2012 International Conference on Industrial and Intelligent Information (ICIII 2012)
IPCSIT vol.31 (2012) © (2012) IACSIT Press, Singapore

Optical flow Estimation using Fractional Quaternion Wavelet Transform

Sanoj Kumar[1+], Sanjeev Kumar[2], N. Sukavanam[3] and Balasubramanian Raman[4]
Department of Mathematics, Indian Institute of Technology Roorkee, Roorkee - 247667, India.
([1]sanojdma,[2]malikfma,[3]nsukvfma,[4]balarfma)@iitr.ernet.in

Abstract. In this paper, a novel phase based approach using fractional quaternion wavelet transform (FrQWT) is proposed for computing disparity as optical flow from a given sequence of images. In the proposed approach, phases are estimated from fractional quaternion analytic signal using the concept of quaternion algebra. Disparity is estimated as the optical flow using the phase difference method. The efficiency of the proposed algorithm is carried out by different experiments on synthetic as well as on realistic image sequences.

Keywords: Fractional Hilbert transform, Fractional Quaternion wavelet transform, Disparity and Optical flow.

1. Introduction

In computer vision, the disparity map between two images is usually computed as a shift to one image pixel when viewed in another image of a 3D scene. The two images can be a stereo image pair or two consecutive frames of a video sequence i.e., $I_1(x,y) = I_1(x,y,t_1)$ and $I_2(x,y) = I_2(x,y,t_2)$. In this case, some time interval $\triangle t = t_2 - t_1$ lies between the frames when $I_1(x,y)$ and $I_2(x,y)$ are taken and $d(x,y)$ can be solved as $d(x,y) = v\triangle t$ where v is the image velocity field or optical flow field.

In literature, there are number of frequency based techniques which compute the shifts between two 1D signals using Fourier phase shifts [1]. Fourier phase is the most appropriate for estimating global signal shifts. The wavelet transform (WT) can provide a better description of the signal [2] instead of Fourier transform and Short term Fourier transform (STFT). Wavelet transform (WT) is able to reveal signal aspects that other analysis techniques miss. Frequency based matching methods, often select phases as matching primitives [3]. To avoid the problems of lighting and noise differences between the two images, band pass filters are used before attempting either a correspondence less or a feature based analysis.

The main advantages of phase based methods in the computation of disparity are: i) phase is amplitude invariant and therefore, these methods are robust even when there exist lighting variations between the two images; ii) phase is predominantly linear in space and therefore the estimation of disparity can be reduced to the displacement of linear functions. Bulow [4] introduced quaternion gabor filter into phase based stereo matching and successfully built 2D signals through Hilbert transforms. Only single-scale phase information was exploited and thereby the disparity maps were sparse and somewhat inaccurate. Corrochano [5] proposed a 2D dual tree filter bank based technique and demonstrated the potential of quaternion phase in 2D disparity estimation. To the early work on stereo matching that phase-based stereo was paid less attention compared to the region-based stereo and feature-based stereo [6] can be traced back.

In our approach, we used the theory of the 2D Fractional Hilbert transform [7], [8] for building an analytic signal. We developed a new transform namely fractional quaternion wavelet transform (FrQWT), where each quaternion wavelet consists of a real DWT and three imaginary parts that are organized according to quaternion algebra. The first two FrQWT phases (ϕ, θ) encode the shifts of image features in

+ Corresponding author. Tel.: + (Sanoj Kumar);
E-mail address: (sanojdma@iitr.ernet.in).

the absolute horizontal and vertical coordinate system, while the third phase ψ does not encode the shift of image features. In this paper, we mainly focus on the use of the magnitude and phase for image analysis. The FrQWT allowed a multi-scale framework for calculating and adjusting local disparities and executing phase unwrapping from coarse to fine scales with linear computational efficiency.

2. Quaternion Wavelet Transform

2.1. Fractional Hilbert Transform (FrHT)

There exits several definitions of the FrHT in the signal processing and optics literature [8]. One is the modification of the standard Hilbert transform which can be seen as a phase shifter with parameter $\tau = n\frac{\pi}{2}$, $n \in R$ and the other definition is based on the fractional Fourier transform (FrFT). In[9], an operator-based formulation is proposed to involving the interpolation of the quadrature identity operator *I* and usual Hilbert transform operator *H*. The authors also defined the FrHT operator H_τ corresponding to the shift parameter $\tau \in R$, as

$$H_\tau = \cos(\pi\tau)I - \sin(\pi\tau)H \tag{1}$$

The fractional Hilbert transform operator along the *x* and *y* coordinate can be defined as

$$\begin{aligned} H_{x,\tau_1} &= \cos(\pi\tau_1)I - \sin(\pi\tau_1)H_x \\ H_{y,\tau_2} &= \cos(\pi\tau_2)I - \sin(\pi\tau_2)H_y \end{aligned} \tag{2}$$

respectively, where H_x and H_y are the usual Hilbert transform operator and $\tau_1, \tau_2 \in R$ are the phase shift parameters along *x*, *y* coordinates respectively.

2.2. Fractional Quaternion Wavelet Transform (FrQWT)

The quaternion wavelet transform is a natural extension of the real and complex wavelet transform, taking into the account of quaternion algebra. By arranging the four quadrature components of a 2D wavelet (the real wavelet and its 2D FrHTs) as a quaternion, we obtain a 2D analytic wavelet and their associated fractional quaternion wavelets transform (FrQWT). Each quaternion wavelet consists of a standard DWT tensor wavelet and three additional real wavelets obtained by 1D fractional Hilbert transforms along either one or both coordinates. The real tensor product wavelet $\psi_{\mathrm{h}}(x)\phi_{\mathrm{h}}(y)$ (for the horizontal sub-band), is defined as

$$\begin{aligned} f &= f(x,y) = \psi_h(x)\phi_h(\mathrm{y}) \\ f_{H_1} &= H_{x,\tau_1}\{\psi_h(x)\phi_h(\mathrm{y})\} \\ f_{H_2} &= H_{y,\tau_2}\{\psi_h(x)\phi_h(\mathrm{y})\} \\ f_{H_3} &= H_{y,\tau_2}H_{x,\tau_1}\{\psi_h(x)\phi_h(\mathrm{y})\} \end{aligned} \tag{3}$$

where each component is computed as a combination of 1D dual-tree complex wavelet. Using quaternion algebra, we can organize the four wavelet components to obtain a fractional quaternion wavelet along the horizontal sub-band as

$$\psi^H(x,y) = f + i\,f_{H_1} + j\,f_{H_2} + k\,f_{H_3} \tag{4}$$

Similar expressions can be obtained for the other two sub-bands namely, vertical sub-band $\phi_h(x)\psi_h(y)$ and diagonal sub-band $\psi_h(x)\psi_h(y)$. To compute the FrQWT coefficients, we used a separable 2D implementation of the dual tree filter bank. During each stage of filtering, we independently apply the two sets of *h* and *g* wavelet filters to each dimension (*x* and *y*) of an image. For instance, applying the set of filters *h* to both dimensions yields the scaling coefficients $W^{\phi}_{hh_{j0,n}}$, and the horizontal, vertical and diagonal wavelet coefficients as $W^{H,\psi}_{hh_{j1,n}}$, $W^{V,\psi}_{hh_{j1,n}}$ and $W^{D,\psi}_{hh_{j1,n}}$, respectively. Therefore, the resulting 2D dual tree implementation combines four independent filter banks, corresponding to each dimensions (*hh, hg, gh, gg*) operating on the same image. The wavelet coefficients of the same sub-band are combined from the output of each filter bank by using quaternion algebra to obtain the FrQWT coefficients. For example, the wavelet

coefficients along horizontal sub-band are given as $W^{H}_{j_1,n} = W^{H,\psi}_{hh_{j_1,n}} + iW^{H,\psi}_{hg_{j_1,n}} + jW^{H,\psi}_{gh_{j_1,n}} + W^{H,\psi}_{gg_{j_1,n}}$. Similarly, the wavelet coefficients along other sub-bands are computed.

3. FrQWT Phase-based Disparity Estimation

In this section, we present an algorithm to estimate the local disparity between the reference image and the target image using the concept of local phase. Given two images $I_1(x,y)$ and $I_2(x,y)$ which are differing by a local displacement $d(x,y) = (d_x, d_y)$, i.e., $I_1(x,y) = I_2(x+d_x, y+d_y)$, where d_x and d_y denote the horizontal and vertical disparity respectively. The range of $d(x,y)$ should be small compared to the domain of $I_1(x,y)$ and $I_2(x,y)$ otherwise, structures visible in $I_1(x,y)$ would lie outside the domain of $I_2(x,y)$, i.e., they would have no visible relation to each other. Knowing the phase of the original spectrum $(\phi_1, \theta_1, \psi_1)$ and that of the translated spectrum $(\phi_2, \theta_2, \psi_2)$ the global image shift [4] can be evaluated as

$$d_x = \frac{\phi_1(x,y) - \phi_2(x,y)}{2\pi u}, \quad d_y = \frac{\theta_1(x,y) - \theta_2(x,y)}{2\pi v} \tag{5}$$

From the range of the quaternionic phase it follows that d_x and d_y can be identified within the range of intervals $\left[-\frac{1}{2u}, \frac{1}{2u}\right[$, and $\left[-\frac{1}{4v}, \frac{1}{4v}\right[$ of length $\frac{1}{u}$ and $\frac{1}{v}$ respectively. The local phase approach to disparity estimation starts with assuming that the phase is varying linearly and the displacements d_x and d_y are defined as

$$d_x = \frac{\phi_1(x,y) - \phi_2(x,y) + n(2\pi + k)}{2\pi u_{local}}, \quad d_y = \frac{\theta_1(x,y) - \theta_2(x,y) + m\pi}{2\pi v_{local}} \tag{6}$$

where u_{local} and v_{local} indicate the local frequencies at the assigned position which are not known a priori. Depending on *m*, *k* is defined as

$$k = \begin{cases} 0, & if \;\; m \;\; is \;\; even \\ 1, & if \;\; m \;\; is \;\; odd \end{cases} \tag{7}$$

where *m* and *n* count the number of phase wrapping circles, and *k* is to keep unchanged. The better approximation of $d(x,y)$ depends strongly on the choice of the local frequencies. There are two methods for calculating the local frequencies. (i) Constant model where u_{local} and v_{local} are chosen the center frequencies of the filter, and (ii) Local model in which the main assumption is that the local phase at corresponding points of the two images will take the same value, i.e., $\Phi_1(x,y) = \Phi_2(x+d_x, y+d_y)$ in two correspondent points of both images. An estimate for $d(x,y)$ is obtained by approximating Φ_2 by a first order Taylor's series expansion about (*x, y*):

$$\Phi_2(x+d_x, y+d_y) \approx \Phi_2(x,y) + d_x \frac{\partial \Phi_2(x,y)}{\partial x} + d_y \frac{\partial \Phi_2(x,y)}{\partial y} \tag{8}$$

where $\Phi = (\phi, \theta)$. Solving (21) for $d = d(x,y)$ we obtain the estimated disparity of the local model. In our experiments, we considered that ϕ varies only along the *x*−direction while θ varies only along the *y*−direction. Using this assumption, the new local frequencies are defined as

$$u_{local} = \frac{1}{2\pi}\frac{\partial \phi_1(x,y)}{\partial x}, \quad v_{local} = \frac{1}{2\pi}\frac{\partial \theta_1(x,y)}{\partial y} \tag{9}$$

and then the disparity $d(x,y) = (d_x, d_y)$ is estimated with the following equations

$$d_x = \frac{\phi_1(x,y) - \phi_2(x,y) + n(2\pi + k)}{\frac{\partial \phi_1(x,y)}{\partial x}}, \quad d_y = \frac{\theta_1(x,y) - \theta_2(x,y) + m\pi}{\frac{\partial \theta_1(x,y)}{\partial y}} \tag{10}$$

4. Experimental Results and Discussions

In this section, we examined our algorithm on synthetic as well as on realistic image sequences for estimating the optical flow. For each image sequence, two consecutive frames of images are used. Optical flow is estimated for different values of τ_1 and τ_2 up to 4 levels. At each level of each pyramid, we obtained 16 images accounting for the four fractional quaternionic outputs (4 for approximation ϕ and 12 for details part i.e., 4 images for each part ψ_1 (horizontal), ψ_2 (vertical) and ψ_3 (diagonal)). After computing the phase values, we proceed to evaluate the disparity images using Eq. (6), where the local frequencies u_{local} and v_{local} are calculated according to Eq. (9). Further, we estimated the optical flow at each point computing to a velocity vector in terms of horizontal and vertical details. Now, we adjusted the final orientation of the velocity vector using the information of the diagonal detail. This procedure starts from higher level to lower level; the resulting matrix of the optical flow is expanded in size equal to the next level. This approach estimates the optical flow at the new level which is compared with one of the expanded previous level. The velocity vectors of the previous level fill the gaps at the new level. In this way, the estimation is refined smoothly and a better optical flow is passed from level to level which increases the confidence of vectors at the finest level. At the finest level, we obtained the optical flow which is closer to the ground truth. In the optical flow, the arrows indicate both the directions and magnitudes of the local shifts, with the magnitudes stretched for better visibility. The results are explained for each image sequence as:

4.1. Translated Tree Image sequence:

The translating Tree image sequence simulates a translational camera movement along the $y-$axis, orthogonal to the line of sight, while capturing a textured, planer surface (the picture of tree), which is slanted with respect to the fronto-parallel plane. It is a synthetic image sequence. This image sequence has 40 frames. We considered two consecutive frames for estimating the optical flow. For this image sequence, we considered frame 18 as reference image and frame 19 as target image. Here we have taken $\tau_1 = 0$ and $\tau_2 = 0$, magnitude threshold is 6 and reliability threshold is 0.2. The results are shown in Fig. 1. The actual optical flow (ground truth) and estimated error optical flow are shown in Fig. 2(a) and 2(b) respectively.

4.2. Rubic's cube Image sequence:

This image sequence shows a Rubic's cube which is rotating counter-clockwise on a turntable. This is an interesting image sequence, because a rotation cannot be captured by a single global translation but can be approximated closely by local translations. It is a real image sequence. This image sequence has 18 frames. We considered two consecutive frames for estimating the optical flow. For this image sequence, we considered frame 8 as reference image and frame 9 as target image. Here we have taken $\tau_1 = 0$ and $\tau_2 = 0$, magnitude threshold is 5 and reliability threshold is 0.2. The results are shown in Fig. 3. There is no actual optical flow field available for real image sequence; it is not possible to evaluate error optical flow for this image sequence.

5. Conclusions

We have introduced a novel FrQWT phase-based technique for estimating the optical flow field. Our development of the FrQWT is based on an alternative definition of the 2D FrHT and 2D analytic signal. Our approach has been tested on various image sequences, both synthetic and realistic. The results are very satisfactory as in case of other phase-based techniques [3]. The main advantage of this technique is that, we obtained the optical flow of a fractional transformed image using phase difference method.

6. Acknowledgement

One of the authors, Sanoj Kumar is thankful to Ministry of Human Resources and Development (MHRD), India for the financial support to carry out this research work.

7. References

[1] J. Weng. Image matching using the windowed Fourier phase. International Journal of Computer Vision. 1993, 11 (3): 211–236.

[2] C. Chui. Wavelets: A Tutorial in Theory and Applications. Academic Press, Boston, 1992.

[3] D. J. Fleet, A. D. Jepson and M. Jenkin. Phase-based disparity measurement. CVGIP: Image Understanding, 1991, 53(2): 198–210.

[4] T. Bulow. Hypercomplex spectral signal representations for the processing and analysis of images. Ph.D. dissertation. Christian Albrechts University, Kiel, Germany. 1999.

[5] E. B. Corrochano. The theory and use of the quaternion wavelet transform. Journal of Mathematical Imaging and Vision. 2006, 24:19–35.

[6] D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. International Journal of Computer Vision. 2002, 47(1): 7–42.

[7] K. N. Chaudhary and M. Unser. On the shift ability of dual-tree complex wavelet transforms. *IEEE Transactions on Signal Processing*. 2010, 58 (1): 221–232.

[8] R. Tao, X. M. Li and Y. Wang. Generalization of the fractional Hilbert transform. *IEEE Signal Processing Letters*. 2008, 15: 365–368.

[9] K. N. Chaudhary and M. Unser. The fractional Hilbert transform and dual-tree Gabor like wavelet analysis. In *Proceedings: ICASSP*, 2009.

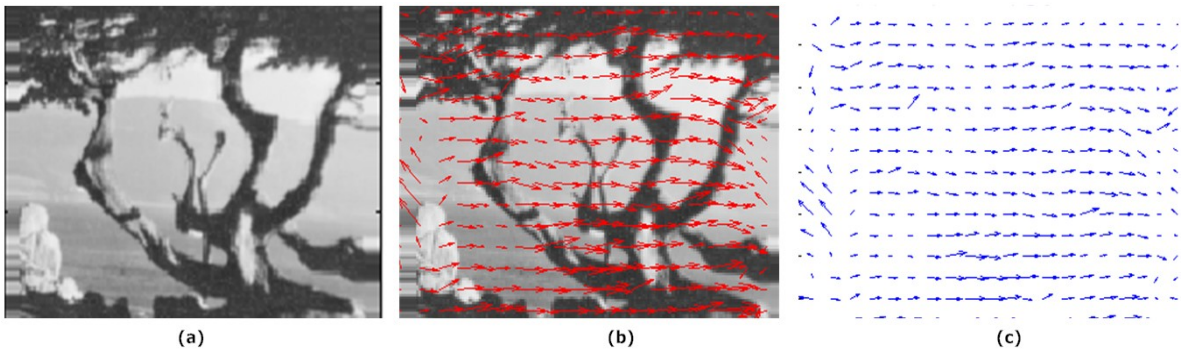

Fig. 1: FrQWT phase-based disparity estimation results for Translated tree image sequence, (a) Reference image, (b) Estimated optical flow plot over reference image, and (c) Estimated optical flow.

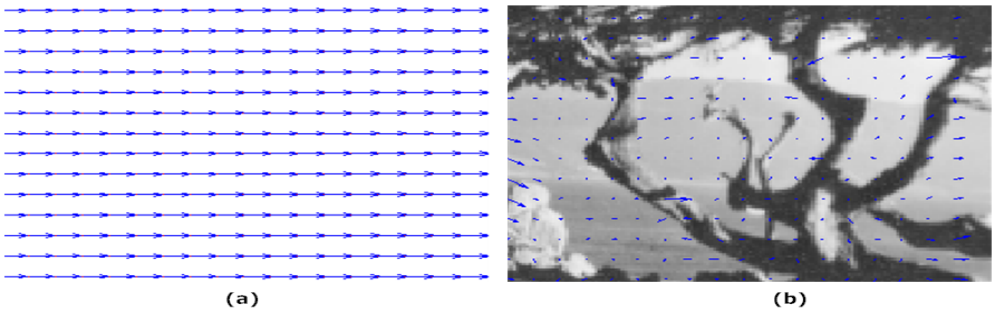

Fig. 2: Error estimation, (a) Actual optical flow (Ground truth), (b) Estimated error optical flow.

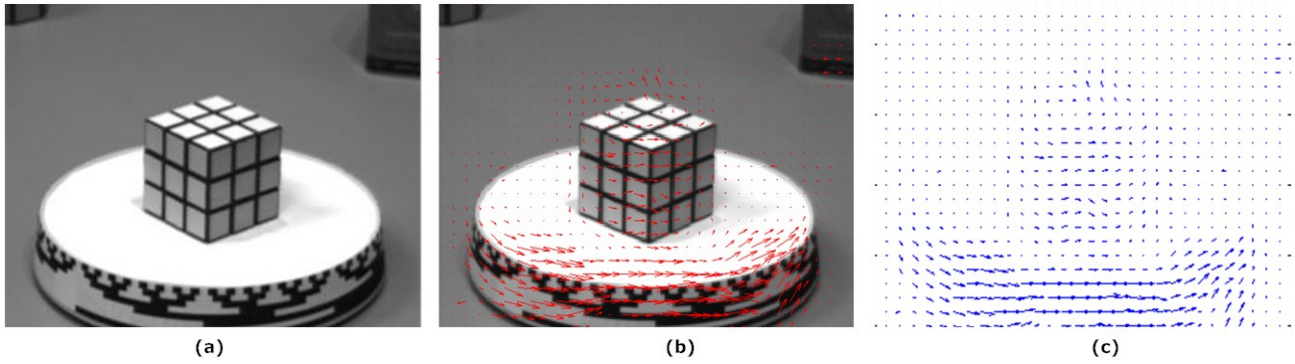

Fig. 3: FrQWT phase-based disparity estimation results for Rubic's cube image sequence, (a) Reference image, (b) Estimated optical flow plot over reference image, and (c) Estimated optical flow.