

SUBSPACE-BASED METHODS FOR IMAGE REGISTRATION AND SUPER-RESOLUTION

Patrick Vandewalle¹, Loïc Baboulaz², Pier Luigi Dragotti², and Martin Vetterli^{3,4}

¹Philips Research, Eindhoven, The Netherlands

²Electrical and Electronic Engineering Department, Imperial College London, UK

³School of Computer and Communication Sciences,

Ecole Polytechnique Fédérale de Lausanne, Switzerland

⁴Department of Electrical Engineering and Computer Science, UC Berkeley, USA

ABSTRACT

Super-resolution algorithms combine multiple low resolution images into a single high resolution image. They have received a lot of attention recently in various application domains such as HDTV, satellite imaging, and video surveillance. These techniques take advantage of the aliasing present in the input images to reconstruct high frequency information of the resulting image. One of the major challenges in such algorithms is a good alignment of the input images: subpixel precision is required to enable accurate reconstruction. In this paper, we give an overview of some subspace techniques that address this problem. We first formulate super-resolution in a multichannel sampling framework with unknown offsets. Then, we present three registration methods: one approach using ideas from variable projections, one using a Fourier description of the aliased signals, and one using a spline description of the sampling kernel. The performance of the algorithms is evaluated in numerical simulations.

Index Terms— Image registration, Image resolution, Image restoration, Spectral analysis, Spline functions.

1. INTRODUCTION

A high resolution input image is required in many imaging applications: recognition of people or license plates, satellite imaging, viewing video material on HD displays, etc. Unfortunately, capturing devices with such a high resolution are often still prohibitively expensive. In some cases, such as satellite imaging, it can also be practically impossible to replace a camera by a higher resolution version when it becomes available. The wide adoption of HD TVs generates an increasing demand for the conversion of legacy SD video material (or even freshly shot camera phone images and movies) to HD resolution. This is exactly the goal of super-resolution algorithms: to generate a high resolution image from a set of low resolution images. Typically, the input images need to be aliased. If multiple images are then taken with small relative motion, the aliasing information from the different images can be used to reconstruct the high frequency part of the high resolution image.

The idea of super-resolution imaging was first introduced by Tsai and Huang in 1984 [4]. They used a frequency domain algorithm to minimize the energy outside the known frequency range of the images. In the past fifteen years, a vast number of algorithms has been presented. Good overviews are given in the special issues on this topic in the IEEE Signal Processing Magazine [5]

Patrick Vandewalle performed his work while at EPFL. Parts of this work were presented earlier in [1, 2, 3].

and the EURASIP Journal of Applied Signal Processing [6]. Super-resolution reconstruction is an ill-posed problem. It is usually turned into a well-posed problem using regularization or an image model. We will consider the second case, where we assume our image (after sampling) to be in the Fourier or spline space. Most algorithms treat image registration (aligning the input images) and reconstruction (combining the aligned images to a single high resolution image) as two separate problems. A standard image registration algorithm is often used for the alignment, and emphasis is put on the reconstruction part of the algorithm [7, 8]. Recently, new algorithms have been developed that specifically consider registration methods for (aliased) input images, and are successfully integrated in the super-resolution algorithm [1, 2, 3, 9, 10]. In this paper, we present three of those registration algorithms in a common framework based on multichannel sampling.

In the next section, we will set up this framework by describing super-resolution as a multichannel sampling problem with unknown offsets. We will then use this setup to describe three algorithms that solve this problem in Section 3: using projections, Fourier analysis, and spline analysis. Simulation results of each method are presented in Section 4, and Section 5 gives some concluding remarks.

2. SUPER-RESOLUTION AS A MULTICHANNEL SAMPLING PROBLEM

We will now analyze the super-resolution setup mathematically. While this description is given in 1D for simplicity, it can straightforwardly be extended to 2D. Let $f(t)$ be a continuous-domain signal in an L -dimensional Hilbert space \mathcal{H} with basis $\{\varphi_l(t)\}_{l=0..L-1}$:

$$f(t) = \sum_{l=0}^{L-1} \alpha_l \varphi_l(t), \quad (1)$$

where α_l is the expansion coefficient corresponding to the l -th basis function. Examples of such signals are truncated Fourier series, wavelets, splines, etc. Without loss of generality, we will only consider the signal on the interval $[0, 1]$ here ($0 \leq t < 1$).

We now take a first set of N uniformly distributed samples (corresponding to the first image) with sampling rate $1/N$ and sampling kernel $\psi(t)$:

$$y_0(n) = \langle f(t), \psi\left(t - \frac{n}{N}\right) \rangle = \sum_{l=0}^{L-1} \alpha_l \langle \varphi_l(t), \psi\left(t - \frac{n}{N}\right) \rangle. \quad (2)$$

Next, we take $K - 1$ other sets of samples (images), with offsets

$\{t_k\}_{k=1..K-1}$ with respect to the first set ($t_0 = 0$):

$$y_k(n) = \langle f(t), \psi\left(t - t_k - \frac{n}{N}\right) \rangle$$

$$= \sum_{l=0}^{L-1} \alpha_l \langle \varphi_l(t), \psi\left(t - t_k - \frac{n}{N}\right) \rangle. \quad (3)$$

The offsets t_k can take arbitrary real values between 0 and 1, and are supposed to have a part that is not a multiple of $1/N$ (subpixel shifts). We can rewrite this in vector notation as

$$\mathbf{y}_k = \Phi_{t_k} \boldsymbol{\alpha}, \quad (4)$$

where \mathbf{y}_k and $\boldsymbol{\alpha}$ are column vectors containing the samples $y_k(n)$ and the expansion coefficients α_l , respectively. The $K \times N$ matrix Φ_{t_k} consists of the basis functions $\varphi_l(t)$ sampled uniformly with a sampling kernel $\psi(t)$, and unknown offset t_k . The K sets of samples can be combined in a single vector, resulting in

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_{K-1} \end{pmatrix} = \begin{pmatrix} \Phi_{t_0} \\ \Phi_{t_1} \\ \vdots \\ \Phi_{t_{K-1}} \end{pmatrix} \boldsymbol{\alpha} = \Phi_{\mathbf{t}} \boldsymbol{\alpha}. \quad (5)$$

This results in K sets of N uniform samples with unknown offsets between the sets of samples, just like in super-resolution imaging, where one takes K pictures of N pixels each, with unknown camera motion between the images (horizontal and vertical shifts or more complex motion). It can be shown that if the total number of samples KN is larger than the total number of unknowns $L + K - 1$ (expansion coefficients and offsets), the problem is well-defined, and has a single solution [11]. We will assume here that this is the case (except in Section 3.3, where no special assumptions are made).

From (5), we can see that this problem is linear in the unknown expansion coefficients α_l , but non-linear in the unknown offsets t_k . Once the offsets are known, we can easily reconstruct the signal coefficients by solving a set of linear equations. We will therefore concentrate in this paper on the computation of the unknown offsets: image registration.

3. SUBSPACE-BASED IMAGE REGISTRATION

In this section, we present a set of algorithms based on a subspace analysis of super-resolution imaging. First, we present a projection-based algorithm, followed by a Fourier-based and a spline-based algorithm. In the first two algorithms, we assume the sampling kernel is a Dirac (or can be factored out after reconstruction), and place restrictions on the (sampled) signal model. The third method uses splines to explicitly model the sampling kernel, and does not restrict the signal model itself.

3.1. Projection-Based Algorithm

From (5), it is clear that our set of images \mathbf{y} belongs to the subspace spanned by the sampled basis functions $\Phi_{\mathbf{t}}$. In practice, the exact sampling points of those basis functions are unknown, due to the unknown shifts. If we take an arbitrary set of shift values $\hat{\mathbf{t}} \neq \mathbf{t}$ and the corresponding subspace $\Phi_{\hat{\mathbf{t}}}$, it does not contain the set of images \mathbf{y} (taken with shifts \mathbf{t}). We can therefore compute the shift values by minimizing the difference between the sample vector \mathbf{y} and its projection $\mathbf{P}_{\Phi_{\hat{\mathbf{t}}}} \mathbf{y}$ onto the subspace $\Phi_{\hat{\mathbf{t}}}$:

$$\mathbf{t} = \arg \min_{\hat{\mathbf{t}}} \|\mathbf{y} - \mathbf{P}_{\Phi_{\hat{\mathbf{t}}}} \mathbf{y}\|. \quad (6)$$

As discussed in Section 2, once the shifts are known, the signal coefficients (and thus a high resolution image) can be reconstructed by solving the set of linear equations from (5). We compute a least squares solution for increased robustness against noise. Note that this algorithm uses ideas similar to separable nonlinear least squares [12] (which was also used with different descriptions in [9, 10]).

3.2. Fourier-Based Algorithm

Let us now assume $f(t)$ can be expressed in a Fourier basis. We can then rewrite (4) as

$$\mathbf{y}_k = \mathbf{F}^* \mathbf{D}_{t_k} \boldsymbol{\alpha}, \quad (7)$$

with \mathbf{F}^* an $N \times L$ inverse Discrete Fourier Transform (DFT) matrix, and \mathbf{D}_{t_k} an $L \times L$ diagonal matrix with elements $\mathbf{D}_{t_k}(l, l) = e^{j2\pi t_k l/N}$. The notation \mathbf{F}^* is used for the Hermitian transpose of the forward transform matrix \mathbf{F} . Note that due to the undersampling, \mathbf{F}^* is an extension of a square $N \times N$ inverse DFT matrix \mathbf{F}_N^* where some columns are repeated, and not a submatrix of an $L \times L$ matrix \mathbf{F}_L^* .

The Fourier transform of (7) can then be written as

$$\mathbf{Y}_k = \mathbf{F}_N \mathbf{y}_k = \mathbf{F}_N \mathbf{F}^* \mathbf{D}_{t_k} \boldsymbol{\alpha}. \quad (8)$$

The vector \mathbf{Y}_k is a phase shifted and aliased version of the original Fourier coefficient vector $\boldsymbol{\alpha}$. If we assume the length L of $\boldsymbol{\alpha}$ is a multiple of N (otherwise we can always add zeros to $\boldsymbol{\alpha}$ to make it a multiple), $\boldsymbol{\alpha}$ can be split in $S = L/N$ blocks $\boldsymbol{\alpha}_i$ of length N . We can then rewrite the Fourier transform as

$$\mathbf{Y}_k = \mathbf{F}_N \left(\mathbf{F}_N^* \mathbf{F}_N^* \cdots \mathbf{F}_N^* \right) \mathbf{D}_{t_k} \boldsymbol{\alpha}$$

$$= \left(\mathbf{I} \mathbf{I} \cdots \mathbf{I} \right) \mathbf{D}_{t_k} \boldsymbol{\alpha} \quad (9)$$

$$= \mathbf{D}'_{t_k} \sum_i e^{j2\pi t_k i} \boldsymbol{\alpha}_i,$$

with \mathbf{D}'_{t_k} the $N \times N$ central part of the matrix \mathbf{D}_{t_k} . The subvectors $\boldsymbol{\alpha}_i$ represent the overlapping parts of the spectrum due to undersampling, and the sum is therefore over each of those S overlapping parts of the spectrum.

From (9), we can see that for each set of samples, the vector $\mathbf{D}'_{t_k} \mathbf{Y}_k$ belongs to the same S -dimensional subspace spanned by the vectors $\boldsymbol{\alpha}_i$. As the diagonal matrix \mathbf{D}'_{t_k} depends on t_k , we can therefore compute the offsets $\{t_k\}$ as the values for which

$$\text{rank} \left(\mathbf{Y}_0 \quad \mathbf{D}'_{t_1} \mathbf{Y}_1 \quad \cdots \quad \mathbf{D}'_{t_{K-1}} \mathbf{Y}_{K-1} \right) = S, \quad (10)$$

where we require $K > S = \lceil L/N \rceil$.

3.3. Spline-Based Algorithm

In many situations the sampling kernel $\psi(t)$ can be estimated or is known a-priori and this knowledge can be used to devise more effective registration techniques.

In the idealized acquisition model of digital cameras, the samples are related to the original view through the point spread function (PSF), which fundamentally models the blur due to the camera lenses and the sensor structure. The PSF of a camera is normally modeled with a Gaussian pulse, however, in this work, we assume that the PSF can be approximated with a B-spline function. We make this choice for two main reasons: first, B-splines of sufficiently high order have a shape which is very close to that of a Gaussian function.

Second, B-splines possess the polynomial reproduction property we can take advantage of in the registration step.

For those not familiar with polynomial splines, let us just recall that a B-spline $\beta_L(t)$ of order L is obtained from the $(L+1)$ -fold convolution of the box function $\beta_0(t)$, that is:

$$\beta_L(t) = \underbrace{\beta_0(t) * \beta_0(t) \dots * \beta_0(t)}_{L+1 \text{ times}}, \quad \text{with } \hat{\beta}_0(\omega) = \frac{1 - e^{-j\omega}}{j\omega}.$$

Here, $\hat{\beta}_0(\omega)$ denotes the Fourier transform of $\beta_0(t)$. Moreover, a linear combination of shifted versions of B-splines of order L can reproduce polynomials up to order L . More precisely, there exist coefficients $c_{m,n}$ such that

$$\sum_{n \in \mathbb{Z}} c_{m,n} \beta_L(t - n) = t^m \quad m = 0, 1, \dots, L. \quad (11)$$

Let us now reconsider the sampling problem of the previous section. In this new acquisition scenario, we have that the measurements $y_k(n)$ observed by the k -th digital camera are given by:

$$y_k(n) = \langle f(t), \beta_L \left(t - t_k - \frac{n}{N} \right) \rangle.$$

We now observe the following:

$$\begin{aligned} \sum_n c_{m,n} y_k(n) &\stackrel{(a)}{=} \langle f(t), \sum_n c_{m,n} \beta_L \left(t - t_k - \frac{n}{N} \right) \rangle \\ &\stackrel{(b)}{=} \int_{-\infty}^{\infty} f(t) (t - t_k)^m dt \\ &\stackrel{\text{def}}{=} \tau_{m,k} \quad m = 0, 1, \dots, L, \end{aligned} \quad (12)$$

where (a) follows from the linearity of the inner product and (b) from the polynomial reproduction formula (11). The above equation is therefore showing that it is possible to retrieve the exact moments $\tau_{m,k}$ of the signal $f(t - t_k)$ from its samples. From the moments it is then almost straightforward to retrieve the offsets t_k . In fact we have that $t_k = (\tau_{1,k} - \tau_{1,0}) / \tau_{0,0}$.

In the case of two-dimensional signals like images, the transformation that relates two signals can be more complicated than a simple translation. For example, the same object in two different images might be related by an affine transformation. One can show that a moment-based registration as described above for a simple shift is also possible for such more complex motion. Since an affine transformation has six degrees of freedom, more moments are required. More precisely, third order moments along the x and y axis are needed in order to retrieve the affine transformation.

Local algorithm - There might be situations where the above approach might become impractical or unstable. In this case, a registration approach based on local features might be more convenient. Many registration algorithms are based on corner detection and then on the matching of the corners of two different images in order to retrieve the transformation between the two images.

In this work, we propose to locate corner points at the intersection of two straight edges. A single edge is parameterized by its angle θ , height ζ and shift γ . Such parameters can be exactly retrieved from the samples using the moment of the derivative of the original function. More precisely, denote with $z(n) = y_0(n) - y_0(n-1)$, one can show the following:

$$z(n) = \left\langle \frac{df(t)}{dt}, \beta_{L+1} \left(t - \frac{n}{N} \right) \right\rangle.$$

Namely, the new samples $z(n)$ are equivalent to those obtained by sampling the derivative of $f(t)$ with a B-spline of order $L+1$ rather than L .

It is then possible to retrieve the step parameters from the new samples $z(n)$. The complete solution for a single step edge is given by:

$$\begin{aligned} \zeta &= -\tau_{0,n}, \\ \tan \theta &= \frac{\tau_{0,n}}{\tau_{1,n+1} - \tau_{1,n}}, \\ \gamma &= \frac{(n+1)\tau_{1,n} - n\tau_{1,n+1}}{\tau_{0,n}}, \end{aligned} \quad (13)$$

where in this case $\tau_{i,n}$ indicates the i -th order moment of $\frac{df(t)}{dt}$ along the n -th row of the sampled image. The derivative $\frac{df(t)}{dt}$ is computed in discrete domain using discrete differences.

The validity of the above schemes will be assessed in the next section.

4. RESULTS

We now apply the presented algorithms on various sets of input images. For the algorithm from Section 3.1, we subsampled an image of 31×31 pixels (without prefiltering, but after applying real-valued shifts), generating 5 randomly shifted and aliased input images of 16×16 pixels. One such image is shown in Figure 1a, and the reconstructed image is shown in Figure 1b. The shift values are estimated up to a precision of 10^{-16} (Matlab precision). Similarly, we tested

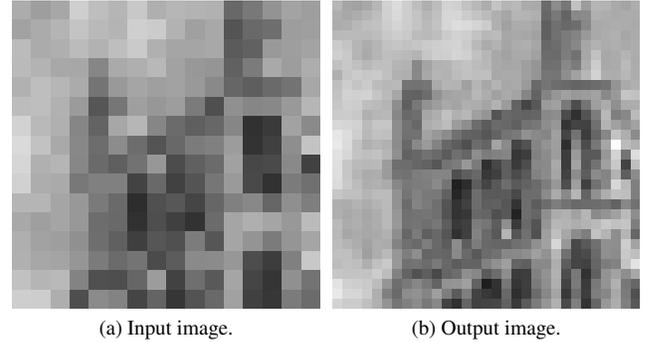


Fig. 1: Results using algorithm from Section 3.1. The high resolution image (b) is perfectly reconstructed from 5 shifted low resolution images (a).

the algorithm from Section 3.2 on a set of 5 shifted and aliased input images of size 32×32 , generated from a 63×63 image. The results are shown in Figure 2. Again, motion parameters are estimated up to working precision.

For the algorithm from Section 3.3, we consider a set of real images as they are acquired by a digital camera. The registration approach considered here is based on continuous moments. Since it takes a sampling point of view, image samples should be modified as little as possible by internal post-processing occurring in a digital camera after acquisition. The set of images is thus acquired by a SLR digital camera (a Nikon D70s) in RAW format. The experiment is presented in Figure 3. 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 and is directly approximated with a cubic B-spline at scale 1. The high resolution image is reconstructed

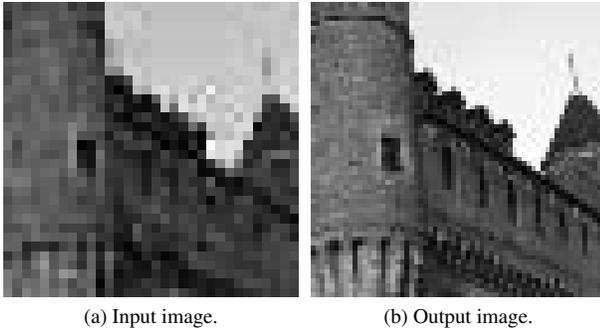


Fig. 2: Results using algorithm from Section 3.2. The high resolution image (b) is perfectly reconstructed from 5 shifted low resolution images (a).

by first calculating the image values on a uniform pixel grid using a cubic interpolation and a Delaunay triangulation, followed by the MRNSD deconvolution algorithm (modified residual norm steepest descent) [13]. Figure 3a shows the moon as acquired by the camera and Figure 3b presents the obtained super-resolved image where details of the moon can be observed.

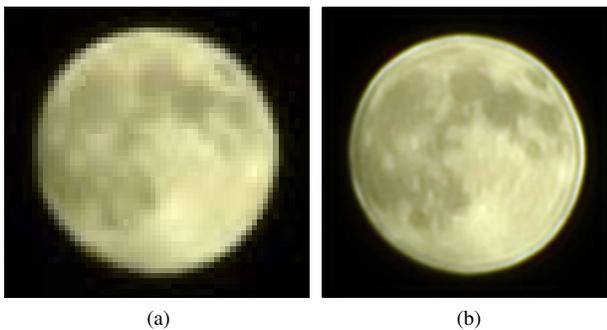


Fig. 3: Real super-resolution of the moon from 60 images acquired with a Nikon D70s SLR camera and a lens (18-70mm, F3.5-4.5) set at a focal length of 38mm (35mm equiv.: 57mm). (a) The moon as acquired by the camera (60x60 px); (b) Super-resolved image of the moon (600x600 px) with MRNSD restoration method.

From the above figures, we can see that all three methods give considerable improvements in image resolution. They also require high computational power. For the algorithms from Sections 3.1 and 3.2, only computer simulations were shown where the original image is perfectly reconstructed. While the memory requirements are higher for the algorithm in Section 3.1, the algorithm from Section 3.2 requires a slightly larger number of input images ($K > S = \lceil L/N \rceil$ instead of $K > (L-1)/(N-1)$). The algorithm from Section 3.3 is tested in an experiment using real images from a digital camera as input. The result can therefore only be evaluated visually. It does not put specific requirements on the minimum number of input images and performs pairwise registration of the images.

5. CONCLUSIONS

We have presented a common framework for some recent subspace-based image registration methods for super-resolution imaging. First, we have shown that super-resolution can be described as a multichannel sampling problem with unknown offsets. Three different solution methods using subspace descriptions were then described: one using ideas from variable projection theory, one using a Fourier analysis of the aliased signals, and finally one based on a spline analysis of the sampling kernel. The performance of the different algorithms is illustrated in numerical simulations.

6. REFERENCES

- [1] P. Vandewalle, L. Sbaiz, J. Vandewalle, and M. Vetterli, "Super-resolution from unregistered and totally aliased signals using subspace methods," *IEEE Transactions on Signal Processing*, vol. 55, no. 7, Part 2, pp. 3687–3703, 2007.
- [2] Loic Baboulaz and Pier Luigi Dragotti, "Distributed acquisition and image super-resolution based on continuous moments from samples," in *Proc. IEEE International Conference on Image Processing*, 2006, pp. 3309–3312.
- [3] Loic Baboulaz and Pier Luigi Dragotti, "Local feature extraction for image super-resolution," in *Proc. IEEE International Conference on Image Processing*, 2007, vol. 5, pp. 401–404.
- [4] R. Y. Tsai and T. S. Huang, "Multiframe image restoration and registration," in *Advances in Computer Vision and Image Processing*, T. S. Huang, Ed., vol. 1, pp. 317–339. JAI Press, 1984.
- [5] "IEEE Signal Processing Magazine, special issue on super-resolution," May 2003.
- [6] "EURASIP journal on applied signal processing, special issue on super-resolution," 2006.
- [7] Sina Farsiu, M. Dirk Robinson, Michael Elad, and Peyman Milanfar, "Fast and robust multiframe super resolution," *IEEE Transactions on Image Processing*, vol. 13, no. 10, pp. 1327–1344, Oct. 2004.
- [8] Tuan Q. Pham, Lucas J. van Vliet, and Klamer Schutte, "Robust Fusion of Irregularly Sampled Data Using Adaptive Normalized Convolution," *EURASIP Journal on Applied Signal Processing*, vol. 2006, 2006, Article ID 83268, 12 pages.
- [9] Dirk Robinson, Sina Farsiu, and Peyman Milanfar, "Optimal Registration of Aliased Images Using Variable Projection with Applications to Superresolution," *The Computer Journal (invited paper)*, April 2007.
- [10] Yu He, Kim-Hui Yap, Li Chen, and Lap-Pui Chau, "Joint image registration and super-resolution using nonlinear least squares method," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, April 2007, vol. 1, pp. 561–564.
- [11] Patrick Vandewalle, *Super-resolution from unregistered aliased images*, Ph.D. thesis, Ecole Polytechnique Fédérale de Lausanne, 2006, EPFL Thesis No. 3591.
- [12] Axel Ruhe and Per Ake Wedin, "Algorithms for Separable Nonlinear Least Squares Problems," *SIAM Review*, vol. 22, no. 3, pp. 318–337, July 1980.
- [13] J. Nagy and Z. Strakos, "Enforcing nonnegativity in image reconstruction algorithms," *Mathematical Modeling, Estimation, and Imaging*, pp. 182–190, 2000.