English Research Express@NCKU Volume 3 Issue 5 - February 29, 2008
[http://research.ncku.edu.tw/re/articles/e/20080229/5.html]

A Region-based Selective Optical Flow Back-Projection for Genuine Motion Vector Estimation

Pau-Choo Chung*[1], Chieh-Ling Huang[1] and E-Liang Chen[2]

[1]Department of Electrical Engineering and Institute of Computer and Communication, National Cheng-Kung University, Tainan, 70101 Taiwan.
[2]Department of Computer Science and Information Engineering, Leader University, Tainan, Taiwan.
Email:pcchung@eembox.ee.ncku.edu.tw

PATTERN RECOGNITION、vol. 40、pp.1066-1077、2007

Motion vector plays one significant feature in moving object segmentation. However, the motion vector in this application is required to represent the actual motion displacement, rather than regions of visually significant similarity. In this research, Region-based Selective Optical Flow Back-projection (RSOFB) which back-projects optical flows in a region to restore the region's motion vector from gradient-based optical flows, is proposed to obtain genuine motion displacement.

Motion perception is an important cognitive element of the visual interpretation of our three-dimensional world. In ideal case, the movement of an object in 3-D space corresponds to a 2-D motion in an image sequence. These projected motions can be represented by a motion vector field in an image plane. The motion vector field is defined as the set of motion vectors that are used to denote the relative displacement of the image intensity values in a time-varying image sequence. Generally speaking, motion estimation algorithms can be classified into (1)non-parametric block-based, (2)parametric motion model-based, and (3)optic-flow model-based. All of these approaches assume that there exists point correspondence between two subsequent frames, which induces dense motion vector field of an image.

Non-parametric block-based method, i.e. block-matching, assumes that the motion field is piecewise translation. Because of its simplicity, fast computation, and relative robustness in visual effect, it is one of the most commonly used motion estimation methods. The weakness of the non-parametric block-based method is its inability to describe rotations and deformations, and the possibility of obtaining motion vectors that completely differ from the "true" motion. This hinders the usefulness of block-based motion vector estimation in MPEG-4 VOP segmentation. The parametric motion model, i.e. affine or perspective model, describes a region in an image with a few parameters, usually the translations, rotations or zooming parameters. Accurate estimation of the parameters required correct feature point correspondences as a pre-requisite, which is also a very challenging task in motion estimation. Moreover, its heavy computation makes it almost impossible to cope with real-time applications. The optic-flow model-based approach, i.e. Horn-Schunck, has the advantage that it does not have to find feature point correspondence. The motion vector field, or the so-called optical flow in gradient-based approach, is estimated based on the instantaneous change in image intensity. However, due to the optical flow constraint, the obtained optical flow does not represent the true motion, but only the motion projection on the direction of image gradient.

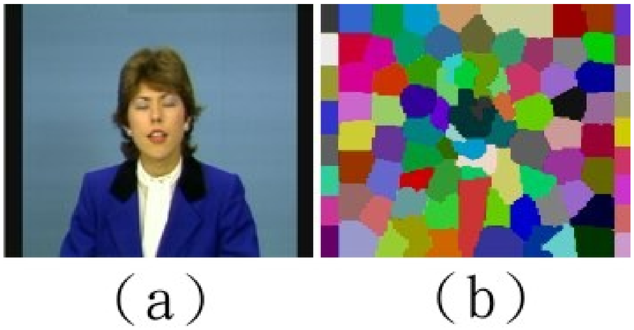


Fig.1. Block clustering result(a)Original image (b)its segmentation result

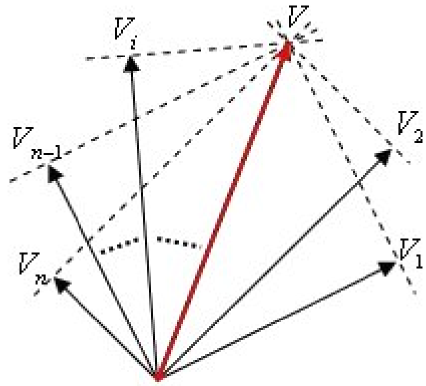


Fig.2. True displacement unveil using several estimated optical flow

In view of the above mentioned reasons, this research proposes Region-based Selective Optical Flow Back-projection (RSOFB) to obtain more reliable motion vectors. As mentioned, due to the optical flow constraint, the obtained optical flow does not represent the true motion, but only the motion projection on the direction of image gradient. Based on this consideration, each image of the sequence is first partitioned by clustering methods such as block clustering into regions where each region is of homogenous features. One segmentation result is shown in Fig. 1. Assume each pixel in a region has the same motion vector. Within each region, optical flow computation is conducted. However, considering computation efficiency and reliability, only pixels of high gradient on each region are calculated for optical flows. As optical flow is the gradient projection of the motion vector, back-projecting the optical flows in an object region in the direction of image gradients would coincide to restore the motion vector of that object region, as shown in Fig. 2, where V_1,V_2,..., and V_n are the obtained optical flows in the image gradient directions and V is the motion vector of back-projecting by optical flows. The back-projection illustrated by Fig. 2 can be performed by the minimization of the distance between gradient projection of the motion vector V and optical flow. Let $V_1\angle\theta_1$, $V_2\angle\theta_2$,... and $V_n\angle\theta_n$ be the obtained optical flows in varied directions (θ_1 , θ_2 ,... and θ_n), according to RSOFB, the genuine motion vector $V\angle\theta$ can be computed by $\theta = \tan^{-1}\left(\frac{\left(\sum_{i=1}^{n} V_i \sin\theta_i\right)\left(\sum_{i=1}^{n}\cos^2\theta_i\right)-\left(\sum_{i=1}^{n} V_i\cos\theta_i\right)\left(\sum_{i=1}^{n}\cos\theta_i\sin\theta_i\right)}{\left(\sum_{i=1}^{n} V_i\cos\theta_i\right)\left(\sum_{i=1}^{n}\sin^2\theta_i\right)-\left(\sum_{i=1}^{n} V_i\sin\theta_i\right)\left(\sum_{i=1}^{n}\cos\theta_i\sin\theta_i\right)}\right)$ and

$V = \frac{\sum_{i=1}^{n}(V_i\cos(\theta-\theta_i))}{\sum_{i=1}^{n}\cos^2(\theta-\theta_i)}$. To be summarized, the procedure of the Region-based Selective Optical Flow Back-Projection is illustrated in Fig. 3.

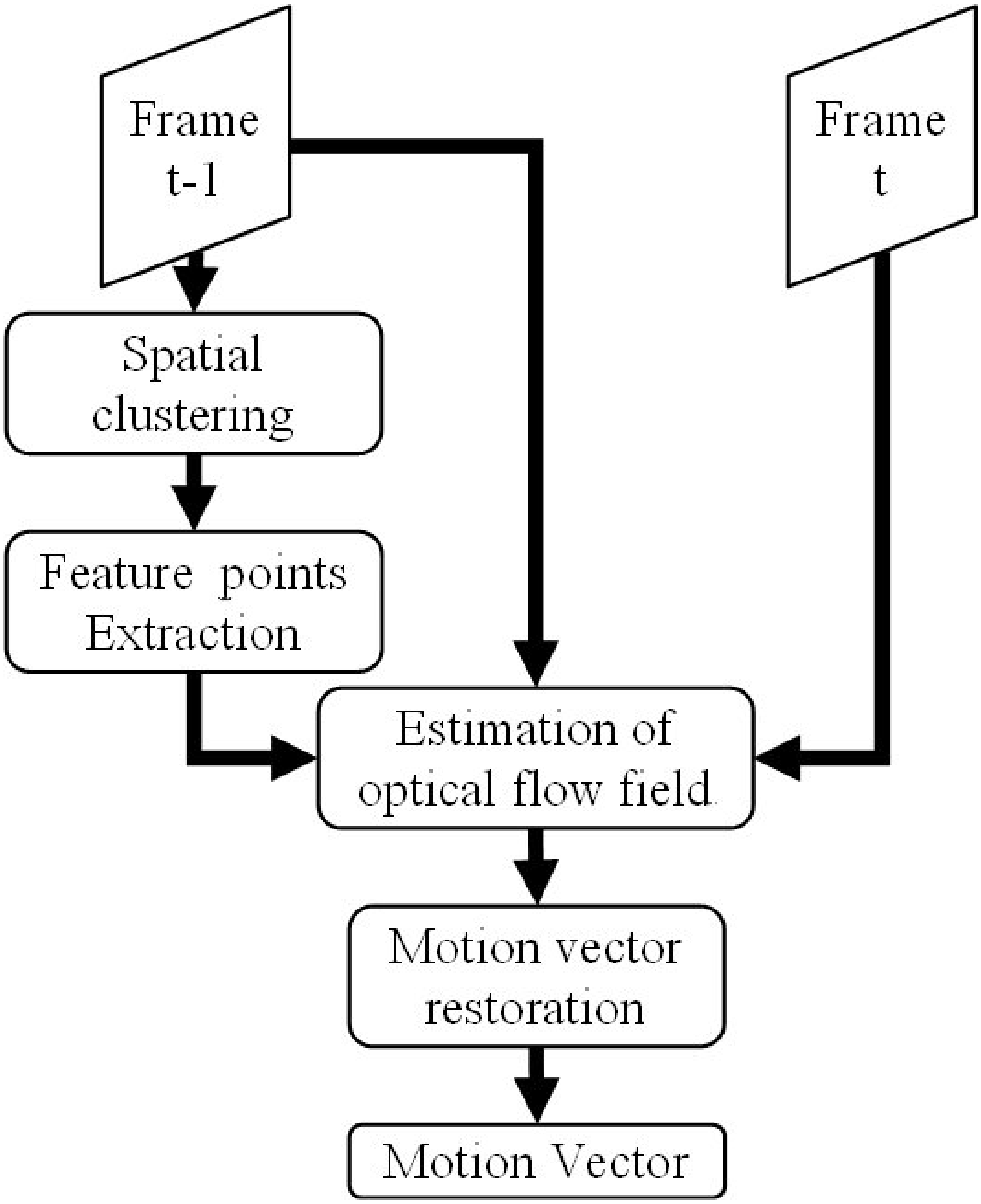
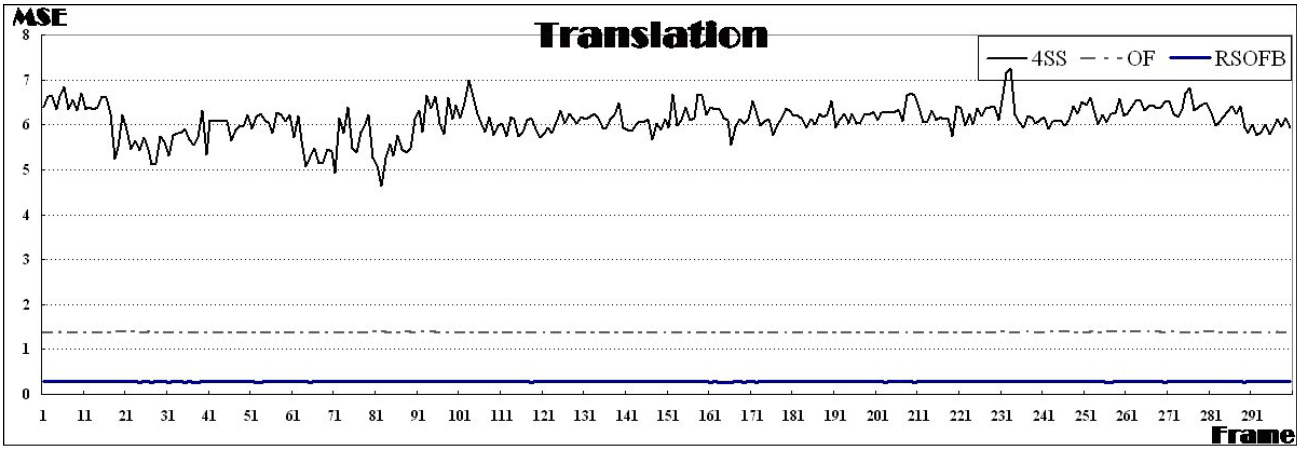
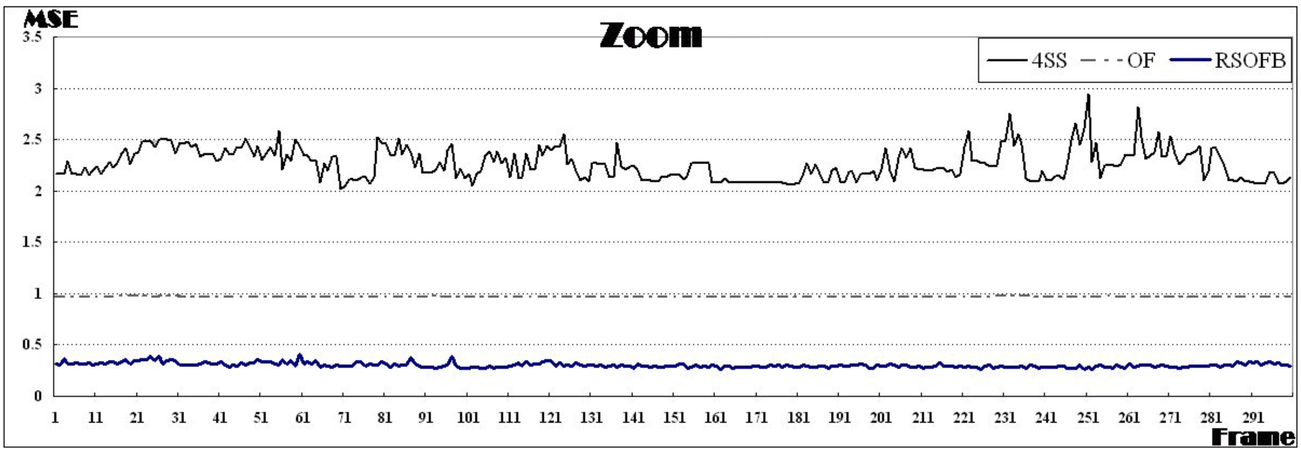


Fig.3. Flow of RSOFB

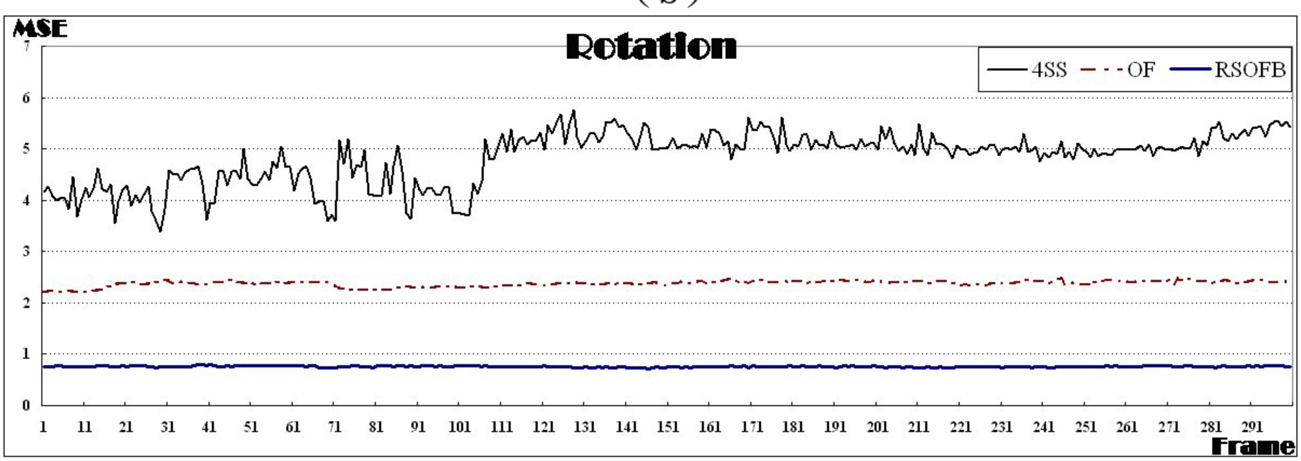
In order to test the accuracy of the proposed RSOFB, a synthetic frame is constructed from a reference frame using perspective motion model. During the testing, three sets of parameters are defined for the perspective transformation, each representing different types of motion: Translation, Zoom, and Rotation. Shown in Figure 4 are the comparisons of the computed motion vectors with the genuine motion vectors when the proposed RSOFB algorithm, the original Horn-Schunck(HS) method, and the traditional four-step search (4SS) are applied, respectively. It is evident that the MSE of RSOFB is less than 4SS and OF in all of the three motion situations. In order to test the effects of lightning to the estimation of motion vectors, we applied gamma transformation on the sequence of images and used the transformed image sequence for the test of our method on motion estimation. Figure 5 shows the Akiyo sequence under gamma transformation with various gamma values. Figure 6 shows the averaged MSE value of each sequence under Gamma transformation. From Fig. 6, we can see that the background light does not cause significant effects to the motion vector estimation as long as the effect is applied on the whole image. Finally, A realistic image sequence Yosemite fly-through sequence, as shown in Fig. 7(a), with more complex camera motion, is also adopted to test the accuracy of RSOFB algorithm. Figure 7(b) shows the MSE for the Yosemite fly-through sequence in Fig. 7(a) when 4SS, OF, and RSOFB approaches are applied. From Fig. 7(b), we can find that the proposed RSOFB algorithm also produces smaller MSE than traditional 4SS and Horn-Schunck method in this Yosemite fly-through sequence.



(a)



(b)



(c)

Fig.4. Performance comparisons over cluster 4SS, OF and RSOFB (a)Translation (b)Zoom (c)Rotation

(a)Gamma=0.1(b)Gamma=0.5(c)Gamma=1(d)Gamma=1.5(e)Gamma=2

Fig. 5. Akiyo sequence with various gamma values

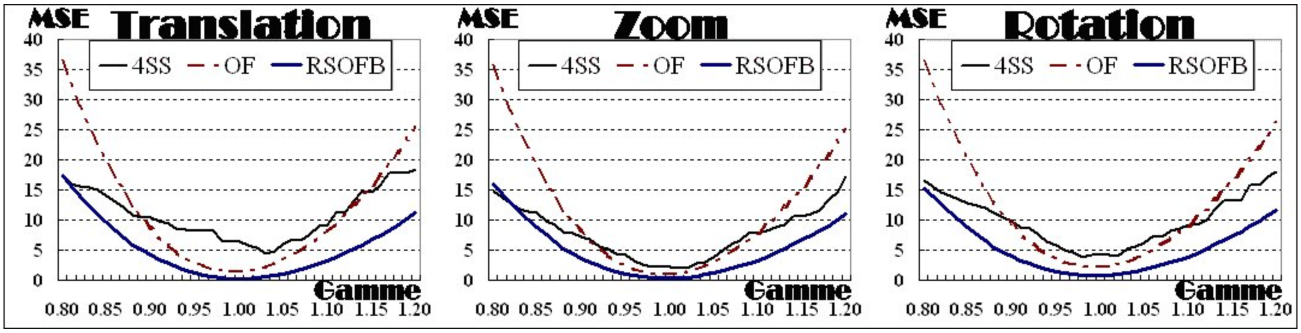
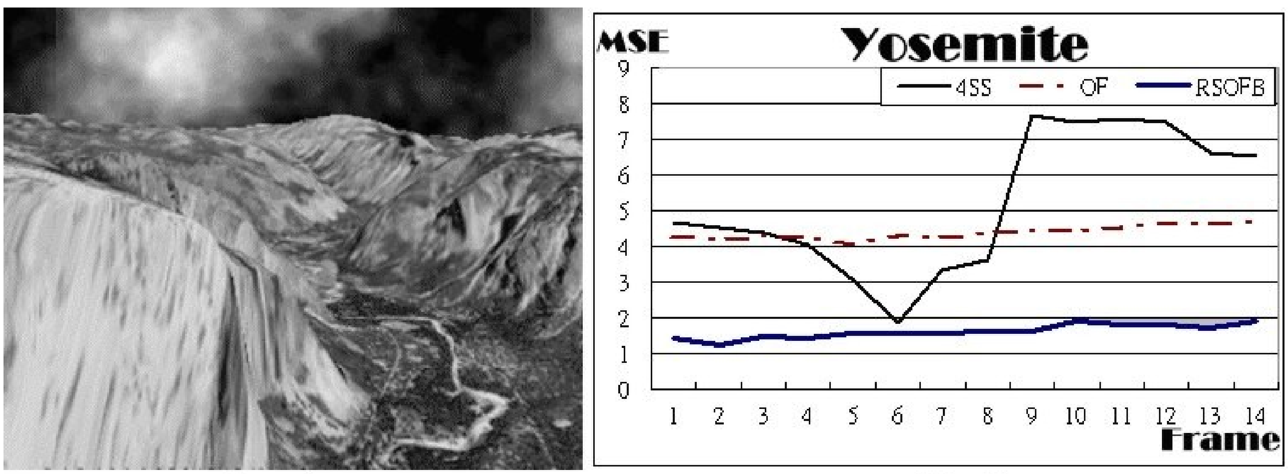


Fig.6. Motion fidelity analysis of various Gamma values (a) Translation (b)Zoom (c)Rotation



(a) (b)

Fig.7. Yosemite sequence results (a)Yosemite sequence (b) MSE

In this research, a region-based selective optical flow back-projection (RSOFB) is proposed for genuine motion vector estimation. The RSOFB back-projects optical flows in a region to restore the genuine motion vector based on the minimization of the projection mean square errors of optical flows on gradient directions. The RSOFB has been compared with the four-step search (4SS) and optical flow (OF) approaches under various situations. Results have shown that the RSOFB can provide the motion vector closer to the true motion vector, showing smaller MSE values. However, the 4SS gives higher visual similarity results, presenting higher PSNR values. This also implies that for applications which require true motion vectors, such as in MPEG-4 VOP segmentation and object tracking, the RSOFB presents as a more promising choice. However, in our experimental experience, when the motion is significantly large, optic-flow model-based method will gradually lost accuracy. Therefore, the optic-flow model-based methods, include we proposed method, are suited for low to medium motion video sequence.