An Optical Flow Computation based on Shearlet Transform: A Mathematical Formulation

Abstract

The optical flow technique describes and computes motion information between two consecutive and adjacent frames. It finds applications in myriad fields, including robotics, video and space applications. Various optical flow methods have been implemented, including Horn–Schunck (HS), Lucas–Kanade (LK), Farneback optical flows, and recently, a wavelet transform-based Optical Flow. The wavelet transform was integrated into the optical flow to solve the problem related to the small capture range and low accuracy (in the face of large displacement) of the existing optical flow methods. However, wavelets are deficient in dealing with multidimensional and multivariate data such as edges, contours and multidimensional singularities. By contrast, shearlet transform is highly efficient in dealing with pointwise and multidimensional singularities that characterize image and video data. Hence, this research proposed a mathematical formulation that showcases the shearlet transform integration with the optical flow method. This is intended to achieve an efficient and robust optical flow in the motion estimation of multidimensional data.

Introduction

ST consists of a multi-scale partition and a directional localization. Pyramid decomposition is used in the multi-scale partition to reduce the sensitivity to the image shift. The shearing filter is employed to partition the frequency plane into a single low-frequency sub-band and multiple trapezoidal high-frequency in directional localization. Consider a 2-D affine system with composite dilations as presented in equation (2.19) (Moussa et al., 2018):

$$A_{DS} = \left\{\psi_{j,k,m}(x) = |det\ D|^{j/2}\psi\left(S^k D^j x - m\right): j, k\epsilon\mathbb{Z}, m\epsilon\mathbb{Z}^2\right\} \tag{2.19}$$

Where,

D is the anisotropic matrix given as $\begin{bmatrix} d & 0 \\ 0 & d^{1/2} \end{bmatrix}$ or $\begin{bmatrix} d^{1/2} & 0 \\ 0 & d \end{bmatrix}$ and d controls the scale

S is the shear matrix given as $\begin{bmatrix} 1 & s \\ 0 & 1 \end{bmatrix}$ or $\begin{bmatrix} 1 & 0 \\ s & 1 \end{bmatrix}$ and s controls the shearlet's direction

The multiscale and multidirectional decomposition can be achieved using the LP and shearing filter (SF) such that a high frequency and low frequency sub-images are generated at every LP decomposition phase, and the high-frequency sub-band can then be decomposed iteratively (Abazari & Lakestani, 2018). Due to its significant advantages, the shearlet transform will be applied to improve the optical flow algorithm to achieve a more efficient motion information estimation.

2.2.6 Optical Flow Techniques

Optical flow is the apparent motion of the image intensity pattern that transmits information about images. Optical flow reveals the 3D structure of objects and the motion information of the observed objects. In addition, the optical flow technique can estimate the field of motion of an image by leveraging relative motion information on continuous frames without prior information. Optical flow, therefore, plays a vital role in computer vision and has uses in object tracking, robot navigation, target detection, segmentation, and other areas. Many works have been done to estimate the motion flow field. Horn–Schunck (HS) developed a global optical flow method to estimate dense optical flow. Lucas–Kanade (LK) combined a local and global approach to estimate a more robust optical flow; however, the large displacement between the two adjacent frames led to low accuracy and robustness. Wavelet approaches, including the work of (Magarey & Kingsbury, 1998), (Demonceaux & Kachi-Akkouche, 2003), (Hillerio et al., 2012) and (Dérian & Almar, 2017) were employed to estimate optical flow in images. However, the abovementioned approaches did not address large motion and small capture range challenges. These problems, if not addressed, will affect the robustness and precision of the optical flow estimates. Hence, this research will adopt the ST approach due to its shift-invariance, multi-scale and multi-resolution features.

The optical flow constraint equation can be expressed in equation (2.20) (Zheng et al., 2019):

$$I(x, y, t) = I(x + \Delta x, y + \Delta y, t + \Delta t) \tag{2.20}$$

Assuming small and approximate constant movement between image frames, the Taylor's series expansion of equation (2.20) result in equation (2.21):

$$I(x + \Delta x, y + \Delta y, t + \Delta t) = I(x, y, t) + \frac{\partial I}{\partial x}\Delta x + \frac{\partial I}{\partial y}\Delta y + \frac{\partial I}{\partial t}\Delta t + H.O.T \tag{2.21}$$

Subtracting $I(x, y, t)$ from both sides of equation (2.21) yields equations (2.22) and (2.23):

$$\frac{\partial I}{\partial x}\Delta x + \frac{\partial I}{\partial y}\Delta y + \frac{\partial I}{\partial t}\Delta t = 0$$

Or

$$\left(\frac{\partial I}{\partial x}\Delta x + \frac{\partial I}{\partial y}\Delta y + \frac{\partial I}{\partial t}\Delta t\right)/\Delta t = 0 \tag{2.22}$$

$$\frac{\partial I}{\partial x}\frac{\Delta x}{\Delta t} + \frac{\partial I}{\partial y}\frac{\Delta y}{\Delta t} + \frac{\partial I}{\partial t} = 0 \tag{2.23}$$

Where,

$\frac{\partial (.)}{\partial (.)}$ denotes partial derivatives

$\frac{\partial I}{\partial x} = I_x, \frac{\partial I}{\partial y} = I_y, \frac{\partial I}{\partial t} = I_t$ respectively being the partial derivatives of grayscale in x, y and t directions

$\frac{\Delta x}{\Delta t} = V_1$ and $\frac{\Delta y}{\Delta t} = V_2$ denote velocity

Therefore, equation (2.23) yields equation (2.24) (Zheng et al., 2019):

$$I_x V_1 + I_y V_2 + I_t = 0 \tag{2.24}$$

Or

$$\nabla I^T . \vec{V} = -I_t$$

In equation (2.24), two unknowns are difficult to resolve and are known as aperture problems of the optical flow algorithm.

The Lucas-Kanade approach has assumed that the displacement of the image content between two adjacent frames is minimal and nearly constant within the region centred at p. The local image velocity must satisfy the equation (2.25):

$$I_x(P_1)V_1 + I_y(P_1)V_2 = -I_t(P_1)$$

$$I_x(P_2)V_1 + I_y(P_2)V_2 = -I_t(P_2)$$

:   :
:   :
:   :

$$I_x(P_n)V_1 + I_y(P_n)V_2 = -I_t(P_n) \tag{2.25}$$

where,
$P_1, P_2\ and\ P_n$ are pixels inside a window
$I_x, I_y\ and\ I_t$ are partial derivatives with respect to positions x, y, t

Equation (2.25) can be summarized into equation (2.26)

$$Qv = X \tag{2.26}$$

where,

$Q =$ $v = \begin{bmatrix} V_x \\ V_y \end{bmatrix},$

$I_x(P_1)$ $I_y(P_1)$
$I_x(P_2)$ $I_y(P_2)$
:   :
:   :
:   :
$I_x(P_n)V_1$ $I_y(P_n)$

$X =$

$-I_t(P_1)$
$-I_t(P_2)$
:
:
:

$-I_t(P_n)$

The Lucas-Kanade method gave a good optical flow results; however, with large displacement between two consecutive frames, the optical flow generated has low robustness and reduced precision.
Multi-resolution wavelet transform approaches were adopted to alleviate these problems, giving stable and computationally efficient optical flow results. However, isotropic Gaussian filtering used in the wavelet approaches gave a blurry output and inaccurate detection of boundary orientation in the presence of sharp changes in the curvature. It has also been widely recognized that conventional wavelets aren't all that successful in dealing with multidimensional signals that comprise distributed discontinuities, such as edges (Moussa et al., 2018).
Wavelet transform performs well in approximating signals with pointwise singularities. On the contrary, the wavelet transform performs poorly in the presence of multidimensional data such as edges, curves and contours. This is because wavelets are isotropic in nature and fail to capture directional features inherent in multivariate data (Moussa et al., 2018). To address this constraint, basis elements with much greater directional sensitivity and various forms must be used to capture the intrinsic geometric characteristics of multidimensional phenomena (Easley et al., 2008).

This research will adopt ST due to its numerous advantages, including anisotropic directional decomposition, computational stability and efficiency, and scale and translation invariances. Consequently, a more stable, computationally efficient, informative, and clearer optical flow output will be achieved.

The discrete shearlet basis can be defined as presented in equation (2.27) (S. Singh et al., 2015):

$$\psi_{j,k,m}^{n}(x,y) = 2^{j\frac{3}{2}}\psi^{n}(S_n^k D_n^k x - m_1, S_n^k D_n^k y - m_2) \tag{2.27}$$

where,
D is the anisotropic matrix is given as $\begin{bmatrix} d & 0 \\ 0 & d^{1/2} \end{bmatrix}$ or $\begin{bmatrix} d^{1/2} & 0 \\ 0 & d \end{bmatrix}$ and d controls the scale
S is the shear matrix given as $\begin{bmatrix} 1 & s \\ 0 & 1 \end{bmatrix}$ or $\begin{bmatrix} 1 & 0 \\ s & 1 \end{bmatrix}$ and s controls the shearlet's direction
j, k and m are scale, direction and shear parameters, respectively.

The inner product of equations (2.23) and (2.27) results in equation (2.28)

$$\langle \frac{\partial I}{\partial x} V_x, \psi_{j,k,m}^{n} \rangle + \langle \frac{\partial I}{\partial y} V_y, \psi_{j,k,m}^{n} \rangle + \langle \frac{\partial I}{\partial t}, \psi_{j,k,m}^{n} \rangle = 0 \quad \forall_n = 1 \dots N \tag{2.28}$$

Since this constant optical flow assumption does not usually hold, the affine model is used here to construct the optical flow vector (Demonceaux & Kachi-Akkouche, 2003), which can be expressed in equations (2.29) and (2.30) (Zheng et al., 2019):

$$v_1(x,y) = ax + by + c \tag{2.29}$$

$$v_2(x,y) = dx + ex + f \tag{2.30}$$

It can be seen from equations (2.29) and (2.30) that the solution to the optical flow's solution can be reformulated as solving(a,b,c,d,e,f).

Based on wavelet theory, $\int_{-\infty}^{\infty} \psi(x)dx = 0$ with equations (2.29) and (2.30) utilizing integration by parts on equation (2.24) and recombining N equations, we obtain equations (2.31) and (2.32) (Zheng et al., 2019):

$$a\left(\langle xI, \frac{\partial\psi_{j,k,m}^n}{\partial x}\rangle + \langle I, \psi_{j,k,m}^n\rangle\right) + b\,\langle yI, \frac{\partial\psi_{j,k,m}^n}{\partial x}\rangle + c\,\langle\frac{\partial I}{\partial x}, \psi_{j,k,m}^n\rangle + d\,\langle x\,\frac{\partial I}{\partial y}, \psi_{j,k,m}^n\rangle +$$
$$e\,\langle\langle yI, \frac{\partial\psi_{j,k,m}^n}{\partial y}\rangle + \langle I, \psi_{j,k,m}^n\rangle\rangle + f\,\langle\frac{\partial I}{\partial x}, \psi_{j,k,m}^n\rangle = \langle\frac{\partial I}{\partial t}, \psi_{j,k,m}^n\rangle,\ \forall n = 1, 2, \dots. N \tag{2.31}$$

$$Q_u v = X_u \tag{2.32}$$

Where,

$Q_u = [\mathrm{A}_1, \mathrm{A}_2, \mathrm{A}_3, \mathrm{A}_4, \mathrm{A}_5, \mathrm{A}_6]$
$v = (a,b,c,d,e,f)^T$
$X_u = -\langle\frac{\partial I}{\partial t}, \psi_{j,k,m}^n\rangle$

$$A_1 = \langle xI, \frac{\partial\psi_{j,k,m}^n}{\partial x}\rangle + \langle I, \psi_{j,k,m}^n\rangle, A_2 = \langle yI, \frac{\partial\psi_{j,k,m}^n}{\partial x}\rangle,\ A_3 = \langle\frac{\partial I}{\partial x}, \psi_{j,k,m}^n\rangle,\ A_4 =$$
$$\langle x\,\frac{\partial I}{\partial y}, \psi_{j,k,m}^n\rangle, A_5 = \langle\langle yI, \frac{\partial\psi_{j,k,m}^n}{\partial y}\rangle + \langle I, \psi_{j,k,m}^n\rangle\rangle,\ A_6 = \langle\frac{\partial I}{\partial x}, \psi_{j,k,m}^n\rangle \tag{2.33}$$

From equation (2.26), we obtain equation (2.34)

$$Q_m^{jk} v_m^{jk} = X_m^{jk} \tag{2.34}$$

Using a Least Square method, we obtain equation (2.35):

$$Q^T Q v = Q^T X$$

Or

$$v = (Q^T Q)^{-1} Q^T X \tag{2.35}$$

Substituting matrix parameters of equations (2.25) into (2.35), it gives:

$$v_m^{jk} = ((Q_m^{jk})^T Q_m^{jk})^{-1} (Q_m^{jk})^T X_m^{jk} \tag{2.36}$$

Where,

Q_m^{jk} denotes the system matrix

X_m^{jk} is the observation matrix

v_m^{jk} is the affine parameter matrix

Conclusion

This paper proposes a mathematical formulation integrating the optical flow technique with the shearlet transform. This proposed method is expected to improve the small capture range associated with the traditional optical flow and extract more descriptive information.

References

Moussa, O., Khlifa, N., & Abdallah, N. Ben. (2018). Video despeckling using Shearlet tensor-based anisotropic diffusion. *Computer Aided Geometric Design*, *67*, 34–46. https://doi.org/10.1016/j.cagd.2018.09.005

Abazari, R., & Lakestani, M. (2018). A hybrid denoising algorithm based on shearlet transform method and Yaroslavsky's filter. *Multimedia Tools and Applications*, *77*(14), 17829–17851. https://doi.org/10.1007/s11042-018-5648-7

Magarey, J., & Kingsbury, N. (1998). Motion estimation using a complex-valued wavelet transform. *IEEE Transactions on Signal Processing*, *46*(4), 1069–1084. https://doi.org/10.1109/78.668557

Demonceaux, C., & Kachi-Akkouche, D. (2003). Optical flow estimation in omnidirectional images using wavelet approach. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, *7*(July). https://doi.org/10.1109/CVPRW.2003.10080

Hillerio, I., Morgan, B. A., Davis, J. E., Posey, J. M., Kinnary, M., & Le, H. (2012). *(s) Jo e O E*. *1*(19), 238–241.

Dérian, P., & Almar, R. (2017). Wavelet-Based Optical Flow Estimation of Instant Surface Currents from Shore-Based and UAV Videos. *IEEE Transactions on Geoscience and Remote Sensing*, *55*(10), 5790–5797. https://doi.org/10.1109/TGRS.2017.2714202

Zheng, J., Wang, H., & Pei, B. (2019). Robust optical flow estimation based on wavelet. *Signal, Image and Video Processing*, *13*(7), 1303–1310. https://doi.org/10.1007/s11760-019-01476-7

Easley, G., Labate, D., & Lim, W. Q. (2008). Sparse directional image representations using the discrete shearlet transform. *Applied and Computational Harmonic Analysis*, *25*(1), 25–46. https://doi.org/10.1016/j.acha.2007.09.003

Singh, S., Gupta, D., Anand, R. S., & Kumar, V. (2015). Nonsubsampled shearlet based CT and MR medical image fusion using biologically inspired spiking neural network. *Biomedical Signal Processing and Control*, *18*, 91–101. https://doi.org/10.1016/j.bspc.2014.11.009