Extension of Phase Correlation to Subpixel Registration

Hassan Foroosh (Shekarforoush), Josiane B. Zerubia, *Senior Member, IEEE*, and Marc Berthod

Abstract—In this paper, we have derived analytic expressions for the phase correlation of downsampled images. We have shown that for downsampled images the signal power in the phase correlation is not concentrated in a single peak, but rather in several coherent peaks mostly adjacent to each other. These coherent peaks correspond to the polyphase transform of a filtered unit impulse centered at the point of registration. The analytic results provide a closed-form solution to subpixel translation estimation, and are used for detailed error analysis. Excellent results have been obtained for subpixel translation estimation of images of different nature and across different spectral bands.

Index Terms—Image alignment, phase correlation, subpixel registration.

I. INTRODUCTION

ANALYSIS and fusion of data in an image sequence often require registration of the images. A wide variety of methods and techniques [4] can be found in the literature for solving this fundamental and challenging problem. Although in many applications pixel-level registration may be adequate, some important problems in remote sensing [13], [22], [25], [26] and biomedical imaging [5], [9] have introduced the requirement for subpixel registration.

In this paper, we are interested in the refinement of coarsely registered images to subpixel accuracy, where by coarse registration we imply pixel-level registration. We will assume that the error in registration at this refinement level is a translation within the bounds of the noise and the error of the imaging system. Therefore we confine our attention only to the problem of subpixel translation estimation. We are also interested in subpixel registration across different spectral bands.

The most commonly used approach to subpixel registration is based on interpolation. Examples include correlation interpolation [6], [28], intensity interpolation [28], phase correlation interpolation [21], [28] and the geometric methods [2], [8]. It is obvious that the accuracy of these methods depends highly on the quality of the interpolation algorithms.

One popular approach for subpixel registration without interpolation is based on the differential properties of image sequences [11]–[13], [17], [28], [30]. The main idea in this approach is to relate the temporal derivatives of the sequence to the spatial derivatives of the reference frame using the so called gradient constraint equation under intensity conservation assumption [10], [11], [17]. Due to the intensity conservation assumption, these methods often require that the inter-frame displacements to be relatively small compared to the intensity gradients of the reference frame. Also, to reduce the sensitivity of the derivative operators to the noise process, in general, some regularization is required and the gradient constraint is combined within some local neighborhood of each pixel. However, the main pitfall of this approach is probably the so called aperture problem [10], where for some patterns such as a gradual curve, the available information in a local neighborhood of a pixel (small aperture) is not sufficient to disambiguate the true registration parameters.

A second possible approach without interpolation is to formulate subpixel registration as an optimization problem. Examples of this approach include [14], [27], [29]. In this approach the problem is formulated as a cost function to be minimized with respect to the registration parameters of the inter-frame transformation model. As in the gradient based methods these methods are highly dependent on image intensity conservation assumption and would fail when applied across multiple spectral bands or when considerable amount of luminance variations are present between frames. This shortcoming can probably be alleviated by an adequate normalization of various spectral bands. Also note that these methods as well as the gradient-based methods implicitly include interpolation by applying local smoothing and regularizing operators.

There are also other techniques that are based on the local normalized correlation [1], polynomial regression [32], the discrete cosine transform [15], and the use of control points using a model-based approach [7].

This work is particularly motivated by some important features of the phase correlation method which make its use attractive for multispectral registration, and hence analytic results have been derived to demonstrate how the method can be extended to subpixel accuracy. A condensed version of part of this work was presented in [23] and [24].

This paper is organized as follows. In the next section, we describe the phase correlation method and outline its important properties. In Section III, we present our extension of the method to subpixel registration. Section IV provides thorough analysis of various sources of error. Experimental results are then given in Section V for different modalities of images including satellite images across different spectral bands. Finally in Section VI some concluding remarks are provided.
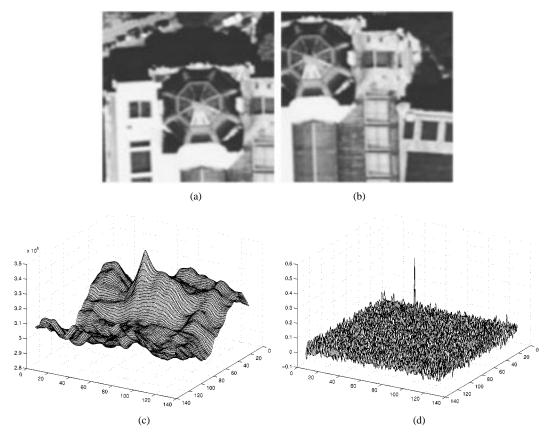
Fig. 1. (a) and (b) Aerial images of Paris with displacements along both axes, (c) standard cross-correlation, and (d) phase correlation.

II. PHASE CORRELATION METHOD

The idea behind this method [16], [21] is quite simple and is based on the Fourier shift property [20], which states that a shift in the coordinate frames of two functions is transformed in the Fourier domain as linear phase differences. This can be described as follows:

Let $f_1(x, y)$ and $f_2(x, y)$ be two functions that are absolutely integrable over $\mathbb{R}^2$. Let also

$$f_2(x, y) = f_1(x - x_o, y - y_o). \tag{1}$$

According to the Fourier shift property

$$\hat{f}_2(u, v) = \hat{f}_1(u, v) \exp(-i(ux_o + vy_o)). \tag{2}$$

Hence the normalized cross power spectrum is given by

$$\frac{\hat{f}_2(u, v)\hat{f}_1(u, v)^*}{|\hat{f}_1(u, v)\hat{f}_1(u, v)^*|} = \exp(-i(ux_o + vy_o)) \tag{3}$$

where $*$ indicates the complex conjugate.

The normalized cross power spectrum may also be viewed as the cross power spectrum of whitened signals. There are two possible ways of solving (3) for (x_o, y_o). One way is to directly work in the Fourier domain. For this purpose, consider a three-dimensional (3-D) Euclidean space whose canonical reference frame is given by the two frequency axes and the phase difference between the two images. In this space $ux_o + vy_o = 0$ defines a plane through the origin perpendicular to the vector $(x_o, , y_o)$ whose slopes along the two frequency axes specify the shifts along the two spatial axes [i.e., (x_o, y_o)]. This approach is rather inaccurate since it requires fitting a plane to noisy phase difference data.

The second possible approach which is more practical and also more robust to noise is to first inverse Fourier transform the normalized cross power spectrum. It is then a simple matter to determine (x_o, y_o), since from (3) the result is $\delta(x - x_o, y - y_o)$ which is a Dirac delta function centered at (x_o, y_o).

In practice, when dealing with images, f_1 and f_2 are specified only in finite size discretized arrays. However, replacing the Fourier transform by the discrete Fourier transform (DFT) and the Dirac delta function by a unit impulse, and also assuming a periodic extension of the images outside their compact support, these results will still hold and work remarkably well [16].

A. Properties of the Method

The most remarkable property of the phase correlation method compared to the classical cross correlation method is the accuracy by which the peak of the correlation function can be detected. Fig. 1 shows an example of two displaced aerial images. The phase correlation method provides a distinct sharp peak at the point of registration whereas the standard cross correlation yields several broad peaks and a main peak whose maximum is not always exactly centered at the right point.

A second important property is due to whitening of the signals by normalization, which makes the phase correlation notably robust to those types of noise that are correlated to the image function, e.g., uniform variations of illumination, offsets in average intensity, and fixed gain errors due to calibration. This property also makes phase correlation suitable for registration across different spectral bands.

Using the convolution theorem, it can be shown that the method can also handle blurred images, provided that the blurring kernel is relatively invariant from one frame to another. One may for instance use this property to register images contaminated with wide-band additive noise, by taking the phase correlation in the low-frequency portion of the spectrum [12].

In the discrete case, however, (3) is valid only if the shift vector (x_o, y_o) is of integer values. Therefore, when applied to discrete images, the method would fail to detect noninteger subpixel shifts. To our best knowledge the only practical approach proposed in the literature for adapting the method to subpixel estimation is the use of some interpolation methods [21], [28].

Also, as is known [16] the phase correlation always contains a single coherent peak at the point of registration corresponding to signal power, and some incoherent peaks which can be assumed to be distributed normally over a mean value of zero. The amplitude of the coherent peak is a direct measure of the degree of congruence between the two images. More precisely, the power in the coherent peak corresponds to the percentage of overlapping areas, while the power in incoherent peaks correspond to the percentage of nonoverlapping areas.

III. SUBPIXEL REGISTRATION

Our motivation is to find a method for subpixel registration with the advantages of the phase correlation method. For this purpose, we will derive analytic expressions to extend the method to subpixel (i.e., noninteger) shift estimation. Our model is based on the assumption that images with subpixel shifts were in fact originally displaced by integer values, which subsequently have reduced to subpixel values due to down-sampling. Therefore, in the following section, we will first examine the normalized cross power spectrum of two downsampled images.

A. Analytic Derivations

Consider two discrete images given by $f_1(x, y) = f(x, y)$ and $f_2(x, y) = f(x - x_o, y - y_o)$, where (x_o, y_o) is an integer valued vector. Let also $\hat{f}_1(u, v) = \hat{f}(u, v)$ and $\hat{f}_2(u, v) = \hat{f}(u, v)\exp(-i(ux_o + vy_o))$ denote their DFTs. Then the corresponding DFTs after down-sampling the images by factors of M and N along x and y axes respectively, will be given by (see [1, pp. 201–211], [29, pp. 194–195])

$$\hat{f}_{d1}(u, v) = \frac{1}{MN} \sum_{m'=0}^{M-1} \sum_{n'=0}^{N-1} \hat{f}\left(\frac{u+2\pi m'}{M}, \frac{v+2\pi n'}{N}\right) \quad (4)$$

and

$$\hat{f}_{d2}(u, v) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \hat{f}\left(\frac{u+2\pi m}{M}, \frac{v+2\pi n}{N}\right)$$
$$\cdot \exp\left(-i\left(\frac{u+2\pi m}{M} x_o, \frac{v+2\pi n}{N} y_o\right)\right). \quad (5)$$

Therefore, the cross-power spectrum of the downsampled images is given by (6), shown at the bottom of the page, where

$$\hat{h}_{mn}(u, v) = \frac{\hat{f}\left(\dfrac{u+2\pi m}{M}, \dfrac{v+2\pi n}{N}\right)}{\displaystyle\sum_{m'=0}^{M-1} \sum_{n'=0}^{N-1} \hat{f}\left(\dfrac{u+2\pi m'}{M}, \dfrac{v+2\pi n'}{N}\right)}. \quad (7)$$

It follows from these results that, the normalized cross power spectrum of two downsampled images is, merely, a polyphase decomposition [31] of a filtered unit impulse, i.e., the cross power spectrum consists of the sum of all polyphase components of a filtered unit impulse.

Despite its misleading appearance, in the absence of aliasing, $\hat{h}_{mn}(u, v)$ does not depend on $\hat{f}(u, v)$, and is in fact an ideal bandpass filter. We shall assume hereafter that aliasing does not occur or is avoided by prefiltering (see also [12]). Under this assumption, we now show that $\{h_{mn}(u, v)\}_{m=0,\ldots,M-1, n=0,\ldots,N-1}$ are ideal bandpass filters for any input $\hat{f}(u, v)$. For this purpose, note that in the absence of aliasing the spectral supports of

$$\left\{\hat{f}\left(\frac{u+2\pi m}{M}, \frac{v+2\pi n}{N}\right)\right\}_{m=0,\ldots,M-1, n=0,\ldots,N-1} \quad (8)$$

do not overlap.[1] This is clear since

$$\left(\frac{u+2\pi m}{M}, \frac{v+2\pi n}{N}\right) \neq \left(\frac{u+2\pi m'}{M}, \frac{v+2\pi n'}{N}\right),$$
$$\text{if } (m, n) \neq (m', n').$$

For instance let u and v be both in $[0, 2\pi[$ let also $M = N = 2$. Then by simply substituting the bounds of u and v (i.e., 0 and 2π), we find that

when $(m, n) = (0, 0)$,
$$\frac{u+2\pi m}{M} \in [0, \pi[\text{ and } \frac{v+2\pi n}{N} \in [0, \pi[$$

[1]The spectral support of a function is the domain outside which the spectrum of the function is identically zero.

$$\hat{C}(u, v) = \frac{\dfrac{1}{MN} \displaystyle\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \hat{f}\left(\dfrac{u+2\pi m}{M}, \dfrac{v+2\pi n}{N}\right) \exp\left(-i\left(\dfrac{u+2\pi m}{M} x_o, \dfrac{v+2\pi n}{N} y_o\right)\right)}{\dfrac{1}{MN} \displaystyle\sum_{m'=0}^{M-1} \sum_{n'=0}^{N-1} \hat{f}\left(\dfrac{u+2\pi m'}{M}, \dfrac{v+2\pi n'}{N}\right)}$$
$$= \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \hat{h}_{mn}(u, v) \exp\left(-i\left(\frac{u+2\pi m}{M} x_o, \frac{v+2\pi n}{N} y_o\right)\right) \quad (6)$$

when $(m, n) = (1, 0)$,

$$\frac{u+2\pi m}{M} \in [\pi, 2\pi[\text{ and } \frac{v+2\pi n}{N} \in [0, \pi[$$

when $(m, n) = (0, 1)$,

$$\frac{u+2\pi m}{M} \in [0, \pi[\text{ and } \frac{v+2\pi n}{N} \in [\pi, 2\pi[$$

when $(m, n) = (1, 1)$,

$$\frac{u+2\pi m}{M} \in [\pi, 2\pi[\text{ and } \frac{v+2\pi n}{N} \in [\pi, 2\pi[$$

which implies that if after down-sampling no spectral fold-over (overlap) occurs, then the supports of the following functions

$$\left\{ \hat{f}\left(\frac{u+2\pi m}{2}, \frac{v+2\pi n}{2} \right) \right\} \qquad m \in \{0, 1\}, \ n \in \{0, 1\}$$

will be disjoint (nonoverlapping). In fact, as seen from the previous discussion, the supports of these functions are found by partitioning the total two-dimensional (2-D) bandwidth $[0, 2\pi[\times [0, 2\pi[$ into $M \times N$ equal rectangular supports along the two frequency axes. This can be readily generalized to any $M > 1$ and $N > 1$.

Now, we sketch a simple proof of the fact that $\hat{h}_{mn}(u, v)$ does not depend on $\hat{f}(u, v)$:

Let u and v both take their values in $[0, 2\pi[$ and denote this 2-D support by B. Let also B_{mn} denote the disjoint rectangular supports of the functions in (8) (i.e., $\hat{f}((u+2\pi m)/M, (v+2\pi n)/N) = 0$ outside B_{mn}, and $\bigcup B_{mn} = B$). Since the spectral supports are disjoint, inside the support B_{mn}, all the terms in the denominator of (7) are zero except for the term corresponding to $(m', n') = (m, n)$. On the other hand outside the support B_{mn} the numerator is zero. Hence

$$\hat{h}_{mn}(u, v) = \frac{\hat{f}\left(\dfrac{u+2\pi m}{M}, \dfrac{v+2\pi n}{N} \right)}{\displaystyle\sum_{m'=0}^{M-1} \sum_{n'=0}^{N-1} \hat{f}\left(\dfrac{u+2\pi m'}{M}, \dfrac{v+2\pi n'}{N} \right)}$$

$$= \begin{cases} \dfrac{\hat{f}\left(\dfrac{u+2\pi m}{M}, \dfrac{v+2\pi n}{N} \right)}{\hat{f}\left(\dfrac{u+2\pi m}{M}, \dfrac{v+2\pi n}{N} \right)} = 1 \\ \quad \text{in } B_{mn} \left(\text{because in } B_{mn} \text{ we have} \right. \\ \quad \hat{f}\left(\dfrac{u+2\pi m'}{M}, \dfrac{v+2\pi n'}{N} \right) = 0, \\ \quad \left. \forall (m', n') \neq (m, n) \right) \\ 0 \quad \text{outside } B_{mn} \left(\text{because outside } B_{mn} \right. \\ \quad \text{we have } \hat{f}\left(\dfrac{u+2\pi m}{M}, \dfrac{v+2\pi n}{N} \right) = 0 \right). \end{cases} \quad (9)$$

In order to visualize better, let us show a schematic view in one dimension too, where a one-dimensional (1-D) signal f has been downsampled by a factor of two resulting in $\hat{f}_d(u) = (1/2) \sum_{m=0}^{1} \hat{f}((u+2\pi m)/2)$. Assuming that $u \in [0, 2\pi[$, we have for $m = 0$, $(u+2\pi m)/2 \in [0, \pi[$ and for $m = 1$,
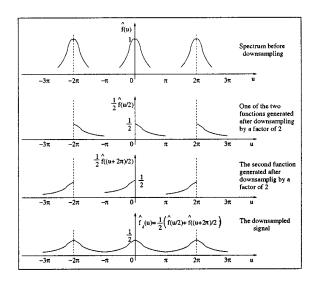

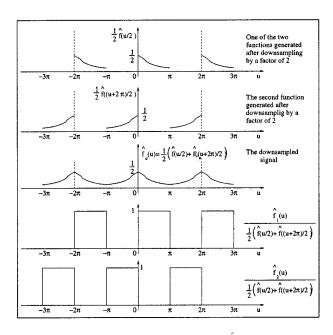Fig. 2. Schematic view of down-sampling by a factor of two for a 1-D signal.

Fig. 3. Schematic view of the filters $\hat{h}$ in 1-D.

$(u+2\pi m)/2 \in [\pi, 2\pi[$. Outside $[0, 2\pi[$ everything is repeated periodically. The spectrum of the downsampled signal is thus the sum of two disjoint (nonoverlapping) spectra $\hat{f}(u/2)$ and $\hat{f}((u+2\pi)/2)$, where $\hat{f}(u/2) \neq 0$ in $[0, \pi[$, and $\hat{f}((u+2\pi)/2) \neq 0$ in $[\pi, 2\pi[$. This is shown in Fig. 2.

As shown in Fig. 3, we can immediately see that by dividing $\hat{f}(u/2)$ or $\hat{f}((u+2\pi)/2)$ by the sum of the two, we get ideal bandpass filters, and this is independently of the form of f. In fact, $\hat{h}_m(u)$ depends only on the support of $\hat{f}((u+2\pi m)/M)$ and not its values within that support.

From the frequency response of the filters in (9), it immediately follows that the cross power spectrum of two downsampled images may be viewed as a downsampled version of a filtered unit impulse, where the filter has a rectangular frequency response. In other words, phase correlation leads to a downsampled 2-D Dirichlet kernel which as shown later is very closely approximated by a 2-D sinc function.
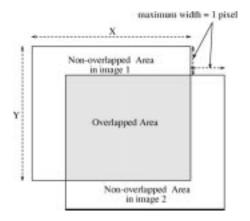
Fig. 4.   Schematic diagram of the overlay of two images.

The main result of this section can, therefore, be summarized as follows:

> *In general, in the discrete case the phase correlation yields a unit impulse. However, in the case of downsampled images the result is a downsampled version of a filtered unit impulse.*

We will show in the following that this result implies that phase correlation yields a downsampled 2-D Dirichlet kernel. Therefore in the case of downsampled images, phase correlation does not contain only a single coherent peak, but rather several coherent peaks with the most eminent ones largely next to each other.

B. Estimating the Sum of Polyphase Components

Equation (6), ideally, provides the normalized cross power spectrum of two downsampled images. However, since for discrete finite-size images the relation holds under periodicity assumption (see Section IV), in practice some additive noise will be introduced due to the nonoverlapped regions (see Fig. 4). This additive noise will be assumed to be distributed normally over a mean value of zero, as in the case of integer displacements [16]. In other words, we will assume to have

$$C_\eta(x, y) = C(x, y) + \eta(x, y) \qquad (10)$$

where $\eta(x, y)$ is a zero-mean Gaussian noise and C_η is the noisy phase correlation.

Using (6), one can readily estimate the sum of the polyphase components in the cross power spectrum. We shall shortly see that this sum is all that is required for estimating the subpixel translations.

The discrete inverse Fourier transform of (6) yields

$$C(x, y) = \frac{1}{WH} \frac{\sin(\pi(Mx - x_o))}{\sin(\pi(Mx - x_o)/W)}$$
$$\cdot \frac{\sin(\pi(Ny - y_o))}{\sin(\pi(Ny - y_o)/H)} \quad (11)$$

where W and H are the image width and height before downsampling.

We will show in the next section that for image sizes used in practice the Dirichlet function in (11) is very closely approxi-

mated by a sinc unction, and that substituting it by a sinc function is unconsequential. Therefore, we have

$$\tilde{C}(x, y) \simeq \frac{\sin(\pi(Mx - x_o))}{\pi(Mx - x_o)} \frac{\sin(\pi(Ny - y_o))}{\pi(Ny - y_o)}. \qquad (12)$$

Substituting from (12) into (10), we get

$$C_\eta(x, y) \simeq \frac{\sin(\pi(Mx - x_o))}{\pi(Mx - x_o)} \frac{\sin(\pi(Ny - y_o))}{\pi(Ny - y_o)} + \eta(x, y). \qquad (13)$$

In general, for subpixel shifts, the noise level is very low and one can easily identify the main peak and the immediate adjacent peaks of the 2-D sinc function by simple inspection. However, in this section, we will suggest a method of estimating an upper bound for the noise variance and a lower bound for the signal variance, which in turn will allow us to separate the noise process by setting a threshold value.

We shall first recall that, since the cross power spectrum has been normalized, we will be only interested in the normalized values of the variances taking values in the interval [0, 1].

Since the nonoverlapped regions are of subpixel width (see Fig. 4) and also since the noise power is given by the percentage of nonoverlapped area in each image, we can estimate the following upper bound for the noise variance in terms of image dimensions after down-sampling, i.e., X and Y:

$$\sigma_\eta^2 < \frac{(X + Y - 1)^2}{(XY)^2}. \qquad (14)$$

Similarly, a lower bound for the total signal power would be

$$\sigma_s^2 \geq \frac{(XY - (X + Y - 1))^2}{(XY)^2}. \qquad (15)$$

These estimations are shown schematically in Fig. 4.

The minimum signal-to-noise ratio is thus given by

$$\text{SNR} = 10 \log \left(\frac{\sigma_s^2}{\sigma_\eta^2} \right) > 10 \log \left(\frac{(XY - (X + Y - 1))^2}{(X + Y - 1)^2} \right). \qquad (16)$$

It is obvious that for subpixel shifts the SNR is very large. Therefore, the choice of a noise threshold value T_η is not a great concern. Nevertheless, in the next section we will provide a rigorous method of choosing it. Thus, the signal is simply detected by using

$$C(x, y) = C_\eta(x, y) \qquad \text{if} \quad C_\eta(x, y) > T_\eta. \qquad (17)$$

Once $C(x, y)$ has been estimated, (x_o, y_o) can be calculated by a straightforward application of (12) to a set of points. Note that results will be more reliable if (12) is applied to points where the signal power is mostly concentrated. In fact, as stated earlier, for two images with subpixel displacements the signal power is largely concentrated in the vicinity of the main peak of the phase correlation (modulo periodicity). To be more precise, for subpixel displacements the signal power in the phase correlation is usually concentrated in a main peak at some coordinates (x_m, y_m) and two side-peaks at (x_s, y_m) and (x_m, y_s) where $x_s = x_m \pm 1 \pmod{X}$ and $y_s = y_m \pm 1 \pmod{Y}$. Using these three points one can then solve for the subpixel shift values.

We will explain the method using an example.

Consider the situation where the signal power is mostly concentrated at a main peak with coordinates (0, 0) and two side

peaks at $(1, 0)$ and $(0, 1)$. By applying (12) to the first two points, we will find

$$\frac{\sin(\pi x_o)}{\pi x_o} \frac{\sin(\pi y_o)}{\pi y_o} = C(0, 0) \qquad (18)$$

$$\frac{\sin(\pi(M - x_o))}{\pi(M - x_o)} \frac{\sin(\pi y_o)}{\pi y_o} = C(1, 0). \qquad (19)$$

Dividing both sides and rearranging, we get

$$\frac{\sin(\pi(M - x_o))}{\pi(M - x_o)} = \frac{C(1, 0)}{C(0, 0)} \frac{\sin(\pi x_o)}{\pi x_o} \qquad (20)$$

or

$$\frac{\pm \sin(\pi x_o)}{(M - x_o)} = \frac{C(1, 0)}{C(0, 0)} \frac{\sin(\pi x_o)}{x_o}. \qquad (21)$$

And after simplifying, we obtain

$$\Delta x = \frac{x_o}{M} = \frac{C(1, 0)}{C(1, 0) \pm C(0, 0)} \qquad (22)$$

where $\Delta x = x_o/M$ is the subpixel displacement along the x-axis.

Note that, no knowledge of the down-sampling rate M is required. Note also that two solutions will be obtained. This ambiguity is due to the isotropic form of $C(x, y)$. However, the correct solution is easy to identify since it is in the interval $[-1, 1]$ and it has the same sign as $x_s - x_m$. Similarly, we can find $\Delta y = y_o/N$ by using $C(0, 0)$ and $C(0, 1)$, which again must be in the interval $[-1, 1]$ with the same sign as $y_s - y_m$.

IV. ERROR ANALYSIS

In this section, we consider the following five sources of error:

- error due to approximating the Dirichlet kernel by a sinc function;
- error due to nonoverlapped regions in the two images;
- error due to aliasing;
- border errors due to periodicity assumption;
- and error due to wide-band random noise.

The absolute pointwise error in the approximation in (12) is given by $|C(x, y) - \tilde{C}(x, y)|$. This error tends to zero as W and H tend to infinity as

$$\lim_{W \to \infty} \lim_{H \to \infty} \left| C(x, y) - \tilde{C}(x, y) \right|$$

$$= \lim_{W \to \infty} \lim_{H \to \infty} \left| \frac{1}{WH} \frac{\sin(\pi(Mx - x_o))}{\sin(\pi(Mx - x_o)/W)} \right.$$
$$\left. \cdot \frac{\sin(\pi(Ny - y_o))}{\sin(\pi(Ny - y_o)/H)} - \tilde{C}(x, y) \right| \qquad (23)$$

$$= \lim_{W \to \infty} \lim_{H \to \infty} \left| \frac{1}{\pi(Mx - x_o)} \frac{\sin(\pi(Mx - x_o))}{\frac{\sin(\pi(Mx - x_o)/W)}{\pi(Mx - x_o/W}} \right.$$

$$\left. \cdot \frac{1}{\pi(Ny - y_o)} \frac{\sin(\pi(Ny - y_o))}{\frac{\sin(\pi(Ny - y_o/H))}{\pi(Ny - y_o/H)}} - \tilde{C}(x, y) \right| \qquad (24)$$
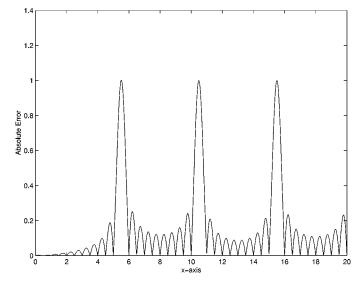

Fig. 5. An example of the variations of the absolute difference between a sinc and a Dirichlet function for $x_o = 1$, $M = 2$, and $W = 10$.
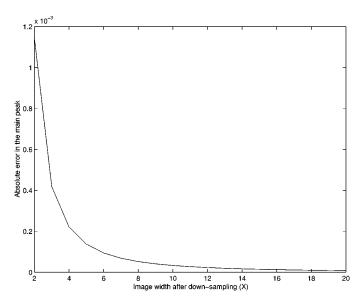

Fig. 6. Error in the main peak of the phase correlation for $x_o = 1$ and $M = 20$ as a function of the downsampled image width X.

$$= \left| \frac{\sin(\pi(Mx + x_o))}{\pi(Mx + x_o)} \frac{\sin(\pi(Ny + y_o))}{\pi(Ny + y_o)} - \tilde{C}(x, y) \right| = 0. \qquad (25)$$

In practice, for finite values of W and H, this error is highly oscillatory and attains its maximum value of 1 at coordinates $((kW + x_o)/M, (lH + y_o)/N) = (kX + \Delta x, lY + \Delta y)$, where $k \neq 0$ and $l \neq 0$ are integers, and X and Y are the image dimensions after down-sampling. Since in practice W and H are much larger than M and N respectively, the maximum error occurs way after the main peak and several side-lobes of the Dirichlet kernel. Fig. 5 shows an example of the variation of this error for a 1-D case.

The worst case would be if the maximum error occurs in the region $[-1, 1] \times [-1, 1]$ which would imply that $|X + \Delta x| \leq 1$ and $|Y + \Delta y| \leq 1$. But this would mean that $X = 1$ and $Y = 1$, i.e., the images to be registered would be 1×1! Fig. 6 shows the variations of the error in the main peak of the phase

correlation as a function of the image width in 1-D or $x_o = 1$ and $M = 20$. This example corresponds to the case when a subpixel accuracy of 0.05 is required. Note that even for an unrealistically small image width of two pixels, the error in the main peak of phase correlation does not exceed ± 0.0012. Note also that this error drops rapidly to extremely small values as X increases. For instance, from the graph in Fig. 6, for a subpixel accuracy of 0.05 even a small image of five-pixels wide would cause an insignificant error of less than 2×10^{-4} in the height of the main peak of the phase correlation, which clearly is unconsequential in the calculation of the subpixel shifts using (13).

For the errors due to nonoverlapping regions, our analyses are similar to [16]. The only difference here, is that the signal power is not concentrated in a single coherent peak, but rather distributed in a set of coherent peaks mostly adjacent to each other.

Since we have assumed a normal distribution $N(0, \sigma_\eta)$ of the noise η, the probability that the noise value exceeds T_η will be given by

$$
\begin{aligned}
P_{T_\eta}(\eta) &= \frac{1}{\sqrt{2\pi}\sigma_\eta} \int_{T_\eta}^{\infty} \exp\left(-\frac{\eta^2}{2\sigma_\eta^2}\right) d\eta \\
&= \frac{1}{\sqrt{2\pi}\sigma_\eta} \int_{T_\eta}^{1} \exp\left(-\frac{\eta^2}{2\sigma_\eta^2}\right) d\eta.
\end{aligned} \tag{26}
$$

The last equality is due to normalization of the cross power spectrum between zero and one.

On the other hand, in this interval, we can write the following inequality:

$$
\begin{aligned}
\frac{1}{\sqrt{2\pi}\sigma_\eta} &\int_{T_\eta}^{1} \exp\left(-\frac{\eta^2}{2\sigma_\eta^2}\right) d\eta \\
&> \frac{1}{\sqrt{2\pi}\sigma_\eta} \int_{T_\eta}^{1} \eta \exp\left(-\frac{\eta^2}{2\sigma_\eta^2}\right) d\eta.
\end{aligned} \tag{27}
$$

This inequality follows from the fact that, in $[T_\eta, 1]$ the integrand on the left hand side is always larger than the integrand on the right hand side. Therefore, substituting (26) into (27) and computing the integral on the right hand side, we will obtain

$$
P_{T_\eta}(\eta) > \frac{\sigma_\eta}{\sqrt{2\pi}} \left(\exp\left(\frac{-T^2}{2\sigma_\eta^2}\right) - \exp\left(\frac{-1}{2\sigma_\eta^2}\right)\right). \tag{28}
$$

And after rearranging and simplifying, we will have

$$
T_\eta^2 > -2\sigma_\eta^2 \ln\left(\frac{\sqrt{2\pi}P_{T_\eta}(\eta)}{\sigma_\eta} + \exp\left(\frac{-1}{2\sigma_\eta^2}\right)\right). \tag{29}
$$

Therefore, for small probability values $P_{T_\eta}(\eta)$, we can remove the noise by thresholding all values below T_η. To choose a value for $P_{T_\eta}(\eta)$, consider the nontrivial case of the inequality (29) given by

$$
\frac{\sqrt{2\pi}P_{T_\eta}(\eta)}{\sigma_\eta} + \exp\left(\frac{-1}{2\sigma_\eta^2}\right) < 1 \tag{30}
$$

or

$$
P_{T_\eta}(\eta) < \frac{\sigma_\eta}{\sqrt{2\pi}} \left(1 - \exp\left(\frac{-1}{2\sigma_\eta^2}\right)\right) \tag{31}
$$

which allows determining the upper bound of $P_{T_\eta}(\eta)$. It is obvious that the right hand side of the previous inequality is always in $[0, 1]$.

Aliasing is another possible source of error, which occurs when the down-sampling factors (i.e., M and N) exceed the critical values (say K and L), above which the Nyquist criterion is not satisfied. In particular, aliasing results in overlaps between adjacent replica. In this paper, we will assume that our image acquisition devices can only cause first-order aliasing. This implies that overlapping in the base-band occurs only with the immediate adjacent replica. Note that first-order aliasing can still cause severe errors.

Analysis of the aliasing error under some general frameworks have already been investigated in the past literature [3], [19]. However in this paper we shall exploit the particular nature of the problem in hand. Let us therefore assume that the images prior to down-sampling were band-limited to $\pm(\pi/K)$ and $\pm(\pi/L)$ along each frequency axis, i.e.,

$$
\hat{f}(u, v) = 0, \qquad \text{for } |u| \geq \frac{\pi}{K} \quad \text{and} \quad |v| \geq \frac{\pi}{L}. \tag{32}
$$

Let also $M > K$ and $N > L$. Clearly then aliasing will occur, and will distort some frequencies adjacent to $\pm \pi$ along each frequency axis. The extent of aliasing depends on the amount by which M and N exceed K and L, respectively.

In our case, we are interested in the effect of aliasing on phase correlation. For this purpose note that the normalized cross power spectrum is merely the convolution of two whitened images, which when under-sampled would have aliasing at high frequencies. Both whitening and convolution involve point-wise operations in the Fourier domain, and hence the process of computing the normalized cross power spectrum does not modify the extent of aliasing. This implies that we can perform our analysis of the aliasing error directly on the cross power spectrum rather than on individual images. The advantage is that the cross power spectrum of two shifted images (as described in Section III-A) has a particular structure, which can be exploited.

Let $M = SK$ and $N = TL$, where S and T are some positive constants larger than one. A natural approach to avoid the aliasing error in the estimated registration parameters is to low-pass filter the cross power spectrum, so that all subbands in the aliased regions are filtered out. Therefore if all high-frequency aliased subbands indexed in the range $[K/2, \ldots, M - (K/2)]$ and $[L/2, \ldots, N - (L/2)]$ are filtered out, then we get

$$
\begin{aligned}
\hat{C}(u, v) = &\sum_{m=0}^{K/2-1} \sum_{n=0}^{L/2-1} \hat{h}_{mn}(u, v) \\
&\cdot \exp\left(-i\left(\frac{(u+2\pi m)}{M} x_o, \frac{(v+2\pi n)}{N} y_o\right)\right) \\
&+ \sum_{m=M-K/2}^{M-1} \sum_{n=N-L/2}^{N-1} \hat{h}_{mn}(u, v) \\
&\cdot \exp\left(-i\left(\frac{(u+2\pi m)}{M} x_o, \frac{(v+2\pi n)}{N} y_o\right)\right)
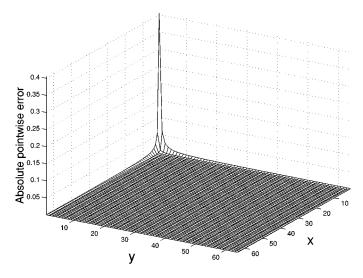\end{aligned} \tag{33}
$$

Fig. 7. Absolute pointwise aliasing error for $S = T = 2$.

Fig. 8. Absolute pointwise aliasing error with 10% spectral overlap.

$$= \sum_{m=0}^{K/2-1} \sum_{n=0}^{L/2-1} \hat{h}_{mn}(u, v)$$
$$\cdot \exp\left(-i\left(\frac{(u+2\pi m)}{K}\frac{x_o}{S}, \frac{(v+2\pi n)}{L}\frac{y_o}{T}\right)\right)$$
$$+ \sum_{m=M-K/2}^{M-1} \sum_{n=N-L/2}^{N-1} \hat{h}_{mn}(u, v)$$
$$\cdot \exp\left(-i\left(\frac{(u+2\pi m)}{K}\frac{x_o}{S}, \frac{(v+2\pi n)}{L}\frac{y_o}{T}\right)\right) \tag{34}$$

which upon inverse transforming gives

$$C(x, y) = \frac{1}{WH} \frac{\sin(\pi(Kx - x_o/S))}{\sin(\pi(Kx - x_o/S)/W)}$$
$$\cdot \frac{\sin(\pi(Ly - y_o/T))}{\sin(\pi(Ly - y_o/T)/H)}. \tag{35}$$

Two comments are in order. First note that the analyses described in Section III still hold after low-pass filtering, since the remaining set of subbands still form a polyphase transform of a filtered unit impulse, except that the bandwidth of the filter is now smaller. Secondly, using our technique, the shift parameters can still be correctly estimated, since they are given by $\Delta_x = (x_o/S)/K = x_o/M$ and $\Delta_y = (y_o/T)/L = y_o/N$. Since the filtered cross power spectrum in (35) provides an error-free estimation of the shift parameters, the aliasing error is simply given by the difference between (35) and an unfiltered version of the cross power spectrum. The absolute pointwise error is therefore given by

$$e(x, y) = \left| \frac{1}{WH} \frac{\sin(\pi(Mx - x_o))}{\sin(\pi(Mx - x_o)/W)} \frac{\sin(\pi(Ny - y_o))}{\sin(\pi(Ny - y_o)/H)} \right.$$
$$- \frac{1}{WH} \frac{\sin(\pi(Kx - x_o/S))}{\sin(\pi(Kx - x_o/S)/W)}$$
$$\left. \cdot \frac{\sin(\pi(Ly - y_o/T))}{\sin(\pi(Ly - y_o/T)/H)} \right|. \tag{36}$$

A close inspection of this error shows that for fixed values of S and T, the error is independent of M, N, K, L, W and H. In particular, this error is highly oscillatory, and for integer values of all parameters attains its maximum value at (x_o, y_o).[2] Unfortunately, this also implies that the aliasing error would mainly affect the coherent peaks of the phase correlation.

Let us now consider the worst scenario for a first-order aliasing. This occurs when $S = T = 2$. Substituting these values in (36) and simplifying, we get

$$e(x, y)$$
$$= \left| \frac{1}{WH} \frac{\sin(2\pi(Kx - x_o/2))}{\sin(2\pi(Kx - x_o/2)/W)} \right.$$
$$\cdot \frac{\sin(2\pi(Ly - y_o/2))}{\sin(2\pi(Ly - y_o/2)/H)}$$
$$- \frac{1}{WH} \frac{\sin(\pi(Kx - x_o/2))}{\sin(\pi(Kx - x_o/2)/W)}$$
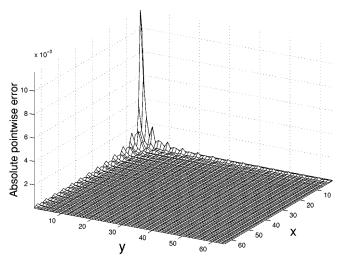$$\left. \cdot \frac{\sin(\pi(Ly - y_o/2))}{\sin(\pi(Ly - y_o/2)/H)} \right| \tag{37}$$
$$= \left| \frac{1}{WH} \frac{\sin(\pi(Kx - x_o/2))}{\sin(\pi(Kx - x_o/2)/W)} \right.$$
$$\left. \cdot \frac{\sin(\pi(Ly - y_o/2))}{\sin(\pi(Ly - y_o/2)/H)} \right|$$
$$\cdot \left| \frac{\cos(\pi(Kx - x_o/2)) \cos(\pi(Ly - y_o/2))}{\cos(\pi(Kx - x_o/2)/W) \cos(\pi(Ly - y_o/2)/H)} - 1 \right|. \tag{38}$$

Fig. 7 shows the variations of $e(x, y)$ for $S = T = 2$ and some arbitrary values of other parameters.

Note that the maximum error for a given pair (x_o, y_o) is quite substantial, and hence would introduce severe error in the estimated registration parameters. In practice, fortunately, the situation is not as bad as it appears to be. Recall that in the previous analysis we had assumed the worst case of $S = T = 2$. This

[2]Note that the parameters M, N, K, L, W, H, x_o, and y_o take only integer values, and hence the interested reader should take into account this constraint for reproducing the results and the graphs in this section.
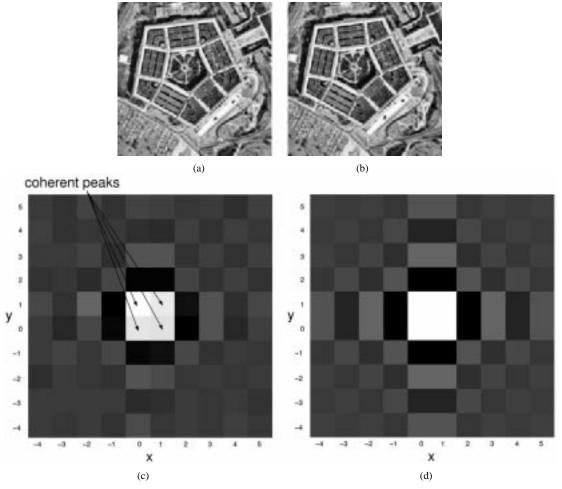
Fig. 9.   (a) and (b) Aerial images of Pentagon with subpixel displacements of $(\Delta_x, \Delta_y) = (0.5, 0.5)$, (c) zoom in the phase correlation in the neighborhood of the main peak, and (d) zoom in the neighborhood of the main peak of a perfect 2-D discrete Dirichlet function with the same values of Δ_x and Δ_y.

corresponds to the situation where the entire baseband is contaminated with first-order aliasing. Clearly, an acquisition device with such a low sampling density would be of little use, if any. In practice, for most image acquisition systems, we have $1 \leq S \ll 2$ and $1 \leq T \ll 2$, which implies that only a small portion of the spectrum at high frequencies is contaminated with aliasing. Also, the error due to aliasing drops very rapidly as S and T become smaller than 2. Fig. 8 shows, for instance, the absolute pointwise error when 10% of the spectrum is aliased. Note that even with a relatively high level of aliasing of 10% the maximum error is of the order of 0.01. This error is clearly small (compared to the amplitude of the coherent peaks in the phase correlation), and hence its influence is effectively negligible in our calculations.

A fourth source of error is introduced by assuming that the signal is repeated periodically. We must emphasize that periodicity is in fact an immediate consequence of using DFT for computing the Fourier spectrum of finite-size discrete signals, and is not an assumption made for the derivation of the results in Section III. In fact, any discrete signal/image processing algorithm which resorts to DFT (and hence to its fast implementations, i.e., FFT) is implicitly making this assumption even if this fact is not mentioned explicitly. This is due to the well-known result that any discrete signal has a 2π-periodic Fourier spectrum. Therefore by applying the duality, if both the signal and

its spectrum are discrete (as in a DFT-pair), then they must be both periodic. Note that compactly supported signals and periodic signals are treated in the same manner in the context of DFT (see [18, Ch. 8]). Of course, the Fourier transform of periodic signals introduces some technical issues (related to its convergence), which is outside the scope of this paper, and hence the reader is referred to classical literature (e.g., [18]).

In this paper, we explicitly referred to this assumption, in order to emphasize the fact that negative shifts are estimated by considering the periodicity of DFT-pairs. We must also emphasize that the proposed method does not rely on compact local operators (e.g., gradient operators), which in general introduce border artifacts. Therefore, the only error that is caused by the periodicity assumption is the introduction of border discontinuities, which can affect the convergence of the Fourier transform. But, since discontinuity is a pointwise concept, and also since we are only concerned with discontinuities at the image borders, these errors would at most affect the nonoverlapping regions at the borders, and hence as pointed out earlier our error analysis for nonoverlapping regions equally applies to border discontinuities. We shall also remind the reader that the convergence of Fourier transform can still be guaranteed for signals that include discontinuities, provided that some local smoothness constraints are assumed in the neighborhood of the discontinuity. We refer the interested reader on this subject to excellent literature such as [20].
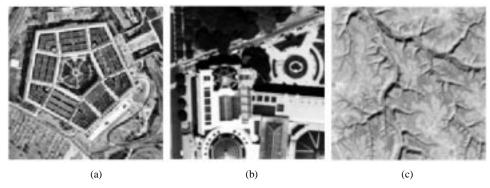
Fig. 10. Some of the images used for experimentation: (a) and (b) aerial images of Pentagon and Paris and (c) a SPOT satellite image.

TABLE I
TABLE OF RESULTS

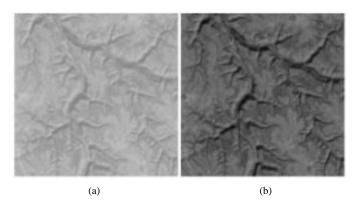| Image	$(\Delta x, \Delta y)$	Estimates
Pentagon	(0.50, -0.50)	(0.48, -0.51)
	(0.25, 0.50)	(0.28, 0.49)
	(-0.25, -0.50)	(-0.25, -0.52)
	(0.0, 0.75)	(0.0, 0.80)
Paris	(0.167, -0.5)	(0.152, -0.49)
	(0.67, 0.25)	(0.69, 0.33)
	(-0.33, -0.167)	(-0.32, -0.15)
	(0.33, 0.33)	(0.325, 0.32)
SPOT	(0.83, 0.25)	(0.84, 0.28)
	(0.33, -0.50)	(0.34, -0.52)
	(-0.83, 0.167)	(-0.83, 0.151)
	(0.25, 0.33)	(0.30, 0.33)

Fig. 11. (a) and (b) Two SPOT satellite images with artificially introduced luminance variations.

TABLE II
TABLE OF RESULTS FOR SIMULATIONS WITH LUMINANCE VARIATIONS

Image	$(\Delta x, \Delta y)$	Estimates
SPOT	(0.83, 0.25)	(0.83, 0.30)
	(0.33, -0.50)	(0.34, -0.53)
	(-0.83, 0.167)	(-0.84, 0.15)
	(0.25, 0.33)	(0.29, 0.34)

A final source of error could be wide-band random noise. Inclusion of this error in our model would in fact introduce a new level of complexity in our problem, which would require thorough analysis. Since wide-band random noise affects mainly the high frequencies, at this stage we have only considered low-pass filtering as a remedy, which has also been put into effect in our experimentations for avoiding aliasing.

V. EXPERIMENTAL RESULTS

In order to verify the algorithm experimentally, some simulations were performed by shifting images according to the following model:

$$g_m = \mathcal{K} * f_m \qquad m = \text{frame number} \qquad (39)$$

where f_m are shifted versions of a high-resolution image convolved by a blurring kernel $\mathcal{K}$ which characterizes image degradations. Each frame g_m is then downsampled at a rate larger than the maximum shifts so that the correspondence between different frames is reduced to subpixel level. Note that, no knowledge of the blurring kernel is required for the registration. The only assumption on $\mathcal{K}$ is that its response is invariant from one frame to another.

The algorithm has been tested on several images of different nature. Fig. 9 shows an example of two aerial images that are coarsely realigned and hence still contain subpixel shifts. Fig. 9(c) shows a zoom in the area where the main peaks occur in the phase correlation of these two images. The actual subpixel shifts in this simulation were $\Delta x = 0.5$ and $\Delta y = 0.5$. A close inspection of the phase correlation within the vicinity of the main peak shows that due to the subpixel nature of the correspondence, there is no single coherent peak (i.e., a single unit impulse as with integer displacements). Instead, the signal power is distributed in the form of polyphase components of a filtered unit impulse (i.e., a 2-D Dirichlet function) centered at the point of correspondence. Fig. 9(d) provides a comparison with a perfect 2-D Dirichlet function with the same values of subpixel shifts.

In Fig. 10, some of the images to which the algorithm was applied have been shown with results summarized in Table I.
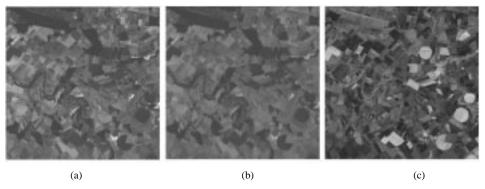
(a)								(b)								(c)

Fig. 12.   SPOT satellite image: (a) XS1, (b) XS2, and (c) XS3.

TABLE III
SIMULATION RESULTS ACROSS DIFFERENT SPECTRAL BANDS, TAKING XS1
AS REFERENCE FRAME

Image	$(\Delta x, \Delta y)$	Estimates
XS2	(0.45, 0.50)	(0.43, 0.51)
	(0.50, -0.50)	(0.48, -0.49)
	(-0.25, 0.0)	(-0.25, -0.03)
	(0.75, 0.75)	(0.76, 0.72)
XS3	(0.33, -0.5)	(0.31, -0.48)
	(0.67, 0.50)	(0.65, 0.49)
	(-0.33, 0.33)	(-0.34, 0.35)
	(0.33, -0.167)	(0.31, -0.163)

Fig. 13.   Frame from a real sequence obtained by a hand-held camera.

TABLE IV
RESULTS FOR THE CAR SEQUENCE. REGISTRATION PERFORMED RELATIVE TO
ONE OF THE FRAMES TAKEN AS A REFERENCE

Image	Estimated shifts for different frames	% power in main peak before sub-pixel alignment	% power in main peak after sub-pixel alignment
Car sequence	(0.46, 0.51)	0.27%	0.77%
	(0.17, -0.45)	0.51%	0.82%
	(0.03, 0.81)	0.55%	0.81%
	(-0.67, -0.1)	0.6%	0.84%

We then applied the method to some images by artificially introducing luminance variations. Results for a test image are shown in Fig. 11 and Table II. These simulations confirm the fact that the method is robust to luminance variations, which only contribute to the magnitude and not the phase response.

A set of experimentations were also performed on satellite images across different spectral bands of SPOT channels (XS1, XS2 and XS3), which were provided by the French space agency (CNES). In these experimentations XS1 was taken as the reference frame and XS2 and XS3 images were registered with respect to XS1. The test images are shown in Fig. 12 and the results are given in Table III.

Further experimentations were performed on a real sequence shown in Fig. 13. Since for these real data the actual registration parameters were unknown, the performance was evaluated by the percentage of power in the main peak (i.e., SNR) of phase correlation after pixel-level and subpixel level alignments. The sequence was obtained by a hand-held camera, and the frames contained both lateral pixel-level translations and some rotations. The image frames were first registered at pixel level accuracy using an optical flow method [10]. This yielded a coarsely aligned sequence, which was then refined for subpixel translational motions using our technique. These experimentations were performed in the context of super-resolution. Results for registration are shown in Table IV.

VI. SOME CONCLUDING REMARKS

We have derived analytic expressions which demonstrate how the phase correlation method can be used to estimate subpixel shifts. The advantages of the method are as follows.

- *High accuracy:* Experimentations show that excellent results can be obtained using the proposed method.
- *Accurate error analysis:* Since closed-form expressions are available, it is possible to perform accurate error analysis and performance evaluation.
- *Multispectral registration:* Due to prewhitening, the method provides a very efficient means of registering images at subpixel accuracy across different spectral bands.

The main drawback of the method is its limitation to translations. Also caution must be exercised when applying the proposed technique. The availability of closed-form solutions might introduce a misleading impression that, compared to

other existing methods, our method can provide the subpixel shifts up to any number of significant figures. However, the "significance" of any estimation beyond certain number of figures is obviously limited by the signal to noise ratio in the phase correlation. Note also that, our classification of existing methods in Section I in terms of the use of interpolation techniques refers to an explicit use of interpolation. For instance one might argue that our technique is implicitly resorting to Shannon's sampling interpolation via the sinc function, and hence involves, in a sense, an infinite order interpolation.

For the aliasing error, we experimented with both prefiltered and unfiltered phase correlation. Our experience shows that for currently used nonexperimental imaging systems the aliasing error is so negligible that the prefiltering stage does not substantially affect the results.

REFERENCES

[1] R. J. Althof, M. G. J. Wind, and J. T. Dobbins, "A rapid and automatic image registration algorithm with subpixel accuracy," *IEEE Trans. Med. Imag.*, vol. 16, pp. 308–316, June 1997.

[2] C. A. Bernstein, L. N. Kanal, D. Lavin, and E. C. Olson, "A geometric approach to subpixel registration accuracy," *Comput. Vis. Graph. Image Process.*, vol. 40, pp. 334–360, 1987.

[3] J. L. Brown, "On mean-square aliasing error in the cardinal series expansion of random processes," *IEEE Trans. Inform. Theory*, vol. IT-24, no. 2, pp. 254–256, 1978.

[4] L. G. Brown, "A survey of image registration techniques," *ACM Comput. Surv.*, vol. 24, pp. 325–376, 1992.

[5] P. A. Van den Elsen, E. J. D. Pol, and M. A. Viergever, "Medical image matching—A review with classification," *IEEE Eng. Med. Biol. Mag.*, vol. 12, pp. 26–39, 1993.

[6] V. N. Dvorchenko, "Bounds on (deterministic) correlation functions with applications to registration," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-5, no. 2, pp. 206–213, 1983.

[7] A. Efrat and C. Gotsman, "Subpixel image registration using circular fiducials," *Int. J. Comput. Geom. Applicat.*, vol. 4, no. 4, pp. 403–422, 1994.

[8] A. Goshtasby, G. C. Stockman, and C. V. Page, "A region-based approach to digital image registration with subpixel accuracy," *IEEE Trans. Geosci. Remote Sensing*, vol. 24, no. 3, pp. 390–399, 1986.

[9] J. V. Hajnal, N. Saeed, E. J. Soar, A. Oatridge, I. R. Young, and G. M. Bydder, "A registration and interpolation procedure for subvoxel matching of serially acquired mr images," *J. Comput. Assist. Tomogr.*, vol. 19, no. 2, pp. 289–296, 1995.

[10] D. J. Heeger, "Model for the extraction of image flow," *J. Opt. Soc. Amer. A*, vol. 4, pp. 1455–1471, 1987.

[11] B. K. P. Horn and B. Schunck, "Determining optical flow," *Artif. Intell.*, vol. 17, pp. 185–203, 1981.

[12] T. S. Huang and R. Y. Tsai, *Image Sequence Analysis: Motion Estimation*. Berlin, Germany: Springer-Verlag, 1981, ch. 1, pp. 1–18.

[13] M. Irani and S. Peleg, "Improving resolution by image registration," *Graph. Mod. Image Process.*, vol. 53, no. 3, pp. 231–239, 1991.

[14] S. P. Kim and W.-Y. Su, "Subpixel accuracy image registration by spectrum cancellation," in *Proc. ICASSP*, 1993, pp. 153–156.

[15] U. Koc and K. J. R. Liu, "Interpolation-free subpixel motion estimation techniques in dct domain," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 8, pp. 460–487, Aug. 1998.

[16] C. D. Kuglin and D. C. Hines, "The phase correlation image alignment method," in *Proc. Int. Conf. Cybernetics Society*, 1975, pp. 163–165.

[17] H. H. Nagel, "On the estimation of optical flow: Relations between different approaches and some new results," *Artif. Intell.*, vol. 33, pp. 299–324, 1987.

[18] A. V. Oppenheim, R. W. Schafer, and J. R. Buck, *Discrete-Time Signal Processing*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1989.

[19] A. Papoulis, "Error analysis in sampling theory," *Proc. IEEE*, vol. 54, no. 7, pp. 947–955, 1966.

[20] ——, *Signal Analysis*. New York: McGraw-Hill, 1977.

[21] J. J. Pearson, D. C. Hines, S. Golosman, and C. D. Kuglin, "Video rate image correlation processor," *Proc. SPIE*, pp. 197–205, 1977.

[22] S. Peleg, D. Keren, and L. Schweitzer, "Improving image resolution using subpixel motion," *Pattern Recognit. Lett.*, vol. 5, pp. 223–226, 1987.

[23] H. Shekarforoush, M. Berthod, and J. Zerubia, "Subpixel image registration by estimating the polyphase decomposition of cross power spectrum," INRIA, France, Res. Rep. 2707, 1995.

[24] ——, "Subpixel image registration by estimating the polyphase decomposition of cross power spectrum," in *Proc. CVPR*, 1996, pp. 532–537.

[25] H. Shekarforoush, M. Berthod, J. Zerubia, and M. Werman, "Sub-pixel bayesian estimation of albedo and height," *Int. J. Comput. Vis.*, vol. 19, no. 3, pp. 289–300, 1996.

[26] H. Shekarforoush and R. Chellappa, "Data-driven multi-channel super-resolution with application to video data," *J. Opt. Soc. Am. A*, vol. 16, no. 3, pp. 481–492, 1999.

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

[28] Q. Tian and M. N. Huhns, "Algorithms or subpixel registration," *Comput. Vis. Graph. Image Process.*, vol. 35, pp. 220–223, 1986.

[29] R. Tsai and T. Huang, "Multiframe image restoration and registration," *Adv. Comput. Vis. Image Process.*, vol. 1, 1984.

[30] A. Verri, F. Girosi, and V. Torre, "Differential techniques for optical flow," *J. Opt. Soc. Amer. A*, vol. 7, pp. 912–922, 1990.

[31] M. Vetterli and J. Kovacevic, *Wavelets and Subband Coding*. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[32] E. W. Zimmers, C. Marathe, and J.-H. Wang, "Subpixel measurement of digital image registration based on polynomial regression," in *Proc. Int. Conf. Automation, Robotics, Computer Vision*, vol. 1, 1994, pp. 400–403.

Hassan Foroosh (Shekarforoush) obtained the B.Eng. and M.Eng. degrees in electrical engineering from Kingston Polytechnic, London, U.K., in 1990, the M.S. degree in robotics and computer vision from the University of Nice and INRIA, Sophia Antipolis, France, in 1993, and the Ph.D. degree in computer vision from the University of Nice and INRIA in 1996.

He was an Assistant Research Scientist and Lecturer at the University of Maryland, College Park, from 1997 to 2000. He is currently a Senior Research Scientist at the University of California, Berkeley. His research interests include computer vision, image processing, signal processing, computer graphics, visualization, and optimization theory. He has published numerous papers on super-resolution, image registration, motion estimation, and 3-D scene reconstruction.

Josiane B. Zerubia (S'81–M'81–SM'99) received the Electrical Engineer degree in 1981, the Dr.Eng. degree in 1986, the Ph.D. degree in 1988, and the "Habilitation" in 1994, all from ENSIEG, Grenoble, France.

She has been a Permanent Research Scientist at INRIA, Sophia Antipolis, France, since 1989. She has been Director of Research since July 1995. She was head of a remote sensing laboratory (PASTIS, INRIA Sophia-Antipolis) from mid-1995 to 1997. Since January 1998, she has been in charge of a new research group working on remote sensing (ARIANA, INRIA-CNRS-University of Nice). She is Adjunct Professor at Sup'Aero (ENSAE), Toulouse, France, since 1999. Previously, she was with the Signal and Image Processing Institute, University of Southern California, Los Angeles, as a Postdoctoral Researcher. She also worked as a Researcher for the LASSY (University of Nice and CNRS) from 1984 to 1988 and in the Research Laboratory of Hewlett-Packard in France and Palo Alto, CA, from 1982 to 1984. Her current interest is image processing (image restoration, image segmentation or classification, line detection, perceptual grouping, stereovision, and super-resolution) using probabilistic models, variational methods or neural networks. She also works on parameter estimation and optimization techniques. She has been a member of the editorial board of the French Society for Photogrammetry and Remote Sensing (SFPT) since 1998.

Dr. Zerubia has been part of the IEEE Signal Processing Society's IMDSP Technical Committee since 1997 and Associate Editor of the IEEE TRANSACTIONS ON IMAGE PROCESSING since 1998.

Marc Berthod received the Ph.D. degree from the University of Paris-6, France, in 1980.

He has been at INRIA since 1972 and headed the Pastis team, dedicated to image understanding and focused on remote sensing from 1982 to 1996. He has authored or coauthored more than 60 papers on knowledge-based and statistical pattern recognition. He has been in charge of many research projects, involving industrial or institutional partners. He has also been consulting on a regular basis for several major firms (GTE Labs, Sagem, Cetia), and a member of several scientific committees. Throughout his career, his research has been on pattern recognition. From 1973 to 1978, he worked mostly on handwritten character recognition—first handprint then cursive writing. Then, he shifted to image understanding, with two main research themes: statistical modeling (hierarchical Markov models, practical optimization algorithms, and automatic parameter estimation), and knowledge based approaches (design of expert systems for automatic classification of patterns in dedicated domains, for generic scene interpretation and for the automatic control of image processing procedures). The main application domains covered are depth recovery from satellite or aerial data (works on stereo, high-definition stereo, synthetic aperture radar imagery, shape from shading), extraction of cartographic information (roads, rivers, urban areas, etc.) from satellite imagery, using high-level and low-level fusion techniques, and extraction of geologically meaningful structures from digital elevation models. He headed INRIA's research unit, Sophia Antipolis, France, from 1996 to 2001, and is presently at INRIA's Science Office.