AN EXTENDED 1-D PHASE CORRELATION ALGORITHM FOR IMAGE REGISTRATION

Haldun Sarnel
Electrical and Electronic Engineering Department,
Dokuz Eylül University,
Buca-Tınaztepe, Izmir, Turkey

ABSTRACT

The 1-D phase correlation algorithm proposed by Alliney for translational image registration is extended by including phase correlations of the two diagonal projections of images to gain robustness against the wide-band noise. The best M points from the phase correlation functions of the vertical and horizontal projections are selected composing $M \times M$ translation candidates. Each candidate translation is assigned a rate on the basis of support given by the best M points from the correlation functions of the diagonal projections. Because of the random distribution of the false translations the true translation only gathers support thus takes the first place after the rating computations. A simple cost function, which favors the true translation in most cases, is offered. Experimental results with $M = 12$ chosen show that the extended algorithm provides a valuable robustness especially for images of wide-band spectrum. The 1-D extended algorithm may also replace the 2-D algorithm when relatively small translations are expected.

1. INTRODUCTION

Geometrically aligning two digital images containing some common scene is known as *image registration* in the image processing literature. The simplest image registration task involves accurate and reliable measurement of the relative displacement between two images when translational difference exists only. This kind of registration task has been the main subject in many practical applications such as guiding a missile to a preselected target by comparing a reference image with a sequence of images taken by a video camera or measuring the velocity of a moving imaging system by using two images acquired within a short and accurate known time interval. Image processing and computer vision applications based on image subtraction (e.g. detection of changes in a scene or automatic inspection of some industrial products) can only be achieved after the successful registration of the two images to be compared.

2-D phase correlation shows remarkable performance when pure planar translations are expected between images. It can even cope with the images corrupted by noise and subjected to changes in illumination or in gain thus making itself a favourite choice for many practical applications [1,2,3]. Phase correlation has also been popular in the field of motion estimation and compensation-based video coding for low bit-rate transmission of video frames [4]. In a relatively recent work [5], it has been shown that following some proper coordinate transforms in the Fourier Transform plane, 2-D phase correlation can be used to register two images not only differing by translation but also by planar orientation and scale change.

In the 1-D phase correlation algorithm [6], horizontal and vertical projections of the two images are first computed. A phase correlation between the vertical projections yields a peak at the point of the relative horizontal translation. In the same manner, a phase correlation between the horizontal projections leads to detection of the relative vertical translation.

When images contain wide-band noise with relatively high variance or when translations are not small, location of the peak in a phase correlation function in the 1-D algorithm is usually not at the correct translation point. In such cases the conventional 1-D phase correlation algorithm fails since it simply searches for the peak in a correlation function for a decision. In spite of the degradation of the correlation function the correct translation point is still among the points with relatively high correlation values. In this paper, the conventional algorithm is extended by phase correlations of the two diagonal projections of images. Only the best M points in each of the usual correlation functions are considered candidates for the horizontal and vertical parts of the true translation. Each translation formed by one candidate point from vertical correlation function and one from horizontal correlation function, is assigned a rate on the basis of support given (if any) by the best M points from the correlation functions of the diagonal projections. The translation with the best rate is decided to be the true translation. The diagonal correlation functions also bear information about the true translation. Therefore, a confirmation is provided by the relevant positions (among those with high correlation values) in the diagonal phase correlation functions.

In the following section the 2-D and 1-D phase correlation algorithms and their performances are summarized. Section 3 explains the implementation of the extended 1-D algorithm in detail. Section 4 presents the performance results of the extended algorithm for noisy images. Finally, a conclusion is given in Section 5.

2. 2-D AND 1-D PHASE CORRELATIONS

2.1 The Conventional Algorithms

Let $f_1(x,y)$ and $f_2(x,y)$ denote two images taken at different times and assume that there is a relative translation between the image contents, that is

$$f_2(x,y) = f_1(x-x_0, y-y_0) \tag{1}$$

Most of the information about the relative displacement between two images is contained in the phase of their cross-power spectrum. The whitened or normalized cross-power spectrum $G(u,v)$ is defined as

$$G(u,v) = \frac{F_2(u,v)\,F_1^*(u,v)}{\left|F_2(u,v)\,F_1^*(u,v)\right|} = e^{-j(ux_0+vy_0)} \tag{2}$$

where F_1 and F_2 are Fourier transforms of f_1 and f_2 respectively, and F_1^* is the complex conjugate of F_1. The inverse Fourier transform of the $G(u,v)$ yields a delta function located at (x_0, y_0) in the spatial domain. Thus relative translations x_0 and y_0 are simply determined by searching for the position of the peak. Since f_1 and f_2 are discrete and do not vanish at the boundaries of the image frames in practice, the delta function is replaced by a unit amplitude narrow pulse.

The phase correlation algorithm is insensitive to brightness and contrast differences between images and copes with the narrow-band noise since all spatial frequencies contribute to the correlation function. The algorithm can even succeed in the presence of high level wide-band noise and perspective distortions.

Unfortunately the conventional phase correlation algorithm is based on the use of 2-D Fourier transforms and its real-time implementation needs costly hardware to overcome the computational load. In order to reduce the computational complexity, an algorithm using 1-D Fourier transformations of the image projections has been proposed [6]. If x and y projections f_x and f_y of an image f are defined as

$$f_x(y) = \int_{-\infty}^{\infty} f(x,y)\,dx \quad , \qquad f_y(x) = \int_{-\infty}^{\infty} f(x,y)\,dy \tag{3}$$

then the projection slice theorem indicates that

$$\mathcal{F}\{f_x(y)\} = F(v) = [F(u,v)]_{u=0} \quad , \tag{4}$$

$$\mathcal{F}\{f_y(x)\} = F(u) = [F(u,v)]_{v=0} \tag{5}$$

Let $f_{1y}(x)$ and $f_{2y}(x) = f_{1y}(x-x_0)$ be the y-projections of $f_1(x,y)$ and $f_2(x,y)$ respectively. Their Fourier transforms satisfies the equation

$$F_{2y}(u) = e^{-jux_0} F_{1y}(u) \tag{6}$$

and x_0 can be determined by locating the peak of the 1-D inverse transform.

$$g_y(x) = \mathcal{F}^{-1}\left\{\frac{F_{2y}(u)\,F_{1y}^*(u)}{\left|F_{2y}(u)\,F_{1y}^*(u)\right|}\right\} = \mathcal{F}^{-1}\{e^{-jux_0}\} = \delta(x-x_0) \tag{7}$$

A similar procedure is used to find y_0. Because only 1-D transformations are required, the number of complex operations is dramatically reduced.

2.2 Performance of the 1-D Algorithm

When the images are multiplied by an appropriate window function $w(x,y)$ before taking the projections, the 1-D algorithm provides satisfactory responses for relatively small translations of the image pair (typically for y_0 and $x_0 < 0.25N$, given $N \times N$ images) [6]. The idea behind the windowing is to reduce the image pixel values near the image boundaries so the disturbances induced by the window finiteness of the real-world images become less effective. Several 2-D window functions monotonically decreasing and symmetrical with respect to the image center such as exponential, Gaussian or raised cosine can be applied. However, the performance of the algorithm also depends on the type of the window function. Although it seems that the 1-D algorithm described above can be used instead of the 2-D algorithm when relatively small translations are expected, its straightforward implementation leads to the poor results in practice because of two major reasons:

- The information available in the image $f(x,y)$ is degraded by the projection.
- The noise and partial dissimilarities in the contents of the images both introduce an array of random phase error in the cross-power spectrum. This results in a random contribution of a certain amount of energy to the correlation function that is shared by N locations in the 1-D algorithm instead of N^2 in the 2-D phase correlation.

To achieve robustness against noise and/or reliable registration performance at larger translations it is essential that more of image information be used in the computations.

3. EXTENDED 1-D ALGORITHM

In this work, extending the conventional 1-D phase correlation algorithm with the correlations of the projections along the directions parallel to the image diagonals is explored in an attempt to gain robustness especially against the wide-band noise.

3.1 Windowing Images

The window function used in this work is in the form of rotationally symmetrical 2-D raised cosine. The 2-D window coefficients $w(x,y)$ are found using the 1-D curve as follows.

$$w(x,y) = W(x)W(y) \tag{8}$$

$$W(n) = 0.5\left(1 - \cos\frac{2\pi n}{N}\right) \qquad n = 0,1,2,\ldots,N-1 \tag{9}$$

The projections of the windowed images $w(x,y)f_i(x,y)$ are used in the algorithm. Here, $N \times N$ images are assumed.

3.2 Projections

Two new variables, one for the main diagonal (from the top left to the bottom right) axis $s = x+y$, and the other for the secondary diagonal (perpendicular to the main diagonal) axis

$t = x - y$ are introduced. The projections of a digital image represented by a small matrice given below are taken as follows.

$$f(x,y) = \begin{matrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \\ m & n & o & p \end{matrix}$$

$$f_y(x) = \{a+b+c+d,\ e+f+g+h,\ i+j+k+l,\ m+n+o+p\}$$
$$f_x(y) = \{a+e+i+m,\ b+f+j+n,\ c+g+k+o,\ d+h+l+p\}$$
$$f_t(s) = \{a+b+e,\ c+f+i+d+g+j+m,\ h+k+n+l+o,\ p\}$$
$$f_s(t) = \{m+i+n,\ e+j+o+a+f+k+p,\ b+g+l+c+h,\ d\}$$

For an $N \times N$ digital image, x, y, s and $t = 0,1,2,\ldots,N-1$

When taking the diagonal projections, lines of projections are combined two by two so that the size of the diagonal projections reduces to N, thus equals the size of the other projections.

3.3 Implementation

By applying the 1-D phase correlation algorithm to the t projections $f_{1t}(s)$ and $f_{2t}(s)$ of the two images the correlation function

$$g_t(s) = \mathcal{F}^{-1}\left\{\frac{F_{2t}(\omega)F_{1t}^*(\omega)}{\left|F_{2t}(\omega)F_{1t}^*(\omega)\right|}\right\} \qquad (10)$$

is obtained. The peak in the correlation function is expected at $s_0 = x_0 + y_0$. Similarly, by applying the algorithm to the s projections of the two images a correlation peak is expected at $t_0 = x_0 - y_0$.

After computing the phase correlation functions $g_y(x)$, $g_x(y)$, $g_t(s)$ and $g_s(t)$, the true translation between the two images can be determined by the following method. From each correlation function the first M points with the highest correlation values are selected and put on four separate lists in descending order by the correlation value. Let us denote the lists by X, Y, S and T, respectively. Then, for each possible translation (x_i, y_j) from the X and Y lists (of total $M \times M$ combinations) a rate $r(i,j)$ based on the available support from the S and T lists are computed as follows:

1. Do Steps 2-5 for all (x_i, y_j), $i = 1,2,\ldots,M$ and $j = 1,2,\ldots,M$
2. Initialise $r(i,j)$ to the largest integer.
3. Do Steps 4-5 for all (s_k, t_l), $k = 1,2,\ldots,M$ and $l = 1,2,\ldots,M$
4. if $|x_i + y_j - s_k| \le 2$ and $|x_i - y_j - t_l| \le 2$, current translation (x_i, y_j) is supported. Compute the cost for the current translation (x_i, y_j): $c_{kl} = i + j + 2(k+l)$
5. if $c_{kl} < r(i,j)$, update the rate: $r(i,j) = c_{kl}$

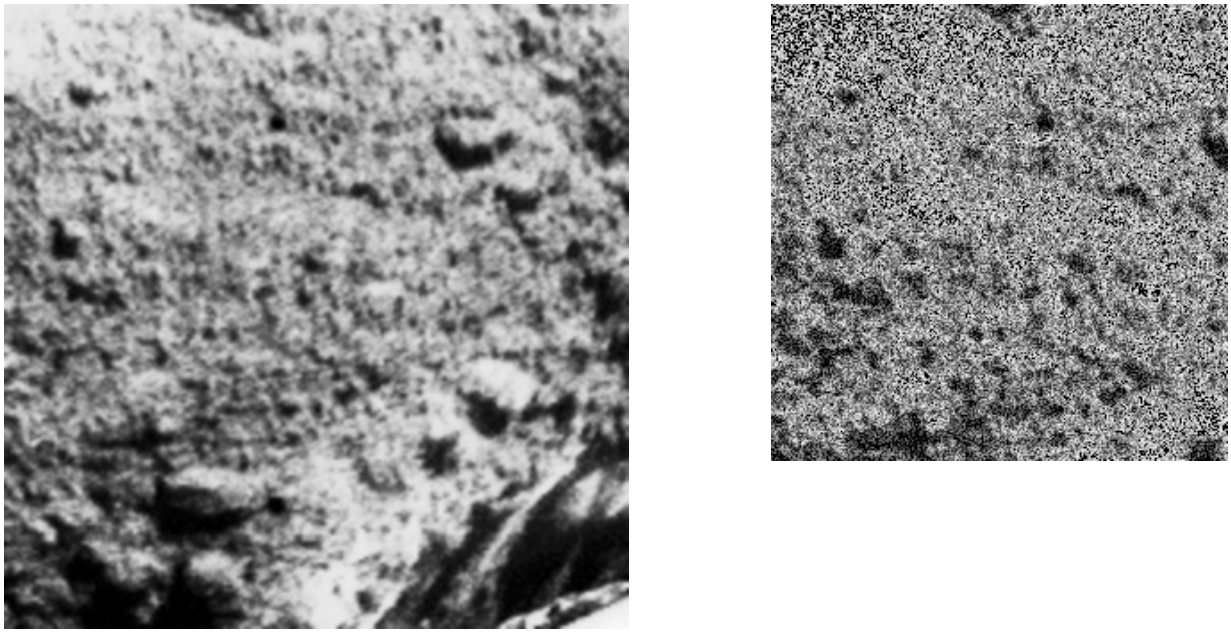
Because of the nearly random distribution of the false translation combinations, the true translation and possibly a few false translations in addition gather support in the rating computations. At the end of the computations, the minimum of the rates is found. Let $r(m,n)$ be the minimum. This indicates that the translation (x_m, y_n) is decided to be the true translation. A cost function c_{kl} is needed in order to grade all translation combinations which gather support from the S and T lists so that the best one can be chosen among them. The simple cost function given in Step 4 is in favor of the translations located at the top positions on a list. This increases the probability of assigning a better cost to the true translation. Positions in the S and T lists give more weight to the suggested cost function than those in the X and Y lists do, because diagonal projections contain more image information compared to the vertical and horizontal projections. However, the suggested cost function does not guarantee that the true translation will always take the minimum rate. Other cost functions having a higher complexity can be investigated to obtain a better performance, however, the suggested function gives satisfactory results when M is relatively small.

3.4 Computational Load

The computational load of the extended algorithm is higher than twice of that of the conventional algorithm since two more phase correlations are added. Computations increase by the value of M and by the complexity of the cost function, but in a typical application the total time required for the algorithm will be less than triple of that of the conventional algorithm. Therefore the computational load still remains very small compared to that of the 2-D algorithm.

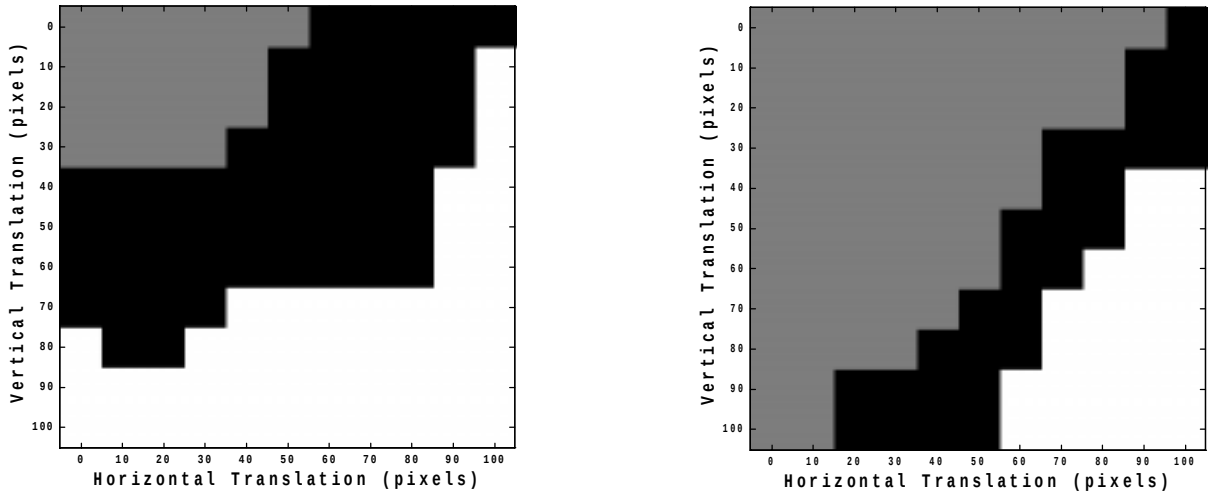
4. EXPERIMENTAL RESULTS

In order to compare the performances of the 1-D conventional algorithm and the extended algorithm some experimental work was done. A test image of 350 x 350 pixel with 256 grey levels was used in the experiments. From the test image a reference image of 256 x 256 pixel, located at the upper-left part, was cut out. Uniformly distributed zero-mean white noise with a standard deviation σ_n was added to the reference image as shown in Fig. 1. A set of 256 x 256 pixel images were also cut out from the same test image at different shifts on a grid spaced by 10 pixels in both directions. For each couple, the noisy reference image and each of the shifted images in the set, the extended 1-D algorithm was applied after windowing the images and taking the projections along the x, y, s and t axes. The number of candidate points, $M = 12$ was chosen. The results were recorded on a *registration performance map* as shown in Fig. 2. Each cell in the map contains the results of both algorithms for the couple between which there is a relative shift specified by its location. A black cell indicates a correct translation found by the extended algorithm at which the conventional algorithm failed. Both algorithms were successful at the translations with a grey cell and failed at those with a white cell. The registration performance map shown in Fig. 2(a) indicates that the extended algorithm provides about %100 increase in the translation range in which a successful registration is obtained when a high noise power is expected in the images. Here, the standard deviations of the reference image and the wide-band noise were about the same. Errors of ± 1 pixel in the translations determined by the algorithm were allowed since accuracy at severe noise inevitably reduces. Another experiment was performed on the same test image at a low noise power in order to see the capability of the extended algorithm when images can be taken in well-controllable



(a) (b)

Figure 1. (a) 350 x 350 *moon* image. (b) The reference image was cut out from (a) then noise with $\sigma_n = 50$ was added.



(a) (b)

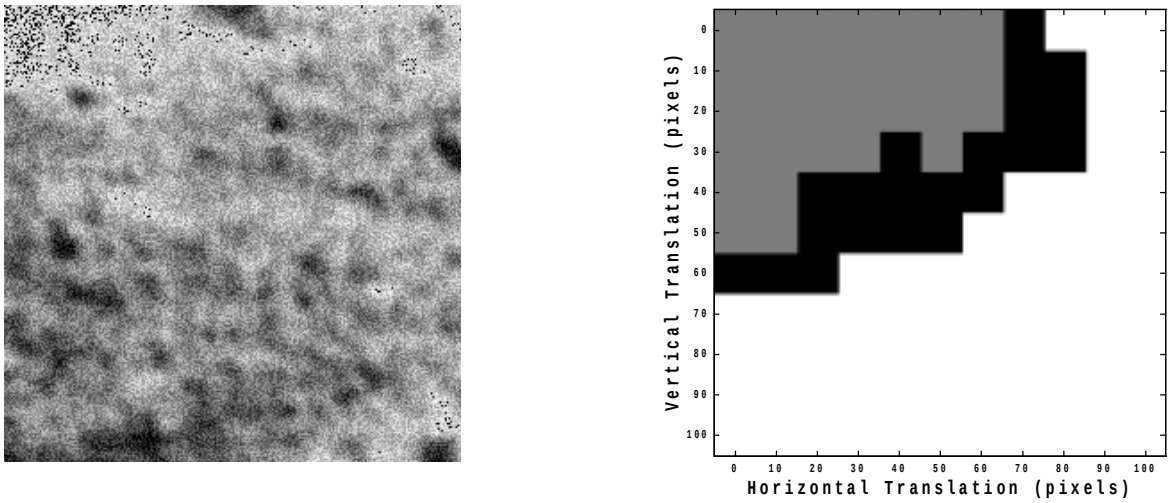
Figure 2. Registration performance maps: (a) $\sigma_n = 50$. (b) $\sigma_n = 5$.

environments. Fig. 2(b) shows that the increase in the translation range provided by the extended algorithm is between %25 and %35 for this case.

In order to see the performance of the extended algorithm on images with a narrow-band spectrum, a heavily blurred version of the test image was created and used in the experiments in the same way. The performance of the extended algorithm is no longer remarkable at a high noise power as shown in Fig. 3(b). This is due to the contribution of the spatial frequencies in the upper band occupied by the noise only. If the bandwidth of the images is known a priori, better results can be obtained by low-pass filtering in the frequency domain.

5. CONCLUSION

A good translational image registration algorithm for real-time applications is desired to have three main features: a low computational load, a large translation range in which the algorithm reliably works and robustness against both wide-band and narrow-band noise. 2-D phase correlation has a large computational load and 1-D phase correlation suffers from the severe wide-band noise due to the insufficient image information. The extended 1-D phase correlation algorithm presented in this paper doubles the image information involved in the computations by the use of diagonal projections. The



(a) (b)

Figure 3. (a) The Reference image cut out from a blurred version of *moon* image after noise with $\sigma_n = 25$ was added (b) Registration performance map for the image in (a).

experimental results show that the extended algorithm is very robust against wide-band noise especially for images of wide-band spectrum. The extended algorithm dramatically outperforms the conventional 1-D algorithm as the power of wide-band noise increases. Even for the images containing a weak noise, the algorithm yields an increased translation range by making itself preferable to the conventional 1-D algorithm. With the low computational load and the robustness to the wide-band noise the extended 1-D algorithm can even be an advantageous alternative to the 2-D phase correlation algorithm when relatively small translations are expected.

6. REFERENCES

[1] J.J. Pearson, D.C. Hines, S. Golosman, and C.D. Kuglin, "Video-rate Image Correlation Processor," in *Proc. SPIE, Appl. of Digital Image Processing* (IOCC 1977), 1977, pp. 197-204.

[2] C.D. Kuglin, A.F. Blumenthal, and J.J. Pearson, "Map-matching Techniques for Terminal Guidance Using Fourier Phase Information," *SPIE (Digital Processing of Aerial Images)*, vol. 186, pp. 21-29, 1979.

[3] C. Morandi, F. Piazza, and R. Capancioni, "Digital Image Registration by Phase Correlation between Boundary Maps," *IEE Proc. E, Comput. and Digital Tech.*, vol. 134, pp. 101-104, 1987.

[4] R.J. Clarke, "Basic Principles of Motion Estimation and Compensation," *IEE Colloquium on 'Applications of Motion Compensation'*, pp. 1-7, 1990.

[5] B.S. Reddy and B.N. Chatterji, "An FFT-Based Technique for Translation, Rotation, and Scale-Invariant Image Registration," *IEEE Trans. Image Processing*, vol. 5, pp. 1266-1271, 1996.

[6] S. Alliney and C. Morandi, "Digital Image Registration Using Projections," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-8, pp. 222-233, 1986.