A PSEUDO 3D MOTION ESTIMATOR FOR MOVING OBJECT ESTIMATION IN NOISY VIDEO SEQUENCES

Christopher L. Topping and Jonathon A. Chambers
Dept. of Electrical and Electronic Engineering,
Imperial College, Exhibition Road
London SW7 2BT, UK
Tel: +44 (0)171 594 6235; fax: +44 (0)171 594 6324
e-mail: c.topping@ic.ac.uk

ABSTRACT

This paper presents a three-dimensional motion estimator for use in cases where we have noisy video sequences containing one moving object on a stationary background. The motion estimator is to be used as part of an image enhancing pre-processing step. A Parallel Extended Kalman Filter (PEKF) developed by J. B. Burl is at the heart of this motion estimator, together with additional data mapping techniques for its expansion to what is referred to as a Pseudo 3D motion estimator. This motion estimator is shown with simulations to have much potential for object motion estimation at low image SNR levels, ($<$ 5dB).

1 INTRODUCTION

A noise free video sequence with small levels of object motion, has relatively high correlation in the temporal direction. In contrast, with the presence of significant observation noise, such temporal correlation is reduced. Space-invariant three-dimensional, (3-D), filters can be used, when there is no motion, for direct noise reduction without losing high spatial frequencies. However, when a moving object is present, there remains strong temporal correlation mostly along the trajectory of the object. Thus, motion compensation is required if spatio temporal filtering is to be successful in preserving much of the high spatial frequency information.

Optic flow [1], feature based [4], and cross correlation based motion estimation techniques have been used to estimate object motion. However, the disadvantages of these techniques are: for optic flow, its noise sensitivity; for the feature based approaches, numerous post processing steps are necessary for feature extraction; and for cross correlation based techniques, the need for interpolation techniques for reasonable sub-pixel accuracy, as well as the need for search techniques. The objective is to overcome these limitations by producing a motion estimator which is robust to noise, whilst maintaining low computational complexity. This paper, presents a motion estimation algorithm that overcomes some of the problems outlined above. Pseudo three-dimensional motion can be estimated without the need for feature extraction, differential methods, or the need for search techniques. The new estimator presented is referred to as a pseudo 3D motion estimator since the true 3D motion parameters are not estimated. However, motion compensation for 3D object motion can be accomplished provided that some motion constraints are satisfied. The noise robustness of the method is demonstrated in simulations. The low computational complexity and lack of heuristic methods needed as compared to most feature based and correlation based techniques, make the algorithm an attractive motion estimator.

2 THE PARALLEL EXTENDED KALMAN FILTER STRUCTURE

The parallel extended Kalman filter,[2], consists of a parallel bank of third order EKF's operating on the Fourier coefficients of the image. These EKF's estimate the actual Fourier coefficients, (real and imaginary), of the image along with a linear combination of the velocity components of the moving object, refered to as the frequency - velocity product. Subsequently, a weighted least squares estimate of this product is fed back into the EKF's for the next iteration. The optimal estimator for the image has a very simple parallel structure in the limit as the velocity estimate approaches the actual velocity.

One of the disadvantages of the PEKF technique, is its limitation to the estimation of object motion constrained parallel to the image plane. Due to the Kalman filter versatility and robustness, slight rotational and scaling, (motion perpendicular to the image plane), can be modelled as object velocity noise. Unfortunately this is not satisfactory for cases where this motion is substantial, (rotation $>$ 1 Degree, scaling $>$ 5%, per frame), since substantial errors are introduced in the 2-D translation estimation. Therefore, it is concluded that for best results in object motion compensation and restoration, the rotational velocity and scaling estimation of the moving object is necessary. Since this technique does not use feature correspondence, obtaining true 3-D object motion is not straight forward, but a pseudo 3-D object motion estimator can be obtained.

3 PSEUDO 3-D MOTION ESTIMATOR

The operation of the PEKF motion estimator, is based on the linear shift theorem of Fourier transforms. The exponential phase component is proportional to the shift that took place, and this is the information used by the EKF's in order to obtain the velocity information. Based on this principle the rotation and scaling of the object can be obtained. It must be noted that it is of great importance to obtain invariance between the estimation of the 2-D translation, scaling and rotation of the moving object. The 3-D pseudo motion estimator algorithm presented here contains a rotation invariant motion estimator. This implies that the object rotational motion estimation is independent of all other motion. The scaling of the object although invariant to 2-D translational motion, is not invariant to rotational motion, therefore 2-D rotational motion compensation should preceed the estimation of scaling. Finally, the 2-D translational motion estimator is variant to all other types of motion and therefore must be used after both compensation for 2-D rotation and scaling has taken place. The result of this process is a 3-D pseudo motion estimator, where, if each motion specific estimator is used in the order stated above, the resulting algorithm can be used to estimate any 3-D motion with no problems of convergence or a need to run the algorithm iteratively as with most correlation based techniques. Computational efficiency and sub-pixel accuracies, are merits of the use of the PEKF structure. This Pseudo 3-D motion estimator is shown in Fig. 1.

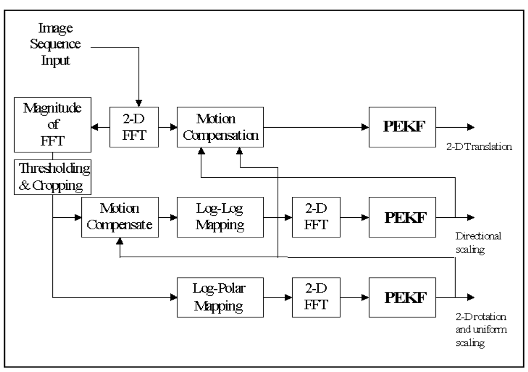

Figure 1: Pseudo 3-D Motion Estimator Structure

3.1 The 2-D Rotational Motion Estimator

In this section, the estimation of 2-D rotational motion, (rotation parallel to the image plane), will be considered. It must be noted that 2-D translation is present, but no scaling is assumed at this point. The presence of translation is to highlight the invariant property of the estimator. Assume that $f_2(x,y)$ is a translated and rotated replica of $f_1(x,y)$ with translation (u_0, v_0) and rotation Θ_0, then,

$$f_2(x,y) = f_1\left(\begin{bmatrix} cos(\Theta_0) & sin(\Theta_0) \\ -sin(\Theta_0) & cos(\Theta_0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} u_0 \\ v_0 \end{bmatrix}\right) \quad (1)$$

Now, taking the 2-D DFT of the translated and rotated image, equation (1), and assuming a square image then,

$$F\{f_2(x,y)\}[\xi_x, \eta_y] = F_1(\xi_u, \eta_v)e^{-2j\pi m(u_0\xi_u + v_0\eta_v)/M}$$
$$F_2(\xi_x, \eta_y) = M_1(\xi_u, \eta_v)e^{j(\phi_1(\xi_u,\eta_v) - 2\pi m(u_0\xi_u + v_0\eta_v)/M)}$$

where

$$\begin{bmatrix} \xi_u \\ \eta_v \end{bmatrix} = \begin{bmatrix} cos(\Theta_0) & sin(\Theta_0) \\ -sin(\Theta_0) & cos(\Theta_0) \end{bmatrix}\begin{bmatrix} \xi_x \\ \eta_y \end{bmatrix} \quad (2)$$

From equation (2), it can be seen that the magnitudes of the DFTs of the two images, $f_1(x,y)$ and $f_2(x,y)$ are translation invariant since the magnitude components are not related to u_0 and v_0 in any way. This has great asset to the principle of rotational motion estimation, since the rotation of the magnitude of the image spectra is equivalent to the rotation of the object in the spatial domain. This helps avoid the need to determine the object centre about which the object is rotating. Also, since we are assuming constant rotational and translational motion, it can be shown [3] that the rotation centre about which the object is rotating, is irrelevant. Thus, whether the object is rotating about an arbitrary point or about its centre, the rotation angle estimation is the same. In order to use the PEKF to obtain a rotational motion estimate it is necessary to map this rotated spectrum such that any rotation will appear as a translation. This translation can then be estimated with the PEKF and hence obtain a result, which is proportional to the rotation of the object. Such a mapping can be obtained with a polar co-ordinate representation of the moving object magnitude spectrum as shown in equation (3),

$$M_1(\rho, \Theta) = M_2(\rho, \Theta - \Theta_0) \quad (3)$$

where ρ and Θ are the magnitude and phase of the polar mapping respectively.

3.2 The 1-D Scaling estimator

Scaling of an object can be considered as the third translation in the 3-D motion of an object. At this point the scaling is assumed to be uniform, meaning that the scaling of the object in both the x and y directions, (horizontal and vertical), will be equal. This is consistent with the object motion containing no rotational motion into the frame, (the remaining two *degrees of freedom* in 3-D object motion). A similar approach in solving this motion will be used as with the 2-D rotational motion. In this case consider that f_1 is the scaled replica of f_2

with factors (s_x, s_y) for the horizontal and vertical directions respectively. According to the Fourier scaling property, the Fourier transforms of f_1 and f_2 are related by,

$$F_1(\xi, \eta) = \frac{F_2(\xi/s_x, \eta/s_y)}{|s_x s_y|} \tag{4}$$

Ignoring the multiplication factor, and assuming that $s_x = s_y = s$, then, by taking the magnitude of equation (4) and polar mapping it, the result of this would be represented mathematically by,

$$M_1(\rho_1, \Theta_1) = M_2(\rho_1/s, \Theta_2) \tag{5}$$

From equation (5) it is shown that, any uniform scaling of the object will result in the scaling of the magnitude axis of the polar mapped magnitude of the image DFT. It is important to note that, the rotation axis is not affected by this scaling. To use the PEKF for the estimation of the image uniform scaling parameter s, this polar magnitude of the polar mapped image must be re-mapped such that the scaling appears as a shift. This is accomplished with the logarithmic property as shown below,

$$M_1(log_\alpha(\rho_1), \Theta_1) = M_2(log_\alpha(\rho_1) - log_\alpha(s), \Theta_2) \tag{6}$$

and substituting $log_\alpha(\rho_1)$ with ρ and $log_\alpha(s)$ with s_0 then,

$$M_1(\rho, \Theta_1) = M_2(\rho - s_0, \Theta_1) \tag{7}$$

where, the base of the logarithm, α, is such that it sets the polar magnitude axis length, e.g. if $\alpha = 1.0386$ and the image size was originally 128 x 128 pixels then the resulting mapped image size would still be 128 x 128. Other values may be used to improves the sub-pixel estimation accuracy if required. From this log-polar mapping technique, a uniform scaling motion estimation algorithm is obtained whos state model has the form,

$$M_1(\rho, \Theta) = M_2(\rho - s_0, \Theta - \Theta_0) \tag{8}$$

where, s_0 and Θ_0 are the uniform scaling and 2-D rotation angle respectively, to be estimated using the PEKF. It is now possible to obtain the 3-D translational motion of the moving object as well as its 2-D rotational motion.

3.3 Pseudo 3-D Rotation

So far, estimation of 3-D translational motion is possible, as well as 2-D rotational motion. However, for the development of this Pseudo 3-D rotation motion estimator, some assumptions are necessary. Firstly, the object motion in the rotational directions which are not parallel to the image plane are small. This can be assumed for vehicles, since this motion is consistent with changing the vehicles direction which cannot be substantial between frames. Secondly, the object is assumed to be at a considerable distance away from the viewing image plane, in order to limit the effects of perspective projection, and therefore it can be assumed that a parallel projection viewing model can be used. These two assumptions make it possible to estimate a measure of the two remaining rotational motions. Since any rotation into the image frame has the effect of directional scaling of the viewed object, by simply considering the object change in scale in both the horizontal and vertical directions, some information on these rotations can be made and eventually compensate for it in the image enhancement step. This is why the method is called a 3-D pseudo motion estimator, since it will not give the actual 3-D motion parameters of the object motion but some partial information of this motion. The state model is very similar to the previous two stated, with the only change being that the magnitude of the image DFT is re-mapped to a log-log plane. This, as with the uniform scaling, results in an object motion state model as shown in equation (9),

$$\begin{aligned} M_1(log_\alpha(\xi), log_\alpha(\eta)) = M_2(&log_\alpha(\xi) - log_\alpha(s_x), \\ &log_\alpha(\eta) - log_\alpha(s_y)) \end{aligned} \tag{9}$$

or

$$M_1(s_x, s_y) = M_2(s_x - s_{x0}, s_y - s_{y0}) \tag{10}$$

where, s_{x0} and s_{y0} are the directional scaling parameters to be estimated by the PEKF algorithm. In this case, directional scaling and 2-D rotational motion are not invariant to each other, and will cause errors in the estimation of the 2-D rotational motion, and vice versa. It is apparent though, that scaling is much more sensitive to rotational motion as compared with the rotation estimates to directional scaling. Therefore, it can be assumed that rotational motion is approximately invariant to directional scaling.

4 MOTION ESTIMATOR RESULTS

Finally, a pseudo 3-D motion estimator is obtained. The order in which each estimator is executed is important to the convergence of this algorithm. It has been determined that 2-D rotational motion estimation is approximately invariant to all other motion, and therefore should be used to compensate for the 2-D rotational motion of the moving object, and then in turn directional scaling estimation can take place. Finally, after compensation for directional scaling and 2-D rotation, translational motion estimation is possible. Some pre-processing steps are however necessary for the correct operation of this motion estimator. The 2D shift matrix in the state equation of the PEKF, must be circulant [2]. This is the case when the object is fully contained in the image boundaries. However, once the polar and log-log mappings have taken place this is no longer true. Therefore, some pre-processing is required, which in this case is a cropping step. This step ensures that no moving object information has contact with the image boundary, therefore maintaining the circulant property of this matrix. However, in the case where image data does have contact with the boundary, only minor errors will be observed according to numerical analysis performed by Burl for the PEKF. The 3-D pseudo rotational motion or directional scaling motion estimator, is tested with the use of an artificial input sequence of a moving square on a black background. Image noise levels of 10dB SNR are demonstrated, where the dotted line represents the true object motion. The number of Fourier components used for the estimation process are, for rotational, directional scaling and translational motion, 4, 4, and 2 components respectively.

5 CONCLUSIONS

From these simulations, the directional scaling motion estimation performance of the algorithm can be seen to be independent of the level of translation present and, to a certain degree, independent of the rotation present. The effects of directional scaling on the 2-D rotation estimation can also be noticed, as well as the biased results obtained for the translational motion. The errors observed in these estimates, are largely dependent on the motion compensation performed. These errors can be minimised as the image size increases, and improved motion compensation steps are considered. Further work to be completed are, improvements on the thresholding and cropping step, and the development of a stationary background removal algorithm. This preprocessing step is required such that the PEKF motion estimator result in accurate motion estimates, [3]. Also extensive tests of this motion estimation method on various real video sequences containing vehicle motion will be performed, for its validation as a practical method for motion compensation of vehicle type motion.

References

[1] G. Adiv. Inherent ambiguities in recovering 3-d motion and structure from noisy flow fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(5):477–489, May 1989.

[2] J. B. Burl. A reduced order extended kalman filter for sequential images containing a moving object. *IEEE Transactions on Image Processing*, 2(3):285–295, July 1993.

[3] J. A. Chambers C. L. Topping. *Moving object enhancement in noisy video sequences.* M.phil to p.hd transfer report, Imperial College of Science, Technology and Medicine, Exhibition Road, London SW7 2BT, ENGLAND, April 1997.

[4] R. Chellappa Y. S. Yao. Tracking a dinamic set of feaure points. *IEEE Transactions on Image Processing*, 4(10):1382–1395, October 1995.

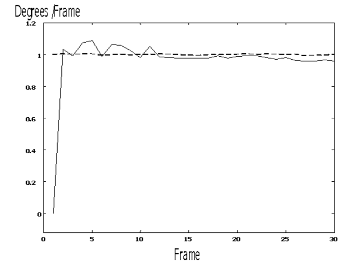

Figure 2: 2-D Rotational Motion Estimate

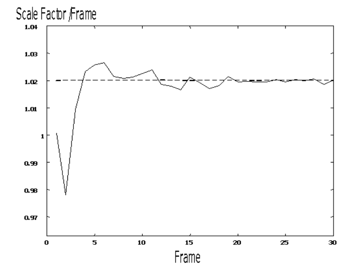

Figure 3: X-Directional Scaling Motion Estimate

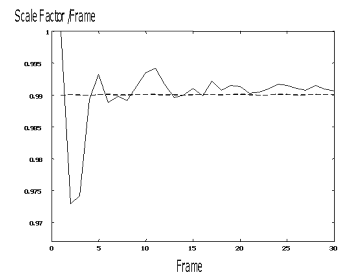

Figure 4: Y-Directional Scaling Motion Estimate

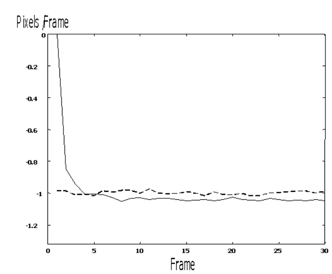

Figure 5: Y-Translational Motion Estimate

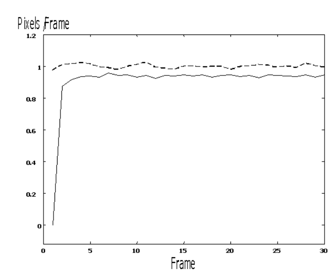

Figure 6: X-Translational Motion Estimate