Subpixel Registration of Images

Harold Stone
NEC Research Institute
Princeton, NJ 08540
hstone@research.nj.nec.com

Michael Orchard
Princeton University
Princeton, NJ 08540
orchard@ee.princeton.edu

Ee-chien Chang
NEC Research Institute
Princeton, NJ 08540
changec@cz3.nus.edu.sg

Abstract

Phase correlation has been studied as a tool for registering pairs of images to subpixel resolution. Shekarforoush et al investigated an algorithm that works well for when observed data are samples of ideally bandlimited images. In fact, images of real scenes captured by modern optics are not ideally bandlimited images. This paper examines how aliasing in sampled images effects the ideal phase relationships between continuous images undergoing translational shifts. We show that the performance of standard phase correlation methods for subpixel image registration can be severely degraded with moderate amounts of aliasing, and we investigate two approaches for modeling aliasing effects in order to improve subpixel image registration accuracy. The first approach, a very simple algorithm, both conceptually and computationally, is based on detecting those frequency components that have become unreliable estimators of shift due to aliasing, and removing the detected components from the shift estimate. The second approach, more ambitious and more complex, attempts to undo the effects of aliasing and to use all de-aliased frequency components in the shift estimator.

1. Introduction

Image registration is an important preprocessing task for many image processing problem that fuse information from multiple images or multiple modalities of imaging. In the image registration problem, 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 (perhaps also disturbed by independent additive noise). This paper addresses the problem of estimating this offset at subpixel accuracy, and using this estimate to register the second image to the grid of the first image.

This problem has been attacked in the signal or pixel domain [1, 6, 7, 9, 10, 11, 14, 15, 16, 17, 18, 19, 20, 22, 24, 23, 27, 30, 31, 33] and in the Fourier domain [2, 3, 4, 5, 7, 8, 12, 13, 21, 25, 26, 28, 29, 32]. Our work follows the latter body, and estimates the shift from basic phase relationships between the Fourier transform of the two images. However, unlike previous work, we do not assume that each observed image represents alias-free samples of an underlying continuous image. In fact, point spread functions of practical optics *almost never* bandlimit the high-frequency components of occluding edges in the scene sharply enough to support this alias-free assumption. Instead, our work explicitly models the aliased frequency components of the two images and predicts how this aliasing affects the phase relationships between their Fourier transforms. We find that aliasing causes some frequency components of the scene to be unreliable for estimating the phase relationships between the two images.

Previous algorithms ignore the effects of aliasing and incorporate these unreliable frequency components into the shift estimates (or time-delay estimates in earlier work), producing biased estimates. We develop two new approaches to subpixel image registration based on modeling the effects of aliasing in each image, and use this model to improve the accuracy of the shift estimate. The first approach identifies a set of frequency components that are sufficiently free from aliasing to avoid biasing an estimate of the shift, and develops an algorithm for estimating subpixel shift from this set of reliable frequency components. This approach leads to a very simple algorithm based on masking out certain frequency components that do not satisfy a simple test condition.

The second approach studied in this paper inverts the effects of aliasing on all frequency components, and then uses all de-aliased components (even those at frequencies higher than the Nyquist bandwidth of the sampled images) to estimate the subpixel shift. The de-aliasing operation can be recognized as equivalent to the problem of interpolating an image at a higher sampling rate based on the two measured images at the lower sampling rate. If the underlying image were bandlimited, standard approaches to interpolation could be applied, but, as noted earlier, this paper is mainly concerned with practical sampled images, which are *rarely* (if ever) alias-free. Accordingly, we propose a novel interpolation method that uses the estimated shift to model each observed Fourier component as an aliasing of two components. (Here we assume that the signal spectral energy is decaying sufficiently fast in frequency Ω, so that the main source of aliasing at each frequency Ω_i comes from only one aliasing component at $\Omega + 2\pi/T$.) By relating two observations of this model (one for each image), we compute an estimate of the underlying aliasfree Fourier components. The resulting estimate of the aliasfree Fourier transform of the image allows more accurate interpolation (and registration) of the second image.

Section 2 sketches the basic mathematical relationships that motivate the approaches of all later sections. Section 3 derives and discusses our algorithm for estimating shift using only a set of "reliable" frequency components extracted from the two images. Section 4 describes our approach for estimating the aliasfree signal components. Finally, Section 5 provides both 1-D and 2-D simulations of our approach, demonstrating significant improvements in registration accuracy for images with aliasing.

2. Mathematical Background

This section reviews basic signal processing relationships that motivate our approach to image registration. To minimize notation clutter, this section characterizes 1-D signals, and later sections take obvious, separable approaches to generalize the 1-D relationships to develop 2-D algorithms.

Let $f_c(x)$ denote a continuous function with Fourier transform $F_c(\Omega)$. The shifted function $g_c(x) = f_c(x - x_0)$ has Fourier transform $G_c(\Omega) = e^{-j\Omega x_0} F_c(\Omega)$. This relationship motivates all Fourier approaches to image registration. Given $f_c(x)$ and $g_c(x)$, related by a shift x_0, the shift can be estimated by comparing the phases of their transforms. In fact, in this idealized setting (with continuous functions and no noise) any single pair of Fourier components $F_c(\Omega')$ and $G(\Omega')$ provides a perfect estimate of x_0, modulo an ambiguity of integer multiples of $(2\pi/\Omega')$.

For our problem, the observed data consists of sampled sequences: $f(n) = f_c(nT)$ and $g(n) = g_c(nT)$, with DTFT's $F(\omega)$ and $G(\omega)$, respectively. If we assume perfect bandlimited signals (e.g. $F_c(\Omega) = 0$, for $|\Omega| > \pi/T$), an arbitrary shift $x_0 \in (0, T)$ is the phase difference of any single pair of DTFT components $F(\omega')$ and $G(\omega')$ (again noting an ambiguity of $(2\pi/\Omega')$). In an ideal setting of bandlimited signal with no noise, there are many different approaches that all perform equally well in correctly estimating x_0. All such methods effectively divide $G(\omega)$ by $F(\omega)$ for some set of ω and derive x_0/T from the resulting phases. However, in the presence of noise and aliasing, such simple divisions do not always yield reliable estimates of shift.

Several estimates of x_0 have been studied. Shekarforoush, et al. [28] suggest computing a cross-power spectrum $R(\omega) = G(\omega)/F(\omega)$ and finding a least-squares fit of its inverse transform $r(n)$ to $\frac{\sin(x-x_0)}{(x-x_0)}|_{x=nT}$. In the noiseless aliasfree case, $R(\omega)$ has unit magnitude for all ω, and thus $\frac{\sin(x-x_0)}{(x-x_0)}|_{x=nT}$ describes the expected form of $r(n)$ for different shift values, x_0. Simple experiments show that if either $F_c(\Omega)$ is not perfectly bandlimited or noise is added, $R(\omega)$ rarely has unit magnitude for all ω. Here we consider the case of nonbandlimited $F_c(\Omega)$. In this case, aliased components with $|R(\omega)| \neq 1$ bias the estimate of Shekarforoush. Aliasing often causes $F(\omega)$ to be very small (i.e. when the vector addition of aliased components cancel), causing $|R(\omega)| >> 1$. In these cases of large $|R(\omega)|$, the phase distortions due to aliasing have a large influence in biasing the shift estimate.

A more common approach used in the time-delay estimation literature is to extract only the phase from $R(\omega)$, forcing all frequency components to unit magnitude. These methods are based on computing the normalized cross power spectrum:

$$\frac{F(\omega)G^*(\omega)}{|F(\omega)||G(\omega)|} = \frac{R(\omega)}{|R(\omega)|}.$$

By normalizing the spectrum, this approach limits the bias introduced by any small set of phase-distorted frequency components on the shift estimate, x_0.

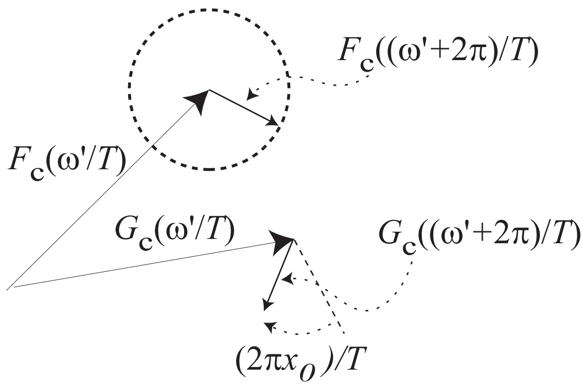

Figure 1. Vector illustration of aliasing

To motivate the approach developed in this paper, we consider the aliasing of a single frequency component at ω' and examine how aliasing biases the phase estimate $R(\omega')$. Figure 1 shows the vectors $F_c(\omega'/T)$ and $G_c(\omega'/T)$, equal in magnitude and offset by phase $(\omega'/T)x_0$, representing the aliasfree components of the continuous transform. They satisfy the expected magnitude and phase relationships perfectly. For clarity, consider the case where only a single aliasing component from $(\omega' + 2\pi)/T$ distorts each of the vectors $F_c(\omega'/T)$ and $G_c(\omega'/T)$. If the two distorting vectors are offset in phase by the same amount, $(\omega'/T)x_0$, the presence of aliasing introduces no bias in the phase estimate $R(\omega')$ since the magnitude and phases of both aliasfree vectors change by the same amount. However, because the aliasing components come from $(\omega' + 2\pi)/T$, the shift of x_0 produces a phase offset of $(\omega' + 2\pi)x_0/T$, or a phase offset of $2\pi x_0/T$ *more* than that of the aliasfree components. In general, the phase of the aliasing component (before shift) is independent of the aliasfree component. We model it as uniformly distributed on the circle shown at the head of vector $F_c(\omega'/T)$ in Figure 1. The vector $F_c((\omega'+2\pi)/T)$ shown in the circle is a sample vector drawn from that distribution. The vector added to $G_c(\omega'/T)$ is then this same sample vector rotated by $(\omega'+2\pi)x_0/T$. The $2\pi x_0/T$ portion of this rotation is what produces both magnitude and phase distortion. In the phase-correlation based methods, the phase distortions bias the shift estimator.

In this paper we consider two approaches to removing bias due to phase distortions. Our first approach is motivated by a very simple observation. Assuming that the aliasing component (before shift) is uniformly distributed in phase, and noting that the aliasing components added to vectors $F_c(\omega'/T)$ and $G_c(\omega'/T)$ differ in phase by $2\pi x_0/T$ more than the phase difference of the vectors themselves, it is highly likely that any phase distortion will be accompanied by magnitude distortion. Note that there are combinations of aliasing vectors and shift values, x_0 that keep the ratio $R(\omega)$ at unity while distorting the phase, but the probability of such combinations is very small. Drawing from this observation, our algorithms evaluate $|R(\omega)|$, and interpret deviations from unity as an indicator of an unreliable phase estimate. The next section develops an algorithm that masks these phase-distorted frequency components from our estimate of the shift, x_0.

Our second approach for dealing with phase distortion due to aliasing attempts to invert the effects of aliasing on all frequency components of the signal. In section 4 we develop an algorithm for reconstructing a continuous signal $F_c(\Omega) = H(\Omega)$, where $H(\Omega)$ models a signal with bandwidth $\Omega \in [-2\pi/T, 2\pi/T]$ twice that of either of the observed sequences $f(n)$ and $g(n)$. (Note: in modeling each sequence as undersampled by only a factor of 2 relative to the Nyquist rate for $H(\Omega)$, we assume that the first aliasing frequency is the dominant source of signal aliasing, and we ignore higher-order aliasing terms.) Our reconstruction algorithm estimates $H(\Omega)$ to be consistent with both observed sequences $f(n)$ and $g(n)$. By consistent we mean that if $h(t)$, the inverse transforms of $H(\Omega)$ sampled at shifts of zero and $\hat{x}$ respectively, are equal to $f(n)$ and $g(n)$, respectively. The value $\hat{x}$ then serves as the estimate of shift, x_0. The main difficulty with this approach is that every arbitrary shift x' has a corresponding $H'(\Omega)$ consistent with the observed sequences $f(n)$ and $g(n)$. We develop an energy-based criterion, relying on the underlying stationarity of the signal, to select among these consistent estimates of shift.

3 Frequency Masking Subpixel Shift Estimation

Motivated by our characterization of aliasing in the previous section, we propose a simple algorithm for estimating subpixel shift based on reliable frequency components of the signal. For our description, we assume that the integer shift has been estimated by some other means. In fact, the estimation of integer shift can be integrated in the various steps of the algorithm described below. For example, a common approach would be to compute the Fourier transform ratios of steps 1) and 2) below, and identify the location of the peak of its inverse Fourier transform. For the purpose of this paper, we assume that such processing has already been completed before estimating subpixel shift.

1) Compute the two Fourier transforms: $F(\omega)$ and $G(\omega)$.

2) Compute the ratio $R(\omega) = \frac{F(\omega)}{G(\omega)}$.

3) Find the largest set, $\mathcal{B} = \{-\omega_c, \omega_c\}$, with cutoff ω_c satisfying the property: $\frac{1}{(1+\rho)} < |R(\omega)| < (1+\rho)$ for all $\omega \in \mathcal{B}$. The value $\rho = 0.1$ is a reasonable choice. To avoid incorporating highly aliased values far from the origin, bound $\mathcal{B}$ to less than 1/8th the image size in each dimension.

4) Find the maximum value of the magnitudes of $R(\omega)$ for ω on the perimeter of $\mathcal{B}$. The perimeter is the set of ω for which $|\omega| = \omega_c$. Let the maximum magnitude value be the *magnitude discriminant* D.

5) Delete from $\mathcal{B}$ all elements ω for which $|R(\omega)| < \alpha D$. The value $\alpha = 0.1$ is a suitable choice.

6) Compute a least-squares linear fit of the phases in set $\mathcal{B}$. Our estimate of x_0 is given by the slope of this line.

The strategy implemented in steps 3)-5) represents a tradeoff between resolution and bias. While it is reasonable to eliminate aliased components from the estimate of subpixel shift, the resolution in time that can be achieved based on a bandlimited data-set with bandwidth ω_c increases with ω_c. Thus, the cost of biasing the estimate with aliased frequency components must be balanced against the cost of reducing time resolution from omitting frequency components. The magnitude discriminant removes additional frequency components whose low relative magnitudes indicate that they are highly aliased.

In our 2-D algorithms tested in section 5, we constrain the set $\mathcal{B}$ to be rectangular (in 2-D) low-pass supported set of frequency components defined by cutoffs frequencies. The cutoffs are indirectly controlled by the parameter ρ, which adjusts our tolerance on the magnitude of the spectral ratio $R(\omega)$.

4 Subpixel Registration Based on De-aliasing

In contrast to the previous section which proposes to estimate phase by ignoring frequency components that are highly aliased, this section proposes to undo the effects of aliasing, allowing all components to be used in the shift estimate. This section outlines our model for the underlying continuous signal transform $H(\Omega)$, and illustrates how aliasing effects can be removed, given the values of the two sequences $f(n)$ and $g(n)$ and the true shift x_0. $H(\Omega)$ has the property that it is bandlimited to $[-2\pi/T, 2\pi/T]$, and, that it is the spectrum of a function whose periodic samples taken T units apart starting at $t = 0$ form the sequence $f(n)$, and those starting at $t = x_0$ form the sequence $g(n)$. Unfortunately, we find that this de-aliased signal is not unique in that, starting from another arbitrary shift x', another continuous signal transform $H'(\Omega)$ can be found that is also bandlimited and also produces $f(n)$ and $g(n)$ when sampled with period T starting at $t = 0$ and $t = x'$. Thus, without either knowing the true shift or applying some other criterion, it is not possible to distinguish the correct de-aliased signal from other candidate de-aliased signals. Recognizing this, we propose a criterion for selecting one of these candidate de-aliased signals, and its corresponding shift, as our signal and shift estimates. Our proposed criterion reflects our belief that the underlying signal is a stationary random process. We present an empirical study of the effectiveness of our criterion for estimating the true shift x_0.

As noted earlier, we model our observed sequences $f(n)$ and $g(n)$ as samples of a continuous signal $h_c(t)$ with sampling period T, whose transform is $H(\Omega)$. $H(\Omega)$ is bandlimited to $[-2\pi/T, 2\pi/T]$. Thus, the components of $f(n)$ and $g(n)$ at frequency ω can be written:

$$\begin{bmatrix} F(\omega) \\ G(\omega) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ e^{-j\omega\frac{x_0}{T}} & e^{-j(\omega+2\pi)\frac{x_0}{T}} \end{bmatrix} \begin{bmatrix} H(\frac{\omega}{T}) \\ H(\frac{\omega-2\pi}{T}) \end{bmatrix},$$

for $0 < \omega < \pi$. The inverse relation gives:

$$\begin{bmatrix} H(\frac{\omega}{T}) \\ H(\frac{\omega-2\pi}{T}) \end{bmatrix} = C \begin{bmatrix} e^{-j(\omega+2\pi)\frac{x_0}{T}} & 1 \\ -e^{-j\omega\frac{x_0}{T}} & 1 \end{bmatrix} \begin{bmatrix} F(\omega) \\ G(\omega) \end{bmatrix}, \quad (1)$$

where $C = 1/(e^{-j(\omega+2\pi)\frac{x_0}{T}} - e^{-j\omega\frac{x_0}{T}})$. The inverse exists for all $0 < x_0 < T$ (i. e., for all subpixel shifts). If $f(n)$ and $g(n)$ are length N real sequences, each specified by $N/2 - 1$ complex and 2 real DFT components, the inverse relation applied to each pair of components produces $2N$ components of $H(\Omega)$. These components can either be viewed as samples of the continuous Fourier transform, or as values of a DTFT, $H_d(\omega)$, specifying a length-2N sequence, $h_d(n)$, with the properties:

- The even samples of $h_d(n)$ equal $f(n)$.
- Perfect bandlimited interpolation of $h_d(n)$ to perform a shift by $2\frac{x_0}{T}$, followed by downsampling by 2, yields the sequence $g(n)$.

The sequence $h_d(n)$ is the sequence that results from sampling the original continuous signal $h(t)$, bandlimited to $[-2\pi/T, 2\pi/T]$, without any shift, with period $T/2$.

Note that the reconstructed $H(\Omega)$ offers a model for the sequences $f(n)$ and $g(n)$ that suffers from no phase distortion in that direct sampling of $h(t)$ exactly matches $f(n)$, and perfect time shift by x_0 followed by sampling exactly matches $g(n)$. (This assumes no additive noise or aliasing from frequencies beyond $2\pi/T$). Unfortunately, the properties above do not characterize $H(\Omega)$ uniquely. In fact, all $x \neq x_0$ produce an $H(\Omega)$ that satisfy the properties. Our question becomes: from all the possible shift/aliasing models of the observed sequences, how do we recognize the correct model?

To answer this question, we propose an additional criterion of merit for the collection of candidate reconstructions and shifts, based on a statistical model for the space of original signals. We assume that the original signal $h(t)$ is a realization from a *stationary* Gaussian random process. Though we believe the criterion is applicable to a much wider class of source signals, this model offers a reasonable model for a broad class of signals, and it greatly simplifies our discussion. Due to stationarity, the Fourier components of $H(\Omega)$ are independent complex random variables, with independent phases. To select among candidate reconstructions and shifts, we propose to find the reconstruction that is most likely to have been generated by a stationary random process. In particular, we propose to measure the energy of each candidate reconstruction spectrum (viewed as a function of the corresponding shift $x \in (0, T)$,) and consider the likelihood that the two sequences $f(n)$ and $g(n)$ come from a stationary process with the specified energy. We defer a complete statistical characterization of this energy function to a more detailed manuscript in preparation, and in this paper present an empirical study of its properties for stationary Gaussian processes.

Let $\mathcal{E}_{f,g}(x) = \frac{1}{2N}(\sum_\omega\{|H_x(\omega)|^2\})$ denote the energy of the candidate reconstruction for a given assumed shift x, where this energy function is parameterized by the observed sequences f and g, which depend, in turn, on the true original function $H(\Omega)$ and true shift x_0. If we let $\mathcal{E}(f, g) = \frac{1}{2N}(\sum_n\{|f(n)|^2 + |g(n)|^2\})$ represent the average energy of the observed sequences, Parseval's theorem implies $\mathcal{E}_{f,g}(T/2) = \mathcal{E}(f, g)$. That is, conditioned on f and g, $\mathcal{E}_{f,g}(T/2)$ does not depend on the specific realization H. In contrast, for $x \in (0, T/2)$, $\mathcal{E}_{f,g}(x)$ is a random variable depending on H, but with mean $\mathcal{E}(f, g)$. To see this, note that $\mathcal{E}_{f,g}(x_0) = \frac{1}{2N}\sum_n\{|h_d(n)|^2\}$ is the average energy of $f(n)$ and a sequence $g_{0.5}(n)$ which is the sequence produced by sampling the shifted original continuous $h(t+T/2)$ with period T. Although $g_{0.5}(n)$ and $g(n)$ do not, in general, have the same energy, due to stationarity, $\frac{1}{N}\sum_n\{|g(n)^2|\}$ is a good, unbiased estimator of the average energy of $g_{0.5}(n)$. Other features of $\mathcal{E}_{f,g}(x)$ are:

- It is convex on the interval $x \in (0, T)$.
- It tends to infinity as $x \to 0$ and $x \to T$.
- It reaches a unique minimum somewhere between $x = T/2$ and $x = x_0$.

Figure 2 shows examples of energy functions $\mathcal{E}_{f,g}(x)$ for H drawn from a stationary Gaussian random process, and $x_0 = 3T/16$, $x_0 = 8T/16$, and $x_0 = 11T/16$. All three curves illustrate the characteristics of $\mathcal{E}_{f,g}(x)$ noted above.

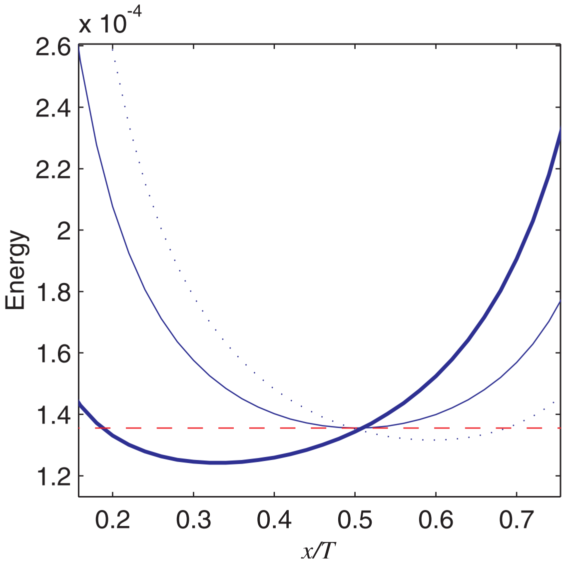

Figure 2. $\Sigma_{H,x_0}(x)$ **for** $x_0 = 3T/16$**,** $x_0 = 8T/16$**, and** $x_0 = 11T/16$

Drawing from these observations, we propose to estimate shift x_0 by finding the value of $x \neq T/2$ satisfying

$$\mathcal{E}_{f,g}(x) = \mathcal{E}(f, g).$$

If no such value exists (i.e. if $x = T/2$ is the only such value), then our estimate is $T/2$. Since $\mathcal{E}_{f,g}(x_0)$ is a random variable with nonzero variance, this estimator is a random variable whose statistics depend on the characteristics of the random process from which $H(\Omega)$ is drawn. The following section presents an empirical evaluation of the proposed estimator for a particular source power spectral density and a particular shift value x_0.

5 Simulations

The simulation of the algorithm of Section 3 generated excellent results for a contrived 2-D image whose frequency spectrum contains a small of amount of aliasing. The image is a circle of radius 249 in a 1024 $\times$ 1024 square. To create two "subpixel" shifts the image, the first image was downsampled by a factor of 8, and the second image was shifted relative to the first up to 4 pixels in x and y, and then downsampled. The down sampling process created each pixel in the reduced coordinate system by averaging pixel intensities over a square region of size 7 $\times$ 7 in the original image.

To examine all distinct relative phases of the two images, the experiment contains all combinations of 81 shifts of the first image (with x and y varying from -4 to + 4 pixels), and for each of those, all 81 shifts of the second image relative to the first. For each pair of shifts, we used the first algorithm to locate the set of spectral components from whose phases we estimated the displacement.

To compute the displacement estimates, let $(phase(\omega, \nu)$ be the phase of the complex ratio $G(\omega, \nu)/F(\omega, \nu)$ at point $(\omega, \nu) \in \mathcal{B}$ as calculated by the algorithm. The slope of the plane that passes through the points of $(phase(\omega, \nu)$ in a phase plot has x and y slopes equal to $2\pi x_0/N$ and $2\pi y_0/N$, respectively, where x_0 and y_0 are the relative displacements of the two images expressed in fractions of a pixel. Since, in general, the phase values do not fall exactly on a plane because of aliasing and other factors, we used a least squares algorithm to estimate the slope. The computation uses the variables $\overline{\omega^2} = (1/|\mathcal{B}|)\sum_i\sum_j \omega_i^2$, $\overline{\nu^2} = (1/|\mathcal{B}|)\sum_i\sum_j \nu_j^2$, $\overline{\omega\nu} = (1/|\mathcal{B}|)\sum_i\sum_j \omega_i\nu_j$, $\overline{\omega phase} = (1/|\mathcal{B}|)\sum_i\sum_j \omega_i phase(\omega_i, \nu j)$, and $\overline{\nu phase} = (1/|\mathcal{B}|)\sum_i\sum_j \nu_j phase(\omega_i, \nu j)$ where the sums are taken over all pairs $(\omega, \nu) \in \mathcal{B}$. Using these, the least squares estimates of the slope of a plane that passes through the origin is:

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

The results of the experiments was the worst case relative error was 1.74% for our algorithm as compared to a worst-case relative error of 2,387% for the SBZ algorithm [28]. If we ignore the outlier worst-case, the next-worst case is 341.5%. The average relative error over the 4096 distinct pairs of shifts was 0.20% for the new algorithm and 23.57% for the SBZ algorithm. (With the outlier removed, the average relative error for the SBZ algorithm drops to 19.9%). For bandlimited images, both algorithms give excellent results. Both algorithms fail to give accurate results when aliasing becomes significant. The simulation reported here shows that techniques that identify and ignore aliased components work well when aliasing effects are not very large.

In a few cases, the value $\mathcal{B}$ discovered by the new algorithm would have grown to maximum size in the absence of bounding to 1/8th of the image dimensions. Using large regions for $\mathcal{B}$ produced relative errors up to 5%. By bounding to a smaller region (in this case to 23×23 out of a possible 128×128) the few cases of relative errors above 2% disappeared.

For the algorithm of Section 4, we generated a discrete stationary Gaussian random process with power spectral density

$$\Phi_H(\omega) = \begin{cases} Ke^{-\omega^2/\gamma} & \text{if } |\omega| \leq 2\pi/32; \\ 0 & \text{if } |\omega| > 2\pi/32. \end{cases}$$

The constant γ is set to make the high-frequency energy roll off slowly to about -18dB at $\pi/32$ and -45dB at $2\pi/32$ (relative to 0dB at $\omega = 0$). (Note: since all $|\omega| > \pi/32$ are aliasing components, this constitutes considerable, though perhaps typical, aliasing.) For each realization $h(n)$, the sequence $f(n)$ was generated by subsampling $h(n)$ by 32, and $g(n)$ was generated by shifting 7 samples and subsampling by 32, thus simulating a shift of $x_0 = 7T/32 = .21875T$. 500 independent realizations were simulated, and the histogram of results found by our proposed algorithm are plotted in Figure 3. The histogram shows the unbaisedness of our estimator, and shows a standard deviation of 0.0139, very low considering the significant aliasing in the observed data sequences.

6 Conclusions

The main results of this paper are the two algorithms for registration that obtain high precision subpixel registration by dealing with aliasing. The first algorithm ignores transform components that are obviously aliased. When aliasing is small, enough components remain to enable high precision phase estimates. As aliasing increases, fewer transform coefficients are available for the estimate, and precision drops substantially.

The second method attempts to estimate the aliasing components, and to remove them to created de-aliased coefficients for the phase estimate. Our results here, though preliminary, show that an energy-based model creates good displacement estimates for 1-D with substantial aliasing energy in a band up to twice the Nyquist frequency. The first algorithm performs poorly on these data because aliasing is too large for it. There are many open questions regarding the second algorithm, particularly its 2-D formulation, the accuracy achievable with this algorithm, and a deeper understanding of the theory that supports the energy model.

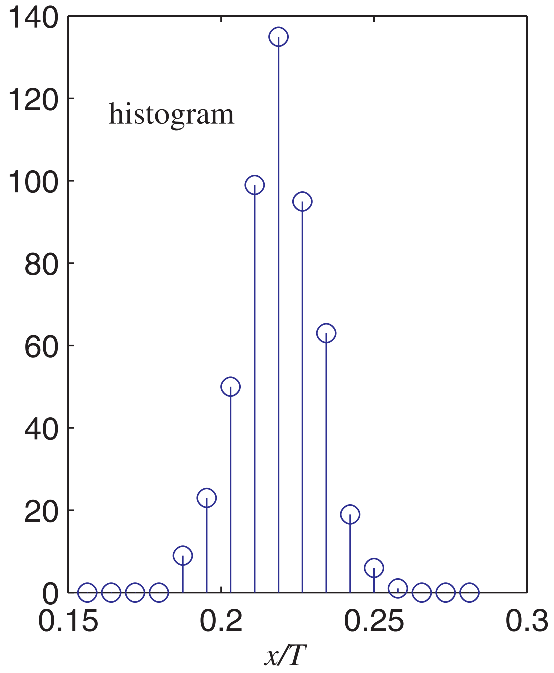

Figure 3. Monte Carlo simulations of De-aliasing Estimator: true shift = .21875T

7 Acknowledgment

The authors are grateful to Stephen Martucci of NEC Research Institute for his helpful advice and careful reading of the manuscript during the preparation of this paper.

References

[1] 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*, 41(12):3536–3541, December 1993.

[2] S. Alliney. Spatial registration of multispectral and multitemporal digital imagery using fast-Fourier transform techniques. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 15(5):499–504, May 1993.

[3] S. Alliney, G. Cortelazzo, and G. A. Mian. On the registrations of an object translating on a static background. *Pattern Recognition*, 29(1):131–141, January 1996.

[4] S. Alliney and C. Morandi. Digital image registration using projections. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 8(2):222–233, March 1986.

[5] P. E. Anuta. Spatial registration of multispectral and multitemporal digital imagery using fast-Fourier transform techniques. *IEEE Trans. on Geoscience Electronics*, 8(4):353–368, October 1970.

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

[7] L. Brown. A survey of image registration techniques. *ACM Computing Surveys*, 24(4):325–376, 1992.

[8] 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*, 16(12):1156–1168, December 1994.

[9] M. Corvi and G. Nicchiotti. Multiresolution image registration. In *Proceedngs 1995 IEEE International Conference on Image Processing*, pages 224–227, 1995.

[10] 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*, pages 66–75, May 1996.

[11] A. P. Cracknell and K. Paithoonwattanakij. Pixel and subpixel accuracy in geometrical correction of AVHRR imagery. *International Journal of Remote Sensing*, 10(4-5):661–667, 1989.

[12] B. Dasgupta and B. N. Chatterji. Fourier Mellin transform based image matching algorithm. *Journal of the Institution of Elect. and Telecommunications Engineers*, 42(1):3–9, 1996.

[13] E. De Castro and C. Morandi. Registration of translated and rotated images using finite Fourier transforms. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, PAMI-9(5):700–703, 1987.

[14] J. P. Djamdji, A. Bijaoui, and R. Maniere. Geometrical registration of images: The multiresolution approach. *Photogrammetric Engineering and Remote Sensing Journal*, 59(5):645–653, 1993.

[15] D. M. Etter and S. D. Stearns. Adaptive estimation of time delays in sampled data systems. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, ASSP-29(3):582–587, June 1981.

[16] 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*, 18(3):348–352, March 1996.

[17] P. L. Feintuch, N. J. Bershad, and F. A. Reed. Time delay estimation using the LMS adaptive filer – Dynamic behavior. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, ASSP-29(3):571–576, June 1981.

[18] C. R. Guarino. A novel method for two-dimensional phase estimation. In *Geoscience and Remote Sensing Symposium, 1994*, volume 4, pages 2279–2281, August 1994.

[19] 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*, IT-19:608–614, 1973.

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

[21] C. D. Kuglin and D. C. Hines. The phase correlation image alignment method. In *IEEE 1975 Conference on Cybernetics and Society*, pages 163–165, September 1975.

[22] 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.

[23] N. M. Namazi and B. Biswal. Variable time-delay estimation in colored and correlated noise environment. *IEEE Transactions on Aerospace and Electronic Systems*, 28(2):597–604, April 1992.

[24] 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*, ASSP-35(12):1649–1660, December 1981.

[25] D. V. Papadimitriou and T. J. Dennis. Stereo disparity using phase correlation. *Electronics Letters*, 30(18):1475–1477, September 1994.

[26] B. S. Reddy and B. N. Chatterji. An FFT-based technique for translation, rotation, and scale-invariant image registration. *IEEE Trans. on Image Processing*, 3(8):1266–1270, August 1996.

[27] F. A. Reed, P. L. Feintuch, and N. J. Bershad. Time delay estimation using the LMS adaptive filer – Static behavior. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, ASSP-29(3):561–571, June 1981.

[28] 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*, pages 532–537, June 1994.

[29] H. S. Stone. Progressive wavelet correlation using Fourier methods. *IEEE Transactions on Signal Processing*, 47(1):97–107, Jan. 1999.

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

[31] P. Thevenaz, U. E. Ruttimann, and M. Unser. A pyramid approach to subpixel registration based on intensity. *IEEE Trans. on Image Processing*, 7:27–41, January 1998.

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

[33] Q. Zheng and R. Chellappa. A computational vision approach to image registration. *IEEE Trans. on Image Processing*, 2(3):311–326, 1993.