DISTRIBUTED ACQUISITION AND IMAGE SUPER-RESOLUTION BASED ON CONTINUOUS MOMENTS FROM SAMPLES

Loïc Baboulaz and *Pier Luigi Dragotti*

Communications and Signal Processing Group, Electrical and Electronic Engineering,
Imperial College London, Exhibition Road, London SW7 2AZ, England.
Email: {L.Baboulaz, P.Dragotti}@imperial.ac.uk

ABSTRACT

Recently, new sampling schemes were presented for signals with finite rate of innovation (FRI) using sampling kernels reproducing polynomials or exponentials [1] [2].

In this paper, we extend those sampling schemes to a distributed acquisition architecture in which numerous and randomly located sensors are pointing to the same area of interest. We emphasize the importance played by moments and show how to acquire efficiently FRI signals with a set of sensors. More importantly, we also show that those sampling schemes can be used for accurate registration of affine transformed and low-resolution images. Based on this, a new super-resolution algorithm was developed and showed good preliminary results.

Index Terms— Moment methods, image registration, image resolution, image sampling, image reconstruction, spline functions, distributed algorithms.

1. INTRODUCTION

Recent developments in sampling theory focus on signals that are not necessarily bandlimited, like signals with finite rate of innovation (FRI) [3]. In [1], perfect reconstruction of FRI signals was demonstrated by using sampling kernels like B-splines or E-splines. Lately, these sampling schemes were extended to 2-D FRI signals [2].

At the same time, in many research areas, a common trend is to replace a single complex acquisition device with an acquisition system made of several cheap and simple devices. For instance, high rate A/D converters can be decomposed into several slow converters operating in parallel and whose synchronization is critical. This can also happen for high resolution digital imaging where a single high resolution camera is replaced by a large array of cheap, low-resolution cameras. A natural question that arises is to understand whether such a system can perform as well as the classical one.

In this paper we investigate the use of the new sampling schemes of [1] and [2] when the single acquisition device has been replaced by a system of multiple devices, each acquiring an undersampled version of the original signal. A particular emphasis is given to the case of 2-D signals and to the case of arrays of cameras. Moreover the scenario of interest to us is when the array contains hundreds of cameras of extremely low resolution. The contribution of the paper is two-fold: (a) for idealized signals and images, the results of [1] and [2] are extended to an acquisition system composed of several sensors; (b) a new super-resolution algorithm operating on real multiview images is presented using the results of (a).

In the next section, we present the sampling setup that models our acquisition system. We also recall important properties concerning the sampling kernels and the moment theory. In Section 3, we propose a distributed acquisition approach for FRI signals and we present simulation results for bilevel polygonal images. In Section 4, we show and illustrate how the previous results can be applied to image registration and to super-resolution of real images. Finally, we conclude in Section 5.

2. PRELIMINARIES AND PROBLEM SETUP

2.1. Image acquisition model

The signal entering the lens of a camera can be thought as the 2-D projection on the image plane of the observed 3-D scene limited to the field of view of the camera. In the classical case of a single acquisition device, see Figure 1(a), the incoming 2-D projection $f(x, y)$ of the 3-D scene is first filtered with a smoothing kernel $\varphi(x, y)$ modeling the point spread function of the lens of the camera. The blurred version $f(x, y) * \varphi(-x/T_x, -y/T_y)$ is then uniformly sampled to produce the set of samples $S_{m,n}$:

$$S_{m,n} = \langle f(x, y), \varphi(x/T_x - m, y/T_y - n) \rangle \qquad (1)$$

with $x, y \in \mathbb{R}$, $m, n \in \mathbb{Z}$, and where $T_x, T_y \in \mathbb{R}^+$ are the sampling periods along x and y respectively. Usually the sampled signal is an accurate discrete version of the original continuous signal (*e.g.* high-resolution image).

In a distributed acquisition system, N cameras P_i, $i = 0, \ldots, N - 1$, are observing the same 3-D scene from different unknown locations. Therefore the incoming 2-D projections $f_i(x, y)$ at each sensor will differ. By choosing a camera of reference (*e.g.* $i = 0$), we can model our distributed acquisition system as depicted in Figure 1(b). We assume that each projection $f_i(x, y)$ is the result of a transformation with $\mathcal{T}_i$ of the projection of reference $f(x, y)$. Examples of transformations $\mathcal{T}$ are translation, rotation or affine transformation according to the observed scene and to the locations of the cameras. Similarly to the single camera case, each sensor outputs a set of samples $S_{m,n}^{(i)}$:

$$S_{m,n}^{(i)} = \langle f_i(x, y), \varphi_i(x/T_x - m, y/T_y - n) \rangle$$

where $\varphi_i(x, y)$ is the sampling kernel of the sensor P_i. With a multiview system as in Figure 1(b), each camera usually has a low resolution and an accurate high-resolution images can be obtained from the set of cameras with a super-resolution algorithm.

2.2. Sampling kernels and theory of moments

Sampling kernels used in this research are 2-D B-splines $\beta_\rho(x, y)$ [4] and 1-D E-splines $\beta_{\underline{\alpha}}(t)$ [5]. A B-spline is characterized by its order ρ and constitutes a basis of the polynomial spline space of

(a) Classical acquisition device.

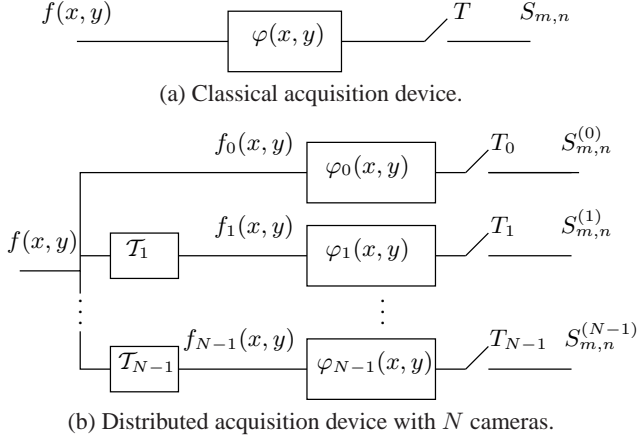(b) Distributed acquisition device with N cameras.

Fig. 1. Classical vs. distributed acquisition device

degree $\rho - 1$. Thus there exists a set of coefficients $\left\{c_{m,n}^{(p,q)}\right\}$ so that polynomial planes $x^p y^q$ can be reproduced:

$$\sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} c_{m,n}^{(p,q)} \beta_\rho (x - m, y - n) = x^p y^q, \qquad (2)$$

where $p, q = 0, \ldots, \rho - 1$.

The E-splines, or exponential B-spline, $\beta_{\underline{\alpha}}(t)$ are a natural extension of the B-splines [5]. An E-spline of order ρ is characterized by its parameters $\underline{\alpha} = (\alpha_0, \ldots, \alpha_\rho) \in \mathbb{C}^{\rho+1}$. They are obtained by successive convolutions of the zeroth order E-splines:

$$\beta_{\underline{\alpha}}(t) = \beta_{\alpha_0} * \ldots * \beta_{\alpha_\rho}(t), \quad \beta_\alpha(t) = \begin{cases} e^{\alpha t}, & 0 \le t < 1, \\ 0, & \text{otherwise.} \end{cases}$$

Notice that if $\underline{\alpha}$ is the null vector, then the E-spline becomes a B-spline. The E-splines also constitute a basis of the exponential spline space with parameter $\underline{\alpha}$. Therefore, similarly to B-spline, there exists a set of coefficients $\left\{c_m^{(\alpha_n)}\right\}$ which satisfies:

$$e^{\alpha_n t} = \sum_{m \in \mathbb{Z}} c_m^{(\alpha_n)} \beta_{\underline{\alpha}} (t - m), \qquad n = 0, \ldots, \rho. \qquad (3)$$

An interesting feature of B-splines is the possibility to retrieve the continuous geometric moments $m_{p,q}$, $p, q \in \mathbb{N}$, of the image $f(x, y)$ from the samples $S_{m,n}$ with a simple linear combination:

$$
\begin{aligned}
m_{p,q} &= \int \int f(x, y) x^p y^q dx dy \\
&\overset{(a)}{=} \int \int f(x, y) \sum_m \sum_n c_{m,n}^{(p,q)} \beta_\rho (x - m, y - n) \, dx dy \\
&= \sum_m \sum_n c_{m,n}^{(p,q)} \int \int f(x, y) \beta_\rho (x - m, y - n) \, dx dy \\
&\overset{(b)}{=} \sum_m \sum_n c_{m,n}^{(p,q)} S_{m,n} \qquad\qquad (4)
\end{aligned}
$$

where (a) and (b) refer respectively to Equations (2) and (1), and $T_x = T_y = 1$ for clarity only. Notice that the coefficients $c_{m,n}^{(p,q)}$ can be calculated and stored beforehand which makes the computation of the continuous moments very easy. Similarly to Equation (4), complex exponential moments h_n are obtained from the samples with E-splines and the coefficients from Equation (3).

Interestingly, some FRI signals are uniquely determined by a finite number of moments. For example, a bilevel polygonal image with K corners can be perfectly reconstructed from the knowledge of $2K - 2$ consecutive complex moments $\kappa_{p,0}$ [6]. In the work of [1] and [2] on sampling of FRI signals, perfect reconstruction is achieved by first computing the necessary moments from the samples with Equation (4). Then by combining those moments, a complete description of the original FRI signal is obtained. As a result, the notion of moments plays an important role in the analysis of FRI signals and we briefly present now the main definitions and important results of 2-D moments that are useful in our context. First we recall the definition of the geometric moments $m_{p,q}$ of order $(p + q)$, $p, q \in \mathbb{N}$ of a continuous function $f(x, y)$:

$$m_{p,q} = \int \int f(x, y) x^p y^q dx dy$$

The central moments $\mu_{p,q}$ are defined at the barycenter $(\overline{x}, \overline{y}) = \left(\frac{m_{1,0}}{m_{0,0}}, \frac{m_{0,1}}{m_{0,0}} \right)$ of $f(x, y)$:

$$
\begin{aligned}
\mu_{p,q} &= \int \int f(x, y) (x - \overline{x})^p (y - \overline{y})^q \, dx dy \\
&= \sum_{k=0}^p \sum_{l=0}^q \binom{p}{k} \binom{q}{l} (-\overline{x})^{p-k} (-\overline{y})^{q-l} m_{k,l}
\end{aligned}
$$

The complex moments $\kappa_{p,q}$ are defined on the complex image plane $z = x + jy, j = \sqrt{-1}$, [7]:

$$
\begin{aligned}
\kappa_{p,q} &= \int \int f(x, y) (x + jy)^p (x - jy)^q \, dx dy \\
&= \sum_{k=0}^p \sum_{l=0}^q \binom{p}{k} \binom{q}{l} j^{p-k+q+l} (-1)^{q-l} m_{k+l, p-k+q-l}
\end{aligned}
$$

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

3. DISTRIBUTED ACQUISITION OF FRI SIGNALS

In a distributed acquisition system, the locations of the sensors are unknown and we want to recover the original signal although each sensor can only retrieve a partial description of it. Consequently, the reconstruction of the observed signal at each sensor is only possible by considering the information of several sensors. Moreover, in order to interpret correctly the data of each sensor, it is first necessary to find the transformations existing between the observed images.

In this section, we are considering FRI signals and their perfect reconstruction from a finite number of continuous moments. We extend to a distributed approach the work of [1] and [2] based on E-splines and B-splines sampling kernels. Calculating all the necessary moments independently as if each sensor were a single acquisition device, would require a complex sampling kernel (*i.e.* of high order ρ) at each sensor. Moreover it would be sub-efficient as it would not take into account the fact that a similar scene is observed by each sensor. Therefore we propose instead to distribute the computation of the required moments among the N sensors so that each of them retrieve only a specific part of the observed scene. By choosing the order of the spline, we can determine which amount of information we want to retrieve from the samples. This allows us to use a simpler sampling kernel (*i.e.* of lower order ρ) at each sensor since fewer moments will be calculated. Once that we have determined the

transformations existing between each observations, it then becomes possible to recover the observed signal at each sensor's location. A major advantage of using simpler sampling kernels is to have less samples per sensor compared to a single acquisition device.

We are now using 2-D B-splines as sampling kernels. Suppose that we have a bilevel image representing a pentagon (five corners) and observed by three sensors. This polygonal image can be completely specified by eight consecutive complex moments [6]. As opposed to E-spline, a B-spline of a given order ρ is unique: if it can reproduce a monomial of degree k, it also reproduces, by construction, all the monomials of degree 0 to $k-1$. Consequently it is not possible to have a sensor providing the moment of order k without also providing moments of order 0 to $k-1$. It is therefore also not possible to distribute equally the computation of moments among the sensors. However we propose an asymmetric distributed acquisition architecture in which exists two types of cameras: a unique main camera and several auxiliary ones. In this setup, the main camera can directly determine all the required moments from its samples and reconstruct perfectly the observed signal similarly to the single acquisition device proposed in [2]. The auxiliary cameras though have very low resolution and only allow to retrieve the unknown transformations between the observed signals. It is therefore usually not possible to reconstruct the FRI signal by only considering the samples of the auxiliary cameras. However, after estimation of the transformation between an auxiliary camera and the main camera, it is possible to apply this transformation to the reconstructed image of the main sensor and to finally obtain the view at the location of this auxiliary camera.

In the case of the pentagon image and assuming an affine transformation between each view, we need a B-spline of order seven at the main camera and a B-spline of order three at each auxiliary cameras (see Section 4 concerning the registration procedure). Figure 2(a) shows the acquired low-resolution sampled images obtained at two auxiliary cameras (left and right) and at the main camera (middle). Figure 2(b) represents the reconstructed view at each corresponding camera after retrieval of the exact affine transformations.

In 1-D, different sensors can only observe different delayed versions of the same signal. By using 1-D E-spline, it is possible to design a symmetric distributed acquisition system. Indeed, since an E-spline is defined by its parameters $\underline{\alpha}$, different sampling kernels of same order ρ can be defined for each sensor. This way, different exponential moments can be obtained at each sensor. The delay $t_{i,j}$ between two sensors can also be retrieved if they share a common parameter $\alpha_n = \alpha$ since it can be shown that $h_n^{(i)} = e^{\alpha_n t_{i,j}} \cdot h_n^{(j)}$. We think that applications of 1-D FRI signals (*e.g.* stream of Diracs) can be useful in electronic (*e.g.* high rate A/D converter) [1].

4. REGISTRATION AND SUPER-RESOLUTION FROM MOMENTS

The estimation of the continuous moments from the samples as described by Equation (4) is not confined to FRI signals but can also be applied to real images. In particular, we propose here a novel approach of image registration based on the knowledge of the continuous moments of an object viewed from different locations. An extensive survey of image registration techniques is provided in [8].

As the resolution of the images to be registered gets lower, the efficiency of algorithms to retrieve salient features (such as corners, edges, etc...) and to determine accurately the transformations becomes poorer and poorer. So far, moments of the sampled image $\widetilde{m}_{p,q} = \sum_m \sum_n S_{m,n} (mT_x)^p (nT_y)^q$ were used in moment-based image analysis as discrete approximations of the original moments.
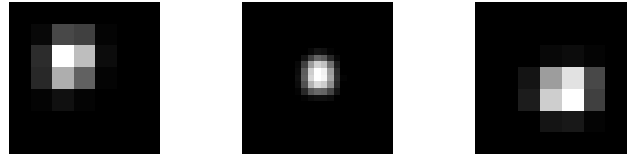
(a) Samples from each of the three cameras.

(b) Reconstructed image from the samples (512x512 px each).

Fig. 2. Asymmetric distributed acquisition with affine registration: left and right: auxiliary camera (7x7 px) - middle: main camera (23x23 px).
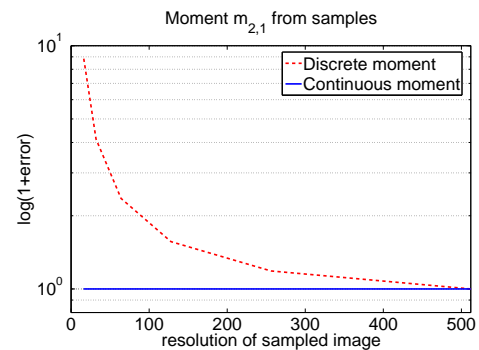

Fig. 3. Average normalized error of the discrete and continuous moment compared to the original moments ($p = 2, q = 1$).

For example, moment invariants have been used for finding correspondences of closed-boundaries whose barycenters were used as control points for registration [9]. This however requires sufficiently high-resolution sampled images so that shapes are well preserved. However, from Equation (4), it is actually possible to find precisely the original moments of any signal from its low-resolution sampled version. Figure 3 shows how the discrete and continuous moments computed from the samples differ from the moments of the original image (for the case of $p = 2$ and $q = 1$). We considered a set of 20 real square images (512x512 pixels) and their sampled versions at different resolution (256x256, 128x128, 64x64, 16x16 and 8x8 pixels). As the resolution decreases, the continuous moments obtained from Equation (4) remain very accurate whereas discrete moments diverge rapidly. The minimum number of required samples to compute the continuous moment depends on the order of the desired moments. The higher the order of the moments is, the larger the number of samples. But since B-splines are the splines with the shortest possible support, they suit particularly well in this framework. Finally, for order as low as 2, B-splines are very similar to a Gaussian pulse which makes them particularly suitable to model camera lenses.

We are assuming in this work that the disparity existing between any two images $f(\mathbf{x})$ and $g(\mathbf{x}')$ can be described by an affine transformation:

$$\mathbf{x}' = \mathbf{A}\mathbf{x} + \underline{t}$$

where $\mathbf{A}$ is a non-singular 2x2 matrix and $\underline{t}$ is 2-D vector. The translation $\underline{t}$ can be recovered by comparing the barycenters of each image. By using a whitening transform, the estimation of $\mathbf{A}$ is reduced
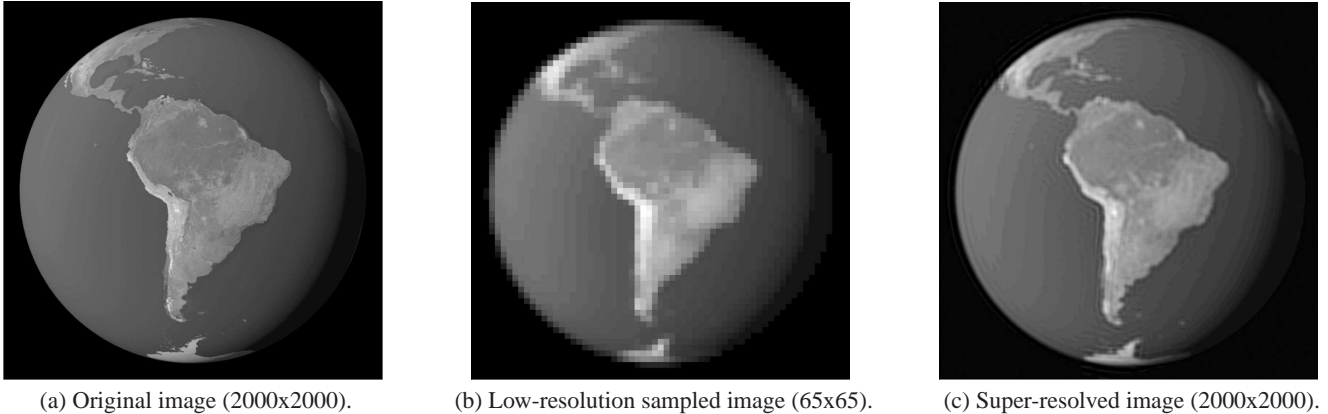
(a) Original image (2000x2000).　　　(b) Low-resolution sampled image (65x65).　　　(c) Super-resolved image (2000x2000).

Fig. 4. Image super-resolution from 100 sensors and based on continuous moments (original image: NASAs Earth Observatory).

to a simpler problem of finding a matrix of rotation $\mathbf{R}$ with a single unknown compared to four with $\mathbf{A}$ [10]:

$$\mathbf{A} = \mathbf{F}_g \mathbf{R} \mathbf{F}_f^{-1} \quad \text{with} \quad \mathbf{F}_{(\cdot)} = \begin{bmatrix} \sqrt{\mu_{2,0}^{(\cdot)}} & 0 \\ \frac{\mu_{1,1}^{(\cdot)}}{\sqrt{\mu_{2,0}^{(\cdot)}}} & \sqrt{\mu_{0,2}^{(\cdot)} - \frac{\mu_{1,1}^{(\cdot)2}}{\mu_{2,0}^{(\cdot)}}} \end{bmatrix}$$

where $(\cdot)$ refers either to the image g or f. In [11], Heikkilä demonstrated how to retrieve the matrix $\mathbf{R}$ from the third order complex moments. The use of the continuous moments, instead of the discrete moments, together with the approach described in [11], allows us to perform an affine registration of very low-resolution sampled images with the accuracy of the original image as it can be seen from Figure 2. An advantage of doing registration based on the moments is that features extraction and features correspondence are not required anymore to determine a transformation as opposed to intensity-, contour-, or point-based methods.

Image super-resolution consists in fusing adequately the information from different low-resolution and blurred images of a same scene in order to obtain a high-resolution image with greater details. A variety of methods have been proposed to solve this problem (see for example [12]). Our simulation results of image super-resolution using one hundred sensors are shown in Figure 4. We assumed that only 2-D translations were present between each sensors. We created one hundred (circularly) shifted views from an original high-resolution image of 2000 x2000 pixels (Figure 4(a)). Each view was sampled so that we obtained a set of low-resolution images of 65x65 pixels (Figure 4(b)). The different offsets were exactly estimated using continuous moments. A 2-D cubic interpolation followed by a restoration with a Wiener filter were then used to obtain a 2000x2000px super-resolved image (Figure 4(c)). Since 947 sensors would be required to obtain a similar amount of information as in the original image, there exists an inevitable discrepancy between the super-resolved image and the original image. However, many details are retrieved in the super-resolved image as it can be seen for example from the coastline of the south American continent.

5. CONCLUSION

We have shown that the sampling of FRI signals can be performed in a distributed fashion by using kernels reproducing either polynomials or exponentials. We also developed a new image registration algorithm by extending the computation of the continuous moments of FRI signals to real images. The registration, which does not require any feature extractions and correspondences, can retrieve up to

an affine transformation and its high accuracy is maintained almost regardless of the image resolution. Good preliminary results have finally been obtained for image super-resolution. In future works, we want to improve the restoration step of our super-resolution algorithm and use real-world data.

6. REFERENCES

[1] 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.

[2] P. Shukla and P.L. Dragotti. Sampling schemes for 2-D signals with finite rate of innovation using kernels that reproduce polynomials. In *Proc. IEEE ICIP*, Genoa, Italy, September 2005.

[3] 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.

[4] M. Unser. Splines: A perfect fit for signal and image processing. *IEEE Signal Processing Magazine*, 16(6):22–38, November 1999.

[5] M. Unser and T. Blu. Cardinal Exponential splines: Part I—Theory and Filtering Algorithms. *IEEE Trans. on Signal Processing*, 53(4):1425–1438, April 2005.

[6] 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.

[7] Y. Abu-Mostafa and D. Psaltis. Recognitive aspects of moments invariants. *IEEE Trans. on Pattern Anal. and Machine Intell.*, PAMI-6(6):698–706, November 1984.

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

[9] 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.

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

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

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