MOTION ESTIMATION USING THE EDGES DETECTION AND WAVELET TRANSFORM

Bing Wang, Rong-Chun Zhao, Ying-Lei Cheng, Guang-Jian Tian
Institute of Computer, Northwestern Polytechnic University, Xi'an Shanxi 710072 China
bingwangx@tom.com

ABSTRACT

This paper presents a novel motion estimation algorithm based on the multiscale edges detection (MED) and continuous wavelet transform (CWT). Multiscale representations are very effective for analyzing the information content of images. First, moving object multiscale edges be detected from the local maxima of a wavelet transform. Then, the CWT is used to map the input signal space to a motion parameters space. It is designed so as to possess simultaneously the properties of MED and CWT. The accuracy, efficiency, and robustness of the new algorithm are demonstrated in comparison testing black matching (BM) motion estimators.

1. INTRODUCTION

Motion estimation (ME) is an important research area in digital video processing and computer vision. Motion estimation is central to many applications where time-varying image data are involved, such as coding of digital video, autonomous navigation of robots, ballistic missile automatic track objects, and automatic traffic monitoring systems.

The problem of designing a robust ME algorithm is however a difficult one. In general, it involves estimating motion in a 3-D scene from a sequence of projections onto the 2-D plane. As such the problem is ill-posed, some additional constraint is required [1]. Other factors, such as the presence of occlusions and noisy data, serve to further increase the complexity of the problem.

This paper addresses the problem of detecting and tracking moving objects in digital image sequences. The purpose for a multiscale edges detection is the fact that moving object found in image sequences is characterized by different resolution displaying different motions; whilst the use of a local energy densities maxima value of CWT allow us to deal with more complex motion, such as rotation, scaling and occlusion [2]. The proposed ME algorithm is thus able to stabilize the MED procedure and to define trajectories for the moving objects present in the scene.

2. MOTION EDGE DETECTION

Automatic object detection and tracking requires the ability to optimally extract the essential features of an object from cluttered environments. ME is based on time domain analysis mainly, while edges detecting based on frequency domain analysis.

According to the definition proposed by Mallat, multisclae edges are special points where the gradient vector modulus is locally maximal in the gradient vector direction [3]. For images, sharp variations are usually located on object contours. We analyze edges with a multiscale approach. This allows us to select edges that belong to details having a certain scale. The scale of the filter determines detail from which we detect edges. We select edges at the characteristic scale of objects. The main advantage of this detection method of edges in comparison with previous ones is a more robust estimation of the moving objects.

We select quadratic B spline function as scale function, that function has a small compact support and uniform regularity[3]. For image edge detection, modulus localized maxima value is fundamentally seek gradient vector modulus maxima. Getting the local maximum value of this wavelet transform modulus can help to locate image edge points [4].

3. THE ME AND TRACKING

A digital video signal as a 3-D space (x, y, t) comprising two spatial dimensional (x, y) corresponding to every image plane and one temporal dimension t . Then in this space, a moving object generates a trajectory. Hence the determination of the moving object amounts to the determination of the motion trajectory by processing globally the image sequence. The CWT is used to map the input signal space to a feature space, which allows for estimate motion parameters.

3.1. The Definition of CWT

The sequence image signal $f(\boldsymbol{x},t)$ is defined in the Hibert space $L^2(R^2 \times R)$, $\boldsymbol{x} \in R^2$, $t \in R$. Wavelet transform of $f(\boldsymbol{x},t)$ is defined as an inner product[2]

$$W_\psi(f, \psi_g) = (c_\psi)^{-1/2} < f, \psi_g > \qquad (1)$$

while $g = (a, \boldsymbol{v}, \theta, \boldsymbol{b}, \tau)$ is CWT parameter space, $\psi_g(\boldsymbol{x},t) \in L^2(R^2 \times R)$ is wavelet basis, and satisfies condition [2]

$$c_\psi = (2\pi)^3 \iint_{R^2 \times R} |\boldsymbol{k}|^{-2} |\omega|^{-1} |\hat{\psi}(\boldsymbol{k},\omega)|^2 \mathrm{d}^2\boldsymbol{k}\mathrm{d}\omega < \infty \quad (2)$$

where $a > 0$ is scale, $\boldsymbol{b} = [b_x, b_y]' \in B$ is the displacement parameter, τ the time variable, $\boldsymbol{v}$ the velocity parameter, B is vicinity region of objects displacement. The spatial integration is constrained to the region B to avoid interference with nearby objects, and distinguished the different objects trajectories. The Fourier transform of $\psi(\boldsymbol{x},t)$ is $\hat{\psi}(\boldsymbol{k},\omega)$, $r_\theta = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$ is the rotate matrix, $\theta \in [0, 2\pi)$ is the direction angle。

The spatio-temporal wavelets are defined in the spatio-temporal domain, as [2]

$$\psi_g(\boldsymbol{x},t) = a^{-3/2}\psi\left(v^{-1/3}a^{-1}\gamma_{-\theta}(\boldsymbol{x}-\boldsymbol{b}), v^{2/3}a^{-1}(t-\tau)\right) \qquad (3)$$

normalize ψ such that $c_\psi = 1$, in this case continuous wavelet transform satisfy Parseval condition for conversation of energy is

$$\zeta = \iint \left|F_\psi(g)\right|^2 \mathrm{d}^2\boldsymbol{x}\mathrm{d}t = \left\|\hat{f}\right\|^2 = 2\pi\left\|f\right\|^2 \qquad (4)$$

while $\left|F_\psi(g)\right|^2$ is energy density of signature f .

The signal $f(\boldsymbol{x},t)$ is perfectly reconstructed from its wavelet transform with

$$f(\boldsymbol{x},t) = (c_\psi)^{-1} \iint (W_\psi f)(g) \psi_g(\boldsymbol{x},t) d^2\boldsymbol{x}dt \qquad (5)$$

It can be seen from (4), (5) that CWT is a kind of energy preserve reversible transform. The CWT is defined as the inner product of the signal of interest and the wavelets of (3), the result is a 6-D representation with parameters directly associated with motion features [2].

3.2. Separable Morlet wavelet

Separable and directional Morlet wavelet is chosen as mother wavelet to detect and trace moving object suitable for motion representation. Morlet wavelet is defined in spatio-temporal domain, by [2]

$$\psi(\boldsymbol{x},t) = \psi(\boldsymbol{x})\psi(t)$$

$$= (\mathrm{e}^{j\boldsymbol{k}_0 \cdot \boldsymbol{A}^{-1}\boldsymbol{x}} \mathrm{e}^{-(1/2)\left|\boldsymbol{A}^{-1}\boldsymbol{x}\right|^2} - \mathrm{e}^{-(1/2)\left|\boldsymbol{A}^{-1}\boldsymbol{x}\right|^2} \mathrm{e}^{-(1/2)\left|\boldsymbol{k}_0\right|^2})$$

$$\times (\mathrm{e}^{j\omega_0 t} \mathrm{e}^{-(1/2)t^2} - \mathrm{e}^{-(1/2)t^2} \mathrm{e}^{-(1/2)\omega_0^2}) \qquad (6)$$

where $\boldsymbol{A} = \begin{bmatrix} 1 & 0 \\ 0 & \varepsilon \end{bmatrix}$ is the anisotropy matrix, $\varepsilon \le 1$ is an anisotropy parameter, $\boldsymbol{k}$ is the wave vector, center frequency is ω_0 。

3.3. Motion Estimation

CWT-based ME algorithm can be viewed as an object tracking algorithm after the initial detection has been performed through MED. The tracking of moving objects relies on extracting local maxima of energy densities in the velocity representation [2],

$$E(\boldsymbol{v},\theta) = \int_{\vec{b}\in B} \left| \left\langle f, \psi_{a_0,\boldsymbol{v},\theta,\boldsymbol{b},\tau_0} \right\rangle \right|^2 d^2\vec{b} \qquad (7)$$

while scale a and the temporal variable τ are fixed. The boundaries of the spatial region B are updated on a frame-by-frame basis to reflect changes in object position. This allows the algorithm to focus on the object of interest, and distinguish different objects.

So, the analysis of motion could be transformed into the analysis of velocity, position and scale energy densities. The energy densities as optimal criteria can be used to derive local estimates of the motion parameters.

Motion parameter can be updated frame to frame, and then the update of each parameter is done separately. The result is a sequence of energy densities peaks showing detected the trajectories of the different objects. The parameter update is done in a specified order: velocity first, then position, and finally scale [2]. We use the Nelder-Mead simple search algorithm which carry out an oriented search and needn't gradients information [5].

4. SIMULATION AND RESULTS

Two simulations were carried out: one to test performance in the presence of rotation and noise, the other to test robustness to object occlusion and detection and tracking of multiple objects. Image sequences is real, the quantity of noise present and occlusion bar width is unknown. It is assumed that the initial conditions such as position and velocity are known for every object in the scenes.

The first simulation was designed to quantify the noise and rotation performance of the new tracking algorithm. Each image of a 100-frames sequence consisted of 352×288 pixels. The object to be tracked has arbitrary shape, and is rotated and scaled at 25 frame/s rate through the image sequences. Background is cloud. Fig. 1 shows 5 frames for a representative range of rotation. Fig. 2 shows object energy densities, and the quantity of random departure is 0～10 pixels. Fig. 3 is edges image of 90~94 frames. Fig. 4 is the object edges energy densities. Energy local maxima value shows the quantity and position of object. Fig. 5 is the trajectory of the 90～100 frames.

The second simulation quantified the occlusion performance of the new ME algorithm compared with the BM algorithm implementation in two cross aircraft scenario. Fig. 6 is the original image of the 6～11 frames of two cross aircraft sequence. Fig. 7 is the sixth ~ eleventh frame edge image.

We compare our tracking results with a block matching approach. This particular implementation is based on an exhaustive search BM algorithm [1]. As is can be seen from Fig. 8, it is the two cross moving object trajectories with block matching algorithm. Fig.9 show the MED and CWT estimation algorithm. Both algorithm are able estimation this position accurately before occlusion. After the occlusion, the BM algorithm is completely off target, while the MED and CWT approach is able to maintain a good level accuracy.

The results show that new algorithm can make use of the motion information of the object and object edge information, and conquer effectively the effect of strong noise, object rotate, distortion and temporary occlusions.

Fig. 1. Original image sequence of 90-94 frames

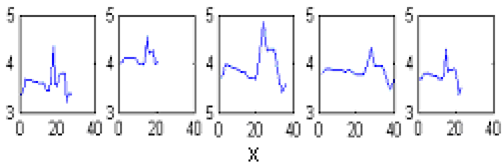

Fig.2. Energy densities of image of 90-94 frames

Fig. 3.   Edges image of 90~94 frames

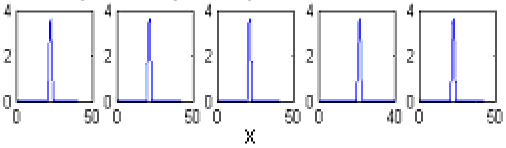

Fig. 4.   Edges energy densities of 90~94 frames

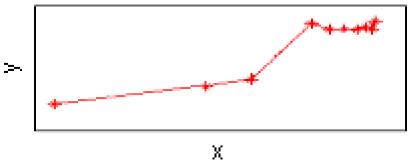

Fig. 5. Motion estimates of 90~100 frames

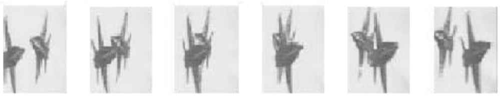

Fig. 6. The sixth ~ eleventh frame original image

Fig. 7. The sixth ~ eleventh frame edges image

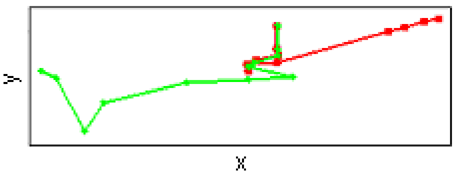

Fig.8.   ME with the BM algorithm

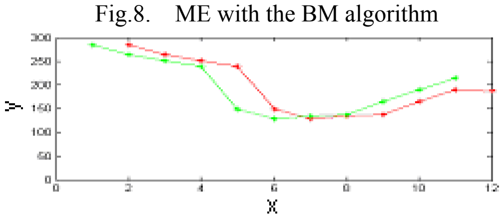

Fig.9. ME with the MED and CWT algorithm

5. CONCLUSIONS

In this paper we examined the effectiveness of MED in separating objects from background in clutter environments. Target edges extracted from a sequence image can be formed into the CWT parameter space, allowing targets to be detected through the energy density local maxima value algorithm described, resulting in the decrease of computational load and noise effect. The method gives superior results compared to the standard generalized block matching and proved to be especially robust under occlusion and noise.

ACKNOWLEDGEMENT

This work was supported by Aeronautical Fund of China (Grant 02I53073) and Doctorate Foundation of Northwestern Polytechnical University (Grant CX200319).

REFERENCES

[1] A. M. Tekalp, *Digital Video Processing*, Tsing Hua and Prentice Hall, Bei Jing, 1998.

[2] F.A. Mujica, J-P. Leduc, R. Murenzi and M.J.T. Smith, "A New Motion Parameter Estimation Algorithm Based on The Continuous Wavelet Transform," IEEE Tran. on Image Processing, 9(5): pp.873-888, 2000.

[3] S. Mallat and S. Zhong, "Characterization of Signals from Multiscale Edges," IEEE Tran. on PAMI, 14(7): pp. 710-732. 1992.

[4] Bing Wang, Rong-Chun Zhao, "A Motion Parameter Estimation Algorithm Based on the Multiresolution Edge Detection," The Second International Conference on Machine Learning and Cybernetics, Xi'an, pp. 2741-2745, November, 2003.

[5] M. D. Destin and R. Murenzi, "Spatio temporal wavelets: Application to the analysis of moving patterns," International Conference on Wavelets Application, Frontieres, pp. 399-408. 1993.