ROTATIONAL AND TRANSLATIONAL MOTION ESTIMATION AND SELECTIVE RECONSTRUCTION IN DIGITAL IMAGE SEQUENCES

Mingqi Kong, Bijoy K. Ghosh

Washington University in Saint Louis
Department of Systems Science and Mathematics
One Brookings Drive, Campus Box 1040
Saint Louis, MO 63130
Email: ghosh@zach.wustl.edu, kong@zach.wustl.edu

ABSTRACT

This paper addresses the problem of motion estimation and selective reconstruction of objects undergoing rotational motion composed with translational motion. The goal is to derive the motion parameters belonging to the multiple moving objects, i.e. the angular velocities and the translational velocities and identify their locations at each time instance by selective reconstruction. These parameters and locations can be used for various purpose such as trajectory tracking, focus/shift attention of robot, etc. The innovative algorithm we have developed is based on angular velocity and translational velocity tuned 2D+T filters. One of the important fact about our algorithm is that it is effective for both spinning motion and orbiting motion, thus unifies the treatment of the two kinds of rotational motion. Also by tuning of the filters, we can derive the translational motion parameters and the rotational motion parameters separately, which has the advantage of making motion estimation faster and more robust comparing to estimating all of them simultaneously. The algorithm is simulated using synthesized image sequences corrupted by noise and shows to be accurate and robust against noise and occlusion.

1. INTRODUCTION

The motion we have studied here is rotational motion composed with translational motion. The rotational motion has been classified into two types according to their rotational axis. If the axis is the center of gravity on the object, it is called spinning. If the axis is outside of the object, it is called orbiting. The signal we wrked on is 2D+T digital image sequences [1].We extend our previous work [9] such that it works fine for both spinning motion and orbiting motion, thus makes no difference in dealing with the two kinds of rotational motion. Besides, it can estimate motion parameters for not only rotational motion, but also rotational and translational motion.

In order to identify motion parameters, we tune the 2D+T filter to angular velocity, translational velocity, spatial scale, spatial orientation, spatial and temporal translations. These matched filters can perform minimum-mean-square error (MMSE) estimation of motion parameters. This approach doesn't involve any point-to-point corresponding problems and doesn't need any statistical models. It is fundamentally different from other techniques proposed in the literature such as those based on gradient-based optical flow, block matching, pel-recursive, Bayesian model and Markov random field (MRF) model [2][3] [4][5] [6]. The angular and translational velocities can be robustly obtained for all the moving objects even after they occlude each other for some time and the sensor is corrupted by noise. It is effective for motion estimation of several moving objects, even when they have the same size, shape and intensity but different motion parameters, which may cause some difficulty for corresponding based algorithms.

The algorithm we proposed separate the translational velocity estimation and the angular velocity estimation, thus requires much less computing time than estimating the angular and translational velocities simultaneously. Also by separating them, the result is more accurate since the peaks are more easy to identify than the simultaneous estimation case, where the true peaks may be disguised by those large valued non peaks close to the other peaks.

The 2D+T digital image sequences we generated in this paper simulated the case when a fixed camera without tilt was directly above several rotating and translating objects. The noise was added to be more close to reality. We have two image sequences, one is two objects with different spinning angular velocities and traveling with different translational velocities, the other is two moving objects with different orbiting angular velocities and different translational velocities. We used our algorithm on both of the image sequences and can correctly derive their motion parameters and perform reconstruction at each time instance.

2. ROTATIONAL AND TRANSLATIONAL VELOCITY TUNED FILTERS

Let $\psi(\vec{x}, t)$ be the 2D+T filter in space and time domain, the translational velocity tuned filters can be written as:

$$\psi_{g_1} = [T_{g1}]\psi(\vec{x}, t) = a^{-1}\psi(a^{-1}R(\theta_0)(\vec{x} - \vec{v}(t-\tau) - \vec{b})) \quad (1)$$

where the parameters of interest in our case are the spatial translation $\vec{b} \in \mathbf{R}^2$, the temporal translation $\tau \in \mathbf{R}$, the translational velocity $\vec{v} \in \mathbf{R}^2$, the scale $a \in \mathbf{R}_+ \backslash \{0\}$, the spatial orientation $\theta_0 \in [0, 2\pi)$, $R(\theta_0) = \begin{pmatrix} \cos\theta_0 & -\sin\theta_0 \\ \sin\theta_0 & \cos\theta_0 \end{pmatrix}$, $g_1 = \{\vec{b}, \tau, \vec{v}, a, \theta_0\}$. We have studied this type of filter and its applications [7][8].

In this paper we introduce the rotational and translational ve-

This material is based upon work supported in part by DOE under grant # $DE-FG02-90ER14140$

locity tuned filters which can be expressed as:

$$\begin{aligned}\psi_{g_2} &= [T_{g2}]\psi(\vec{x},t) \\ &= a^{-1}\psi(a^{-1}R(\theta_0+\theta_1 t)(\vec{x}-\vec{v}(t-\tau)-\vec{b}))\end{aligned} \quad (2)$$

Where θ_1 is the angular velocity and $g_2 = \{\vec{b},\tau,\vec{v},a,\theta_0,\theta_1\}$. When $\theta_1 = 0, \psi_{g_2} = \psi_{g_1}$.

The $\psi(\vec{x},t)$ we chose is the Morlet function which behaves as an bandpass filter, its 2D+T version in frequency domain is given by:

$$\begin{aligned}\hat{\psi}(\vec{k},\omega) &= \hat{\psi}(\vec{K}) \\ &= |det(D)|^{\frac{1}{2}}\left(e^{-\frac{1}{2}(\vec{K}-\vec{K}_0)^T D(\vec{K}-\vec{K}_0)}\right. \\ &\quad \left. -e^{-\frac{1}{2}\vec{K}_0^T D\vec{K}_0} e^{-\frac{1}{2}\vec{K}^T D\vec{K}}\right)\end{aligned} \quad (3)$$

where the hat $\hat{}$ stands for Fourier transform. $\vec{K} = (\vec{k},\omega)^T \in \mathbf{R}^2 \times \mathbf{R}$, $\vec{k},\omega$ are spatial and temporal coordinates in frequency domain. $\vec{K}_0 = (k_{0x},k_{0y},\omega_0)^T$, D is a positive definite matrix. For $2D+T$ signals, $D = \begin{pmatrix}\epsilon_x & 0 & 0\\ 0 & \epsilon_y & 0 \\ 0 & 0 & \epsilon_t\end{pmatrix}$, where the ϵ factors introduce anisotropy in the filter shape.

To match filters with the signal, take inner product of the tuned filters ψ_{g_2} and the 2D+T signal s, which can be written in frequency domain as:

$$\begin{aligned}< \hat{\psi}_{g_2} \mid \hat{s} > &= < \hat{\psi}_{\vec{b},\tau,\vec{v},a,\theta_0,\theta_1} \mid \hat{s} > \\ &= \int_{\mathbf{R}^2\times\mathbf{R}} d^2\vec{k}d\omega[T_{g_2}\hat{\psi}]\left(\vec{k},\omega\right)\bar{\hat{s}}\left(\vec{k},\omega\right)\end{aligned} \quad (4)$$

where the overbar $^-$ stands for the complex conjugate. When $\theta_1 = 0$, the above equation becomes $< \hat{\psi}_{g_1} \mid \hat{s} >$.

3. MOTION ESTIMATION AND SELECTIVE RECONSTRUCTION ALGORITHM

The algorithm is based on three steps. It allows us to derive the motion parameters and performs selective reconstruction robustly, accurately and quickly. we treat the spinning and the orbiting the same except considering that they may have different scale.

The first step consist in applying ψ_{g_1} to the 2D+T signal $s(\vec{x},t)$. Let we fix the scale $a = a_n$, the estimated translational velocity can be obtained as

$$\vec{v}^* = \arg\max_{\vec{v}} \sum_{\theta_0}\sum_{\vec{b}}\sum_{\tau} | < \hat{\psi}_{\vec{b},\tau,\vec{v},a=a_n,\theta_0} | \hat{s} > |^2 \quad (5)$$

If there are multiple objects, their translational velocities are the velocities corresponding to those several maxima.

At the second step, in order to derive the angular velocity, we use ψ_{g_2} as the tuning filters and each time fix $v = v^*$, where v^* is obtained in step one.

$$h(\theta_1) = \sum_{\theta_0}\sum_{\vec{b}}\sum_{\tau} | < \hat{\psi}_{\vec{b},\tau,\vec{v}=\vec{v}^*,a=a_n,\theta_0,\theta_1} | \hat{s} > |^2 \quad (6)$$

From $h(\theta_1)$ we get a series of peaks corresponding to the multiples of the actual angular velocity θ_1^*, which are $n\theta_1^*, n = 0, 1, ...M$. $M = \left[\frac{2\pi}{\theta_1^*}\right] - 1$, where [ ] means taking the integer part. We can obtain the θ_1^* by taking n=1, i.e, the angular velocity corresponding to the second peak.

Finally, to reconstruct image at a selected time instance $\tau = \tau_n$, with rotational and translational motion $\vec{v}^*, \theta_1^*$, the algorithm is as followed:

$$I(\vec{b}) = \sum_{\theta_0} | < \hat{\psi}_{\vec{b},\tau=\tau_n,\vec{v}=\vec{v}^*,a=a_n,\theta_0,\theta_1=\theta_1^*} \mid \hat{s} > |^2 \quad (7)$$

The location of the object corresponds to the peaks in $I(\vec{b})$.

4. ROTATIONAL AND TRANSLATIONAL MOTION ANALYSIS

Let $s(\vec{x})$ be the still 2D signal, its Fourier transform is $\hat{s}(\vec{k})$. Its rotating and translating version can be expressed as $\tilde{s}(\vec{x},t) = s(R(\theta_1^* t)(\vec{x} - \vec{b}^* - \vec{v}^* t))$, where θ_1^* is the angular velocity, $\vec{v}^*$ is the translational velocity and $\vec{b}^*$ is the position of the rotational axis, it determines whether the rotation is a spinning or an orbiting. If $\vec{b}^*$ lies outside the object, it is orbiting, if it is on the object, it is spinning.

The Fourier transform of the $\tilde{s}(\vec{x},t)$ is

$$\hat{\tilde{s}}(\vec{k},\omega) = \int_{\mathbf{R^2}\times\mathbf{R}} d^2\vec{x}dt\ s(R(\theta_1^* t)(\vec{x} - \vec{v}^* t - \vec{b}^*))e^{-i(\vec{k}\cdot\vec{x}+\omega t)}$$

$$= e^{-i\vec{k}\cdot\vec{b}^*}\int_{\mathbf{R^2}\times\mathbf{R}} d^2\vec{x}'dt\ s(R(\theta_1^* t)\vec{x}')e^{-i(\vec{k}\cdot\vec{x}'+(\vec{k}\cdot\vec{v}^*+\omega)t)}$$

$$= e^{-i\vec{k}\cdot\vec{b}^*}\int_{\mathbf{R^2}\times\mathbf{R}} d^2\vec{y}dt\ s(\vec{y})\ e^{-i(\vec{k}\cdot R^T(\theta_1^* t)\vec{y}+(\vec{k}\cdot\vec{v}^*+\omega)t)}$$

$$= e^{-i\vec{k}\cdot\vec{b}^*}\int_{\mathbf{R}} dt\ \hat{s}(R(\theta_1^* t)\vec{k})e^{-i(\vec{k}\cdot\vec{v}^*+\omega)t}$$

We can see no matter it is spinning or orbiting motion, the expression is the same for both cases in our algorithm. Also, $\hat{s}(R(\theta_1^* t)\vec{k})$ is the Fourier transform of the 2D still signal rotating with angular velocity θ_1^*. A special case is when $\theta_1^* = 0, \hat{\tilde{s}}(\vec{k},\omega) = e^{-i\vec{k}\cdot\vec{b}^*}\hat{s}(\vec{k})\delta(\vec{k}\cdot\vec{v}^*+\omega), \vec{k}\cdot\vec{v}^*+\omega = 0$ is the velocity plane. For general θ_1^*, in order to integrate out with respect to t, we can use polar coordinate in frequency domain. i.e, let $\vec{k} = (\rho\cos\phi, \rho\sin\phi)^T$, and denote $\hat{s}(\vec{k}) = f(\rho,\phi)$,

$$\hat{\tilde{s}}(\rho,\phi,\omega) = e^{-i\vec{k}\cdot\vec{b}^*}\int_{\mathbf{R}} dt\ f(\rho,\phi+\theta_1^* t)e^{-i(\vec{k}\cdot\vec{v}^*+\omega)t}$$

$$= \frac{e^{-i\vec{k}\cdot\vec{b}^*}}{\theta_1^*}\int_{\mathbf{R}} dt\ d\phi' f(\rho,\phi')e^{-i(\vec{k}\cdot\vec{v}^*+\omega)\frac{(\phi'-\phi)}{\theta_1^*}}$$

$$= \frac{e^{-i\vec{k}\cdot\vec{b}^*}}{\theta_1^*}e^{i\frac{(\vec{k}\cdot\vec{v}^*+\omega)\phi}{\theta_1^*}}\hat{f}(\rho, \frac{\vec{k}\cdot\vec{v}^*+\omega}{\theta_1^*})$$

5. SIMULATION RESULT AND ANALYSIS

We have tested our algorithm for both spinning translating motion (Figures 1 and 2) and orbiting translating motion(Figures 7 and 8) corrupted by noise. Occlusion also occurs for some time in both image sequences.

Using our algorithm, we can first derive the translational velocities. For spinning and translating motion, the two peaks locate at $\vec{v} = (1, 0)$ and (0,1) (Figure 3). For orbiting and translating motion,the two peaks locate at $\vec{v} = (1, 0)$ and (1,1) (Figure 9), which are exactly the translational velocities of the objects.

The next step is to obtain the angular velocity for fixed translational velocity. For example, in spinning and translating case, fix velocity $\vec{v}^* = (1, 0)$ in the filter ψ_{g_2}, we can estimate the angular velocity θ_1. From Figure 4, we can derive a series of peaks corresponding to $n * \frac{\pi}{4}$ within $[0, 2\pi)$, n=0,1,2,...7. The actual angular velocity corresponds to the 2nd peak, i.e, $\theta_1^* = \pi/4$. If fix translational velocity $\vec{v}^* = (0, 1)$, from Figure 5, we derive the the peaks corresponding to $n * \frac{\pi}{3}$ between $[0, 2\pi)$, n=0,1,2,...5. So $\theta_1^* = \pi/3$. It is the same with the orbiting and translating motion. Tune ψ_{g_2} to $\vec{v}^* = (1, 0)$, the peaks in Figure 10 are $n * \frac{\pi}{4}$ within $[0, 2\pi)$, n=0,1,...7. The angular velocity θ_1^* can be derived as $\frac{\pi}{4}$. If tune the filter to $\vec{v}^* = (1, 1)$, the peaks in Figure 11 are $n * \frac{\pi}{6}$ within $[0, 2\pi)$, where n=0,1,...11. the angular velocity θ_1^* can be obtained as $\frac{\pi}{6}$.

In order to obtain the location of the objects at each time instance, we can reconstruct the objects undergoing the motions (Figures 6 and 12), thus enable us to obtain its location by detecting the peaks of the intensity in the reconstructed images. This step can be performed for all the τ, which enable us to derive a trajectory of the moving objects. It can be used for tracking and other purpose.

6. CONCLUSION

In this paper, we are developing a new method to achieve motion estimation for rotational motion composed with translational motion. The method is original in the sense that it is based on spatio-temporal angular and translational velocity tuned filters that provide angular and translational velocities, location and orientation of all the discontinuities embedded in the digital image sequences. It is effective for both orbiting and spinning motion. The technique has been shown to be robust against image noise, temporary occlusion. Further studies will develop matched filter method to optimally estimate, track, segment and reconstruct moving object according to their actual motion parameters in 3-D+T made of translational and rotational components from their projections on the 2D+T image sequences.

7. REFERENCES

[1] A. Murat Tekalp "Digital Video Processing", *Prentice-Hall, 1995.*

[2] J. L. Barron, D. J. Fleet, and S.S. Beauchemin "Systems and experiment: Performance of optical flow techniques ", *International Journal of Computer Vision, vol. 12, no. 1, pp.43-77, 1994.*

[3] J.-M. Odobez and P. Bouthemy "MRF-based Motion Segmentation exploiting a 2D motion model robust estimation", *Proceedings, ICIP'95, p.3 vol(xliii+664+666+672), 628-31, vol3, Oct., 1995.*

[4] N. J. Konard and E. Dubois "Bayesian estimation of Motion vector fields ", *IEEE Transaction on Pattern Analysis and Machine Intelligence. vol. 14, no. 7, pp. 910-927, sep. 1992.*

[5] J. N. Driessen, L. Boroczky, and J. Biemond "Pel-recursive motion field estimation from image sequences ", *Journal of Visual Communication and Image Reproduction. vol. 2 no. 3, pp. 259-280, May. 1991.*

[6] B. Liu and A. Zaccarin "New fast algorithms for the estimation of block motion vectors ", *IEEE Transactions on Circuits and Systems for Video Technology, vol. 3 no. 2, pp. 148-157, Apr. 1993.*

[7] M. Kong, J.-P. Leduc, B. K. Ghosh, V. Wickerhauser "Spatio-Temporal continuous wavelet transforms for motion-based segmentation in Real image sequences", *Accepted by ICIP-98, Chicago, Oct. 1998.*

[8] M. Kong, B. K. Ghosh, J.-P. Leduc "Image segmentation with velocity tuned filters", *MTNS'98, Mathematical Theory of Network and Systems, Italy, July 6-10, 1998.*

[9] M. Kong, J.-P. Leduc, B. K. Ghosh, J. Corbett and V. Wickerhauser "Wavelet Based Analysis of Rotational Motion in Digital Image", *Proceedings of ICASSP-98, Seattle, Vol. 5, pp. 2777-2780, May, 1998.*

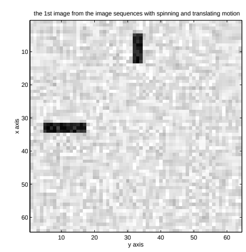

Figure 1: The first image from image sequences with two objects undergoing spinning and translating motion. One object with initial spinning axis at(8,33) its translational velocity $\vec{v}^* = (1, 0)$ and angular velocity $\theta_1^* = \pi/4$, the other's initial spinning axis is at (33,10), it has $\vec{v}^* = (0, 1)$ and $\theta_1^* = \pi/3$. Noise is added.

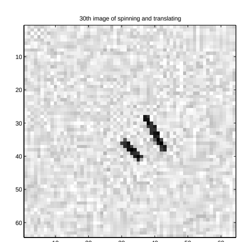

Figure 2: The 30th image of the spinning and translating motion image sequences

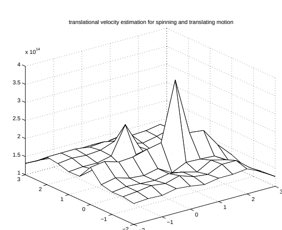

Figure 3: The translational velocity estimation, the two peaks correspond to $\vec{v}^* = (1, 0)$ and $\vec{v}^* = (0, 1)$

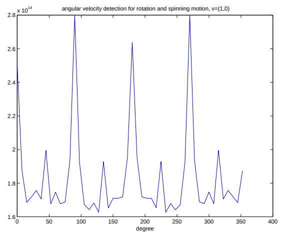

Figure 4: The angular velocity θ_1 corresponding to $\vec{v}^* = (1, 0)$ for the image sequences. The unit of the x-axis is degree. The eight peaks correspond to $n * \pi/4$(i.e, $45°n$) within $[0, 2\pi)$,n=0,1,...7.

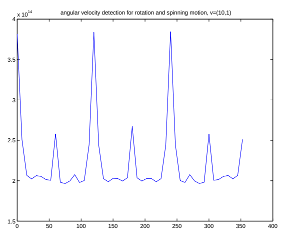

Figure 5: The angular velocity θ_1 corresponding to $\vec{v}^* = (0, 1)$ for the image sequences. The unit of the x-axis is degree. The six peaks correspond to $n * \pi/3$(i.e, $60°n$) within $[0, 2\pi)$,n=0,1,...5.

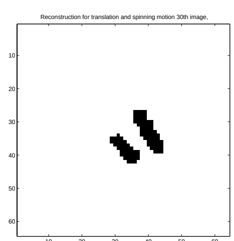

Figure 6: The reconstructed image for the 30th image from the image sequences

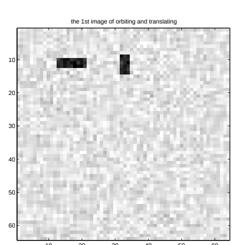

Figure 7: The first image from the image sequences with two objects undergoing orbiting and translating motion. One has initial orbiting center at(16,33), its translational velocity $\vec{v}^* = (1, 0)$, and its angular velocity $\theta_1^* = \pi/4$. The other's initial orbiting center is at (10,10), its $\vec{v}^* = (1, 1)$ and its $\theta_1^* = \pi/6$. Noise is added.

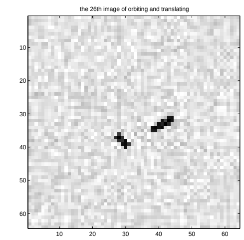

Figure 8: The 26th image of the orbiting and translational motion image sequences .

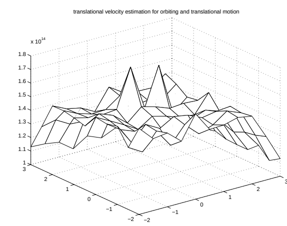

Figure 9: The translational velocity estimation.The two peaks correspond to $\vec{v}^* = (1, 1)$ and $\vec{v}^* = (1, 0)$

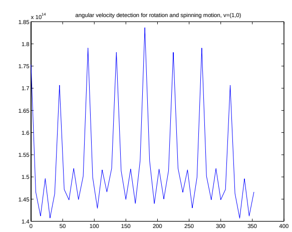

Figure 10: The angular velocity θ_1 corresponding to $\vec{v}^* = (1, 0)$, The unit of the x-axis is degree. The eight peaks correspond to $\pi/4 * n$(i.e, $45°n$) within $[0, 2\pi)$, n=0,1,...,7.

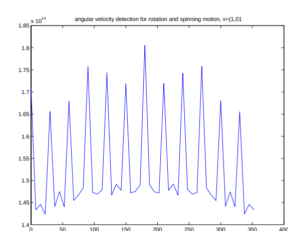

Figure 11: The angular velocity θ_1 corresponding to $\vec{v}^* = (1, 1)$, The unit of the x-axis is degree. The twelve peaks correspond to $\pi/6 * n$(i.e, $30°n$) within $[0, 2\pi)$, n=0,1,...,11.

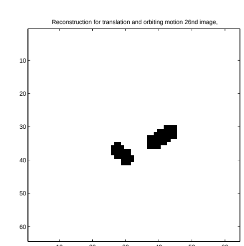

Figure 12: The reconstructed 26th image for orbiting and translating motion from the image sequences.