LOCAL FEATURE EXTRACTION FOR IMAGE SUPER-RESOLUTION

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

The problem of image super-resolution from a set of low resolution multiview images has recently received much attention and can be decomposed, at least conceptually, into two consecutive steps as: registration and restoration. The ability to accurately register the input images is key to the success and the quality of image super-resolution algorithms. Using recent results from the sampling theory for signals with Finite Rate of Innovation (FRI), we propose in this paper a new technique for subpixel extraction from low resolution images of local features like step edges and corners for image registration. By exploiting the knowledge of the sampling kernel, we are able to locate exactly the step edges on synthetic images. We also present results of full frame super-resolution of real low resolution images using our registration technique. We obtain super-resolved images with a much improved visual quality compared to using a standard local feature detection approach like a subpixel Harris corner detector.

Index Terms— Image super-resolution, Image edge analysis, Image registration, Spline functions, Image restoration.

1. INTRODUCTION

Many techniques in image processing rely on successful extractions of image features like edges, corners or ellipses. When a given scene is observed by several cameras situated at unknown locations, registering the different images as accurately as possible is often the very first task to complete and is critical to the output quality of any subsequent processing.

Algorithms for image super-resolution aim at constructing a single detailed image with a higher resolution after registration, fusion and restoration of a set of low resolution images taken at different unknown locations. Because it is an inverse ill-conditioned and sometimes underdetermined problem, the output quality of the restoration step depends heavily on the accuracy of the registration. Many works on image super-resolution concentrate on the restoration problem and often assume that input images are already accurately registered beforehand. Furthermore as pointed out in [1], the correct registration of images at a subpixel precision is the most crucial and difficult part of restoration problems like image super-resolution.

There exists various methods to register images. In a previous study, we showed how to obtain moments that can provide accurate registration of affine related low resolution images [2]. Another approach proposed in [1] jointly registers images and reconstructs a super-resolved image using Maximum Likelihood and Maximum A Posteriori rules. Their work focuses in particular on images related by translations. In [3], interest points like corners are extracted in each image to a sub-pixel accuracy with the Harris-Plessey corner detector. Corner correspondences are first obtained using a correlation matching algorithm and are then refined with a RANSAC algorithm by estimating the transformation that fits the largest number of corresponding corners. Using local features instead of global features allows to estimate more complicated transformations between images (*e.g.* projective transformation). In [3], the set of images to register was captured by a digital video camera with the usual resolution. This allows a large number of feature points (several hundreds) to be extracted which, by an averaging effect, allows to register accurately the images. Their super-resolution results are then obtained for regions of interest of 40x25 pixels where fusion and restoration are accomplished with either a Maximum Likelihood estimation or a Maximum A Posteriori estimation.

The underlying goal of this study is to carry out "full frame" image super-resolution where the only available images have very low resolution (*e.g.* of size 64x64 pixels) and where the entire observed view is super-resolved. As opposed to [2] where we retrieved global features, here we want to retrieve very accurately local features in images of low-resolution so that a precise registration is obtained. As the resolution decreases, the local two-dimensional structure of an image degrades and two-dimensional features like corners get more difficult to track and to locate. However, since only one coordinate is sufficient to define its location (like in $y = f(x)$), an edge is mainly one-dimensional and its structure is therefore more resistant to downsampling. We are thus interested in this research in retrieving the parameters of step edges and in inferring the location of possible corners at edge intersections for registration purposes. To achieve this efficiently at very low resolution, we assume to know the point-spread function (the optical blur) of the lens that acquired the images. Note that this assumption is not needed in the work of [3] and [1]. The novelty of this paper is first of all to show that it is possible to retrieve from the samples the exact parameters of simple edges and this is demonstrated on synthetic images. Secondly we show that this approach can be extended to real images with application to image registration. We finally provide results of full frame image super-resolution of a real scene using the proposed image registration approach.

The next section introduces the model of our sampling setup for image acquisition as well as the model of the considered type of edges. In Section 3, we show how a continuous edge can be exactly retrieved from a low resolution sampled image by using results from the sampling theory of signals with Finite Rate of Innovation (FRI). Section 4 shows applications of our subpixel edge location algorithm for image registration and image super-resolution. We conclude the discussion in Section 5.
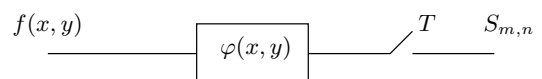

Fig. 1. Image acquisition model

2. PROBLEM SETUP

2.1. Image acquisition model

In Figure 1, we model a digital camera. The incoming signal entering its lens is the continuous high resolution irradiance light-field $f(x, y)$ limited to the field of view of the camera and parallel to the image plane. When going through the lens, this signal is filtered with the point spread function $\varphi(x, y)$ of the lens which makes it blurred. The smoothed version $f(x, y) * \varphi(-x/T_x, -y/T_y)$ is then uniformly sampled at the sensor of the image plane to produce the set of image samples $S_{m,n}$:

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

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. The standard basis for $\mathbb{R}^2$ centered in the middle of the image plane is chosen for x and y, while m and n refer respectively to the columns and the rows of the discrete image.

2.2. Sampling kernel

The function $\varphi(x, y)$ is thereafter referred to as the *sampling kernel* and is assumed known. The optical blur introduced by a lens is often modeled in the literature by a Gaussian function whose support is infinite. In our camera model, we consider the sampling kernel to be a B-spline function $\beta^\rho(x, y)$ of a order ρ and of compact support. Motivations are that B-splines are very similar to a Gaussian function for order as low as 2 and the B-spline of infinite order is equal to a Gaussian function.

The family of B-spline functions enjoys many interesting properties. First, a B-spline is uniquely characterized by its order ρ and is obtained from successive convolutions of the box B-spline [4]:

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

The 2-D B-spline $\beta^\rho(x, y)$ is obtained by tensor product of two 1-D B-splines $\beta^\rho(x)$ and $\beta^\rho(y)$ and is therefore a variable separable function.

In addition, B-spline functions form a Riesz basis and satisfy Strang-Fix conditions [5]:

$$\beta^\rho(x) \Leftrightarrow \mathbf{B}^\rho(f) = [\text{sinc}(f)]^{\rho+1} \rightarrow \begin{cases} \mathbf{B}^\rho(0) \neq 0 \\ D^j \mathbf{B}^\rho(2\pi k) = 0. \end{cases}$$

where $k \in \mathbb{Z}$, $j = 0, 1, \ldots, \rho$, $\mathbf{B}^\rho(f)$ is the Fourier transform of $\beta^\rho(x)$ and D^j is the differential operator. Strang-Fix conditions ensure that a linear combination of B-splines are reproducing polynomials up to degree ρ. Thus there exists a set of coefficients $\left\{ c_m^{(p)} \right\}$ such that:

$$\sum_{m \in \mathbb{Z}} c_m^{(p)} \beta^\rho(x - m) = x^p, \tag{2}$$

where

$$c_m^{(p)} = \langle x^p, \widetilde{\beta}^\rho(x - m) \rangle$$

with $p = 0, \ldots, \rho$. The function $\widetilde{\beta}^\rho(x)$ is the dual B-spline associated with $\beta^\rho(x)$ [4]:

$$\widetilde{\beta}^\rho(x) = \sum_{k=-\infty}^{\infty} (b^{2\rho+1})^{-1}[k] \cdot \beta^\rho(x - k)$$

where $b^\rho[k]$ is the discrete B-spline of order ρ:

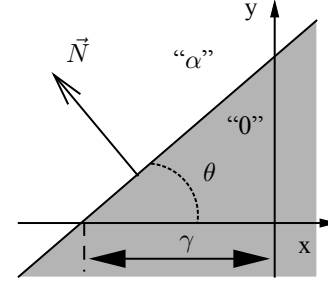$$b^\rho[k] = \beta^\rho(x)\big|_{x=k}$$

Fig. 2. Step edge model

2.3. Step edge model

There exists a variety of types of edges that can be found on natural images. Examples of edges are roof edges, bar edges or step edges. We focus on ideal straight step edges as shown in Figure 2. This type of edge can be described by three parameters, namely its amplitude α, its orientation θ with respect to the x-axis and its offset γ on the x-axis. Let $\vec{N}$ be the unit vector normal to the orientation of the edge and let define $\vec{d}$ for any point (x, y) in $\mathbb{R}^2$:

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

A straight continuous step edge $h(x, y)$ can be mathematically expressed using the Heaviside function U as:

$$h(x, y) = \alpha U\left(\vec{d} \cdot \vec{N} \right) \tag{3}$$

Although such function is not bandlimited because it exhibits a singularity, it has a finite rate of innovation (FRI) as three parameters completely describe it.

3. EDGE LOCATION

Edge location refers to the problem of finding as accurately as possible the parameters of a continuous edge from its digital representation in an image. This problem of finding edge parameters is therefore different from the problem of edge detection where a binary solution is looked for showing whether or not an edge is present at a pixel precision (*e.g.* with a Canny edge detector). Moment-based approaches have been widely used for subpixel edge location, see *e.g.* [6] for a recent study. In [7], an overcomplete set of step functions called wedgelets was designed to find edges at subpixel accuracy. The sampling theory of FRI signals (see [8]) aims at reconstructing perfectly from the samples signals which are not bandlimited but which present a finite number of degree of freedom per unit of time. As seen in Section 2.3, straight step edges belong to the class of FRI signals. We therefore propose to apply recent developments in this theory to the problem of edge parameters extraction.
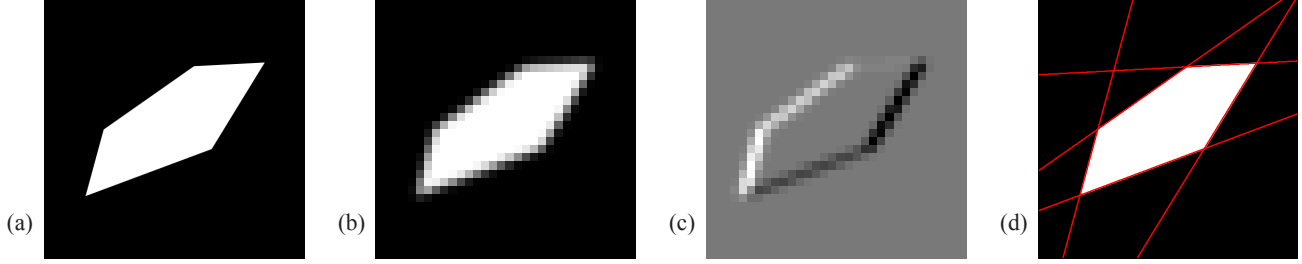
Let $h(x, y)$ be the function describing an ideal step edge as in Equation (3) and consider its sampled version $h_{m,n}$ obtained as in Equation (1):

$$h_{m,n} = \langle h(x, y), \varphi(x - m, y - n) \rangle$$

We now filter the set $h_{m,n}$ with a finite difference operator and get a new set of differentiated samples denoted $d_{m,n}$:

$$d_{m,n} = h_{m+1,n} - h_{m,n} \tag{4}$$

The discrete differentiation is intimately related to the derivation in the continuous domain with the introduction of a modified sampling

Fig. 3. (a) original image before acquisition (512x512 pixels); (b) acquired samples $h_{m,n}$ (32x32 pixels); (c) differentiated samples $d_{m,n}$; (d) retrieved edges plotted on the original image.

kernel[8]. It follows that:

$$d_{m,n} = \langle \frac{dh(x,y)}{dx}, \varphi(x-m, y-n) * \beta^0(x-m) \rangle$$

Moreover we have:

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

where δ is the Dirac function. Since the sampling kernel $\varphi(x,y)$ is a 2-D B-spline of order ρ, it follows that:

$$d_{m,n} = -\alpha sin\theta \cdot \langle \delta\left(\vec{d} \cdot \vec{N}\right), \beta^{\rho+1}(x-m)\beta^\rho(y-n) \rangle$$

Let $\tau_n^{(p)}$ be the linear combination of samples $d_{m,n}$ along the row n only with the coefficients c_m^p from Equation (2):

$$\tau_n^{(p)} = \sum_m c_m^p d_{m,n} \qquad (5)$$

It can then be proved that:

$$\tau_n^{(0)} = -\alpha \quad \text{and} \quad \tau_n^{(1)} = -\alpha\left(\frac{2n+1}{2tan\theta} + \gamma\right)$$

Thus by considering two consecutive rows and computing $\tau_n^{(0)}$, $\tau_n^{(1)}$ and $\tau_{n+1}^{(1)}$, the parameters of the step edge can be retrieved exactly:

$$\alpha = -\tau_n^{(0)}, \quad tan\theta = \frac{-\alpha}{\tau_{n+1}^{(1)} - \tau_n^{(1)}}, \quad \gamma = -\frac{\tau_n^{(1)}}{\alpha} - \frac{2n+1}{2tan\theta}.$$
$$(6)$$

Based on this model, we implemented a local algorithm to extract edges with a very high accuracy in low-resolution images. The pseudo-code is given in Algorithm 1. The measure of similarity of two edges is based on the similarity of their parameters α, $tan\theta$ and γ. For better results, Algorithm 1 is run on the sampled image as well as on the transposed sampled image and both results are merged and averaged again.

Figure 3 shows the results of the edge location on a synthetic image where edge parameters were exactly retrieved. The "continuous" observed scene is shown in Figure 3(a) and its sampled version is in Figure 3(b). Figure 3(c) shows the differentiated samples $d_{m,n}$ and Figure 3(d) shows the retrieved edges plotted on the original scene for a visual appreciation of the proposed method.

4. IMAGE REGISTRATION AND SUPER-RESOLUTION

4.1. Image registration

Real-world images cannot obviously be considered as a FRI signal. However, locally, the images can exhibit a FRI structure that can be retrieved precisely using the the sampling theory of FRI signals. These local FRI regions can, in particular, represent important features in image registration, such as corners or edges. Retrieving these features with the highest possible accuracy can therefore improve greatly the efficiency of registration algorithms and subse-

---

Algorithm 1 Subpixel Edge Location
───────────────────────────────
Define $weight = 1$,
Compute the differentiated samples using Equation (4),
Run a Canny-like edge detector on sampled image,
for all position (i,j) detected as an edge **do**
    find the differentiated samples in the neighborhood of (i,j),
    calculate $\tau_j^{(0)}$, $\tau_j^{(1)}$ and $\tau_{j+1}^{(1)}$ using Equation (5),
    calculate α, $tan\theta$ and γ using Equations (6),
    store $[\alpha \quad tan\theta \quad \gamma \quad weight]$ as a candidate edge
end for
while there exists similar edges **do**
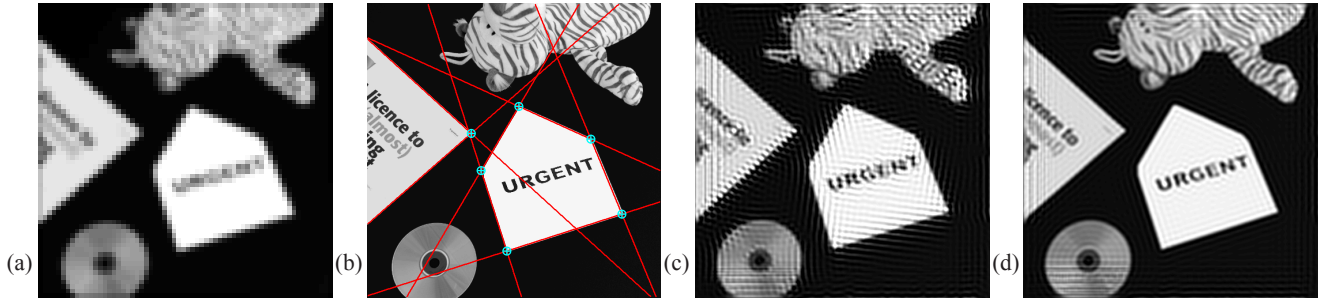   Merge similar edges *i.e.* average $[\alpha \quad tan\theta \quad \gamma]$ and add *weight* together,
end while
Discard edges having a too small weight.
───────────────────────────────

quent processing (like image super-resolution). In [9], a local algorithm for corner detection was proposed based on directional derivatives and the sampling theory of FRI signals. Although exact corner location was obtained theoretically, the method was not practically satisfying due to the complicated nature of the modified sampling kernel. Our proposed approach requires a single horizontal differentiation which makes the modified sampling kernel easier to compute. Besides the retrieved parameters are averaged along edges to gain in stability and robustness. Edge intersections are used as feature points. Only intersections close to the location where edges were estimated are considered as corner points. In addition to its position, a corner can also be described by the angle of its two generative edges.

Once the feature points are extracted, we use a similar approach to [3] to find corner correspondences (inliers) across images: putative correspondences are obtained using correlation matching of the samples in the neighborhood of each corners. Given this correspondence, a first estimation of the transformations between images is calculated. Note that correlation matching methods are not efficient for images with low resolutions and the transformation estimation is often not good. However corner correspondences and transformation estimation can then be jointly refined using the RANSAC algorithm. The RANSAC algorithm operates only on the extracted feature points and therefore does not depend on the resolution of the available images. Thus, the lack of precision in the feature locations can be overcome by the number of extracted features so that accurate image transformations can be computed (as in [3]). However, because we are dealing with low-resolution images, only a small number of feature points can be extracted, it is therefore of utmost importance to know their precise subpixel location to register as accurately as possible the set of images.

Fig. 4. (a) Acquired low resolution image (64x64); (b) Retrieved edges and corners (⊕) plotted on high resolution image; (c) Super-resolved image using Harris corner detector (512x512); (d) Our result: super-resolved image using features extracted with Algorithm 1 (512x512).

4.2. Image super-resolution

To assess the accuracy of the proposed feature extraction method, we implemented an image super-resolution algorithm using the proposed image registration approach. We acquired a high resolution image of a simple scene presenting step edges, text and textures. From this image, twenty shifted images were created by cropping an area of 512x512 pixels at different randomly-generated locations. Each of these images was then downsampled by a factor 8 to 64x64 pixels with a 2-D quadratic B-spline. This new set of low resolution images constitutes the input of our image super-resolution algorithm. The super-resolved image has size 512x512 pixels (zoom factor = 8). Note that with twenty images, the system is underdetermined.

In order to compare our results, we implemented a simplified image super-resolution algorithm inspired by [3] using the Matlab functions provided by [10]. Corners were extracted using a traditional subpixel Harris corner detector and their correspondences obtained using correlation matching and a RANSAC algorithm. Due to the low resolution, around twenty feature points only were extracted on each image. During the simulations, no corner correspondence error were observed that would have led to bad registration results. For both approaches, differences between corresponding corners were calculated and averaged to give the estimated subpixel translation between images. We then applied a bicubic interpolation procedure to estimate missing samples on the high resolution grid. For the image restoration step, we chose to use the Wiener filter as the deconvolution procedure for both case as well. The choice of using a Wiener filter was dictated by the simplicity and the efficacy of this method as it has also been reported in the recent experiments of [1].

Figure 4 shows the obtained results. Figure 4(a) presents the acquired low resolution image used as reference (64x64 pixels). Figure 4(b) shows the edges and corners extracted with Algorithm 1 and plotted on the high resolution image. Figure 4(c) is the super-resolved image using the subpixel Harris corner detector approach (PSNR = 14.95dB). Figure 4(d) shows our result of image super-resolution using the Algorithm 1 for feature extraction (PSNR = 15.96dB). This latter image presents a much better visual quality than Figure 4(c). With our approach, artifacts are visible essentially on the border of the super-resolved image. This is due to the fact that the low resolution images are all overlapping with the center of the scene (the envelope). Thus, after registration, there exists a higher density of samples in the middle than at the border of the image (*e.g.* the CD) where artifacts appear after restoration. Exact registration was achieved using Algorithm 1 whereas the average registration error with the Harris corner detector is 0.0923 pixels. Although this error seems relatively small, the reconstruction quality is severely degraded given the same restoration procedure. This shows how much accurate registration schemes are essential and needed for obtaining good image super-resolution results.

5. CONCLUSION

In this paper, we developed new theoretical results for step edge location from a sampled image using the sampling theory of signals with finite rate of innovation. We also implemented and applied these new results to real images and proposed a novel approach for image registration when step edges occur. This new registration approach was used for full frame image super-resolution and the experiments showed that exact registration was possible by using the proposed approach. We finally obtained super-resolved images of a better quality when compared with a similar image super-resolution algorithm based on the Harris corner detector for registration. Further works should investigate the possibility of extracting different features having a finite number of parameters, and analyze the robustness of the method to the noise level.

6. REFERENCES

[1] 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.

[2] L. Baboulaz and P.L. Dragotti. Distributed acquisition and image super-resolution based on continuous moments from samples. *Proc. IEEE ICIP*, pages 3309–3312, 2006.

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

[4] M. Unser, A. Aldroubi, and M. Eden. B-Spline signal processing: Part I—Theory. *IEEE Trans. on Signal Processing*, 41(2):821–833, February 1993.

[5] 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.

[6] I. Popovici and D. Withers. Locating thin lines and roof edges by custom-built moments. *Proc. IEEE ICIP*, pages 753–756, 2006.

[7] D. Donoho. Wedgelets: Nearly minimax estimation of edges. *Annals of Statistics*, 27(3):859–897, 1999.

[8] 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*, July 2006.

[9] P. Shukla and P.L. Dragotti. Sampling schemes for multidimensional signals with finite rate of innovation. *IEEE Trans. on Signal Processing*, 2006.

[10] P. Kovesi. Peter's functions for computer vision. *Available at http://www.csse.uwa.edu.au/~pk/research/matlabfns/index.html*, January 2007.