones or webcams usually introduce a strong blur due to the low quality of the lens. In that scenario, the support of the corresponding sampling kernel is large although the image sensor has a good resolution. If we assume that the sampling period is reduced by an integer factor M, then the image samples are in this case given by:

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

The samples can be divided into their polyphase components $\mathbf{g}^{(i,j)}$, $i = 0, \ldots M-1$ and $j = 0, \ldots M-1$:

$$\begin{aligned}\mathbf{g}^{(i,j)}\left[m,n\right] &= \left\langle f\left(x,y\right), \varphi\left(x/T - m - i/M, y/T - n - j/M\right)\right\rangle \\ &= \left\langle f\left(x,y\right), \varphi^{(i,j)}\left(x/T - m, y/T - n\right)\right\rangle,\end{aligned}$$

where $\varphi^{(i,j)}\left(x/T - m, y/T - n\right) = \varphi\left(x/T - i/M, y/T - j/M\right)$. Each polyphase component is treated independently and the corresponding coefficients $c^{(p,q)}_{Mm+i,Mn+j}$ are obtained using Equations (4) with the sampling kernel $\varphi^{(i,j)}$. This polyphase decomposition will be used in Section V-B.

C. Multiview Images and Registration Problem

Assuming that each camera of a multiview system has the same intrinsic parameters and the same sampling kernel $\varphi\left(x,y\right)$, then the acquired image by the i-th sensor is expressed as:

$$\mathbf{g}_i\left[m,n\right] = \left\langle f_i\left(x,y\right), \varphi\left(x/T - m, y/T - n\right)\right\rangle. \tag{6}$$

By neglecting border effects, it is often assumed that the different observations f_i can be related to a single observation of reference (*e.g.* f_1) via a geometric transformation $\mathcal{T}_i$ of the coordinates x, y:

$$f_i\left(x,y\right) = f_1\left(\mathcal{T}_i\left(x,y\right)\right), \tag{7}$$

where $\mathcal{T}_1$ is the identity matrix. The transformations $\mathcal{T}_i$ can be of various types depending on the complexity of the scene. Transformations range from simple translation to complex nonlinear transformations. We consider here linear transformations, *i.e.* translation, rotation, affine transformation and projective transformation. The imaging situation corresponds to an observed scene that is flat, or to a 3D scene that is observed from a distance much greater than the distance between cameras so that parallax effects are negligible [7]. Combining Equations (6) and (7), we obtain the following popular model:

$$\mathbf{g}_i[m,n] = \langle f_1(\mathcal{T}_i(x,y)), \varphi(x/T - m, y/T - n)\rangle. \tag{8}$$

The goal of image registration is thus to find the different transformations $\mathcal{T}_i$ as accurately as possible given the acquired images $\{\mathbf{g}_0, \mathbf{g}_1, \ldots\}$. As the resolution of the images decreases, less information is available and a correct estimation of the $\mathcal{T}_i$ gets more and more difficult. Image registration based on (local or global) features therefore relies on the accurate extraction of features in $\{f_i\}$ given the information provided by $\{\mathbf{g}_i\}$.

III. Image Registration with Continuous Moments from Samples

This section describes how the continuous moments of an image can be calculated exactly from a discrete image when the assumptions related to the imaging model considered are satisfied. From these results, an exact registration can be achieved.

A. Continuous moments

Since the first work of Hu [22], functions of moments have been extensively used in pattern recognition to build features that are invariant to a given transformation [21]. In image reconstruction, Milanfar *et al* showed how a convex bilevel polygonal shape can be perfectly reconstructed from a finite number of moments [30] [13]. Finally, moments have been used in various ways to perform image registration [17][27][48].

A moment is defined by its type (geometric, central or complex $\ldots$) and its order. The basic moments are the geometric moments $m_{p,q}$, of order $p+q$, which are obtained in 2D by inner product between polynomial planes and the function of interest $f(x,y)$:

$$m_{p,q} = \iint f(x,y)\, x^p y^q dxdy, \quad p,q \in \mathbb{N}. \tag{9}$$

The barycenter (x_b, y_b) of $f(x,y)$ is defined as $(x_b, y_b) = \left(\frac{m_{1,0}}{m_{0,0}}, \frac{m_{0,1}}{m_{0,0}}\right)$. The central moments $\mu_{p,q}$ are then expressed as:

$$\mu_{p,q} = \iint f(x,y)\,(x-x_b)^p\,(y-y_b)^q\,dxdy = \sum_{k=0}^{p}\sum_{l=0}^{q}\binom{p}{k}\binom{q}{l}(-x_b)^{p-k}\,(-y_b)^{q-l}\,m_{k,l}. \tag{10}$$

The complex moments $C_{p,q}$ are defined on the complex image plane $z = x + jy$, $j = \sqrt{-1}$ as:

$$C_{p,q} = \iint f(x,y)\,(x+jy)^p\,(x-jy)^q\,dxdy = \sum_{k=0}^{p}\sum_{l=0}^{q}\binom{p}{k}\binom{q}{l} j^{p-k+q+l}\,(-1)^{q-l}\,m_{k+l,p-k+q-l}. \tag{11}$$

As shown above, the various types of moments can be obtained by a linear combination of the geometric moments which therefore constitute the basic elements of moment-based analysis.

With an image acquisition system, 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 (9):

$$\underline{m}_{p,q} = \sum_{m,n} \mathbf{g}[m,n] (mT)^p (nT)^q . \tag{12}$$

When the resolution of $\mathbf{g}$ gets low, the discrete moments $\underline{m}_{p,q}$ do not provide a good approximation of the continuous moments and this discrepancy can degrade the performance of any moment-based techniques dramatically. An alternative solution might be to deconvolve each image first and then evaluate the discrete moments on the deconvolved samples. This approach may improve the end result but does not solve the problem when the resolution is low as indicated in Figure 3.

In [10] and [38], new sampling results were proposed for 1D and 2D FRI signals. In particular, it is shown that it is possible to compute the exact moments of an FRI signal from its sampled version, provided that the sampling kernel satisfies the Strang-Fix conditions. In this paper, we propose to use these results on real images in order to extract the true *continuous moments* of a real object f from its samples $\mathbf{g}$. The continuous moments are obtained by linear combination of the samples with the coefficients $c_{m,n}^{(p,q)}$ as follows:

$$\begin{aligned} m_{p,q} &= \int\int f(x,y)\, x^p y^q dxdy \\ &\overset{(a)}{=} \int\int f(x,y) \sum_m \sum_n c_{m,n}^{(p,q)} \varphi(x/T - m, y/T - n)\, dxdy \\ &= \sum_m \sum_n c_{m,n}^{(p,q)} \int\int f(x,y)\, \varphi(x/T - m, y/T - n)\, dxdy \\ &\overset{(b)}{=} \sum_m \sum_n c_{m,n}^{(p,q)} \mathbf{g}[m,n] , \end{aligned} \tag{13}$$

where (a) and (b) follow respectively from Equations (2) and (1). Thus the proposed combination of the samples with $c_{m,n}^{(p,q)}$ allows the extraction of the exact moments from a sampled version of the observed continuous scene. Once the continuous geometric moments are obtained, other types of continuous moments (*e.g.* central or complex) can be calculated using Equations like (10) and (11).

Figure 3 (a) and (b) show how the estimation of the moments using Equation (12) or (13) changes when the resolution decreases. Figure 3 (a) is the case when there is no noise on the samples and Figure 3 (b) is when a white Gaussian noise is added to the samples (SNR=18dB). For this experiment, 20 standard images (*e.g.* Lena, Goldhill, Peppers, Mandril) of size 512x512 are artificially blurred and downsampled 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 (12) or (13) are compared to the true moments $m_{p,q}$ of the

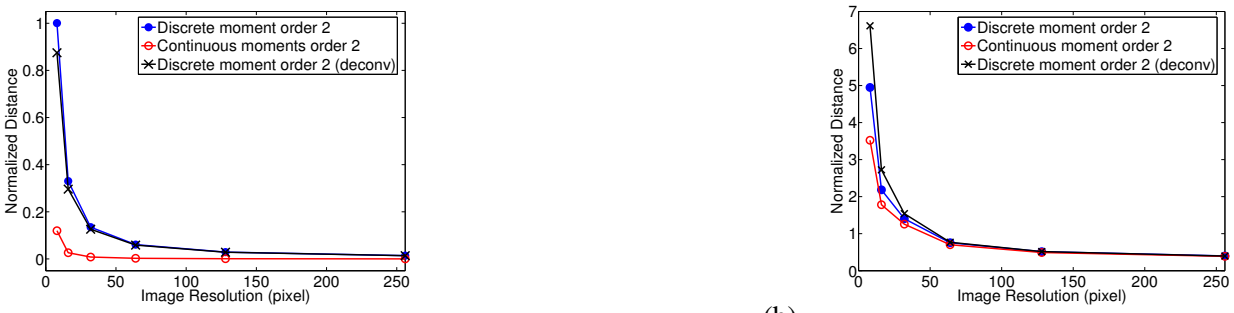


Fig. 3. Estimation of image moments of order 2 in function of the image resolution. Lines represent the normalized distances between the true moments and the discrete moments obtained from Equation (12); the continuous moments obtained from Equation (13); the discrete moments obtained from Equation (12) using the deconvolved samples. (a) Noiseless case; (b) Noisy case (SNR = 18dB).

original image by calculating the normalized distance between them:

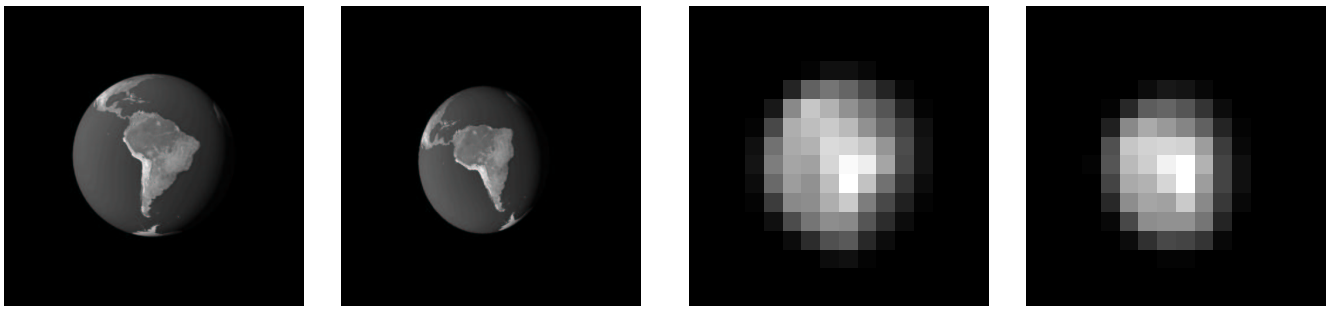
$$d_k = \sum_{i+j=k} \frac{(\hat{m}_{i,j} - m_{i,j})^2}{m_{i,j}^2},$$

where k defines the order of the moments considered. Figure 3 shows the variation of the average normalized distance d_2 with respect to the resolution of the sampled images for moments of order 2. For completeness, we also plot the normalized distance of the moments estimated using Equation (12) and the samples after deconvolution. When the sampling kernel is known and reproduces polynomials, the moments obtained with Equation (13) provide much more accurate results than those obtained with Equation (12) (with or without deconvolution before computation) and this is true even in the presence of noise.

B. Registration Method and Simulation Results

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 a global information on the image. Thus, in order to use the moments of two different images 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. If the background is uniform only in the neighborhood of the object of interest, then background subtraction and segmentation techniques can be used to extract the objects of interest from the background and treat each object as if the background was globally uniform. This was shown in [8].

Let $\mathbf{g}_1$ and $\mathbf{g}_2$ be two acquired images of the views f_1 and f_2 obtained as in Equation (6). 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 (7)). 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



(a) (b) (c) (d)

Fig. 4. (a)-(b) The two original views f_1 and f_2 (512x512 pixels each); (c)-(d) The two corresponding acquired low-resolution images $\mathbf{g}_1$ and $\mathbf{g}_2$ (16x16 pixels each). Original image "Blue Marble" by NASA's Earth Observatory.

$(\mathrm{X}_{shear}, \mathrm{Y}_{shear})$:

$$\mathcal{T}_2(x,y) = \mathbf{A}\begin{bmatrix} x & y \end{bmatrix}^T + \mathbf{t} \quad \text{with} \quad \mathbf{A} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \cdot \begin{bmatrix} \mathrm{X}_{scale} & \mathrm{X}_{shear} \\ \mathrm{Y}_{shear} & \mathrm{Y}_{scale} \end{bmatrix}. \tag{14}$$

It is shown in [40] that if f_1 and f_2 are transformed into their canonical forms (*i.e.* their covariance matrix is equal to the identity matrix), then the affine problem simplifies to finding a rotation angle which can then be retrieved using complex moments. This can be achieved by applying a whitening transform as in [21] where the expressions relating the central moments of the original signals and the complex moments of the signals in their canonical forms can also be found. Thus, since Equation (13) provides us with the exact continuous moments in the absence of noise, it is possible in theory to register exactly low-resolution images.

To measure the accuracy of the estimated transformation, the average and maximum geometric registration errors ε and $\varepsilon_{\max}$ are calculated as in [47]:

$$\varepsilon = \frac{1}{N^2}\sum_{x,y} \|\mathcal{T}(x,y) - \tilde{\mathcal{T}}(x,y)\|_2, \quad \text{and} \quad \varepsilon_{\max} = \max_{x,y} \|\mathcal{T}(x,y) - \tilde{\mathcal{T}}(x,y)\|_2, \tag{15}$$

where $\tilde{\mathcal{T}}$ is the calculated estimation of the exact affine transformation $\mathcal{T}$, and N is the size of the considered images f_1 and f_2.

In Figure 4 (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 angle $\theta = 36°$, 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 have:

$$\mathcal{T}_2(x,y) = \mathbf{A}\begin{bmatrix} x & y \end{bmatrix}^T + \mathbf{t} \quad \text{with} \quad \mathbf{A} = \begin{bmatrix} 0.706 & -0.4481 \\ 0.3893 & 0.7869 \end{bmatrix} \quad \text{and} \quad \mathbf{t} = \begin{bmatrix} -12 \\ 7 \end{bmatrix}.$$

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 (c)-(d). Given these two low-resolution images,

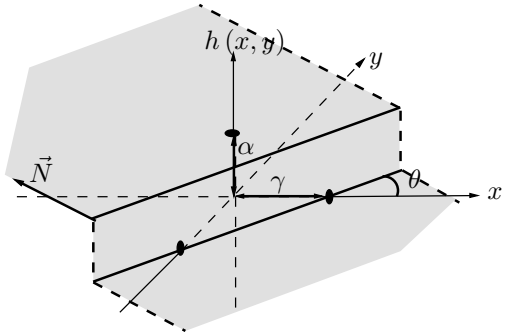


Fig. 5. Step edge model. A straight step edge is described by 3 parameters: its amplitude α, its orientation θ and its offset γ.

we apply the registration method with the continuous moments to estimate $\mathbf{A}$ and $\mathbf{t}$. The calculated affine transformation $\tilde{\mathcal{T}}_2$ 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 $\varepsilon = 0.0413$ pixels and $\varepsilon_{\max} = 0.148$ pixels which are the same errors obtained from the true moments of f_1 and f_2. 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:

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

with an average and maximum error of $\varepsilon = 11.2$ pixels and $\varepsilon_{\max} = 40.4$ pixels. The improvement of the average registration error is by a factor 280 in this simulation.

IV. Image Registration with Local Features

A. Step Edge Extraction

In this Section, the features considered are now 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. We 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 edge is presented in Figure 5. We now demonstrate how to retrieve the exact parameters from the samples.

Let $\vec{N} = (-\sin\theta, \cos\theta)^T$ be the vector normal to the edge and $\vec{d} = (x - \gamma, y)^T$ the vector of any point (x, y) in $\mathbb{R}^2$. Given $\vec{N}$ and $\vec{d}$, a step edge function $h(x, y)$ can be expressed as:

$$h(x, y) = \alpha H\left(\left\langle \vec{d}, \vec{N} \right\rangle\right), \tag{16}$$

where $H(t)$ is the unit step function whose value is 1 if $t \geq 0$ and 0 if $t < 0$. We assume that the signal $h(x,y)$ is sampled using a 2-D B-spline sampling kernel $\varphi(x,y)$. Therefore, following Equation (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 $\mathbf{d}[m,n]$ which are referred to as the differentiated samples:

$$\mathbf{d}[m,n] = \mathbf{h}[m+1,n] - \mathbf{h}[m,n]. \tag{17}$$

It is shown in [10] that the samples $\mathbf{d}[m,n]$ are the samples that would have been directly obtained by 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 also reproduce polynomials with one degree higher 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 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 $\varphi_1(x/T - m) * \beta_0(x/T - m)$:

$$\tau_{p,n} = \sum_{m \in \mathcal{S}_n} c_m^{(p)} \mathbf{d}[m,n], \tag{18}$$

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

$$\tau_{p,n} = -\alpha \sum_{j=0}^{p} \binom{p}{j} \frac{m_{p-j}}{(\tan\theta)^{p-j}} \left(\gamma + \frac{n}{\tan\theta}\right)^j , \tag{19}$$

where m_j are the moments of the sampling kernel: $m_j = \int_{\infty}^{\infty} t^j \varphi_2(t)\, dt$. Since $\varphi_2(t)$ is known, its moments m_j can be computed numerically once and stored. Besides, $m_0 = 1$ because $\varphi_2(t)$ satisfies partition of unity.

Algorithm 1 Subpixel Edge Location

Define $weight = 1$,
Run a Canny-like edge detector on sampled image,
Compute the differentiated samples using Equation (17),
for all position (i, j) detected as an edge **do**
    find the differentiated samples in the neighborhood of (i, j),
    calculate $\tau_{0,j}$, $\tau_{1,j}$ and $\tau_{1,j+1}$ using Equation (18),
    calculate α, $\tan\theta$ and γ using Equations (22),
    store $[\alpha \quad \tan\theta \quad \gamma \quad weight]$ as a candidate edge
end for
while there exists similar edges **do**
    Merge similar edges *i.e.* average $[\alpha \quad \tan\theta \quad \gamma]$ and add $weight$ together,
end while
Discard edges having a too small weight.

Also, for symmetric functions like B-splines, the odd order moments are equal to zero. 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 $p = 0, 1, 2$, formula (19) becomes:

$$\tau_{0,n} = -\alpha m_0; \quad \tau_{1,n} = -\alpha \left[m_0 u_n + \frac{m_1}{\tan\theta} \right]; \quad \tau_{2,n} = -\alpha \left[m_0 u_n^2 + 2\frac{m_1}{\tan\theta} u_n + \frac{m_2}{(\tan\theta)^2} \right]. \tag{20}$$

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:

$$\tan\theta = \frac{\tau_{0,n}}{\tau_{1,n+1} - \tau_{1,n}}. \tag{21}$$

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

$$\alpha = -\tau_{0,n}, \quad \tan\theta = \frac{\tau_{0,n}}{\tau_{1,n+1} - \tau_{1,n}}, \quad \text{and} \quad \gamma = \frac{(n+1)\,\tau_{1,n} - n\tau_{1,n+1}}{\tau_{0,n}}, \tag{22}$$

where it has been assumed that $m_0 = 1$ and $m_1 = 0$. Thus Equations (22) allows the calculation of the exact parameters of a step edge from its sampled version using only two consecutive rows and measurements $\tau_{k,n}$ up to order one.

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:

$$h(x, y) = \sum_{k=1}^{K} \alpha_k H\left(\left\langle \vec{d_k}, \vec{N} \right\rangle \right), \tag{23}$$

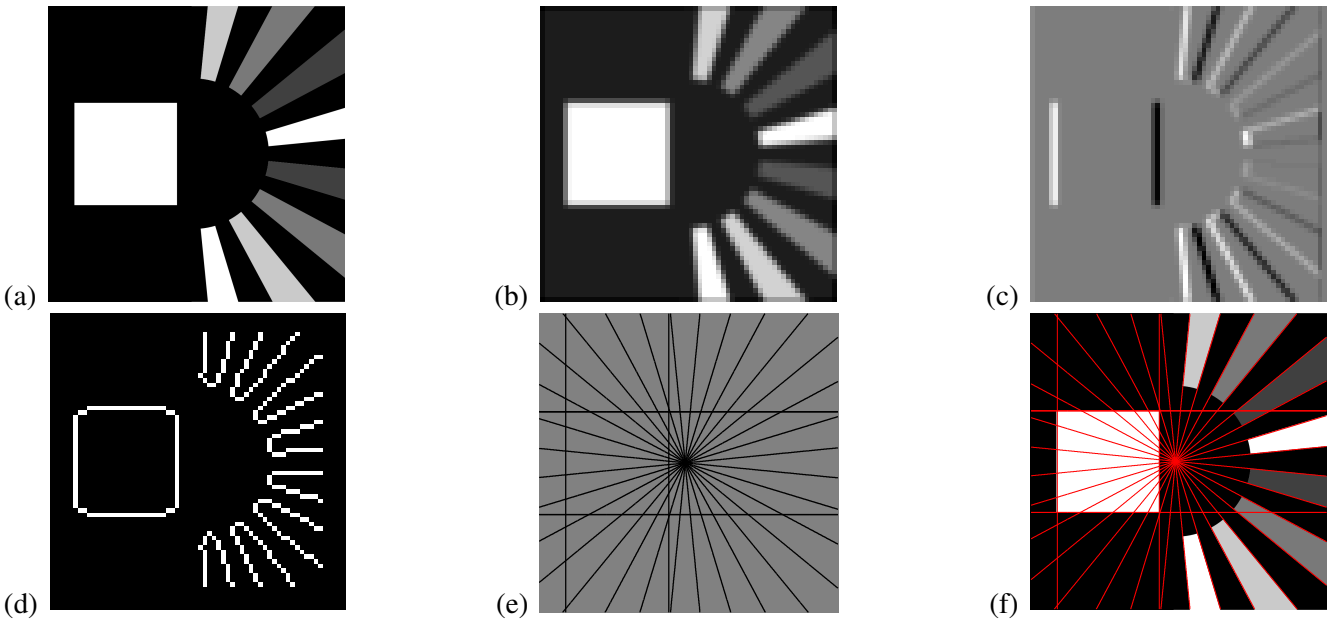


Fig. 6. (a) Original image (1024x1024pixels); (b) Sampled image with a B-spline (64x64pixels); (c) Horizontally differentiated samples; (d) Canny edge detection; (e) Retrieved edges with Algorithm 1; (f) Retrieved edges plotted against the original image.

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 (21). Moreover, by defining the quantity $\widehat{\tau_{p,n}}$ as follows:

$$\widehat{\tau_{p,n}} = \begin{cases} \tau_{0,n}, & p = 0, \\ \tau_{p,n} - \sum_{j=0}^{p-1} \binom{p}{j} \frac{m_{p-j}}{m_0 (\tan\theta)^{p-j}} \widehat{\tau_{j,n}}, & p > 0, \end{cases} \tag{24}$$

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

$$\widehat{\tau_{p,n}} = \sum_{k=1}^{K} \lambda_k \cdot (u_{k,n})^p \qquad p = 0, 1, \ldots, M-1, \tag{25}$$

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 $M \geq 2K$. For more details on the annihilating filter method, we refer to [46], [10], [41].

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 (22). 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 in the end a weight equal to k. Finally, edges with a weight below a given threshold are discarded 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 (17), horizontal edges cannot be extracted. In practice, Algorithm 1 is run on the sampled image and

on the transposed image in order to find all possible step edges, and both results are fused again. The rational of this approach is to find the largest possible region where a step edge occurs by discarding outliers so that averaging of the parameters improves stability and robustness to noise (see the Cramér-Rao bounds).

Figure 6 shows how the edge extraction is performed with Algorithm 1 on a synthetic image presenting different step edges with various orientations, amplitudes and offsets. Figure 6(a) shows the original scene before acquisition (1024x1024 pixels). Figure 6(b) is the acquired image of size 64x64 pixels obtained with a quadratic B-spline sampling kernel. Figure 6(c) is the differentiated samples $\mathbf{d}[m, n]$. Figure 6(d) shows the position of potential step edges using the Canny Edge detector. The retrieved edges are presented in Figure 6(e) and are also plotted against the original scene in Figure 6(f).

It is of interest to see how the estimation of the parameters of the step edge degrades when noise is present on the samples. We therefore consider the Cramér-Rao lower bound which provides on average the best estimation of a set of parameters using any unbiased algorithm. In [24], Kakarala *et al* have calculated the Cramér-Rao bound for the problem of edge localization. However, as opposed to our work, they do not take into account the effect of sampling. Our proposed approach can provide exact results even for heavily downsampled signals and the Cramér-Rao bound derived in this case provides a more realistic bound than the one given in [24]. With no loss of generality, we assume a zero mean white Gaussian noise with variance σ^2. The Cramér-Rao lower bound is derived in the case of our step edge detector in Appendix C. Figure 7 (a) and (b) show the scatterplots of the step edge parameters ($\theta = \pi/4$ and $\gamma = 0$) compared to the Cramér-Rao bounds for SNR levels from 0 to 50 dB for a sampled image of size 8x8 pixels ($N = 8$). Similarly, the average standard deviations of these parameters are also plotted versus the Cramér-Rao bound in Figure 7 (c) and (d). Simulations show that the proposed algorithm behaves well down to noise level of 13 dB.

B. Image Registration

We follow a similar registration approach to the one of Capel and Zisserman used for super-resolution in [7]. In their work, the features used are corners which are extracted with the Harris corner detector. This detector can generally achieve sub-pixel accuracy only up to 1/3 to 1/4 of a pixel by using quadratic fitting in the neighborhood of the local maxima. However, it is still possible to achieve a precise registration provided that a large number of feature can be extracted (several hundreds). Thus, in this case, the registration error due to the approximate location of each corner tends to diminish as the number of extracted features increases.

On images with low resolution, only a small number of features, say between 10 to 20, can usually be extracted in each image and matched. Thus, very accurate extraction of features is essential to obtaining a precise image registration. For this reason, the step edge extraction technique described in Algorithm 1 is used to find possible contours in the low-resolution images. Intersections of edges are then computed to locate

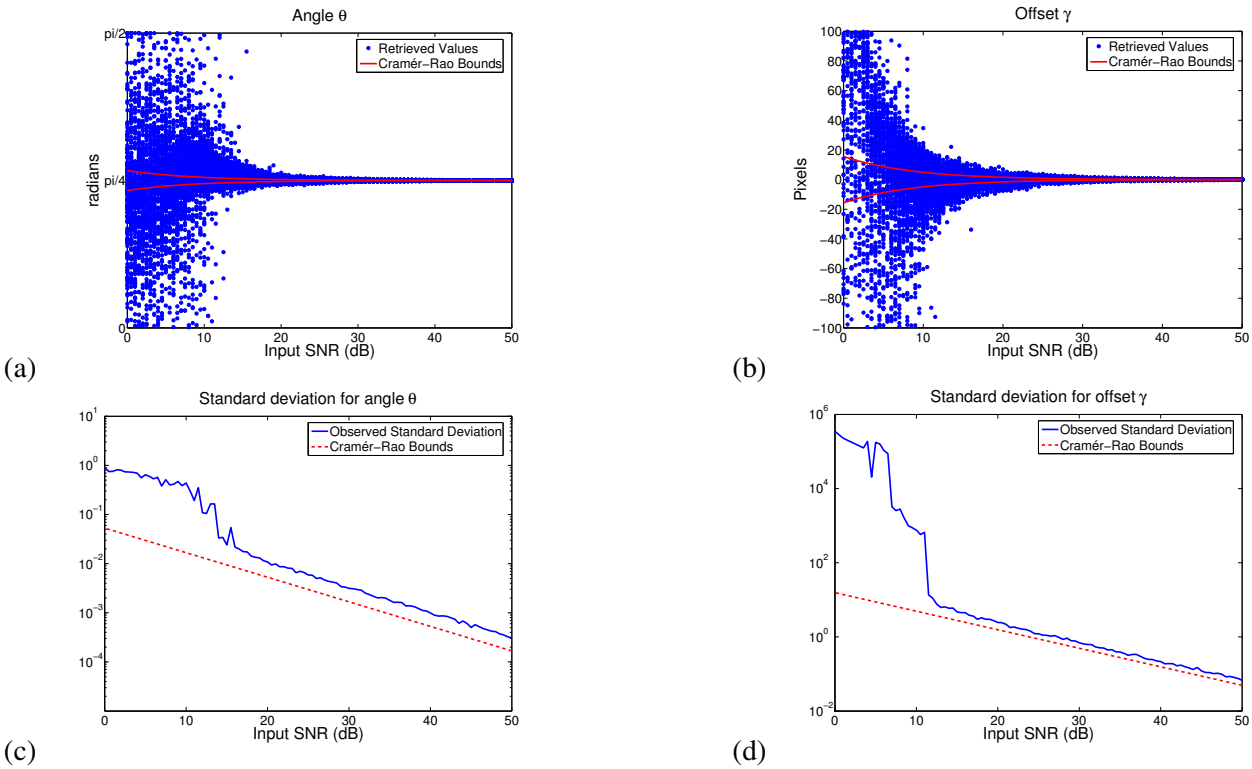


Fig. 7. Retrieval of a step edge with parameters ($\theta = \pi/4$ and $\gamma = 0$) in an 8x8 pixels image under noisy condition. (a)-(b) Scatterplots of angle and offset versus Cramér-Rao bound; (c)-(d) Standard deviation (averages over 10000 realizations) versus Cramér-Rao bound.

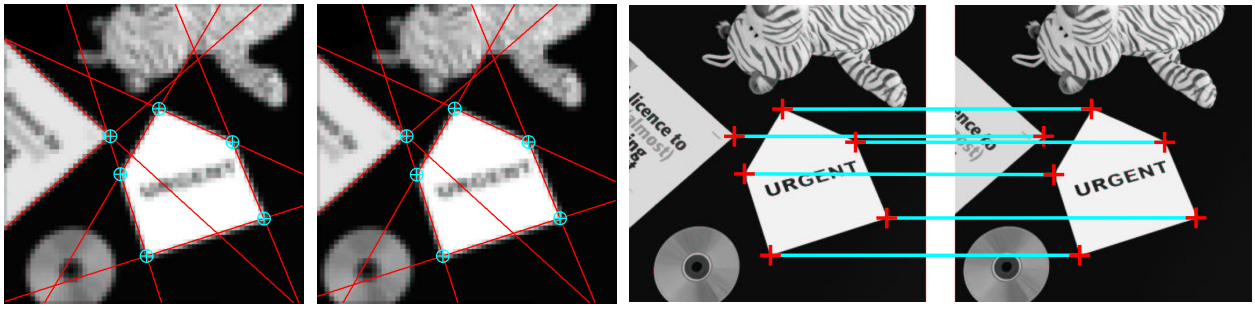
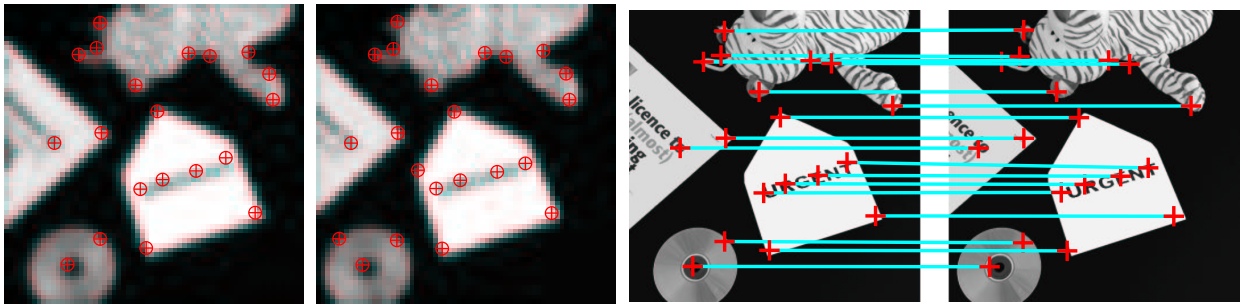


Fig. 8. (a)-(b) Extracted edges and corners with our approach on two low-resolution images of size 64x64 pixels each; (c)-(d) 6 corners are matched with correlation+RANSAC methods plotted against the high resolution images (512x512 pixels).

possible corners in the image. A first putative transformation is calculated from corresponding features obtained by using a correlation-based matching algorithm. The estimated transformation is then iteratively refined using a RANSAC procedure similarly to [7].

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 experiment. A high resolution image of a simple scene, favorable to both our method and the Harris corner detector, is first acquired. The scene contains steps edges, sharp corners, text and textures. The acquired picture is then cropped at two different known locations to create two pictures f_1 and f_2 with different fields of view and size 512x512 pixels. The transformation (a



(a) (b) (c) (d)

Fig. 9. (a)-(b) Extracted corners using Harris method on two low-resolution images of size 64x64 pixels each; (c)-(d) 18 corners are matched using correlation+RANSAC methods plotted against the high resolution images (512x512 pixels).

single translation) between each picture is known exactly and is given by: $\mathbf{t} = [0 \quad 28]^T$. Each image is then artificially downsampled with a quadratic B-spline of scale 8 giving two images of size 64x64 pixels each. Features are then extracted from g_1 and g_2 using either the Harris corner detector or our step edge detector. The functions used in this research for the subpixel Harris corner detector, the correlation matching and the RANSAC fitting are available from [26].

The features extracted with our approach are shown in Figure 8(a) and (b). Six corners in total have been successfully located from the extracted edges and matched across the two images using correlation and RANSAC techniques. The matched features are plotted against f_1 and f_2 for a visual appreciation of the subpixel accuracy (see 8(c) and (d)). The feature points returned by the Harris corner detector on each image are presented in Figure 9(a) and (b). Eighteen corners have been matched successfully across images and are plotted against f_1 and f_2 on Figure 9(c) and (d).

Each set of features is independently used to estimate the translation $\mathbf{t}$ between the two images. The estimated translation and registration errors with our approach are:

$$\tilde{\mathbf{t}} = [-0.15 \quad 28.13]^T \quad \text{and} \quad \varepsilon = \varepsilon_{\max} = 0.039 \quad \text{pixel.}$$

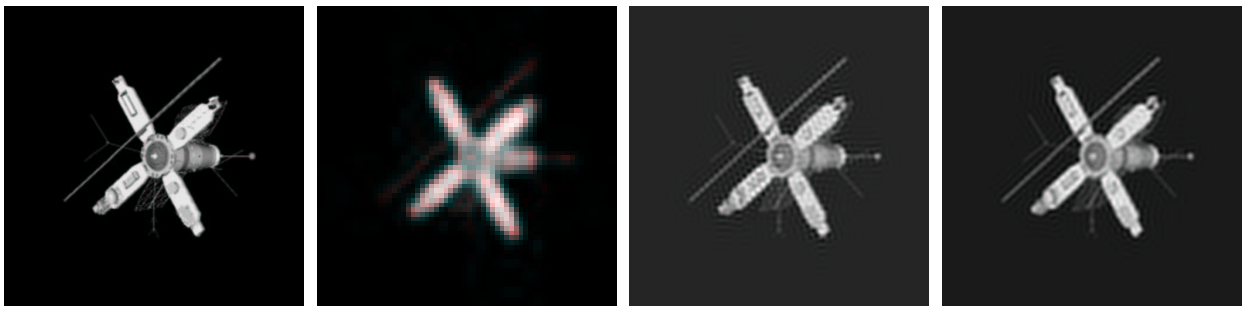
Similarly, the estimated translation and registration errors with the Harris features are:

$$\tilde{\mathbf{t}} = [-0.94 \quad 26.93]^T \quad \text{and} \quad \varepsilon = \varepsilon_{\max} = 2.04 \quad \text{pixel.}$$

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

V. Application to Image Super-resolution

The goal of image super-resolution is to construct a single, detailed, high-resolution image using a set of low-resolution images of the same scene. The problem of image super-resolution can be conceptually divided into



(a) (b) (c) (d)

Fig. 10. 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.

two sub-problems known as image registration and image reconstruction. Image registration aims at finding the disparity between the low-resolution images whereas image reconstruction consists in fusing the set of registered images into a single image and removing any blur and noise introduced during acquisition ([11])[1]. We consider in this work two restoration techniques: the Wiener deconvolution approach and the iterative Modified Residual Norm Steepest Descent (MRNSD) [31](*a.k.a.* EMLS [25])[2]. The Matlab function `deconvwnr` is used for the Wiener deconvolution taking the known PSF as input parameter. The MRNSD technique is an Expectation-Maximization iterative algorithm which forces the solution to be nonnegative.

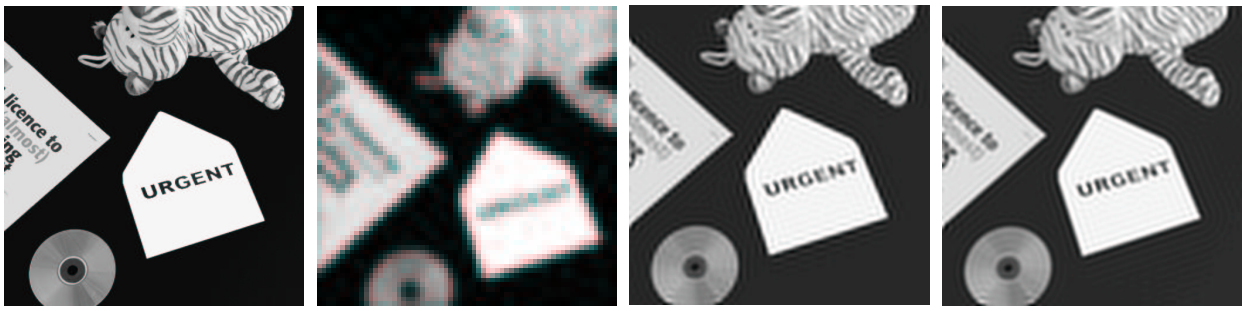
An overview and tutorial of super-resolution techniques can be found in [32][14]. Most earlier works focused on the restoration stage assuming that traditional registration methods provided a sufficiently accurate solution. However, as also observed in [36], the quality of the restoration in super-resolution problems depends heavily on the accuracy of the registration. It is therefore quite natural to test our registration algorithms in the context of image super-resolution.

A. Image Super-resolution: Simulations

1) Moment-based Registration: The moment-based registration method is applied to the case of image super-resolution. The first experiment is shown in Figure 10. As in the registration experiments, we use a single high resolution image (Satellite image, 512x512 pixels) shown in Figure 10(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 10(b)). This set of low-resolution images is then used as input for

[1]Notice that it is correct to divide the super-resolution problem into the two aforementioned sub-problems only when the motion is linear and the PSF is spatially and rotationally invariant [12].

[2]Other methods based on Total Variation ([9] [15]) are not considered here.



(a) (b) (c) (d)

Fig. 11. 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 with the Harris corner detector and Wiener filter, 512x512 pixels, PSNR = 15.1 dB; (d) Super-resolved image with the proposed edge detector and Wiener filter, 512x512 pixels, PSNR = 15.6 dB.

super-resolution. The translations are retrieved from the discrete and 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 obtained is exact to machine precision. For a fair comparison, the Wiener filter is used in both cases for restoration as it is fast and does not involve iterations. The super-resolved images are shown in Figure 10(c) and (d). The image in Figure 10(c) is obtained after registration with the discrete moments and the final PSNR is equal to 16.8dB. The super-resolved image shown in Figure 10(d) results from the utilization of the continuous moments for registration. 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.

2) Edge-based Registration: In this second experiment, we consider the registration based on the extraction of step edges. As in the previous section, we generated 20 images of scene (Figure 11(a)) by cropping a high resolution image at different locations. The 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 11(b)). In this simulation, two feature extraction methods are considered: the proposed step edge extractor and the subpixel Harris corner detector. As previously, in order to do a fair comparison, we do not use an iterative method but consider the Wiener filter as restoration method in both cases. The super-resolution algorithm based on the Harris features is inspired by the work of Capel *et al* in [7] where a similar approach is considered. With Harris features, the average 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 the maximum registration error observed in the 20 images is $\bar{\epsilon}_{\max} = 0.14$

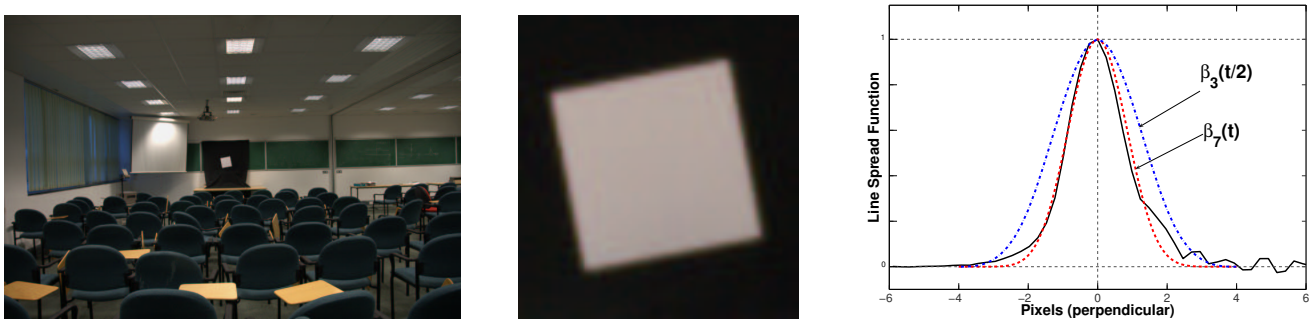


Fig. 12. Estimation of the PSF with the Slanted Edge Method; (a) Image of a slanted white square with step edges for PSF estimation (acquired with a Nikon D70s digital camera); (b) Zoom on the target; (c) Measured Line Spread Function (solid line). Its support has length 8 ranging from -4 to 4. The sampling kernel $\varphi(t)$ can either be modeled by a B-spline of degree 7, $\beta_7(t)$ (dashed line) or by a B-spline of degree 3 scaled by 2, $\beta_3(t/2)$ (dash-dot line).

pixels. Six features are matched in any pair of images in the set and the registration is improved on average by a factor 10. The super-resolution results are similar to the previous case where global features were used. Figure 11(c) presents the super-resolved image obtained with Harris features. This image shows more artefact (PSNR = 15.1 dB) compared to the super-resolved image obtained with the proposed registration technique which delivers a better visual quality and a PSNR = 15.6 dB. This simulation, thus, better highlights the fact that more accurate registration leads to better super-resolution.

B. Image Super-resolution: Real-case Scenario

In this section, we consider the case of image super-resolution from real images acquired with a digital camera (Nikon D70s). The following experiments thus naturally include the natural noise that occurs on the samples during acquisition. The registration approach considered here is based on the continuous moments. Since it takes a sampling point of view, we don't want our image samples to be modified by internal post-processing that is usually applied inside a digital camera right after acquisition (*e.g.* edge sharpening or noise attenuation). For this reason, the set of images is acquired in RAW format. In a first experiment, pictures are taken in a classroom with a focal length at 18mm (35mm equivalent: 27mm) and other settings at F16, 1/60s and ISO 200. The slanted edge method is used to estimates the PSF indirectly by measuring first the Edge Spread Function (ESF)[6]. By differentiating along the edge's normal direction, the Line Spread Function (LSF) can be obtained and represents a cross-section of the PSF. Only one LSF is necessary to characterize a circularly symmetric and spatially invariant PSF. The estimation of the PSF is presented in Figure 12. The acquired image is shown in Figure 12(a) and the target used for the PSF estimation is shown in Figure 12(b). The estimated LSF is the solid line in Figure 12(c). It can be observed that its support is approximately 8 pixels, ranging from -4 to 4. The PSF is therefore modeled by two different B-splines with support 8: a B-spline of degree 7, $\beta_7(t)$ (dashed

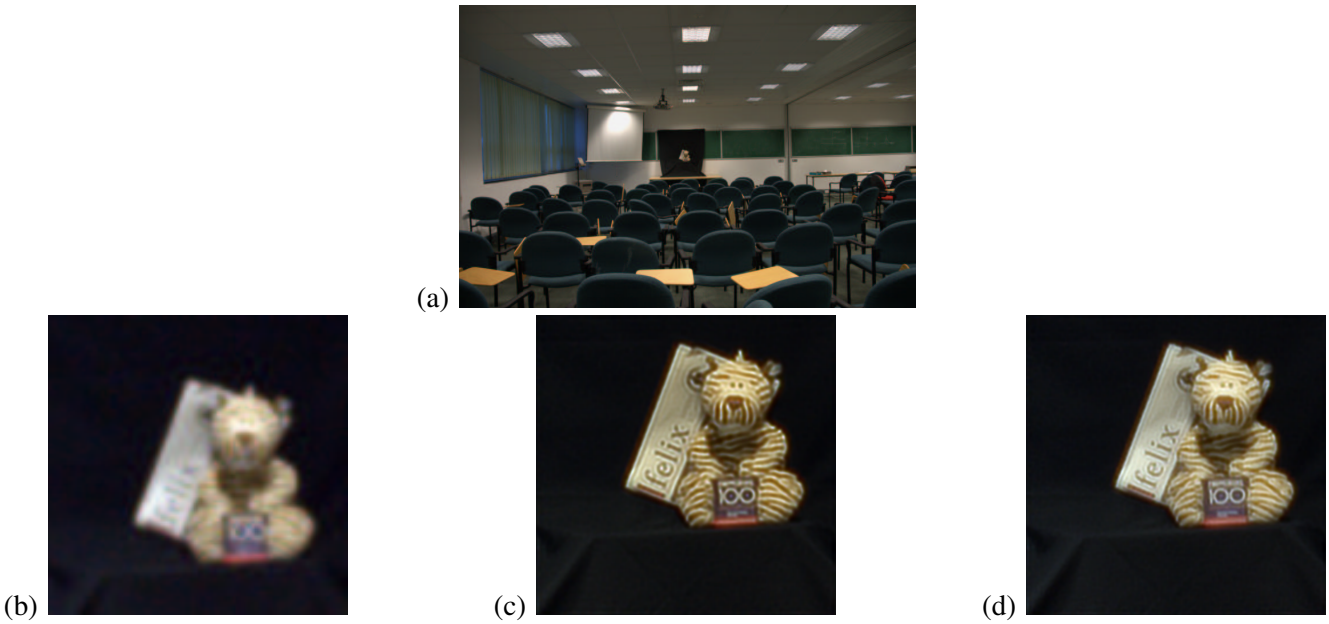


Fig. 13. Real image super-resolution from 40 images acquired with a Nikon D70s SLR camera; (a) One of the 40 images acquired by the camera; (b) Region of interest (128x128 pixels) used for super-resolution; (c) Super-resolved image of size 1024x0124 pixels (SR factor = 8). The PSF in this case is modeled 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 are restored with 60 iterations of the MRNSD method.

line), and a B-spline of degree 3 scaled by 2, $\beta_3(t/2)$ (dash-dot line).

The target used for PSF estimation is then replaced with a Tiger plush and a newspaper (see Figure 13(a))). Keeping the camera settings unchanged, 40 images are taken from random positions by moving the camera horizontally and vertically between each acquisition. It should be noted that the distance between foreground objects (desks, chairs) and background objects (blackboard) is large and traditional registration techniques that considers the whole image would not achieve good subpixel accuracy unless segmentation is first applied. Since the proposed feature extraction techniques can be exact on low-resolution image, we consider directly the regions of interest for registration and thus estimate an accurate local motion. The same region of interest of size 128x128 pixels is selected in each image (Figure 13(b)) and only this region is used for registration and restoration.

The sampling kernel is first modeled by the B-spline of degree 7. We register the images using continuous moments and the fused image is restored with the MRNSD method. The obtained super-resolved image (SR factor =8) is shown in Figure 13(c). In the second case, the sampling kernel is modeled by a B-spline of degree 3 scaled by two. Because of the scaled kernel, the device is oversampling by a factor two with respect to our model. The sampling period is consequently reduced by a factor two and the samples can be written as: $\mathbf{g}[m,n] = \langle f(x,y), \varphi(x/T - m/2, y/T - n/2)\rangle$. Two consecutive samples are now distant by $T/2$ instead

(a)  (b) 

Fig. 14. 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 (35mm 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.

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 [10]:

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

For each region of size 128x128 pixels (Figure 13(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. 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 13(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 less saturated than the image obtained with B-spline of level 7.

Another experiment is presented in Figure 14. Sixty pictures of the Moon are taken with a digital SLR camera and a lens with a focal length at 38mm (35mm equivalent: 57mm) and settings: F16, 1/60s, ISO 200. The PSF in this case is not estimated as previously and is directly approximated with a cubic B-spline at scale 1. The MRNSD algorithm is used as restoration method. Figure 14(a) shows the Moon as acquired by the camera and Figure 14(b) presents the obtained super-resolved image where new details of the Moon can now be observed.

VI. Conclusion

We have presented in this paper two novel approaches for feature extraction that take maximum advantage of the a priori knowledge of the acquisition filter and that are based on the basic principles behind the sampling of FRI signals. The first proposed method allows the exact retrieval of the continuous moments of an object from its sampled image. The second method retrieves the exact location of local image features such as step edges or parallel edges. These are then used to retrieve the exact location of corner points which are utilized

for the exact registration of low resolution images like in the context of image super-resolution. Experimental results on artificially sampled images and natural images shows the efficiency of the proposed feature extraction methods and the validity of the proposed acquisition model.

Although the use of these new methods was demonstrated only for registration, these techniques can also be used for pattern recognition, camera calibration and photogrammetry. This is part of our on-going research.

APPENDIX A

We now prove Equation (19): $\tau_{p,n} = -\alpha \sum_{i=0}^{p} \binom{p}{i} \frac{m_i}{(\tan\theta)^i} \left(\gamma + \frac{n}{\tan\theta}\right)^{p-i}$.

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 $f(t-T)$:

$$m'_p = \int f(t-T)\, t^p dt = \int f(x)\,(x+T)^p\, dx = \sum_{k=0}^{p} \binom{p}{k} T^k \int f(x)\, x^{p-k} dx = \sum_{k=0}^{p} \binom{p}{k} T^k m_{p-k} \qquad (26)$$

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

$$\begin{aligned}
\tau_{p,n} &= \sum_{m \in \mathcal{S}_n} c_m^{(p)} \mathbf{d}[m,n] \\
&= -\alpha \sin\theta \sum_{m \in \mathcal{S}_n} c_m^{(p)} \left\langle \delta\left(\vec{d}\cdot\vec{N}\right), \varphi(x-m, y-n) * \beta_0(x-m) \right\rangle \\
&= -\alpha \sin\theta \left\langle \delta\left(\vec{d}\cdot\vec{N}\right), \left(\sum_{m \in \mathcal{S}_n} c_m^{(p)} \varphi_1(x-m) * \beta_0(x-m) \right) \otimes \varphi_2(y-n) \right\rangle \\
&= -\alpha \sin\theta \left\langle \delta\left(\vec{d}\cdot\vec{N}\right), x^p \otimes \varphi_2(y-n) \right\rangle \\
&= -\alpha \sin\theta \int_{n-L}^{n+L} \varphi_2(y-n) \int \delta(-x\sin\theta + y\cos\theta + \gamma\sin\theta)\, x^p dx dy \\
&= -\alpha \sin\theta \int_{n-L}^{n+L} \varphi_2(y-n) \int \delta(-t + y\cos\theta + \gamma\sin\theta) \left(\frac{t}{\sin\theta}\right)^p \frac{dt}{\sin\theta} dy \\
&= -\alpha \int_{n-L}^{n+L} \varphi_2(y-n) \left(\frac{y\cos\theta + \gamma\sin\theta}{\sin\theta}\right)^p dy \\
&= -\alpha \int_{-L}^{L} \varphi_2(t) \left(\frac{t+n}{\tan\theta} + \gamma\right)^p dt \\
&= \frac{-\alpha}{(\tan\theta)^p} \int_{-L}^{L} \varphi_2(t)\,(t + n + \gamma\tan\theta)^p\, dt,
\end{aligned}$$

then applying Equation (26), we obtain: $\tau_{p,n} = \frac{-\alpha}{(\tan\theta)^p} \sum_{i=0}^{p} \binom{p}{i} m_{p-i} (n + \gamma\tan\theta)^i$, where $m_i = \int t^i \varphi_2(t)\, dt$ which finally yields:

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

■

APPENDIX B

Let $\mathbf{d}[m,n]$ be the differentiated samples of the signal $h(x,y)$ composed of K parallel step edges $\{\alpha_k, \gamma_k, \theta\}$. We have $h(x,y) = \sum_{k=1}^{K} \alpha_k H\left(\left\langle \vec{d_k}, \vec{N} \right\rangle\right)$, with $\vec{d_k} = (x - \gamma_k, y)^T$.

- We now prove that for K parallel step edges, Equation (21) is true: $\tan\theta = \frac{\tau_{0,n}}{\tau_{1,n+1} - \tau_{1,n}}$

  Proof: From the previous Appendix, it is straightforward to show that:

$$\begin{aligned}
\tau_{p,n} &\equiv \sum_{m \in \mathcal{S}_n} c_m^{(p)} \mathbf{d}[m,n] \\
&= \sum_{k=1}^{K} -\alpha_k \sin\theta \sum_{m \in \mathcal{S}_n} c_m^{(p)} \left\langle \delta\left(\vec{d_k} \cdot \vec{N}\right), \varphi(x-m, y-n) * \beta_0(x-m) \right\rangle \\
&= \sum_{k=1}^{K} -\alpha_k \sum_{i=0}^{p} \binom{p}{i} \frac{m_{p-i}}{(\tan\theta)^{p-i}} \left(\gamma_k + \frac{n}{\tan\theta}\right)^i \\
&= \sum_{i=0}^{p} \binom{p}{i} \frac{m_{p-i}}{(\tan\theta)^{p-i}} \left(\sum_{k=1}^{K} -\alpha_k (u_{k,n})^i\right), \qquad (27)
\end{aligned}$$

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

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

  which finally leads to the desired result. ■

- We now prove Equation (25) with $\lambda_k = -m_0 \alpha_k$ and $u_{k,n} = \gamma_k + \frac{n}{\tan\theta}$:

$$\widehat{\tau_{p,n}} = \sum_{k=0}^{K} \lambda_k (u_{k,n})^p, \quad \text{where} \quad \widehat{\tau_{p,n}} = \begin{cases} \tau_{0,n}, & p = 0, \\ \tau_{p,n} - \sum_{j=0}^{p-1} \binom{p}{j} \frac{m_{p-j}}{m_0 (\tan\theta)^{p-j}} \widehat{\tau_{j,n}}. & p > 0, \end{cases}$$

  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 (27), 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:

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

We can then conclude with the desired result:

$$\widehat{\tau_{i,n}} = \sum_{k=1}^{K} -m_0\alpha_k \left(u_{k,n}\right)^i = \sum_{k=1}^{K} \lambda_k \left(u_{k,n}\right)^i .$$

■

APPENDIX C
DERIVATION AND EVALUATION OF THE CRAMER-RAO BOUNDS

The signal we consider is made of K step edges, each step edge is determined by the three parameters a_i, θ_i, γ_i. We form a vector of the unknown parameters as follows:

$$\Theta = (a_0, a_1, ..., a_K, \theta_0, \theta_1, ..., \theta_K, \gamma_0, \gamma_1, ..., \gamma_K)^T.$$

We aim to retrieve Θ from the measured samples

$$\hat{y}_{n,m} = \langle f(x,y), \varphi(x/T - n, y/T - m)\rangle + \epsilon_{n,m}, \qquad n, m = 0, 1, ...N-1,$$

where $\epsilon_{n,m}$ is i.i.d. additive Gaussian noise with zero mean and variance σ^2. For simplicity we denote $\hat{y}_{n,m}$ as follows:

$$\hat{y}_{n,m} = f(\Theta, n, m) + \epsilon_{n,m}.$$

The performance of any unbiased estimator $\hat{\Theta}$ is lower bounded by the Cramer-Rao bound: $\text{var}(\hat{\Theta}) \geq I^{-1}(\theta)$, where $I(\Theta)$ is the Fisher Information Matrix (FIM) defined as $I(\Theta) = E\left(\nabla l(\Theta) \nabla l(\Theta)^T\right)$ and $l(\Theta)$ is the log-likelihood function.

First notice that $p_{\hat{y}}(\hat{y}_{n,m}|\Theta) = p_\epsilon(\hat{y}_{n,m} - f(\theta, n, m))$, where

$$p_\epsilon(\epsilon_{n,m}) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{\epsilon_{n,m}^2}{2\sigma^2}\right).$$

Hence, using independency of the noise samples we have:

$$l(\theta) = \ln P(\hat{y}_{0,0}, \hat{y}_{1,0}, ..., \hat{y}_{N-1,N-1}|\Theta) = \ln \prod_{n=0}^{N-1}\prod_{m=0}^{N-1} p_{\hat{y}}(\hat{y}_{n,m}|\Theta) = \sum_{n=0}^{N-1}\sum_{m=0}^{N-1} \ln p_\epsilon(\hat{y}_{n,m} - f(\Theta, n, m)).$$

Next, we compute the partial derivative of the log-likelihood with respect to the parameters θ_i. We obtain:

$$\frac{\partial l(\Theta)}{\partial \theta_i} = \frac{1}{\sigma^2}\sum_{n=0}^{N-1}\sum_{m=0}^{N-1} \epsilon_{n,m} \frac{\partial f(\Theta, m, n)}{\partial \theta_i}$$

$$\nabla l(\Theta) = \frac{1}{\sigma^2}\sum_{n=0}^{N-1}\sum_{m=0}^{N-1} \epsilon_{n,m} \nabla f(\Theta, n, m).$$

We can now determine the Fisher information matrix:

$$\begin{aligned} I(\Theta) &= E\left(\nabla l(\Theta)\nabla l(\Theta)^T\right) \\ &= E\left(\frac{1}{\sigma^4}\sum_n\sum_m\sum_l\sum_k \epsilon_{n,m}\epsilon_{l,k}\nabla f(\Theta,n,m)\nabla f(\Theta,l,k)^T\right) \\ &\overset{(a)}{=} \frac{1}{\sigma^4}\sum_n\sum_m\sum_l\sum_k E(\epsilon_{n,m}\epsilon_{l,k})\nabla f(\Theta,n,m)\nabla f(\Theta,l,k)^T \\ &\overset{(b)}{=} \frac{1}{\sigma^2}\sum_{n=0}^{N-1}\sum_{m=0}^{N-1}\nabla f(\Theta,n,m)\nabla f(\Theta,n,m)^T, \end{aligned}$$

where (a) follows from the linearity of the expectation and (b) from the fact that the noise is uncorrelated (independent). The Cramér-Rao bound is thus given by:

$$\text{CRB}(\Theta) = \sigma^2\left(\sum_{n=0}^{N-1}\sum_{m=0}^{N-1}\nabla f(\Theta,n,m)\nabla f(\Theta,n,m)^T\right)^{-1}.$$

The evaluation of the FIM is not straightforward, however, if we assume that the edges are sufficiently apart, then we can treat each edge independently. Moreover, since our main interest is the localization of the edge, we assume that the edge has fixed known amplitude $a_0 = 1$ so that the parametric space is reduced to $\Theta = (\theta, \gamma)$. We assume that the signal $f(x, y)$ is defined over the domain $[0, \tau] \times [0, \tau]$ with the sampling period $T = \tau/N$. The sampling kernel is $\varphi(x, y) = \beta_0(x)\beta_0(y)$, where $\beta_0(x) = 1$ for $x \in [0, 1]$. We thus obtain:

$$\begin{aligned} \frac{\partial f(\Theta,n,m)}{\partial\gamma} &= \frac{\partial}{\partial\gamma}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\beta_0(x/T-n)\beta_0(y/T-m)dxdy \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{\partial}{\partial\gamma}u(-x\sin\theta+y\cos\theta+\gamma\sin\theta)\beta_0(x/T-n)\beta_0(y/T-m)dxdy \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\sin\theta\,\delta(-x\sin\theta+y\cos\theta+\gamma\sin\theta)\beta_0(x/T-n)\beta_0(y/T-m)dxdy \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\sin\theta\,\delta(-t+y\cos\theta+\gamma\sin\theta)\beta_0(t/(T\sin\theta)-n)\beta_0(y/T-m)\frac{dt}{sin\theta}dy \\ &= \int_{-\infty}^{\infty}\beta_0\left(\frac{y\cos\theta+\gamma\sin\theta}{T\sin\theta}-n\right)\beta_0\left(\frac{y}{T}-m\right)dy \\ &= \int_{mT}^{(m+1)T}\beta_0\left(\frac{y}{T\tan\theta}+\frac{\gamma}{T}-n\right)dy \end{aligned}$$

and

$$\begin{aligned} \frac{\partial f(\Theta,n,m)}{\partial\theta} &= \frac{\partial}{\partial\theta}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\beta_0(x/T-n)\beta_0(y/T-m)dxdy \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{\partial}{\partial\theta}u(-x\sin\theta+y\cos\theta+\gamma\sin\theta)\beta_0(x/T-n)\beta_0(y/T-m)dxdy \\ &= \int\int(-x\cos\theta-y\sin\theta+\gamma\cos\theta)\delta(-x\sin\theta+y\cos\theta+\gamma\sin\theta)\beta_0(x/T-n)\beta_0(y/T-m)dxdy \\ &= \int\int(\frac{-t}{\tan\theta}-y\sin\theta+\gamma\cos\theta)\delta(-t+y\cos\theta+\gamma\sin\theta)\beta_0(t/(T\sin\theta)-n)\beta_0(y/T-m)\frac{dt}{\sin\theta}dy \end{aligned}$$

$$= \frac{1}{\sin\theta}\int_{-\infty}^{\infty}\left(-\frac{y\cos\theta+\gamma\sin\theta}{\tan\theta}-y\sin\theta+\gamma\cos\theta\right)\beta_0\left(\frac{y\cos\theta+\gamma\sin\theta}{T\sin\theta}-n\right)\beta_0\left(\frac{y}{T}-m\right)dy$$
$$= \int_{mT}^{(m+1)T}-\frac{y}{\sin^2\theta}\beta_0\left(\frac{y}{T\tan\theta}+\frac{\gamma}{T}-n\right)dy.$$

Given the above two equations, it is possible to evaluate (at least numerically) the Fisher information matrix for specific values of γ and θ. Disregarding the trivial cases $\theta = 0$ and $\theta = \pi/2$, an interesting scenario is when $\gamma = 0$ and $\theta = \pi/4$. In this case we have:

$$\frac{\partial f(\Theta,n,m)}{\partial\gamma} = \begin{cases} T & \text{for } n=0,1,...,N-1 \quad \text{and} \quad m=n. \\ 0 & \text{Otherwise} \end{cases}$$

and

$$\frac{\partial f(\Theta,n,m)}{\partial\theta} = \begin{cases} -(2n+1)T^2 & \text{for } n=0,1,...,N-1 \quad \text{and} \quad m=n. \\ 0 & \text{Otherwise} \end{cases}$$

This leads to the following Fisher information matrix:

$$I(\Theta)_{11} = \frac{1}{\sigma^2}\sum_{n=0}^{N-1}\left(\frac{\partial f}{\partial\theta}\right)^2 = \frac{T^4}{\sigma^2}\sum_{n=0}^{N-1}(2n+1)^2,$$
$$I(\Theta)_{22} = \frac{1}{\sigma^2}\sum_{n=0}^{N-1}\left(\frac{\partial f}{\partial\gamma}\right)^2 = \frac{NT^2}{\sigma^2},$$
$$I(\Theta)_{12} = I(\Theta)_{21} = \frac{1}{\sigma^2}\sum_{n=0}^{N-1}\frac{\partial f}{\partial\gamma}\frac{\partial f}{\partial\theta} = \frac{-N^2T^3}{\sigma^2}.$$

REFERENCES

[1] 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.

[2] 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.

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

[4] T. Blu and M. Unser. Quantitative Fourier analysis of approximation techniques: Part I - Interpolators and Projectors. *IEEE Trans. on Signal Processing*, 47(10):2783–2795, October 1999.

[5] L. G. Brown. A survey of image registration techniques. Survey, Dept of Computer Science, Columbia University, January 1992.

[6] P. Burns. Slant edge analysis tool SFRMAT 2.0. $http://www.i3a.org/downloads_iso_tools.html$, December 2007.

[7] D. Capel and A. Zisserman. Computer vision applied to super-resolution. *IEEE Signal Processing Magazine*, pages 75–86, May 2003.

[8] V. Chaisinthop and P. L. Dragotti. A new approach to distributed video coding using sampling of signals with finite rate of innovation. *Picture Coding Symposium (PCS)*, November 2007.

[9] 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.

[10] 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.

[11] M. Elad and A. Feuer. Super-resolution reconstruction of continuous image sequences. *IEEE Int. Conf. on Image Processing*, 3:459–463 vol.3, 1999.

[12] M. Elad and Y. Hel-Or. A fast super-resolution reconstruction algorithm for pure translational motion and common space invariant blur. *IEEE Transactions on Image Processing*, 10(8):1187–1193, 2001.

[13] 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.

[14] S. Farsiu, M. Elad, and P. Milanfar. A practical approach to superresolution. *Visual Communications and Image Processing*, 6077(1):607703, 2006.

[15] 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.

[16] 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.

[17] 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.

[18] 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.

[19] A. Goshtasby. *2-D and 3-D Image Registration,*. John Wiley and Sons, New York, Feb. 2005.

[20] C. Harris and M. Stephens. A combined corner and edge detector. In *4th Alvey Vision Conference*, pages 147–151, 1988.

[21] J. Heikkilä. Pattern matching with affine moment descriptors. *Pattern Recognition*, 37(9):1825–1834, March 2004.

[22] M. K. Hu. Visual pattern recognition by moments invariants. *IRE Trans. on Information Theory*, 8:179–187, 1962.

[23] M. Irani and S. Peleg. Improving resolution by image registration. *Computer Vision Graphics and Image Processing*, 53(3):231–239, 1991.

[24] R. Kakarak and H. A. O. On achievable accuracy in edge localization. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 20(3):19–20, May 2003.

[25] L. Kaufman. Maximum likelihood, least squares, and penalized least squares for PET. *IEEE Trans. on Medical Imaging*, 12(2):200–214, June 1993.

[26] 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/>.

[27] 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.

[28] L. Lucchese and G. M. Cortelazzo. A noise-robust frequency domain technique for estimating planar roto-translations. *IEEE Trans. on Signal Processing*, 48(6):1769–1786, June 2000.

[29] 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.

[30] 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.

[31] 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.

[32] S. C. Park, M. K. Park, and G. K. Moon. Super-resolution image reconstruction: A technical overview. *IEEE Signal Processing Magazine*, pages 21–36, May 2003.

[33] M. Peyrovian and A. Sawchuk. Image restoration by spline functions. *Applied Optics*, 17:660, 1978.

[34] J. Pluim, J. Maintz, and M. Viergever. Mutual-information-based registration of medical images: a survey. *IEEE Trans. on Medical Imaging*, 22(8):986–1004, August 2003.

[35] W. K. Pratt. Correlation techniques of image registration. *IEEE Trans. on Aerospace and Electronic Systems*, AES-10(3):353–358, May 1974.

[36] 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.

[37] C. Schmid, R. Mohr, and B. C. Evaluation of interest point detectors. *Int. Journal on Computer Vision*, 37(2):151–172, June 2000.

[38] 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.

[39] S. Smith and J. Brady. SUSAN – a new approach to low level image processing. *Int. J. on Computer Vision*, 23:45–78, 1997.

[40] J. Sprinzak and M. Werman. Affine point matching. *Pattern Recognition Letters*, 15:337–339, April 1994.

[41] P. Stoica and R. Moses. *Introduction to spectral analysis*. Prentice-Hall, 2000.

[42] 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.

[43] 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.

[44] 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.

[45] 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.

[46] 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.

[47] M. Xia and B. Liu. Image registration by super-curves. *IEEE Trans. on Image Processing*, 13(5):720–732, May 2004.

[48] 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.

[49] B. Zitova and J. Flusser. Image registration methods: a survey. *Image and Vision Computing*, 21:977–1000, October 2003.