Estimating Translation/Deformation Motion through Phase Correlation *

Filiberto Pla† and Miroslaw Bober‡

†Dept. of Computer Science.
University Jaume I, 12071 Castelló, Spain.
‡Dept. of Electrical and Electronic Engineering.
University of Surrey, Guilford GU2 5XH, UK.
pla@uji.es m.bober@ee.surrey.ac.uk

Abstract. Phase correlation techniques have been used in image registration to estimate image displacements. These techniques have been also used to estimate optical flow by applying it locally. In this work a different phase correlation-based method is proposed to deal with a deformation/translation motion model, instead of the pure translations that the basic phase correlation technique can estimate. Some experimentals results are also presented to show the accuracy of the motion paramenters estimated and the use of the phase correlation to estimate optical flow.

Key Words : Motion, Optical Flow, Image Registration, Phase Correlation.

1 Introduction

Estimating visual motion is a valuable information for many machine vision applications, like traffic monitoring, surveillance, image coding, etc. The work presented here has been aimed at developing some techniques for accurate enough visual motion estimation and optical flow. Optical flow estimation methods measure the velocity and displacement vectors perceived from the time-varying image intensity pattern, that is, they measure the apparent 2D image plane motion which is the projected 3D motion of the objects of the scene in the image plane. Most of existing techniques for optical flow estimation can be classified into *Gradient-based techniques*, that is, methods based on the so-called *optical flow equation* [7][12]; *Block matching-based techniques* try to overcome the aperture problem by assuming that all pixels in a block undergo the same motion [6][11]; *Spatiotemporal Energy-based techniques* which exploits the equivalent optical flow equation in the frequency domain [9]; *Bayesian techniques* utilize probabilistic smoothness constraints, usually in the form of a Gibbs ramdom field [10][8].

According to Tekalp's classfication [13], the above mentioned methods correspond to the *non parametric models*. *Parametric models* try to work out the

* This work has been partially supported by the *JP/ij-estades/96* grant from the *Generalitat Valenciana*, and ESPRIT project EP21007 IOTA.

motion model of the ortographic or perspective projection of the 3D motion in the image plane. Motion models are also used in block matching techniques, thus, in order to look for the best matching block in the other image, some motion model has to be assumed. There are generalized 2D image motion models. One of the most used is the affine coordinate transformation,

$$x' = a_1 x + b_1 y + c1 \ ; \qquad y' = a_2 x + b_2 y + c2 \tag{1}$$

with (x', y') being the image transformed coordinates, and $a_1, b_1, c_1, a_2, b_2, c_2$ the motion parameters of the affine model. The affine tranformation corresponds to the ortographic projection of a 3D rigid motion of a planar surface. A particular instance of this model is a translation/deformation model, which takes into account tranlations and scale changes in the image plane,

$$x' = a_1 x + c1 \ ; \qquad y' = b_2 y + c2 \tag{2}$$

In the work we are presenting here, a block matching technique is used to estimate motion and optical flow. This block matching technique will be analyzed and extended from the translational model to the deformation/translation motion model (equation 2). Next section will introduce the basics of the phase correlation technique, which will be further analyzed and extended in a subsequent section. Some experimental results are also presented to shown the accuracy of the extended method.

2 Phase correlation techniques

Another type of techniques which could be included in the block matching techniques are the *phase correlation* techniques. Phase corelation techniques have been used in image registration [1][5]. Although in image registration phase correlation techniques are applied to the entire image, the phase correlation method has also been used in optical flow estimate applying it locally, using a small window around the point of interest where the image flow velocity is being estimated [13].

The basic phase correlation method estimates the relative shift between two image blocks by means of a normalized cross-correlation function computed in the 2D spatial Fourier domain. It is also based on the principle that a relative shift in the spatial domain results in a linear phase term in the Fourier domain. Therefore, let $f_k(x, y)$ be the image intensity function at time k, and let $f_{k+1}(x, y)$ be the image intensity function at time $k + 1$.

If we assume that $f_k(x, y)$ has undergone a translation (x_0, y_0), then $f_{k+1}(x, y) = f_k(x - x_0, y - y_0)$. Taking the Fourier tranformation of both functions, the image displacement (x_0, y_0) can be calculated by means of the normalized cross-power spectrum funtion, as

$$\overline{C}_{k,k+1}(u, v) = \frac{F_{k+1}(u, v) F_k^*(u, v)}{|F_{k+1}(u, v) F_k^*(u, v)|} \tag{3}$$

Since $F_k(u,v)$ and $F_{k+1}(u,v)$ are related as $F_{k+1}(u,v) = e^{-j(ux_0+vy_0)}F_k(u,v)$, the inverse of the above equation results in $\overline{c}_{k,k+1}(x,y) = \delta(x-x_0, y-y_0)$ which is the cross correlation function consisting of an impluse whose location gives the displacement vector. Ideally, we would expect to find a single impulse in the phase correlation function, but in practice several factors contribute to the degeneration of the phase correlation function.

On the other hand, the phase correlation method has some advantages. One important feature is that the phase correlation method is relatively insensitive to changes in illumination, because variations in the mean value or multiplication for a constant do not affect the Fourier phase. Since the phase correlation function is normalized with the Fourier magnitude, the method is also insensitive to any other Fourier-magnitude-only degradation.

The use of the DFT to compute the phase correlation function has some undesirable efects:

Boundary effects. To obtain a perfect impulse, the shift in the spatial domain has to be ciclic. Since things appearing at one end of the block (window) generally do no appear at the other end, the impulse degenerate into a peak. Further, since the 2D DFT assumes periodicity in both directions, discontinuities from left to right boundaries, and from top to bottom, may introduce spurious peaks.

It is well known that the boundary effects due to the finitness of the image (block) frame become less relevant if the image function has small values near the frame boundaries. Therefore, the rectangular window representing the framing process, may be substituted by a weighting window $w(x,y)$ that produces the decay of the image function values near the boundaries. In this case, we have addopted a Gaussian-like windowing function to our approach [1],

Spectral leakage. In order to observe a perfect impulse, the components of the displacement vector must correspond to an integer multiple of the fundamental frequency. Otherwise, the impulse degenerates into a peak due to the well known spectral leakage phenomenom. Thus, if we assume that the peak values are normally distributed around its maximum, then the actual maximum would be the mean of this distribution.

Range of displacement estimates. Since the 2D DFT is periodic with the block size (N,M), only displacements (x_0,y_0) can be detected if they satisfy that $-N/2 \leq x_0 \leq N/2$ and $-M/2 \leq y_0 \leq M/2$ due to the wrapping effect of the Fourier transform. Therefore, in order to estimate diplacements in the range $(-d,d)$ along a spatial direction, the block size has to be theoretically at least of $2d$ size in this spatial direction.

3 Extending the phase correlation technique to estimate translation/deformation motion

So far, the phase correlation function has been used to estimate displacements assuming a pure translation model within the image or block. Some work has been also done in image registration to estimate pure rotation of images by means of an iterative-search procedure [5]. In order to extend the phase correlation techniques to a different motion model and taking into account some properties of the Fourier transform, let us analyze what would happen if the motion undergone for the pixels in the block is not a pure translation.

Let us consider that an image (block or window) $f_k(x, y)$ at instant k undergoes a translation/deformation motion as shown in equation 2. Therefore, the corresponding gray level values $f_{k+1}(x, y)$ of that block at instant $k+1$ are related with $f_k(x, y)$ as $f_{k+1}(x, y) = f_k(a_1 x + c1, b_2 y + c_2)$. Hence, taking the definition of the Fourier Transform of the above expression, and making the transformation $x' = a_1 x + c1$ and $y' = b_2 y + c_2$, the expression for $F_{k+1}(u, v)$ becomes

$$F_{k+1}(u, v) = \frac{1}{|a_1 b_2|}\ e^{j(\frac{uc_1}{a_1} + \frac{vc_2}{b_2})} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_k(x', y')\ \ e^{-j(\frac{ux'}{a_1} + \frac{vy'}{b_2})} dx' dy'$$

which is a combination of the *shift* and *similarity* theorems of the Fourier transform. In the above expression, only positive values of a_1 and b_2 will be considered, since negative values would mean a deformation plus a symmetry transformation in the image, situation that cannot occur in a real moving scene. Hence, and rewriting the above equation as a function of the Fourier transform of $f_k(x, y)$,

$$F_{k+1}(u, v) = \frac{1}{a_1 b_2}\ e^{j(\frac{uc_1}{a_1} + \frac{vc_2}{b_2})} F_k(\frac{u}{a_1}, \frac{v}{b_2}) \tag{4}$$

In order to calculate the parameters of the motion undergone by $f_k(x, y)$, note that from the above equation, the magnitude of $F_{k+1}(u, v)$ is $|F_{k+1}(u, v)| = \frac{1}{a_1 b_2}\ |F_k(\frac{u}{a_1}, \frac{v}{b_2})|$. Thus, from the above equation we can obtain the values of the deformation parameters a_1 and b_2 as follows. Let us calculate the weighted energy of the spectrum for F_{k+1} as

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |u\ F_{k+1}(u, v)|\ dudv = \frac{1}{a_1 b_2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |u\ F_k(\frac{u}{a_1}, \frac{v}{b_2})|\ dudv$$

making the transformation $u' = u/a_1$ and $v' = v/b_2$ yields

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |u\ F_{k+1}(u, v)|\ dudv = a_1 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |u'\ F_k(u', v')|\ du' dv'$$

Solving for a_1, and analogously for b_2, and exchanging integrals by summations for the discrete case, we reach to the expression

$$a_1 = \frac{\sum_{u=-N/2}^{N/2} \sum_{v=-M/2}^{M/2} |u\ F_{k+1}(u, v)|}{\sum_{u=-N/2}^{N/2} \sum_{v=-M/2}^{M/2} |u\ F_k(u, v)|}\ ; \qquad b_2 = \frac{\sum_{u=-N/2}^{N/2} \sum_{v=-M/2}^{M/2} |v\ F_{k+1}(u, v)|}{\sum_{u=-N/2}^{N/2} \sum_{v=-M/2}^{M/2} |v\ F_k(u, v)|} \tag{5}$$

Given $F_k(u, v)$ and $F_{k+1}(u, v)$, equations 5 provide the values for a_1 and b_2. On the other hand, the translation parameters, c_1 and c_2, can be obtained as follows. Given $F_{k+1}(u, v)$, let us transform it into $G_{k+1}(u, v)$ using the estimated a_1 and b_2 as $G_{k+1}(u, v) = F_{k+1}(a_1 u, b_2 v)$. Applying now equation 4, yields

$$G_{k+1}(u,v) = \frac{1}{a_1 b_2} e^{j(uc_1+vc_2)} F_k(u,v)$$

Note that if we now calculate the normalized cross power spectrum of $G_{k+1}(u,v)$ and $F_k(u,v)$ we obtain $\overline{C}_{k,k+1}(u,v) = e^{j(uc_1+vc_2)}$ whose inverse Fourier transform will give an impulse function at $(-c_1, -c_2)$, corresponding to the translation parameters.

Therefore, given the estimated motion parameters for a pixel, to estimate the optical flow field we can compute the velocity vectors (v_x, v_y) at every pixel as $(v_x, v_y) = (x - (a_1 x + c_1), y - (b_2 y + c_2))$.

4 Experimental Results

Series	$\sigma(a_1)$	$\sigma(b_2)$
1	0.00197139	0.00750315
2	0.0128401	0.00464964
3	0.015433	0.00788849

Table 1. Standard deviations of estimated deformation parameters.

In order to test the proposed approach, two type of experiments are shown in this section. The first experiments are directed to show the accuracy of the estimated parameters. A second set of experiments will compare the optical flow estimation made by the proposed method with respect to other methods.

To compute the DFT we use the Fast Fourier Transform (FFT) algorithm. Therefore, for image block sizes not multiple of 2, the zero padding technique is used to complete a power 2 block size. Moreover, we utilized the nearest neighbour technique as interpolation function, but calculating $F_{k+1}(u,v)$ in a mesh k times the size of the block, usually $k = 2$.

Figure 1.a shows the *Tree* image used in these experiments, the same as used in [2][3] and it will be used later for comparison purposes. This image was transformed using three series of deformation parameters. The transformed images $g(x,y)$ were calculted from the original image $f(x,y)$ using the ideal interpolator function. The series of transformed images were calcuated as follows:

1. Fixing $a_1 = 1$, $c_1 = 0$ and $c_2 = 0$ and varying b_2 from 0.9 to 1.1 incrementing it in steps of 0.02.
2. The same as in the above serie, but fixing $b_2 = 1$ and varying a_1 in the same way.

3. The same as in the above serie, but now varying a_1 and b_2 in the same way as in the series 2 and 1, respectively.

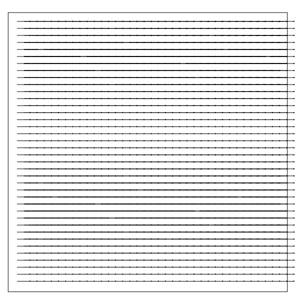
The motion parameters were estimated using the whole 150×150 image as a block. In table 1 we can see the standard deviations of the estimated values for a_1 and b_2. Note that the maximum standard deviation is 0.015433, so the expected accuracy estimate is around this order.

	Translating			Diverging		
Technique	Av. Error	Std	Density	Av. Error	Std	Density
Lucas [12]	1.75	1.43	40.8%	3.05	2.53	49.4%
Barron [2]	0.36	0.41	76%	1.08	0.52	64.3%
Bober [3]	0.33	0.25	100%	3.69	4.39	100%
defor. model	1.79	1.10	100%	9.00	3.65	100%

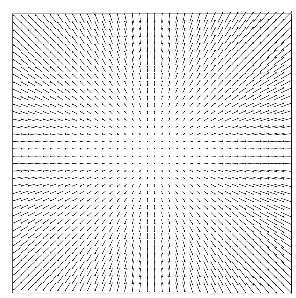
Table 2. Errors for the optical flow of the *Translating* and *Diverging Tree* (Ref. Bober et al., 1994).

Another experiment to test the accuracy of the method is related to the estimation of dense optical flow. Figure 1.b represents the ground optical flow field of the *Translating Tree*, and Figure 1.c represents the optical flow field of the *Diverging Tree* of figure 1.a. Both examples were used in [2][3]. The error measure utilized is the one defined in [2], and the same as the one used in the above mentioned works.

(a) Original image



(b) Ground translating flow



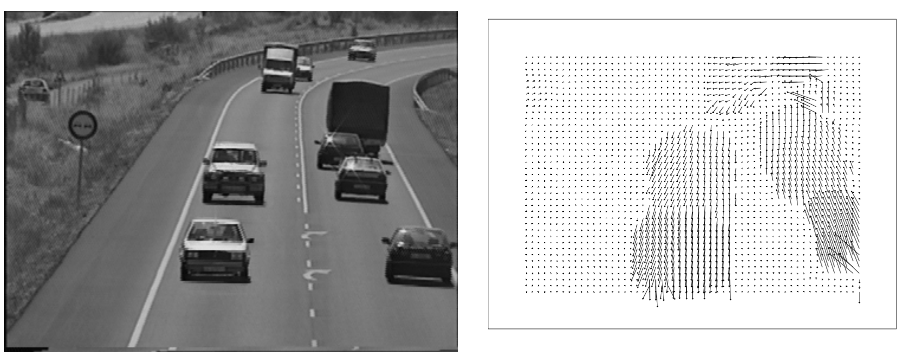
(c) Ground diverging flow

Fig. 1. The ground flow fields of the *Tree* image.

Table 2 shows the results presented in [3] comparing serveral optical flow methods on the same *Translating* and *Diverging Tree* images, where only the 3 best results for each example are shown, together with the results obtained by the proposed method. The flow estimated by all these variations in table 2 have been computed using a 32x32 window centered at every pixel.

Although errors from the proposed method are bigger with respect to the errors shown in table 2, note that errors provided by Lucas [12] and Barron [2] are estimated over 40% and 76% of the flow field for the *Translating Tree*, and 49% and 64% for the *Diverging Tree* respectively. Bober's method [3] has got the advantage that it is a combined flow estimation and segmentation method, so it also uses global information to estimate the velocity at each point, however, the approach presented here estimates the flow using only local information. In addition, neither smoothing or any type of filtering have been applied as a postprocessing to the raw flow field.

Figure 2 shows an example of the flow field computed where the background (road) is approximately stationary and the cars move in both directions. Looking at the figures we can notice that the optical flow discontinuities are tackled quite satisfactorily. The same applies to figure ??, where we can notice even the different motions undergone from different parts of the car, for example, the wheel. In this example, the car is almost stationary with respect to the camera and the background moves, however the rotation of the wheels can be appreciated with respect to the camera.



(a) Original image (b) Estimated flow field

Fig. 2. The *n234* sequence.

5 Conclusions

In this paper, the phase correlation method has been combined with a deformation/translation motion model, and some techniques have been developed to estimate the motion parameters assuming this motion model. Experimental results have shown the accuracy of the estimated deformation paramenters by the proposed method.

The developed phase correlation techniques have also been applied to estimate optical flow in image sequences providing satisfactory and encouraging results. A more accuarte subpixel estimation method would be necessary to obtain better optical flow estimates.

Futur extensions of the present work are directed to several issues. First of all, work is being carried out to try to extend the proposed deformation/translation motion model using phase correlation techniques to an affine motion model, which can deal with rotation, deformation and translation.

References

1. Alliney, S. and Morandi, C.; "Digital image registration using projections", *IEEE Trans. Patt. Anal. Mach. Intel.*, vol. PAMI-8, no. 2, pp. 222-233, 1986.
2. Barron, J.; Fleet, D. and Beauchemin, S.; "Performance of optical flow techniques", *Int. J. Comp. Vision*, vol. 12, pp. 43-77, 1994.
3. Bober, M. and Kittler, J.; "Robust motion analysis", Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, Seatle, pp. 947-952, 1994.
4. Bracewell, R. N.; *The Fourier transform and its applications*, McGraw Hill, 1986.
5. De Castro, E. and Morandi, C.; "Registration of translated and rotated images using finite Fourier transforms", *IEEE Trans. Patt. Anal. Mach. Intel.*, vol. PAMI-9, no. 5, pp. 700-703, 1987.
6. Ghraravi, H. and Mills, M.;"Block-matching motion estimation algorithms: New results", *IEEE Trans. Circ. and Syst.*, vol. 37, pp. 649-651, 1990.
7. Horn, B.K.P. and Schunk, B. G.; "Determining optical flow", *Artificial Intelligence*, vol. 5, no. 3,pp. 276-287, 1993.
8. Iu, S. L.; "Robust estimation of motion vector fields with discontinuity and occlusion using local outliers rejection", *SPIE*, vol. 2094, pp. 588-599, 1993.
9. Jacobson, L. and Wechsler, H.; "Derivation of optical flow using a spatiotemporal-frequency approach", *Computer Vision, Graphics and Image Processing*, vol. 38, pp. 29-65, 1987.
10. Konrad, J. and Dubois, E.; "Comparison of stochastic and deterministic solution methods in Bayesian estimation of 2D motion", *Image and Vision Computing*, vol. 9, pp. 215-228, 1991.
11. Liu, B. and Zaccarin, A.;"New fast algorithms for the estimation of block motion vectors", *IEEE Trans. Circ. and Syst. Video Tech.*, vol. 3, no. 2, pp. 148-157, 1993.
12. Lucas, B. D. and Kanade, T.; "An iterative image registration technique with an application to stereo vision", *Proc. DARPA Image Understanding Workshop*, pp. 121-130, 1981.
13. Tekalp, A. M.; *Digital video processing*, Prentice Hall, 1995.