A Fast Direct Fourier-Based Algorithm for Subpixel Registration of Images

Harold S. Stone, Michael Orchard, Ee-Chien Chang, Stephen Martucci *†

July 18, 2000

Abstract

This paper presents a new direct Fourier-based algorithm for image registration to subpixel accuracy for images that differ by translation and uniform changes of illumination. The algorithm detects those frequency components that have become unreliable estimators of shift due to aliasing, and removes them from the shift-estimate computation. In

*H. Stone and S. Martucci are with NEC Research Institute, Princeton, NJ 08540. E-mail: hstone,martucci@research.nj.nec.com. M. Orchard is with Princeton University, Princeton, NJ 08540. E-mail: orchard@ee.princeton.edu. E.-C. Chang is with National University of Singapore, E-mail: changec@cz3.nus.edu.sg

†An early version of this paper appeared in the *Proceedings of the 1999 Asilomar Conference on Signals, Systems and Computers.*, ©IEEE, 1999

the presence of aliasing, the average precision is a few hundredths of a pixel.

Several authors have proposed subpixel algorithms that work well when observed data are samples of ideally bandlimited images. However, images of real scenes captured by modern optics are not ideally bandlimited, and thus contain aliasing artifacts. Experimental data presented here show that the new algorithm yields superior precision in the presence of aliasing when compared to several earlier methods, and comparable precision to the iterative method of Thévenaz *et al.*

1 Introduction

Image registration is an important preprocessing operation that aligns the pixels of one image to corresponding pixels of a second image. Registration is the primary tool for comparing two or more images to discover the differences in the images or to fuse multiple modalities to create a composite that reveals information not easily accessible within individual images. It is used in the remote sensing community to study satellite images of the earth, and in the medical community to enhance the diagnostic capability of radiological imagery.

Registration algorithms typically assume that images differ by some transformation from a given family, and they find the transform within that family that optimizes a particular criterion. Transformation families include rigid transforms (translation, rotation, and rescaling), linear and affine (skewed and perspective transforms), and nonlinear warping. Optimization criteria include minimizing sum of squares of pixel differences,

maximizing normalized correlation coefficient, and maximizing mutual information. For this paper, we assume that two observed sampled images represent the same scene sampled on identical grids, but offset from each other by an unknown translational shift, as well as differing by a uniform change of intensity, perhaps also disturbed by independent additive noise.

Townshend *et al.* [1] make a case for subpixel accuracy of registration in their studies of vegetation changes because the changes that they are trying to measure are on the order of the errors introduced by misregistration by less than a pixel. The literature contains registration algorithms of various precision that operate in the pixel (signal) domain [2-22], and in the Fourier domain [2,5,23-36].

Among the authors who claim subpixel precision, Shekarforoush *et al.* [34] describe an algorithm that uses the cross-power spectrum of the Fourier transforms of two ideally bandlimited images. The Fourier inverse of the cross-power spectrum is a sinc function displaced from the origin by the amount of the translation. Abdou [2] describes three algorithms that use various interpolation schemes to find the translational difference between images. Kim and Su [30] present a Fourier-based algorithm that estimates translation changes by modifying the phase of one Fourier transform to make it as similar as possible to a second Fourier transform. The phase change corresponds to the translational difference. To eliminate aliasing effects their algorithm relies on the low frequency components of the transforms. Thévenaz *et al.* [21] report an elegant pixel-based iterative algorithm that is able to register to high precision, and can deal with rotation, translation,

changes of scale, and illumination changes. Of the schemes mentioned here, it is the most general.

Algorithms reported to have subpixel precision for ideally bandlimited images typically have reduced precision in the presence of aliasing. Kim and Su's algorithm [30] treats aliasing explicitly, but the others listed do not.

The main result of this paper is the development of a direct algorithm for image registration that achieves high precision in the presence of small amounts of aliasing. This paper models the aliased frequency components of the two images and predicts how this aliasing affects the phase relationships between their Fourier transforms. Aliasing causes some frequency components of the scene to be unreliable, and the new algorithm masks them out of the registration process. The new algorithm resembles [30] because it eliminates certain frequency components from the calculations, but it eliminates more than do Kim and Su. For our experimental data, the frequency-masking algorithm achieves a worst-case precision of a few hundredths of a pixel, and an average precision of less than a hundredth of a pixel for a broad range of nonideal prefilters. The iterative algorithm of [21] achieves comparable precision experimentally, even though the algorithm does not deal with aliasing explicitly. All other algorithms compared in this paper have poorer precision.

Section 2 reviews theoretical issues related to aliasing and its effect on Fourier spectra, particularly on the phase of Fourier transforms. Section 3 describes the new algorithm, and how it deals with aliasing. The ground-truth model and experimental results appear in Section 4. A summary and

open research issues is in the final section.

2 Mathematical Background

The basic idea behind the new registration algorithm is that the phase of the Fourier spectra of an image pair contains sufficient information to determine the translation offset difference of the images. This section reviews how this can be exploited in the absence of aliasing, as described by [34], and then shows why aliasing causes problems that reduce the precision of this method. The analysis gives us a powerful tool for detecting the frequency components most likely to be corrupted by aliasing.

Given a two-dimensional image $f(x, y)$ and a translated version of the image $g(x, y) = f(x - x_0, y - y_0)$, we wish to find an efficient algorithm that gives the displacement vector (x_0, y_0). In the remainder of the discussion, we reduce the problem to one dimension and note that all the results generalize to two-dimensions straightforwardly.

Let $f_c(x)$ denote a continuous function with Fourier transform $F_c(\Omega)$. The Fourier transform of the shifted function $g_c(x) = f_c(x - x_0)$ is $G_c(\Omega) = F_c(\Omega)e^{-j\Omega x_0}$. This is the key relationship for all Fourier-based image-registration algorithms. In the ideal case of continuous transforms of noiseless images, for any Ω' the Fourier components $F_c(\Omega')$ and $G_c(\Omega')$ provide a perfect estimate of x_0, to within an integer multiple of $(2\pi/\Omega')$.

Now consider digital images, which are sampled versions of continuous images. Assume that $f_c(x)$ is continuous and bandlimited to frequencies less than Ω_0. Sampled versions of this image are $f(n) = f_c(nT)$ and $g(n) =$

$f_c(nT - x_0)$, where n is an integer, T is the sampling interval, and x_0 is an arbitrary real-valued displacement with a magnitude less than 0.5. In this formulation, the value of x_0 is the residual that remains after registering images to the nearest pixel by using a standard algorithm from the literature. Our goal is to find x_0. Because $|x_0| \leq 0.5T$, the phase ambiguity of $2\pi/\Omega'$ is not present.

If the sampling period T is less than π/Ω_0, then the Sampling Theorem states that

$$f_c(x) = \sum_{n=-\infty}^{+\infty} f(n)\,\mathrm{sinc}(x/T - n) \tag{1}$$

for any x and for integer values of n. Recall that $\mathrm{sinc}(x) = \sin(\pi x)/(\pi x)$.

The registration method of [34] uses discrete Fourier transforms of finite sampled images for which the reconstruction of Eq. (1) holds approximately as shown below. The Sampling Theorem requires an infinite number of samples of $f(n)$, whereas only N samples are available. The discrete Fourier transform implies a periodicity for $f(n)$ beyond the boundaries of the N samples, whereas the actual value of the $f(n)$ is unknown beyond the boundaries. Because of these deviations from the underlying image source, even though an exact relation exists between the continuous $g(x)$ and $f(x)$, the corresponding relation between their discrete counterparts is only an approximation. To obtain that approximation, limit the summation in (1) to the range $0 \leq n \leq N-1$ and replace $\mathrm{sinc}(x)$ by $\mathrm{dirichlet}_N(2\pi x/N)$. This leads directly to

$$g_c(x) = f_c(x - x_0) \approx \sum_{n=0}^{N-1} f(n)\,\mathrm{dirichlet}_N((2\pi/N)((x - x_0)/T - n)). \tag{2}$$

The dirichlet function is a periodic approximation to sinc, and is defined to

be $\text{dirichlet}_N(x) = \sin(Nx/2)/(N\sin(x/2))$ for x not a multiple of 2π, and is $\text{sign}(\cos(x))$ otherwise. The dirichlet function is very close to sinc for the interval in the summation in Eq. (2).

For integer frequencies ω, $-N/2 \leq \omega \leq N/2 - 1$, let $F(\omega)$, $G(\omega)$, and $D(\omega)$ be the respective discrete Fourier Transforms of $f(x)$, $g(x)$, and $\text{dirichlet}_N((2\pi/N)((x - x_0)/T))$. Then the convolution in Eq. (2) corresponds to

$$G(\omega) = F(\omega)D(\omega). \tag{3}$$

This explains why the inverse Fourier transform of the spectral ratio $D(\omega) = G(\omega)/F(\omega)$, which is called the *cross-power spectrum*, is approximately a sinc function displaced from the origin by an amount x_0 [34]. The approximation error is due to the truncation of an infinite sum to a finite sum in Eq. (2) and to the approximation of sinc by dirichlet. If $f_c(x)$ is sampled below its Nyquist rate then its continuous transform has energy in frequencies higher than Ω_0, in which case Eq. (2) is not guaranteed to be a good approximation because the perfect reconstruction of Eq. (1) no longer holds.

A related registration algorithm uses *phase correlation* defined to be the Fourier inverse of $phase(G(\omega)\hat{F}(\omega))$ where the function $phase(F(\omega))$ is the unit magnitude complex function whose phase is equal to the phase of $F(\omega)$. When shift distance x_0 is an integral number of pixels, the function $phase(G(\omega)\hat{F}(\omega))$ equals the cross-power spectrum phase. The inverse transform of both the phase correlation and the cross-power spectrum at these shifts is $\delta(x - x_0)$. The functions differ when x_0 is a fractional shift

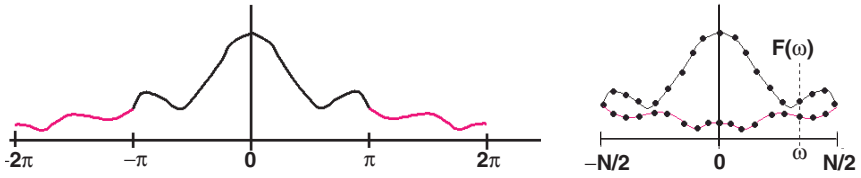
because phase correlation is an inverse transform of a unit-magnitude function, but the cross-power spectrum does not have unit magnitude in general.

At this point we turn to the problem of estimating the phase of an aliased cross-power spectrum when $f_c(x)$ is sampled below its Nyquist rate. To do this, we make an assumption about the aliasing that corrupts digital images. We assume that the optical system or other prefilter is not an ideal low-pass filter, and that the spectral leakage is most likely to be in the frequencies just above the sampling cutoff frequency. Also, the attenuation increases strongly with frequencies higher than sampling cutoff, with essentially no energy at frequencies above twice the sampling cutoff. Figure 1(a) illustrates this situation. The figure shows a Fourier spectrum plotted against normalized continuous frequency $\Omega, -2\pi \leq \Omega < 2\pi$, which represents the spectrum in a one-dimensional image after prefiltering according to our assumptions. Sampling the image at a sampling interval of $T = 2\pi/\Omega_0$ produces the N samples as shown in Fig. 1(b). The sprecta for $F(\omega)$ and $G(\omega)$ in Fig. 1(b) are sums of phase-shifted components of frequencies $F_c(\Omega)$ in Fig. 1(a). For integer ω in the interval $-N/2 \leq \omega < 0$ we have

$$\begin{aligned} F(\omega) &= F_c(2\pi\omega/N) + F_c(2\pi(\omega+N)/N) \\ G(\omega) &= F_c(2\pi\omega/N)e^{-2\pi j\omega x_0/N} + F_c(2\pi(\omega+N)/N)e^{-2\pi j((\omega+N)x_0/N)} \end{aligned} \quad (4)$$

For ω in the interval $0 \leq \omega \leq N/2 - 1$, replace $(\omega + N)$ by $(\omega - N)$ in the second term of Eq. (4).

The aliasing terms in Eq. (4) create $F(\omega)$ and $G(\omega)$ such that the vector ratio $G(\omega)/F(\omega)$ is not equal to $e^{-2\pi j\omega x_0/N}$, as required for Eqs. (1) and (2). A graphic illustration of this situation appears in Fig. 2. The complex



(a) (b)

Figure 1: (a) Spectrum of image after prefiltering. (b) Downsampling and resulting aliasing of prefiltered image.

vector $F(\omega)$ is shown to be the sum of two complex components of $F_c(\Omega)$. The complex vector $G(\omega)$ is the complex sum of those same components, one of which is rotated by $2\pi\omega x_0/N$, and the other rotated by $2\pi(\omega + N)x_0/N$. The vector $G(\omega)$ is not a rotation of $F(\omega)$ by $2\pi\omega x_0/N$ because of the excess phase of $2\pi x_0$ in the second component of the sum. In general, the aliasing caused by the excess rotational phase leaves $G(\omega)$ with an amplitude different from that of $F(\omega)$ whereas bandlimited images experience no change of amplitude in the frequency domain after translation.

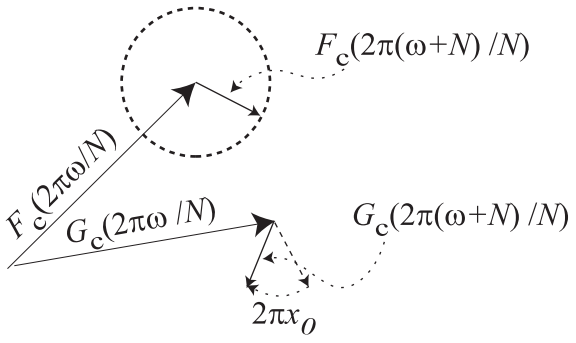


Figure 2: Vector illustration of aliasing

Our problem is to estimate the phase $2\pi\omega x_0/N$ as a function of ω from observations of the vector sums shown in Fig. 2. The components of F_c are

not directly observable. The next section shows how we can eliminate badly aliased frequency components from further consideration, and thereby estimate x_0 from frequency components whose phases yield accurate estimates of x_0.

3 Frequency Masking Subpixel Shift Estimation

We propose to estimate x_0 from the phase of the ratio $G(\omega)/F(\omega)$ for a selected set of ω. Examination of Fig. 2 shows that observed relative phase of $G(\omega)/F(\omega)$ for a specific value of ω is likely to be a good estimate of x_0 if the magnitude of the alias component $F_c(2\pi(\omega + N)x_0/N)$ is small compared to the magnitude of the in-band component $F_c(2\pi\omega x_0/N)$. Fig. 1(b) shows that this is likely to occur under our assumptions at frequencies near the origin because of the attenuation of aliasing magnitude with increasing frequency by the prefilter. Hence, we should limit the frequency range to frequencies near the origin [30].

But this is not sufficient to attain high precision, as indicated by the experimental data later in the paper. There usually exist highly aliased frequency components near the origin. Using these frequencies greatly reduces the precision of the estimate of x_0. The frequencies that are most likely to be corrupted are those for which the spectral magnitude is small. Therefore, the algorithm masks out contributions from spectral components whose magnitudes are small relative to the rest of the magnitudes, regardless of whether they occur at low or high frequencies.

The full algorithm for two-dimensional data is very simple and consists

of these major steps:

1. Use any image registration algorithm to find a translation that registers the two images to the nearest integral pixel coordinates.

2. Apply a Blackman or Blackman-Harris window in the pixel domain to eliminate image-boundary effects in the Fourier domain [37].

3. Calculate the discrete Fourier transforms of $f(x, y)$ and $g(x, y)$.

4. Mask out spectral components that lie outside a radius of R from the central peak. A suitable value of R is $0.6N/2$ where N is the minimum of the number of samples in the x and y dimensions.

5. Mask out spectral components for which either $F(\omega, \nu)$ or $G(\omega, \nu)$ have magnitudes less than a specified threshold α.

6. Using the frequencies that remain after masking, find a least-squares estimate of (x_0, y_0).

The windowing operation is well-known and eliminates the spurious introduction of high-frequency spectral energy due to edge effects. We found that a separable Blackman window (as well as a separable Blackman-Harris window) worked quite well [37]. We also tested a radially symmetric Blackman window and several other windows that are flatter than the Blackman window in the middle of the image. The radially symmetric window gave results comparable to the separable window, but is more complex to create. The flatter windows use more information from the center of the image, but they tend to be less effective in eliminating spurious high frequency energy

from the edges of the images. The separable Blackman and Blackman-Harris windows yielded the best results for the least computation.

The use of radius $R = 0.6N/2$ constrains the frequencies to be close to the origin. We found that the constant factor 0.6 can be as small as 0.5 or as large as 0.7 without materially affecting the algorithm.

The choice of threshold α warrants a brief discussion. The algorithm sorts the frequencies by magnitude and retains the M largest in the spectrum, for some value of M. In the absence of other information, a good way to choose M is to vary M over a range of values, and observe the estimated translation as a function of M. Experimental data in the next section show that there is a region where the estimated translation is virtually independent of M. The displacement estimate produced by M in this region is the one to use.

To estimate translation displacements from the Fourier spectra, let $\mathcal{B}$ denote the set of frequency coordinate pairs (ω, ν) that survive the masking operations. Let $phase(\omega, \nu)$ be the phase of the complex ratio $G(\omega, \nu)/F(\omega, \nu)$ at point $(\omega, \nu) \in \mathcal{B}$. In the absence of aliasing, $phase(\omega, \nu)$ has x and y slopes equal to $2\pi x_0/N$ and $2\pi y_0/N$, respectively. The least squares estimate of the slope of a plane that passes through the origin is:

$$\begin{aligned} x_0 &= \left(\frac{N}{2\pi}\right)\left(\frac{\overline{\omega\nu}\ \overline{\nu\ phase} - \overline{\nu^2}\ \overline{\omega\ phase}}{\overline{\omega\nu}^2 - \overline{\omega^2}\ \overline{\nu^2}}\right) \\ y_0 &= \left(\frac{N}{2\pi}\right)\left(\frac{\overline{\omega\nu}\ \overline{\omega\ phase} - \overline{\omega^2}\ \overline{\nu\ phase}}{\overline{\omega\nu}^2 - \overline{\omega^2}\ \overline{\nu^2}}\right) \end{aligned}$$

where the quantities with overbars are the means of the respective products taken over all frequency pairs that survive the masking operation. Note that

any uniform change of intensity of the form $g' = Ag + B$ for constants A and B has no effect on the registration process because change of illumination does not change the value of $phase(\omega, \nu)$.

This completes the description of the registration algorithm.

4 Experimental Data

4.1 Ground truth

In order to measure the precision of the registration algorithm, we prepared ground truth using a scheme described in [34]. The idea is to use a single high-resolution image to represent the actual scene, and to create an image pair from this scene by filtering and downsampling the high-resolution image in two different ways. The downsampled images are shifted with respect to each other by integer amounts x_0 and y_0 in the high-resolution grid. After downsampling by M in each dimension, the relative shifts are x_0/M and y_0/M, respectively. After downsampling, one image optionally has its intensity values rescaled uniformly.

For the detailed analysis of the frequency-masking algorithm, we used 52 aerial photos of various urban and agricultural landscapes for this study. Each of the images is 1024×1024 pixels, and was downsampled by a factor of $M = 8$ in each dimension. Each of the two images was shifted by integer amounts ranging from -4 to +3 pixels in each dimension, thereby creating 64 possible shifts for each image and 4096 total pairs of shifts for each image, and over 200,000 registrations overall. This presented the algorithm

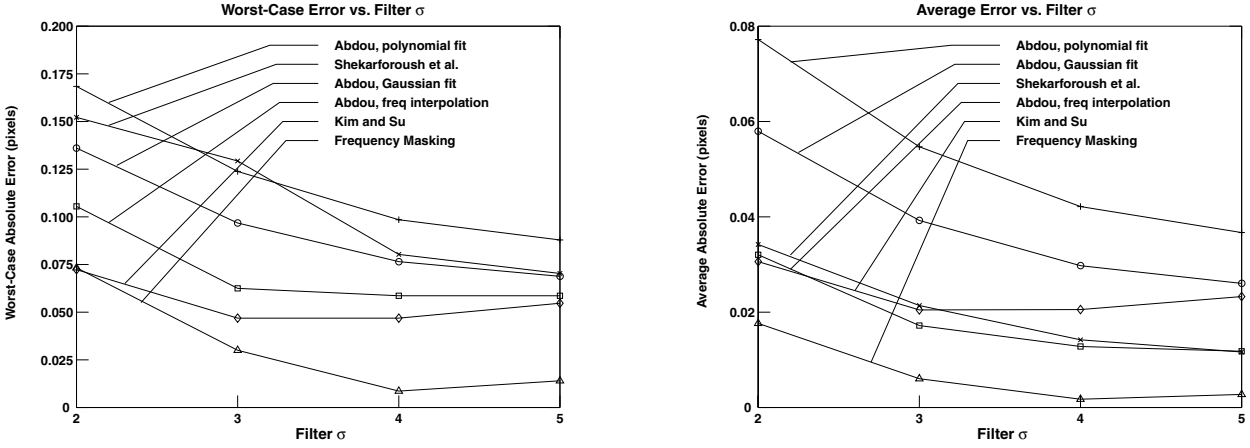
with every possible relative phase shift of the images with respect to each other and with respect to the Blackman window. The results show that the algorithm exhibits virtually no bias with respect to the position of the Blackman window relative to the images for our data. For the comparisons of the frequency-masking algorithm with other algorithms, we used a single aerial photo in 4096 relative pairs of shifts.

To control the amount of aliasing, we used a Gaussian filter prior to downsampling. The filter was characterized by a support area and a σ. Since downsampling by 8 compresses an 8×8 region of pixels into a single pixel, a bandlimiting filter must have a central peak at least 8×8 to keep aliasing small. Our experiments changed σ from 2 to 5 in steps of 1. The size of the central peak is approximately 2σ, so that $\sigma = 2$ produces substantial aliasing and $\sigma = 5$ produces a very small amount of aliasing. The support of the filter was 17×17 for the data reported here. We also explored other ranges of support and σ to confirm that the registration algorithm behaves as expected as we move outside the parameter region studied in detail, and found that the results were consistent with the data reported here.

4.2 Experimental results

The first experiment is a comparison of the precision of our algorithm and other algorithms as a function of σ. For these experiments, we used a single image, and no change of illumination. These results appear in Fig. 3 for worst-case errors over both individual coordinates and average errors per coordinate direction. The mean-square error in two-dimensions is approxi-

mately $\sqrt{2}$ times the average error per coordinate. Figure 4 shows the image used to compare the algorithms.



(a) (b)

Figure 3: (a) Comparison of worst-case error per coordinate for six direct subpixel registration algorithms. (b) Comparison of average error per coordinate for the same six algorithms.

The Kim/Su algorithm is most like ours because it restricts its attention to low frequencies in the Fourier domain. It seeks a phase for which the sum of squares differences between the spectrum of the first image and the phase shifted spectrum of the second image is minimal. Our implementation tuned the algorithm somewhat from the published description. To eliminate spectral leakage, we used a Blackman window, and affirmed that windowing made the registration robust with respect to boundary effects. The data in [30] were taken without a window, although the paper did indicate that windowing should be considered. When we tried this algorithm without win-

Figure 4: Aerial photograph used in the comparisons. (Courtesy of Positive Systems, Inc.)

dowing, the results tended to be more accurate than with windowing, but were highly dependent on how the boundaries of the ground truth were prepared for the experiment. Windowing removes this dependence and removes artifacts of spectral leakage. The frequency region to which we restricted the algorithm had a size 89×89 for a frequency domain of size 128×128. We found experimentally that this size yields the most precise results for the experiment. Our implementation of the algorithm uses iterative hill-climbing to locate the optimal value of the Fourier phase.

The Shekarforoush algorithm estimates the displacements by fitting points to a sinc function, but when the points are corrupted by aliasing, the algorithm produces poor estimates. We found a slight instability in the implementation when the ratio $G(\omega)/F(\omega)$ became large because of a small value of $F(\omega)$, which we removed by artificially setting $F(\omega)$ to unity at this

frequency.

The plot shows three different algorithms proposed by Abdou. The first fits a polynomial curve through points that lie on a correlation peak. The second fits a Gaussian through those points, and the third linearly interpolates the spectra at integer offsets to find the closest approximation to the observed spectrum. The first two algorithms are direct algorithms. The third algorithm iterates a search over the interpolation coefficients. Because the third algorithm operates in the Fourier domain, we used a Blackman window on the images prior to registration. We did not window the data when applying the first two algorithms.

Abdou's spectral interpolation algorithm required special treatment and some modifications. This algorithm computes the spectra of four copies of one image, each displaced relative to the other by displacements that lie on the corners of a unit square in the pixel-domain grid. The algorithm interpolates the cross products of each of these with the spectrum of the second image and finds the interpolation closest to the cross product of the spectra of the two images. Since the central peak dominates the spectral magnitudes, virtually the entire estimate of displacement is due to interpolation of the central peaks. To avoid this problem, we zeroed out the central peaks in the spectra, which is equivalent to normalizing the images to the same average intensities. Nevertheless, the spectral magnitudes near the central peak are very large compared to magnitudes elsewhere in the frequency plane, so that the interpolation depends on relatively few of the frequency coefficients.

All of the algorithms except the Kim/Su algorithm produce higher precision as aliasing decreases.

The iterative algorithm of [21] was studied somewhat differently from the other algorithms. The important aspect of this algorithm is that it drives toward a minimum sum of squared pixel differences between two images by performing a sequence of spline interpolations in the image domain. The spline interpolations are very close to sinc interpolations, and therefore they tend to affect the phase of the corresponding Fourier spectra without changing the magnitudes. Hence, the interpolations do not remove aliasing artifacts that may exist in one or both images as they interpolate one image into the other. For this reason, the iterations are unlikely to be able to drive the differences in the images to zero. In the absence of aliasing, it is clear that the iterations can reduce the sum of squares to near zero, and the point at which this occurs corresponds to the subpixel translation difference of the images. In the presence of aliasing, it is not clear that the subpixel coordinates of the minimum sum of squared pixel differences are the same as those that minimize the subpixel translational difference of the images of the images.

We studied this question experimentally by computing the position of the minimum of the sum of squares of pixel differences for various ground truth shifts in the presence of aliasing. The experiments showed that the minimum occurs at a point consistent with our ground truth and with a precision to that of our frequency masking algorithm. Hence, the iterative algorithm can achieve comparable precision if it can drive interpolations

to the minimum sum of squares value. This was tested and confirmed by Thévenaz et al. who graciously ran their code on a sample image pair supplied by the authors. The observed precision for an image pair filtered with σ of 3.0 was approximately 0.005 pixels, which is approximately the same as the frequency masking algorithm on the same image pair. The question remains open as to why this algorithm achieves this precision with aliased images.

To check the robustness of our algorithm, we tested it on a suite of 52 images with a $\sigma = 3$. In this study, two of the 52 images were outliers, and had insufficient detail in their respective centers to give good worst-case registrations. The Blackman window weighs the center of the images very strongly, and for these images, not enough detail was left after windowing to yield very precise results at the subpixel level. The worst case registrations for these two images had errors that lay between 0.1 and 0.2 pixels in at least one dimension. The remaining 50 had worst-case errors that did not exceed 0.067 pixels in either dimension. The average error in each dimension was on the order of 0.0055 pixels. We repeated the experiment with uniform changes of intensity applied to one of the images, and obtained essentially the same results. Hence, the algorithm is insensitive to such changes of intensity, as the theory suggests it should be.

To deal with the question of how to set thresholds, Fig. 5 plots the displacement estimate as a function of threshold for the registration of a specific image for one offset with $\sigma = 3$. The figure verifies that there is a region of the threshold parameter space for which the displacement is

almost independent of threshold. In this case, we thresholded magnitudes by eliminating frequency components whose magnitudes fell below $\alpha \cdot p_{RMS}$, where p_{RMS} is the RMS magnitude of this spectrum at frequencies that lie in a 5×5 region around the central peak. (Any threshold function that orders the magnitudes and accepts the M largest can be substituted.)

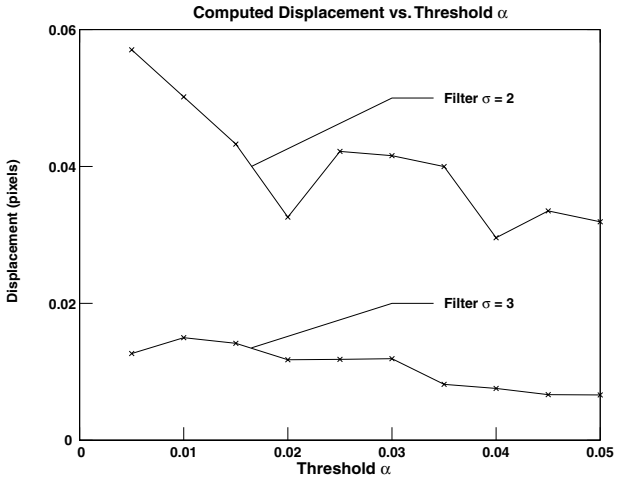
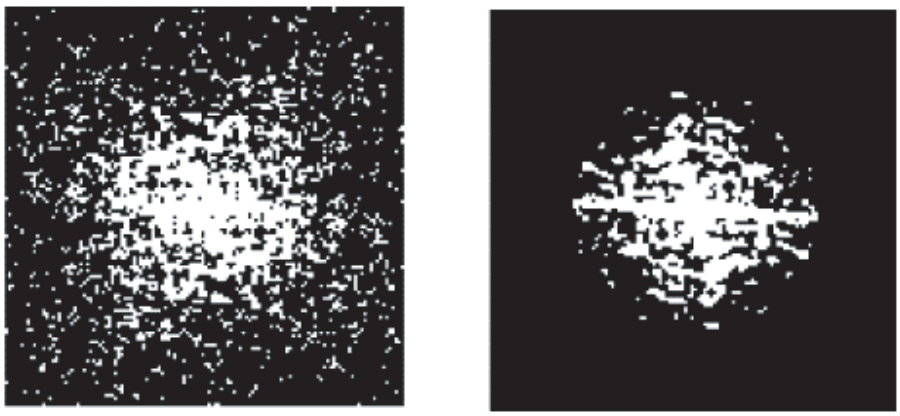


Figure 5: Plot of displacement as a function of threshold variable α.

Fig. 6(a) shows a disparity map in the frequency plane that illustrates how aliasing effects the phase of $G(\omega, \nu)/F(\omega, \nu)$ as a function of (ω, ν). A pixel is colored white in this plane if the phase is very close to the phase for an alias-free image pair, otherwise the pixel is colored black. Note how the white pixels tend to lie near the center of the image. Fig. 6(b) shows the map of frequencies actually used in the registration process. Note that almost all of the white frequencies in Fig. 6(b) are white in Fig. 6(a), which indicates that the algorithm successfully ignores frequency components corrupted by aliasing.

The algorithm also performs well in the presence of noise. Figure 7 plots the precision of the algorithm for an image filtered with $\sigma = 3$ in the



(a) (b)

Figure 6: (a) Disparity map in the frequency domain that shows alias-free components (in white). (b) Frequency components used for registration (white).

presence of additive white Gaussian noise for various signal-to-noise ratios. The horizontal lines in the plot indicate the precision in the absence of noise. The performance is excellent for both the average and worst-case errors. This performance is comparable to the performance reported for the iterative algorithm in [21].

The data presented above indicates that the algorithm is quite robust and accurate over a range of nonideal low pass filters.

5 Conclusions and Open Research

The subpixel registration presented here enjoys very high precision in the presence of aliasing. It is conceptually simple to implement and is very efficient because its complexity is essentially that of two Fourier transforms.

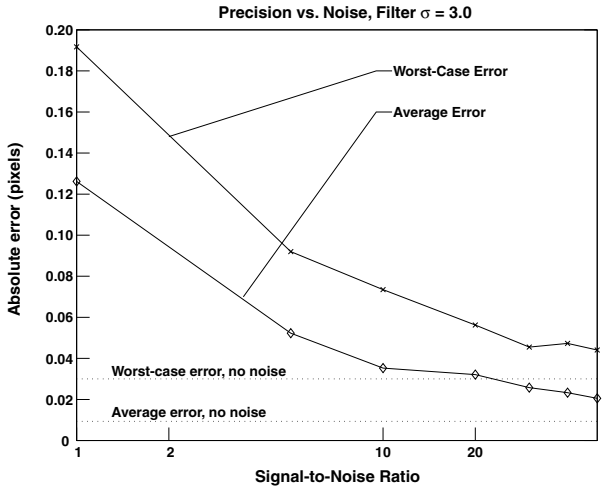


Figure 7: Plot of displacement as a function of signal-to-noise ratio for additive white Gaussian noise.

The other direct algorithms appear to be more sensitive to aliasing, which leads to precision reduction. The iterative algorithm of [21] has the advantage of dealing with rotations and scale changes as well as translations and illumination changes, and appears to yield comparable precision in the presence of aliasing. Our algorithm can be adapted to deal with rotations and scale changes by using Fourier-Mellin invariants as described in [33], but to do so requires interpolations either in the Fourier or in the image domain. The algorithm becomes interative in that setting rather than direct. The additional complexity may not produce an algorithm that is inherently more accurate or more efficient than the algorithm in [21]. Nevertheless, the algorithm is useful in applications in which illumination and translational differences between images have to be discovered accurately and efficiently. Also, the algorithm can be used to find an initial state for driving the iterative algorithms in [2, 21, 30] .

Among the questions that remain open is the question of creating a faster

version of this algorithm by taking smaller Fourier transforms. Since the algorithm uses essentially only the central quadrant of the Fourier domain, it would be useful to produce only those coefficients instead of filling the entire Fourier domain before discarding most of the frequency data.

Another interesting question is to investigate the sum of squares pixel differences criteria for registering two images, and to determine why that criterion seems to give correct results in the presence of aliasing. When registering aliased images in general, is it sufficient to interpolate images during registration with spline or sinc, which do not alter aliasing? Or is it necessary to do a more sophisticated interpolation calculation that removes aliasing artifacts?

Acknowledgment

The Authors thank Positive Systems for the use of the images that formed the basis of the experiments. Also, the authors thank M. Unser, P. Thévenaz, and T. Blu at EPFL for suggestions and comments during the preparation of this paper. We also thank I. E. Abdou and H. Shekarforoush for useful information regarding their algorithms.

References

[1] J. R. G. Townshend, C. O. Justice, C. Gurney, and J. McManus, "The impact of misregistration on change detection," *IEEE Trans. on Geo-*

science and Remote Sensing, vol. 30, no. 5, pp. 1054–1060, September 1992.

[2] I. E. Abdou, "Practical approach to the registration of multiple video images," in *SPIE Conf. On Visual Communications and Image Processing '99*, January 1999, vol. 3563, pp. 371–382.

[3] R. L. Allen, F. A. Kamangar, and E. M. Stokely, "Laplacian and orthogonal wavelet pyramid decompositions in coarse-to-fine registration," *IEEE Trans. on Signal Processing*, vol. 41, no. 12, pp. 3536–3541, December 1993.

[4] M. Berthod, H. Shekarforoush, M. Werman, and J. Zerubia, "Reconstruction of high resolution 3D visual information," in *Computer Vision and Pattern Recognition '94*, June 1994, pp. 654–657.

[5] L. Brown, "A survey of image registration techniques," *ACM Computing Surveys*, vol. 24, no. 4, pp. 325–376, 1992.

[6] M. Corvi and G. Nicchiotti, "Multiresolution image registration," in *Proceedngs 1995 IEEE International Conference on Image Processing*, 1995, pp. 224–227.

[7] I. J. Cox, M. L. Miller, S. W. Omohundro, and P. N. Yianilos, "Target testing and the PicHunter Bayesian multimedia retrieval system," in *Proceedings of Advances in Digital Libraries*, May 1996, pp. 66–75.

[8] A. P. Cracknell and K. Paithoonwattanakij, "Pixel and sub-pixel accuracy in geometrical correction of avhrr imagery," *International Journal of Remote Sensing*, vol. 10, no. 4-5, pp. 661–667, 1989.

[9] J. P. Djamdji, A. Bijaoui, and R. Maniere, "Geometrical registration of images: The multiresolution approach," *Photogrammetric Engineering and Remote Sensing Journal*, vol. 59, no. 5, pp. 645–653, 1993.

[10] D. M. Etter and S. D. Stearns, "Adaptive estimation of time delays in sampled data systems," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-29, no. 3, pp. 582–587, June 1981.

[11] C. M. Fan, N. M. Namazi, and P. B. Penafiel, "A new image motion estimation algorithm based on the EM technique," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 18, no. 3, pp. 348–352, March 1996.

[12] P. L. Feintuch, N. J. Bershad, and F. A. Reed, "Time delay estimation using the LMS adaptive filter – dynamic behavior," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-29, no. 3, pp. 571–576, June 1981.

[13] C. R. Guarino, "A novel method for two-dimensional phase estimation," in *Geoscience and Remote Sensing Symposium, 1994*, August 1994, vol. 4, pp. 2279–2281.

[14] W. R. Hahn and S. A. Tretter, "Optimum processing for delay-vector estimation in passive signal arrays," *IEEE Transactions on Acoustics,*

Speech, and Signal Processing, vol. IT-19, pp. 608–614, 1973.

[15] D. Keren, S. Peleg, and R. Brada, "Image sequence enhancement using sub-pixel displacements," in *Computer Vision and Pattern Recognition '88*, June 1988, pp. 742–746.

[16] J. Le Moigne, W. J. Campbell, and R. F. Cromp, "An automated parallel image registration technique of multiple source remote sensing data," submitted to *IEEE Trans. on Geoscience and Remote Sensing*, 1998.

[17] N. M. Namazi and J. A. Stuller, "A new approach to signal representation with the emphasis on variable time delay estimation," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-35, no. 12, pp. 1649–1660, December 1981.

[18] N. M. Namazi and B. Biswal, "Variable time-delay estimation in colored and correlated noise environment," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 28, no. 2, pp. 597–604, April 1992.

[19] F. A. Reed, P. L. Feintuch, and N. J. Bershad, "Time delay estimation using the LMS adaptive filter – static behavior," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-29, no. 3, pp. 561–571, June 1981.

[20] J. A. Stuller and N. Hubing, "New perspectives for maximum-likelihood time-delay estimation," *IEEE Trans. on Signal Processing*, vol. 45, pp. 513–525, March 1997.

[21] P. Thévenaz, U. E. Ruttimann, and M. Unser, "A pyramid approach to subpixel registration based on intensity," *IEEE Trans. on Image Processing*, vol. 7, pp. 27–41, January 1998.

[22] Q. Zheng and R. Chellappa, "A computational vision approach to image registration," *IEEE Trans. on Image Processing*, vol. 2, no. 3, pp. 311–326, 1993.

[23] S. Alliney, "Digital analyis of rotated images," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 15, no. 5, pp. 499–504, May 1993.

[24] S. Alliney, G. Cortelazzo, and G. A. Mian, "On the registrations of an object translating on a static background," *Pattern Recognition*, vol. 29, no. 1, pp. 131–141, January 1996.

[25] S. Alliney and C. Morandi, "Digital image registration using projections," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 8, no. 2, pp. 222–233, March 1986.

[26] P. E. Anuta, "Spatial registration of multispectral and multitemporal digital imagery using fast-fourier transform techniques," *IEEE Trans. on Geoscience Electronics*, vol. 8, no. 4, pp. 353–368, October 1970.

[27] Q.-S. Chen, M. Defrise, and F. Deconinck, "Symmetric phase-only matched filtering of Fourier-Mellin transforms for image registration and recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 12, pp. 1156–1168, December 1994.

[28] B. Dasgupta and B. N. Chatterji, "Fourier Mellin transform based image matching algorithm," *Journal of the Institution of Elect. and Telecommunications Engineers*, vol. 42, no. 1, pp. 3–9, 1996.

[29] E. De Castro and C. Morandi, "Registration of translated and rotated images using finite Fourier transforms," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. PAMI-9, no. 5, pp. 700–703, 1987.

[30] S. P. Kim and W. Y. Su, "Subpixel accuracy image registration by spectrum cancellation," in *Proc. ICASSP '93*, April 1993, vol. V, pp. 153–156.

[31] C. D. Kuglin and D. C. Hines, "The phase correlation image alignment method," in *IEEE 1975 Conference on Cybernetics and Society*, September 1975, pp. 163–165.

[32] D. V. Papadimitriou and T. J. Dennis, "Stereo disparity using phase correlation," *Electronics Letters*, vol. 30, no. 18, pp. 1475–1477, September 1994.

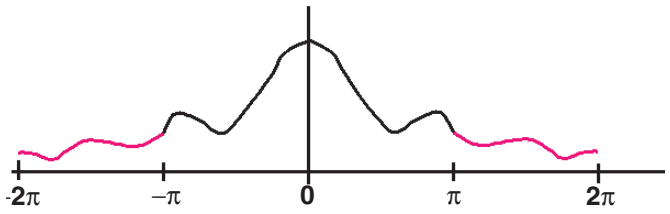
[33] B. S. Reddy and B. N. Chatterji, "An FFT-based technique for translation, rotation, and scale-invariant image registration," *IEEE Trans. on Image Processing*, vol. 3, no. 8, pp. 1266–1270, August 1996.

[34] H. Shekarforoush, M. Berthod, and J. Zerubia, "Subpixel image registration by estimating the polyphase decomposition of cross power spectrum," in *Computer Vision and Pattern Recognition '96*, June 1994, pp. 532–537.

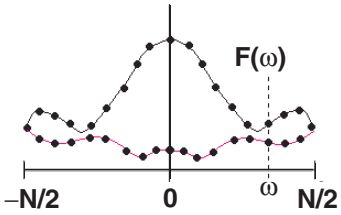
[35] H. S. Stone, "Progressive wavelet correlation using Fourier methods," *IEEE Transactions on Signal Processing*, vol. 47, no. 1, pp. 97–107, Jan. 1999.

[36] J. W. Wong and E. L. Hall, "Scene matching with invariant moments," *Computer Graphics and Image Processing*, vol. 8, pp. 16–24, 1978.

[37] F. J. Harris, "On the use of windows for harmonic analysis with the discrete Fourier transform," *Proceedings of the IEEE*, vol. 66, no. 1, pp. 51–83, January 1978.



(a)



(b)

Figure 1: (a) Spectrum of image after prefiltering. (b) Downsampling and resulting aliasing of prefiltered image.

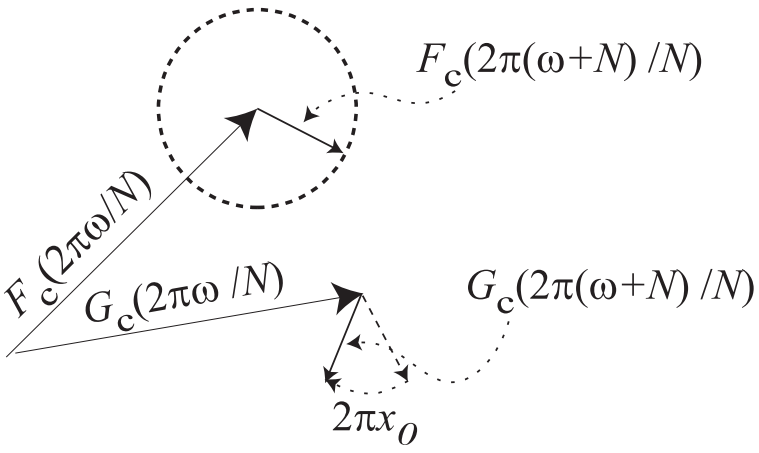
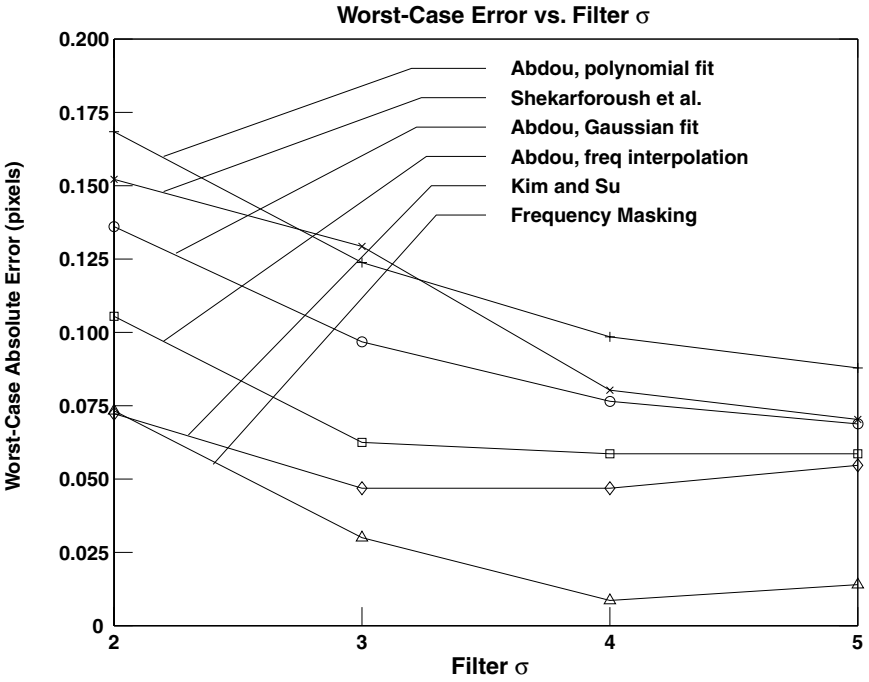
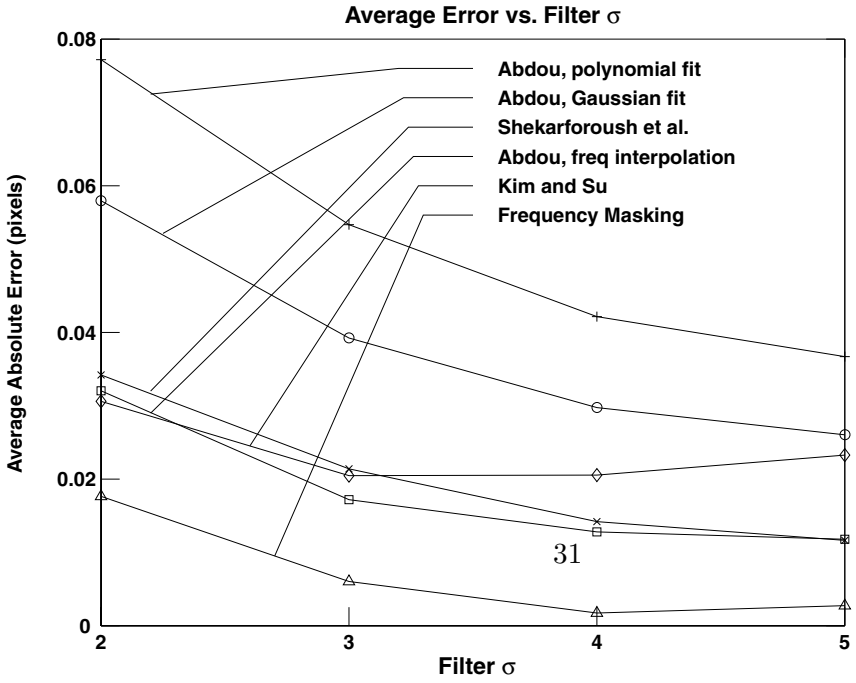


Figure 2: Vector illustration of aliasing



(a)



(b)

Figure 3: (a) Comparison of worst-case error per coordinate for six direct subpixel registration algorithms. (b) Comparison of average error per coordinate for the same six algorithms.

Figure 4: Aerial photograph used in the comparisons. (Courtesy of Positive Systems, Inc.)

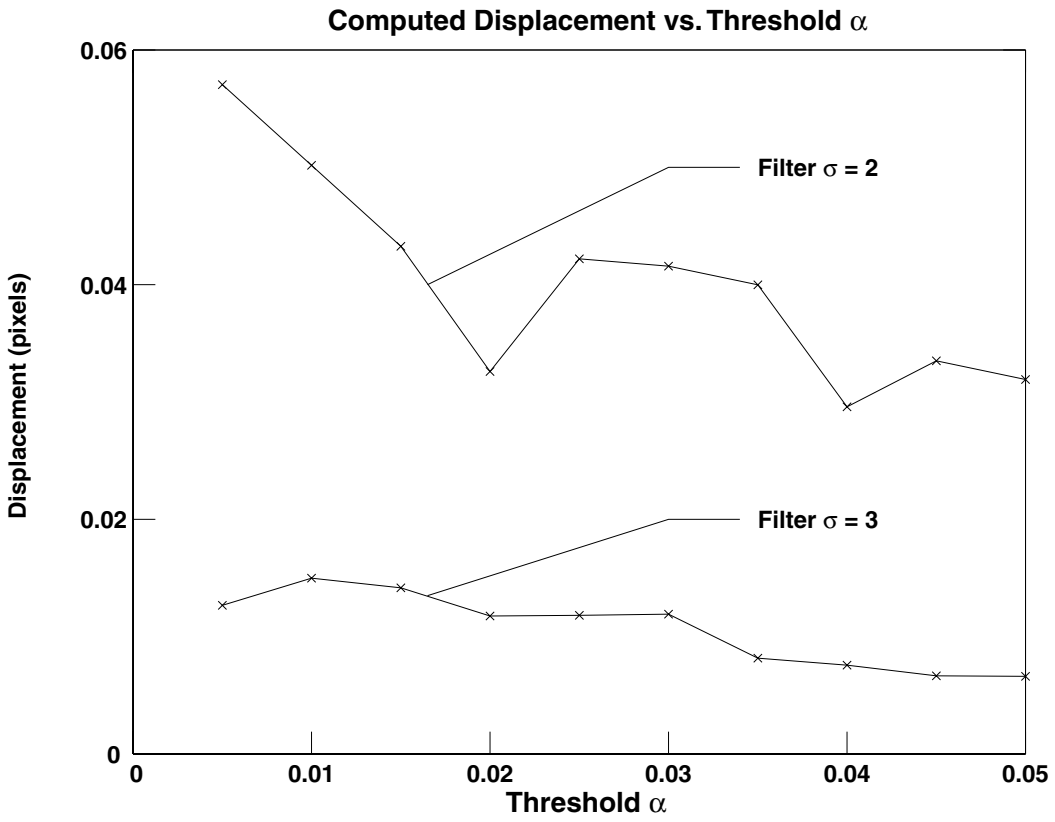
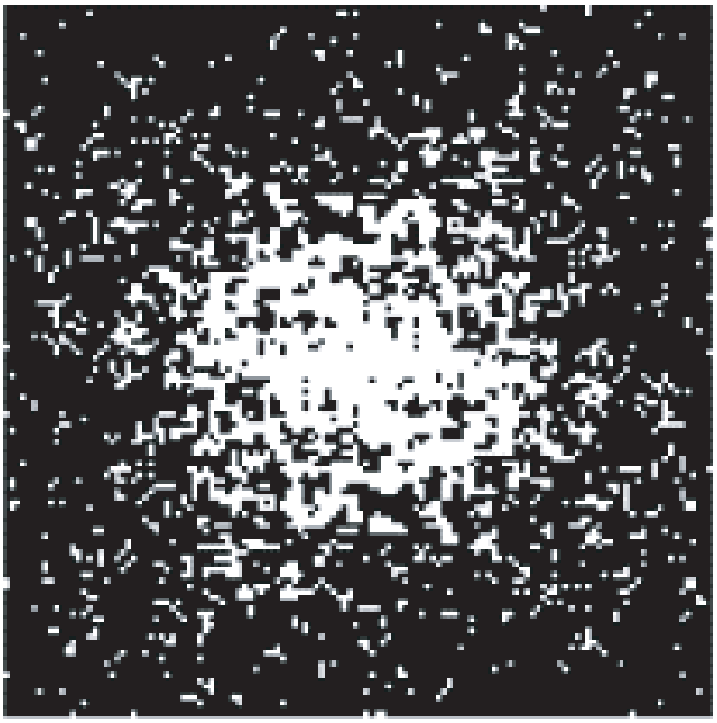
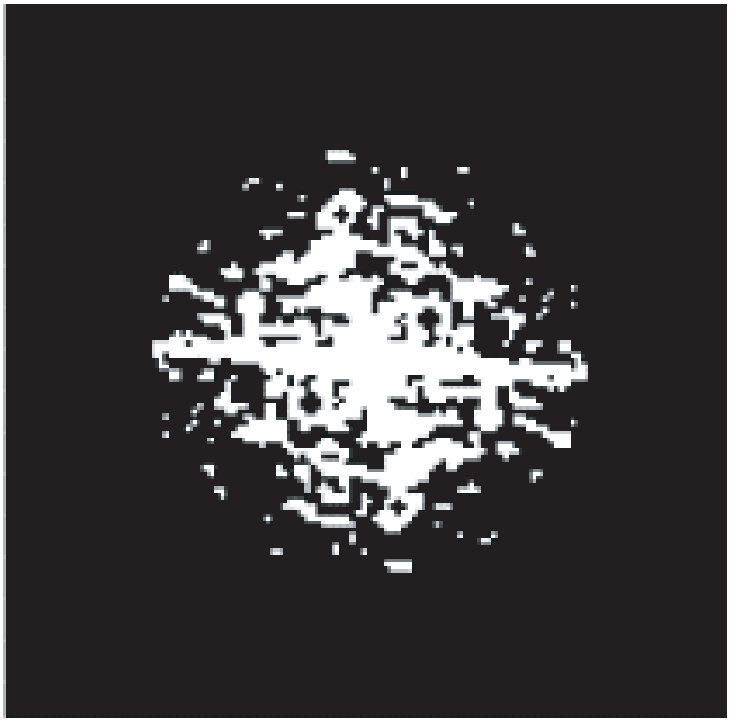


Figure 5: Plot of displacement as a function of threshold variable α.



(a)



(b)

Figure 6: (a) Disparity map in the frequency domain that shows alias-free components (in white). (b) Frequency components used for registration (white).

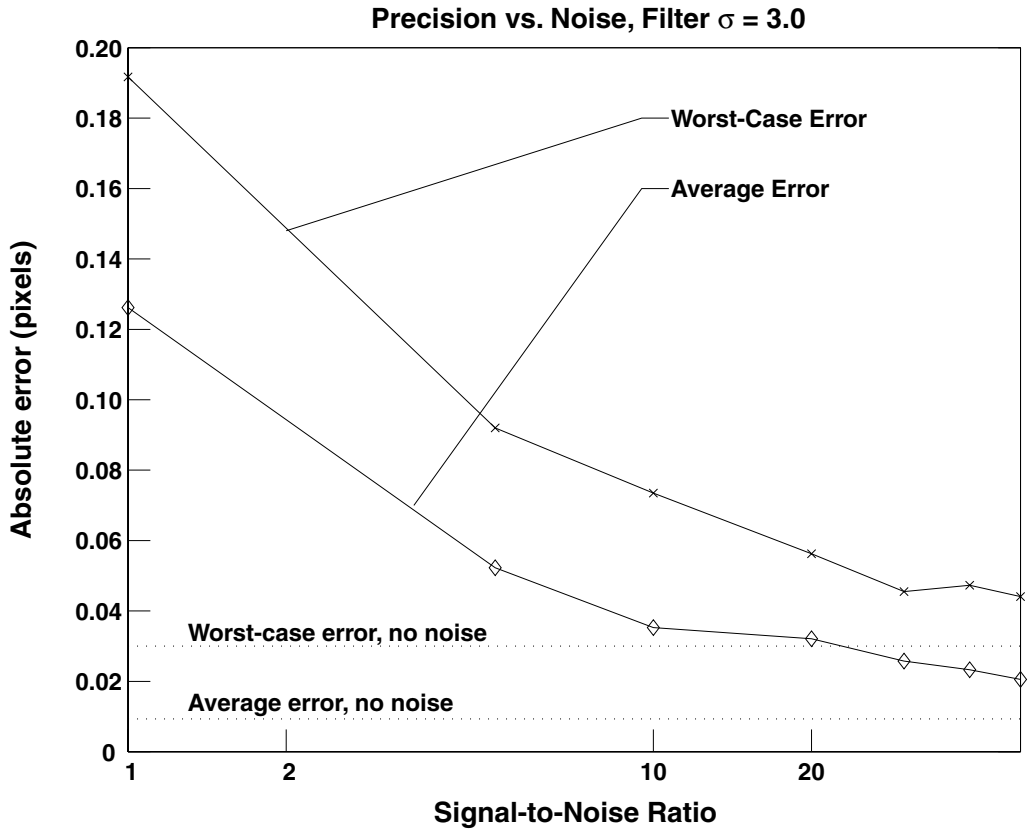


Figure 7: Plot of displacement as a function of signal-to-noise ratio for additive white Gaussian noise.