AFFINE TRANSFORMATIONS OF IMAGES : A LEAST SQUARES FORMULATION

Michael Unser, Matthew A. Neimark, Chulhee Lee

Biomedical Engineering and Instrumentation Program, NCRR, Bldg. 13, Room 3W13
National Institutes of Health, Bethesda, MD 20892 USA

ABSTRACT

We present a general framework for the design of discrete geometrical transformation operators, including rotations and scaling. The first step is to fit the discrete input image with a continuous model that provides an exact interpolation at the pixel locations. The corresponding image model is selected within a certain subspace $V(\varphi) \subset L_2(R^p)$ that is generated from the integer translates of a generating function φ; particular examples of this construction include polynomial spline and bandlimited signal representations. Next, the geometrical transformation is applied to the fitted model, and the result is re-projected onto the representation space. This procedure yields a solution that is optimal in the least squares sense. We show that this method can be implemented exactly using a combination of digital filters and a re-sampling step that uses a modified sampling kernel. We then derive explicit implementation formulas for the piecewise constant and cubic spline image models. Finally, we consider image processing examples and show that the present method compares very favorably with a standard interpolation that uses the same model.

1. INTRODUCTION

The standard approach for implementing geometrical transformations (translation, rotation, scaling, etc...) is to fit the image with a continuous model (image interpolation), and then re-sample this two-dimensional function on a new sampling grid[1, 2]. The most commonly used methods are nearest neighbor and bilinear interpolation, which correspond to zero and first order spline models, respectively. Unfortunately, these simple approaches have several problems associated with them : (a) they frequently introduce noticeable image distortions (blocking artifact, smoothing), especially when lower order models are used, and (b) they are not well suited for image reduction because of potential aliasing problems.

An attractive solution for improving the quality of these reconstructions is to reformulate the problem as an optimization task where the goal is to minimize an error criterion that is entirely specified in the continuous domain. The demonstration that such a least squares approach can result in significant performance improvement for image scaling (expansion and reduction) was provided recently in [3]. In this paper, we extend this formulation to the more general class of affine transformations (including rotations and scaling), and present general computational solutions. In addition, we consider a more general class of continuous image models and provide an explicit (non-tensor-product) formulation for multi-dimensional signals defined over R^p.

Specifically, the problem that we address is the following. Assuming that the input image $s_1(\boldsymbol{x})$ is continuously defined, we want to find a good digital approximation of its affine geometrical transformation

$$(A_T s_1)(\boldsymbol{x}) = s_1(\boldsymbol{T}(\boldsymbol{x}-\boldsymbol{t})), \tag{1}$$

where $\boldsymbol{T}$ is a given $p \times p$ non-singular transformation matrix (e.g., a combination of rotations and scalings), and $\boldsymbol{t}$ a translation vector. Since the output needs to be represented on a digital grid, we have to expect some potential loss of information. Our goal here will be to minimize this error and improve the standard interpolation solution, which computes the digital approximation by mere sampling: $s_2[\boldsymbol{k}] = (A_T s_1)(\boldsymbol{x})\big|_{\boldsymbol{x}=\boldsymbol{k}}$.

1.1 Mathematical notations

$L_2(R^p)$ is the vector space of measurable, square-integrable p-dimensional functions $f(\boldsymbol{x})$, $\boldsymbol{x} = (x_1, \cdots, x_p) \in R^p$. It is a Hilbert space whose metric $\|\cdot\|$ (the L_2-norm) is derived from the L_2-inner product $\langle f_1, f_2 \rangle = \int_{\boldsymbol{x} \in R^p} f_1(\boldsymbol{x}) f_2^*(\boldsymbol{x}) d\boldsymbol{x}$. The Fourier transform $\hat{f}$ of f is $\hat{f}(\boldsymbol{\omega}) = \langle f, e^{j\boldsymbol{\omega} \boldsymbol{x}} \rangle$ where $\boldsymbol{\omega} = (\omega_1, \cdots, \omega_p)$ is the p-dimensional frequency variable.

$l_2(Z^p)$ is the space of square summable p-dimensional sequences (or discrete signals) $a(\boldsymbol{k}), \boldsymbol{k} \in Z^p$.

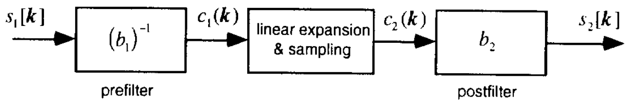

Fig.1 : General block diagram of the least squares transformation algorithm.

2. PRELIMINARIES

The continuous multi-dimensional image model $s(\boldsymbol{x})$ is selected within a certain subspace $V(\varphi)$ of $L_2(R^p)$. This subspace is generated from the integer translates of the *generating* function φ:

$$V(\varphi) = \left\{ s(\boldsymbol{x}) = \sum_{\boldsymbol{k}\in Z^p} c(\boldsymbol{k})\varphi(\boldsymbol{x}-\boldsymbol{k}) : c(\boldsymbol{k}) \in l_2(Z^p) \right\}. \quad (2)$$

The only constraint on φ is that there exist two strictly positive constants A and B such that

$$A \le \sum_{\boldsymbol{k}\in Z^p} \left|\hat{\varphi}(\boldsymbol{\omega} + 2\pi\boldsymbol{k})\right|^2 \le B, \ a.e \quad (3)$$

where $\hat{\varphi}$ is the Fourier transform of φ. The admissibility condition (3) insures that $V(\varphi)$ is a well-defined subspace of $L_2(R^p)$ with $\{\varphi(\boldsymbol{x}-\boldsymbol{k})\}_{\boldsymbol{k}\in Z^p}$ as a Riesz basis [4]. In other words, any function in $V(\varphi)$ has a unique and stable representation of the form given by (2). Note that the basis is orthonormal if and only if A=B=1 in (3).

The present signal model turns out to be quite general and covers many special cases that have been considered in the literature. For instance, if we choose $\varphi(\boldsymbol{x}) = \mathrm{sinc}(\boldsymbol{x})$, then $V(\varphi)$ is the traditional space of bandlimited functions. Another very relevant family of functions that fall into this framework are the polynomial splines of degree n [5, 6]. In the simplest case of tensor-product splines, we have $\varphi(\boldsymbol{x}) = \beta^n(x_1)\beta^n(x_2)\cdots\beta^n(x_p)$, where $\beta^n(x)$ is the univariate B-spline of degree n, which can be constructed from the $(n+1)$-fold convolution of unit rectangular pulse.

The general sampling theory for the approximation and representation of functions in $V(\varphi)$ was formulated in [4] for the univariate case, and can easily be extended for higher dimensional signals. A key result is that the expansion coefficients in (2) of the least squares (minimum L_2-norm) approximation in $V(\varphi)$ of any function $r(\boldsymbol{x}) \in L_2(R^p)$ can be obtained by simple inner-product

$$c(\boldsymbol{k}) = \langle r, \ddot{\varphi}(\boldsymbol{x}-\boldsymbol{k})\rangle \quad (4)$$

where $\ddot{\varphi} \in V(\varphi)$ is the *dual* generating function, which is defined as follows

$$\ddot{\varphi}(\boldsymbol{x}) \xleftrightarrow{Fourier} \frac{\hat{\varphi}(\boldsymbol{\omega})}{\sum_{\boldsymbol{k}\in Z^p}\left|\hat{\varphi}(\boldsymbol{\omega}+2\pi\boldsymbol{k})\right|^2}. \quad (5)$$

Specific results for polynomial spline approximations can be found in [7].

3. THE ALGORITHM

Our least squares (LS) transformation method is schematically summarized in Fig. 1. It uses the three basic steps that are described below.

3.1 Image interpolation

We start by fitting our digital image with a continuous function $s_1 \in V(\varphi_1)$ of the form

$$s_1(\boldsymbol{x}) = \sum_{\boldsymbol{k}\in Z^p} c_1(\boldsymbol{k})\varphi_1(\boldsymbol{x}-\boldsymbol{k}). \quad (6)$$

The requirement at this stage is that the model provides an exact interpolation of the initial pixel values $s[\boldsymbol{k}] \in l_2(Z^p)$. It is not difficult to show that the corresponding expansion coefficients $c_1[\boldsymbol{k}]$ (one per pixel) can be determined by digital filtering (pre-filtering)

$$c_1(\boldsymbol{k}) = (b_1)^{-1} * s[\boldsymbol{k}] \quad (7)$$

where $b_1(\boldsymbol{k}) := \varphi_1(\boldsymbol{x})\big|_{\boldsymbol{x}=\boldsymbol{k}}$ is the sampled generating function, and where $(b_1)^{-1}$ denotes the corresponding convolution inverse. In the B-spline case, the pre-filtering in (7) can be performed very efficiently using the recursive algorithm described in [8].

3.2 Transformation and approximation

The next step is to apply the affine transformation operator $A_T : L_2(R^p) \mapsto L_2(R^p)$ defined by (1) to the fitted image model $s_1(\boldsymbol{x})$. The transformed model $(A_T s_1)(\boldsymbol{x})$ is then approximated in the least square sense by another continuous function $s_2 \in V(\varphi_2)$ of the form

$$s_2(\boldsymbol{x}) = \sum_{\boldsymbol{k}\in Z^p} c_2(\boldsymbol{k})\varphi_2(\boldsymbol{x}-\boldsymbol{k}). \quad (8)$$

Using the results in Section 2, we show that the optimal image coefficients are given by

$$\begin{aligned} c_2(\boldsymbol{l}) &:= \langle A_T s_1(\boldsymbol{x}), \ddot{\varphi}_2(\boldsymbol{x}-\boldsymbol{l})\rangle, \\ &= \sum_{\boldsymbol{k}\in Z^p} c_1(\boldsymbol{k})\langle A_T\varphi_1(\boldsymbol{x}-\boldsymbol{k}), \ddot{\varphi}_2(\boldsymbol{x}-\boldsymbol{l})\rangle \\ &= \sum_{\boldsymbol{k}\in Z^p} c_1(\boldsymbol{k})\cdot\xi_T(\boldsymbol{x}-\boldsymbol{t}-\boldsymbol{T}^{-1}\boldsymbol{k})\big|_{\boldsymbol{x}=\boldsymbol{l}} \end{aligned} \quad (9)$$

where ξ_T is the modified sampling kernel defined by

$$\xi_T(\boldsymbol{y}) = \int_{\boldsymbol{x}\in R^p} \varphi_1(\boldsymbol{T}\boldsymbol{x})\overset{o}{\varphi}_2(\boldsymbol{x}-\boldsymbol{y})d\boldsymbol{x}. \quad (10)$$

The result given by (9) is quite general but this form may not always be very appropriate for an explicit numerical evaluation. The problem is that the summation for a given index $\boldsymbol{l}$ may involve an infinite number of terms, unless, of course, ξ_T is compactly supported. This property can be enforced by selecting the generating functions φ_1 and φ_2 appropriately. For the polynomial spline model, the most judicious choice is $\varphi_1(\boldsymbol{x}) = \overset{o}{\varphi}_2(\boldsymbol{x}) = \beta^n(x_1)\beta^n(x_2)\cdots\beta^n(x_p)$, since the B-splines are compactly supported by construction. With this particular choice, the evaluation of (4) will provide us with the expansion coefficients of s_2 in the so-called dual spline representation [9].

The other practical difficulty is that we will also need an explicit formula for the modified kernel (10), which can be rather challenging to obtain analytically since it depends heavily on the transformation matrix $\boldsymbol{T}$. There is another more pragmatic approach (not pursued here) which is to pre-compute a digital approximation of the convolution product (10) on a multi-dimensional grid with step size $\Delta \ll 1$. This computation can be done efficiently using FFT-based convolution techniques. The values of $\xi_T(y)$ lying in-between grid points can then be approximated as required using bilinear interpolation.

3.3 Model resampling

The last step is to display the resulting approximation $s_2(\boldsymbol{x})$. The pixel values $s_2[\boldsymbol{k}]$ are obtained from the $c_2(\boldsymbol{k})$ by discrete convolution with the sampled kernel $b_2(\boldsymbol{k}) := \varphi_2(\boldsymbol{x})\big|_{\boldsymbol{x}=\boldsymbol{k}}$:

$$s_2[\boldsymbol{k}] = b_2 * c_2(\boldsymbol{k}). \quad (11)$$

In our experiments, we have chosen φ_2 to be the dual B-spline which corresponds to the post-filtering kernel

$$b_2(\boldsymbol{k}) = \left(b^{2n+1}\right)^{-1} * b^n(\boldsymbol{k}), \quad (12)$$

where b^n denotes the separable discrete B-spline of degree n. This digital filter can also be implemented recursively using the same technique as before [3].

4. RESULTS

We have implemented this least squares procedure for computing rigid 2D image transformations (translation + rotation + scaling) using both zero order and cubic spline image models (LS-n, n=0, 3, resp.). Note that the standard interpolation algorithms (INT-n, n=0,1,3) can be described by a formula similar to (9), with the difference that the sampling kernel ξ_T is replaced by φ_1 (the interpolation function).

For the zero-order case ($\varphi_1 = \varphi_2 = \overset{o}{\varphi}_2 = \text{rect}(x)\cdot\text{rect}(y)$), we were able to derive the analytical form of the modified sampling kernel in (5) as a function of the rotation angle θ and the scaling factor a. Despite the simplicity of the model, this kernel turns out be a rather involved piecewise bilinear function (first order spline) with up to 41 different subregions. No prefiltering and post-filtering is necessary since both φ_1 and φ_2 are true interpolating kernels. It can also be shown that the zero-order algorithm (LS-0) is equivalent to a bilinear interpolation (INT-1) in the case of a simple translation ($\boldsymbol{T} = \boldsymbol{I}$).

For the cubic spline case ($\varphi_1 = \overset{o}{\varphi}_2$: bicubic B-spline), we used a separable (and isotropic) Gaussian approximation of the convolution product in (10) of the form $\xi_{\theta,a}(\boldsymbol{x}) \cong \xi_a(x_1)\xi_a(x_2)\cdots\xi_a(x_p)$ where

$$\xi_a(x) \cong \begin{cases} \dfrac{a}{\sqrt{2\pi}\sigma_a}\exp\left\{\dfrac{-x^2}{2\sigma_a^2}\right\}, & |x| < \dfrac{n+1}{2}(1+a) \\ 0, & \text{otherwise} \end{cases} \quad (13)$$

and

$$\sigma_a = \sqrt{\frac{n+1}{12}\left(1+a^2\right)} \quad (14)$$

with n=3. This approximation formula is based on the fact that the cubic B-spline is extremely close to a Gaussian and that the convolution of two Gaussians is also a Gaussian whose variance is the sum of the variances of the individual components. This argument is also supported by the central limit theorem. In fact, this approximation is valid for higher order splines and its accuracy improves with increasing n. The corresponding pre- and post-filters in Fig. 1 were all implemented recursively. In any case, digital filtering is not the time consuming part of the algorithm; most of the computational effort is spent in the evaluation of the 2D summation formula (9).

To assess the performance of a given algorithm, we followed the initial transformation (rotation by θ and reduction by a factor of a) by its inverse using the same algorithm; a signal-to-noise ratio was then computed on the central 128×128 part of the image. Some of those results for the test image Lena are summarized in Table 1. The results obtained with other test images (Mandrill, MRI) were qualitatively very similar. In our experiments (θ=0, 15, 30, 45°; a=$2^{-1/2}$, 2^{-1}, 2^{-2}), LS-0 was consistently superior to INT-0 (nearest neighbor) with improvements of the order of 2 to 5 dBs, depending on the image, and the reduction factor. The most promising approach was undoubtedly LS-3 which outperformed all other standard methods (including INT-1) by several dBs. In most cases, the performance of LS-0 and INT-1 (bilinear interpolation) was

TABLE I
COMPARISON OF GEOMETRIC TRANSFORMATION ALGORITHMS FOR THE TEST IMAGE LENA

	$\theta = 0°$	$\theta = 15°$	$\theta = 30°$
$a = 2^{-1/2}$			
INT-0	22.01 dB	22.23 dB	22.17 dB
LS-0	25.26 dB	25.29 dB	25.31 dB
INT-1	25.41 dB	25.35 dB	25.39 dB
INT-3	27.38 dB	27.04 dB	26.74 dB
LS-3	28.07 dB	27.84 dB	27.73 dB
$a = 2^{-1}$			
INT-0	19.65 dB	19.94 dB	20.33 dB
LS-0	23.14 dB	23.14 dB	23.29 dB
INT-1	23.22 dB	23.34 dB	23.39 dB
INT-3	22.99 dB	23.27 dB	23.15 dB
LS-3	24.79 dB	24.72 dB	24.70 dB
$a = 2^{-2}$			
INT-0	17.30 dB	17.40 dB	17.64 dB
LS-0	20.02 dB	19.90 dB	19.96 dB
INT-1	19.34 dB	19.96 dB	20.01 dB
INT-3	18.89 dB	19.19 dB	19.24 dB
LS-3	21.14 dB	21.10 dB	21.14 dB

essentially equivalent. However, we found a slight advantage of LS-0 over INT-1 for images with a strong high frequency content and for large reduction factors; that is, when the effect of aliasing is quite significant. In any case, for a given image model, the least squares algorithm always outperformed the corresponding interpolation and produced images of much better quality. In addition, the transformed images did not exhibit the characteristic artifacts (blocking, or excessive smoothing) that are usually associated with low order interpolation methods (INT-0 and INT-1). Finally, we found that the proposed LS techniques would usually also result in some improvement for image magnification and zooming.

5. CONCLUSION

In this paper, we proposed a new least squares formulation for the geometrical transformation of images. The resulting procedure is similar to a standard interpolation, except that it uses a modified sampling kernel which depends explicitly on the transformation parameters. The method also requires an additional post-filtering step. For a given image model, the LS algorithm generally outperforms the corresponding interpolation solution in the sense that the transformed image is a more faithful copy of the original. Advantages of the method include lesser visual distortions and an improved suppression of aliasing artifacts. Its only drawback is that it requires more computations because of the increased size of the local neighborhood that needs to be considered for the evaluation at a particular image location.

As a future direction of research, we intend to develop faster sub-optimal versions of these algorithms, and consider the application of these techniques to the registration of volumetric data in medical imaging (PET and MRI).

Acknowledgments : We thank Dr. Ronald Levin for his support and encouragements.

References

[1] J.A. Parker, R.V. Kenyon and D.E. Troxel, "Comparison of interpolating methods for image resampling," *IEEE Trans. Med. Imaging*, vol. MI-2, no. 1, pp. 31-39, 1983.

[2] W.K. Pratt, *Digital image processing*. New York: Wiley, 1978.

[3] M. Unser, A. Aldroubi and M. Eden, "Enlargement or reduction of digital images with minimum loss of information," *IEEE Trans. Image Processing*, to appear.

[4] A. Aldroubi and M. Unser, "Sampling procedures in function spaces and asymptotic equivalence with Shannon's sampling theory," *Numer. Funct. Anal. and Optimiz.*, vol. 15, no. 1&2, pp. 1-21, February 1994.

[5] I.J. Schoenberg, *Cardinal spline interpolation*. Philadelphia, PA: Society of Insdustrial and Applied Mathematics, 1973.

[6] C.K. Chui, *Multivariate splines*. Philadelphia, PA: Society for Industrial and Applied Mathematics, 1988.

[7] M. Unser, A. Aldroubi and M. Eden, "Polynomial spline signal approximations : filter design and asymptotic equivalence with Shannon's sampling theorem," *IEEE Trans. Information Theory*, vol. 38, no. 1, pp. 95-103, January 1992.

[8] M. Unser, A. Aldroubi and M. Eden, "Fast B-spline transforms for continuous image representation and interpolation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, no. 3, pp. 277-285, March 1991.

[9] M. Unser, A. Aldroubi and M. Eden, "The L_2 polynomial spline pyramid," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 15, no. 4, pp. 364-379, April 1993.