움직이는 카메라에 의한 변화하는 환경하의 강인한 배경 획득 및 유동체 검출

김태호O 조강현
울산대학교 지능시스템 연구실
{thkimO, jkh2005}@islab.ulsan.ac.kr

Robust background acquisition and moving object detection from dynamic scene caused by a moving camera

Taeho KimO Kang-Hyun Jo
Intelligent systems lab. Uinv. of Ulsan

Abstraction

A background is a part where do not vary too much or frequently change in an image sequence. Using this assumption, it is presented a background acquisition algorithm for not only static but also dynamic view in this paper. For generating background, we detect a region, where has high correlation rate compared within selected region in the prior pyramid image, from the searching region in the current image. Between a detected region in the current image and a selected region in the prior image, we calculate movement vector for each regions in time sequence. After we calculate whole movement vectors for two successive images, vector histogram is used to determine the camera movement. The vector which has the highest density in the histogram is determined a camera movement. Using determined camera movement, we classify clusters based on pixel intensities which pixels are matched with prior pixels following camera movement. Finally we eliminate clusters which have lower weight than threshold, and combine remained clusters for each pixel to generate multiple background clusters. Experimental results show that we can automatically detect background whether camera move or not.

1. Introduction

One of the main tasks in video surveillance, auto moving mobile robot is segmentation of moving objects. Most common approach to identifying the objects is background subtraction. Moreover, background subtraction powerfully detect moving object from stationary camera. In recent decades, many researchers try to reconstruct robust background model but it still has difficult to strongly identify moving objects especially when camera move. If we can generate robust background even though camera moves, it has quite big merit. The meaning of having robust background under the case of camera move is we know the environment without object. Moreover we also can expect the same backgrounds like a building, tree or other artificial objects are located in matched position with camera moves. Therefore it can be successfully used to detect moving object and localize obvious landmark for map generation by automobile robot. For this reason, we present background reconstruction algorithm which can overcome the effect of camera movement in this paper.

Generating background has many methods. One of the methods is mean-background model in time sequences which model is estimated by averaging image sequence during time. It is quite simple method and useful when few object are existed in image sequence. However, many objects move in the sequence or a few objects belong in the image long time, the objects have heavy effect to generate wrong background model. Another algorithm is pixel intensity classification(PIC) proposed by Hou et al.[12]. Main idea in PIC is intensity of background has high probability in time sequence. Also their frequency of intensity is higher than any other, so it is enough to select for background pixel. It also useful method to estimating background model but if a frequency of existing object has same or higher than real background, it is unfortunately selected to a background. Using single Gaussian background model was presented by Wren et al.[2]. It is statistical and probability background but it is difficult to deal when background are multimodal. The problem which single Gaussian model has are solved by Stauffer et al.[7] called mixture of Gaussian(MOG). This algorithm detects objects even though background is multimodal. It also segments objects when image sequence has slowly changing illumination, repeated movement. However MOG has to estimate its parameters computationally. Therefore badly estimated parameters are brought harmful effect for MOG. Elgammal et al.[10] adapted a multimodal background model for each pixel. They use a Gaussian function for background model and the width of the Gaussian function is estimated by calculating median value of absolute differences between related image frames. The merit of their idea is this method work under dynamic situations even though the image sequence includes small movements. The other algorithm is median filter. Using this method, we define background model to the median value for each pixel from image buffer which included whole image sequence. It is simple idea to estimating background however it is similar method with pixel intensity classification method. In this algorithm, a pixel will be belonged to a background if a median value of the pixel will be exiting over half frames in the sequence. Xiao et al.[14] proposed multi-background images which is organized by classification of clusters for each pixel. It is simple and effective algorithm to generate background model. Generated background is statistical and predictable. Moreover their algorithm has high performance to estimate background almost same as MOG with low cost. However, multi-background images are difficult to deal when camera moves because this algorithm also organized into pixel base.

In this paper, we propose a background reconstruction method which effectively constructs the background for dynamic scene and dynamic view. The proposed method works well for both static and dynamic scenes

2. Detecting movement of camera

If we generate robust background, we easily detect moving object. As we find in recent research[2, 6-8, 10-13], a number of algorithms to generate background without camera movement are already existed. Unfortunately, most algorithm are weaken when camera has unexpected effect like shaking, waving or moving for a long time. In this paper, we present pixel based multiple background algorithm that is calculating relation between current region to prior region which has high similarity calculated by correlation rate. Detection of camera movement is described in Fig. 1.

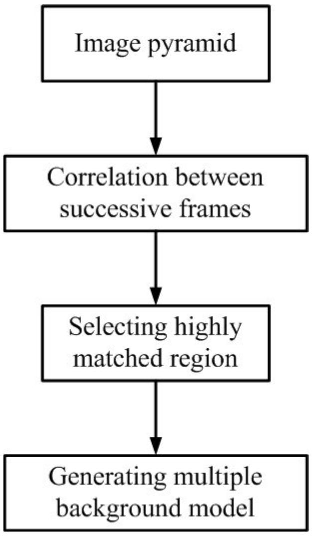


Fig. 1. Block diagram for detecting movement of camera

2.1. Image Pyramid

When camera has unexpected movement, it makes difficult to generate robust background.

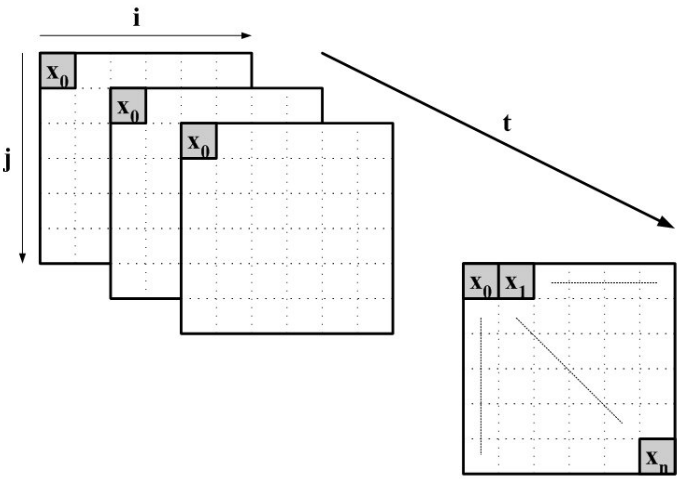


Fig. 2. Image sequence

$$\boldsymbol{I}_t(p)=\{I_t(x_1),I_t(x_2),\cdots,I_t(x_n)\},\quad p=\{x_1,x_2,\cdots,x_n\} \qquad (1)$$

To overcome this problem, we calculate the similarity by correlation rate between current and prior image. However, correlation process needs high cost when searching area is quite large. Moreover small changing of intensities in the image possibly generates lower similarity. Therefore we use image pyramid to reduce calculating cost and small noise. In experimentation, we use 320×240 resolution image and 0 to 2 times pyramid image are used to detect camera movement. Mean value in 2×2 region is used to make pyramid image. $\boldsymbol{I}_t^{\delta}(p)$ is a δ times pyramid image and it is calculated by eq. (3)

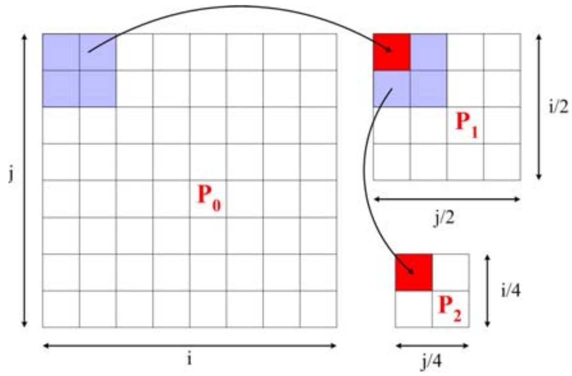


Fig. 3. Image pyramid

$$\boldsymbol{I}_t^{\delta}(p)=\{I_t^{\delta}(x_1),I_t^{\delta}(x_2),\cdots,I_t^{\delta}(x_\chi)\},\quad \chi=\frac{n}{4^{\delta}} \qquad (2)$$

$$I_t^{\delta}(x_\chi)=I_t^{\delta}(i_{new},j_{new})=\frac{1}{4^{\delta}}\sum^{l+1,m+1}I_t^{\delta-1}(l,m) \qquad (3)$$

$$\begin{cases} i_{new}=l/2,\ j_{new}=m/2,\ \delta=1,2 \\ \text{Interval}:l=l+1,\ m=m+1 \end{cases}$$

2.2. Detecting Matched Region Compare within Prior Image

Not only the algorithm of multi-background images[7, 10, 11, 13, 14] but also most of recent background generation algorithms are organized on pixel base. The common weak point in those algorithms is it easily gets harmful effect to organize background by moving camera. We use correlation rate between current and prior images to overcome common weak point. Each pixel in image sequence is defined in eq.(1) and pyramid images are described in Fig.3 and it is defined in eq.(2), (3).

Checking the similarity, we calculate correlation rate for 10×10 region on current image within 30×30 searching region in previous image where includes the same pixel positions with current region.

$$\begin{aligned}\rho_t^{I_t^{\delta},I_{t-1}^{\delta}}(k) &= \frac{\operatorname{cov}(I_t,I_{t-1})}{\sigma_{I_t}\cdot\sigma_{I_{t-1}}} \\ &= \frac{E\left((I_t-\mu_{I_t})(I_{t-1}-\mu_{I_{t-1}})\right)}{\sigma_{I_t}\cdot\sigma_{I_{t-1}}}\end{aligned} \qquad (4)$$

In eq. (4) k is searching areas which size are 10×10. After we calculate similarity for whole searching region, we select best matched region by eq. (5).

$$k_t = \arg\max_{\rho}\left(\rho_t^{I_t,I_{t-1}}(k) > 0.9\right) \quad (5)$$

When we find the best matched region k_t, we easily compose moving vector based on eq. (6). The magnitude and angle for moving vector, which are defined to υ_k and θ_k respectively in eq. (6), are described the movement of each best matched region. (i_c, j_c) is left-top pixel in current region and (i_p, j_p) is left-top pixel in best matched region within prior image. Among movement vectors, we choose a vector which has high density compare with others.

$$\begin{aligned} |\upsilon_k| &= \sqrt{(i_p - i_c)^2 + (j_p - j_c)^2} \\ \theta_k &= \tan^{-1}\frac{j_p - j_c}{i_p - i_c} \end{aligned} \quad (6)$$

3. Background Reconstruction

Fig. 4 shows three typical distributions for pixel belonged in background. Whole intensity distributions are taken during 100 frames for each pixel with gray scale. Data 1 is an intensity distribution which has almost frequently varied background like a neon sign board. Among 100 frames, data 1 has heavy variation different with data 2 and data 3. Data 2 is a case of high building's wall which is difficult to occlude by moving objects. As we intuitively know on the curve of data 2, it is a pixel which has typical distribution for background having almost statically small variation for long frames. Data 3 is a intensity distribution which is easily occluded by moving objects. On the curve of data 3, we also find similar intensity distribution with data 2 among 0^{th} to 14^{th}, 24^{th} to 46^{th} and 51^{st} to 100^{th} frame. Those three distributions are different but all of those have background. To generate background which include those distribution altogether, we should use multiple background model

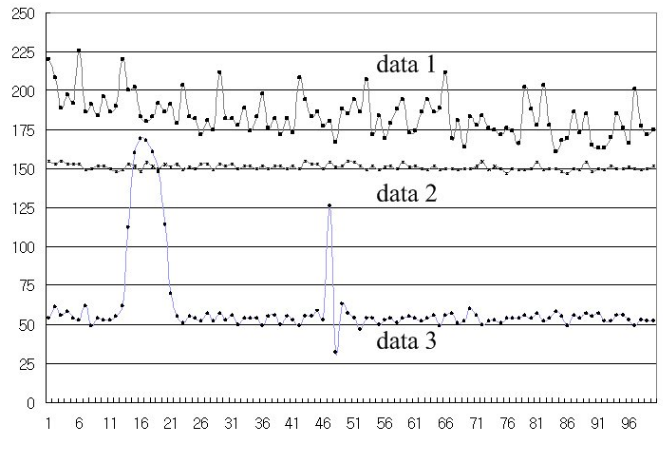


Fig. 4. Variations of intensity in 100 frames

3.1. Algorithm of background reconstruction

Even static background which has few intensity variation in the natural environment, sometimes it is also failed to generate background model because of unexpected luminance changing, moving foreground or noise by camera etc.

In this paper, we protect to be blended background model by multiple clusters for each pixel. When we have best matched region by correlation rate, we classify clusters for each pixel based on camera flow. Selected N frame pixels from whole image sequence are described in eq. (7) with clusters belong in each pixel have frame members and their mean value.

$$I_t(p) = (I_1, I_2, \cdots\cdots, I_N), \quad t = 1, 2, \cdots\cdots, N \quad (7)$$

$M^i(p)$: Member of each cluster

$C^i(p)$: Mean of cluster

Our approach need to initial condition to classifying clusters on each pixel. Therefore the pixels in first frame must be a initial clusters. From very next frames, multi-background clusters are automatically classified by the algorithm which described in [14]. When multi-background clusters are successfully estimated, even though we don't expect but a few clusters include moving objects. Data 3 in Fig. 4 shows unexpected case which includes objects in multi-background clusters. However we intuitively know that weight of each cluster, that is the key to solve this error, because unexpected clusters existed by moving object must have small weight compare with background. Therefore we try to find weight for each cluster to reducing error clusters.

$$W^i(p) = \frac{M^i(p)}{\sum_{t=1}^{G(p)} M^i(p)}, \quad i = 1, 2, \cdots, G(p) \quad (8)$$

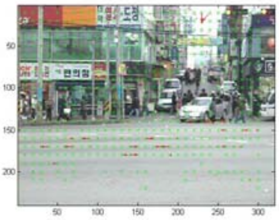
$W^i(p)$ is a weight for i^{th} clusters in a pixel p. The next step is eliminating small clusters among $G(p)$ depend on $W^i(p)$. Probabilistic threshold ζ is used in this process. Finally we obtain $S(p)$ clusters from $G(p)$ which has high weight than threshold ζ. Using obtained cluster $S(p)$, we decide multiple background with their weight and numbers by eq. (9), (10)

$$B^i(p) = C^i(p), \quad i = 1, 2, \cdots, S(p) \quad (9)$$

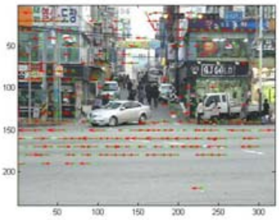
$$W^i(p) = \frac{M^i(p)}{\sum_{t=1}^{G(p)} M^i(p)} \quad (10)$$

4. Experimental Result

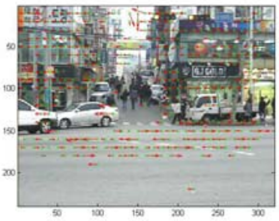
In experiment, we use initial value followed from [14]. Because we use dynamic scenes caused by moving camera, however, we need to determine camera dynamics. Finding camera dynamics, we use histogram of moving vectors which vectors are calculated by eq. (6).



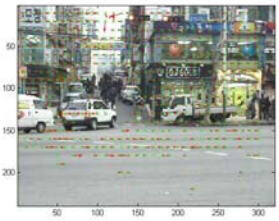
(a) frame 117



(b) frame 136



© frame 139



(d) frame 148

Fig. 5. Samples of camera flow

(a) frame 217

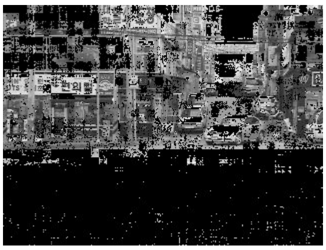
(b) frame 241

© frame 275

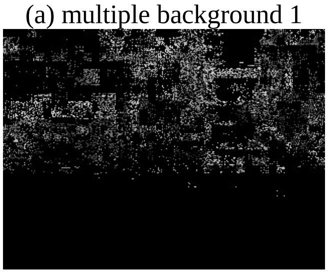
(d) frame 311

Fig.6. Sample of original sequence in first 100 frames.

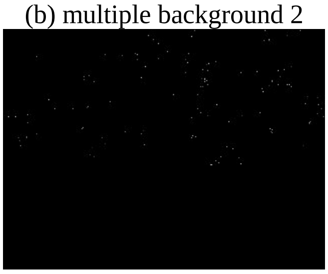
(a) multiple background 1



(b) multiple background 2

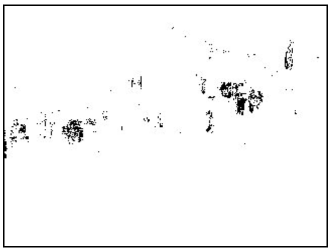


(c) multiple background 3

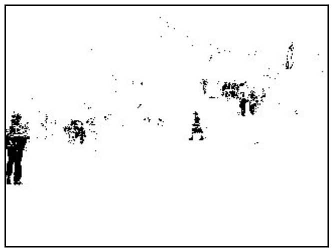


(d) multiple background 4

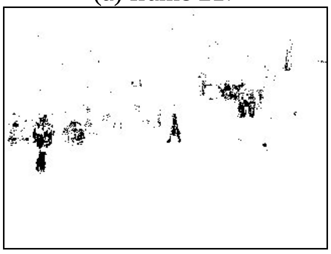
Fig.7. Result of multiple background.



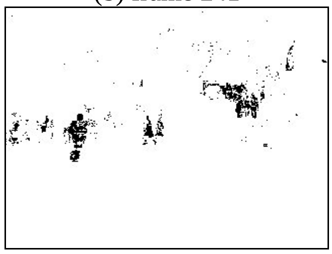
(a) frame 217



(b) frame 241



(c) frame 275



(d) frame 311

Fig.8. Detecting result of moving object by multiple background.

When we choose one of high density vector described to a camera movement from the histogram, we start generating process for multiple background. In this process, if a pixel contains over 100 existences in image sequence, it is started to generate multiple background for the pixel. Generated multiple background model can be bigger than image size because of moving camera. However it can not be the background, even though camera moved, if the existence of pixel can not over 100 times. Fig. 6 shows sample image from original sequences and we see multiple background in Fig 7. In original image sequence, lots of moving objects are existed in the image but from Fig. 7 and their extracted result in Fig. 8 shows detected result of moving object by multiple background model.

5. Conclusions

This paper is concerned with robust background acquisition and object detection. Robust background is useful to separate moving object and it is used in many application like a video surveillance, auto moving mobile robot etc.. Especially we try to reconstruct multiple background even though a camera moves. Correlation rate is used to find a best matched region in successive frames. Using matched region we reconstruct multiple background which contain robust background under moving camera. Experimental results show that the algorithm handles cases where the image sequence contains dynamic scenes caused by a moving camera.

Acknowledgment

The authors would like to thank to Ulsan Metropolitan City and MOCIE and MOE of Korean Government which partly supported this research through the NARC and post BK21 project at University of Ulsan..

References

[1] B. Gloyer, H. K. Aghajan, K. Y. Siu and T. Kailath, *"Video-based freeway monitoring system using recursive vehicle tracking"*, Proc. SPIE, vol.2421, pp.173-180, 1995.

[2] C.R. Wren, A. Azarbayejani, T. Darrell and A.P. Pentland, *"Pfinder: Real-time tracking of the human body"*, IEEE Transactions on Pattern Analysis and Machine Intelligence, ver.19, iss.7, pp.780-785, 1997.

[3] R. Cutler, and L. Davis, *"View-based detection and analysis of periodic motion"*, Proc. Fourteenth International Conference on Pattern Recognition, pp.495-500, 1998.

[4] I. Haritaoglu, D. Harwood and L. Davis, *"Who? When? Where? What? A Real Time System for Detecting and Tracking People"*, Int. Conf. of Automatic Face and Gesutre Recognition, pp.222-227, 1998.

[5] I. Haritaoglu, D. Harwood, and L. Davis, *"A real time system for detection and tracking people"*, Int. J. of Image and Vision Computing, 1999

[6] P. Kornprobst, R. Deriche and G. Aubert, *"Image sequence analysis via partial difference equations"*, Int. J. of Mathematical Imaging and Vision, ver.11, iss.1, pp.5-26, 1999.

[7] C. Stauffer and W.E.L. Grimson, *"Adaptive background mixture models for real-time tracking"*, Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp.246-252, 1999.

[8] K. Toyama, Krumm J., Brumitt B. and Meyers B., Wallflower, *"Principles and practice of background maintenance"*, Int. Conf. on Computer Vision, Kerkyra Greece, pp.255-261, 1999.

[9] I. Haritaoglu, R. Cutler, D. Harwood and L. Davis, *"Backpack: Detection of people carrying objects using silhouettes"*, Int. Conf. of Computer Vision and Image Understanding, pp.385-397, no.3, 2001.

[10] A. Elgammal, R. Duraiswami, D. Harwood and L.S. Davis, *"Background and foreground modeling using nonparametric kernel density estimation for visual Surveillance"*, Proc. of the IEEE, ver.90, iss.7, pp.1151-1163, 2002.

[11] Liyuan Li, Weimin Huang, Gu, I.Y.H., Qi Tian, *"Foreground object detection in changing background based on color co-occurrence statistics"*, Sixth IEEE Workshop on Applications of Computer Vision, pp.269- 274, 2002

[12] Z.Q. Hou, and C.Z. Han, *"A Background Reconstruction Algorithm based on Pixel Intensity Classification in Remote Video Surveillance System"*, Proc. of the Seventh International Conference on Information Fusion, pp.754-759, 2004.

[13] Liyuan Li, Weimin Huang, Irene Yu-Hua Gu, Qi Tian, *"Statistical modeling of complex backgrounds for foreground object detection"*, IEEE Transactions on Image Processing, vol.13, no.11, pp.459-1472, 2004

[14] Mei Xiao, Chongzhao Han, Xin Kang, *"A background reconstruction for dynamic scenes"*, Information Fusion, pp.1-7, 2006

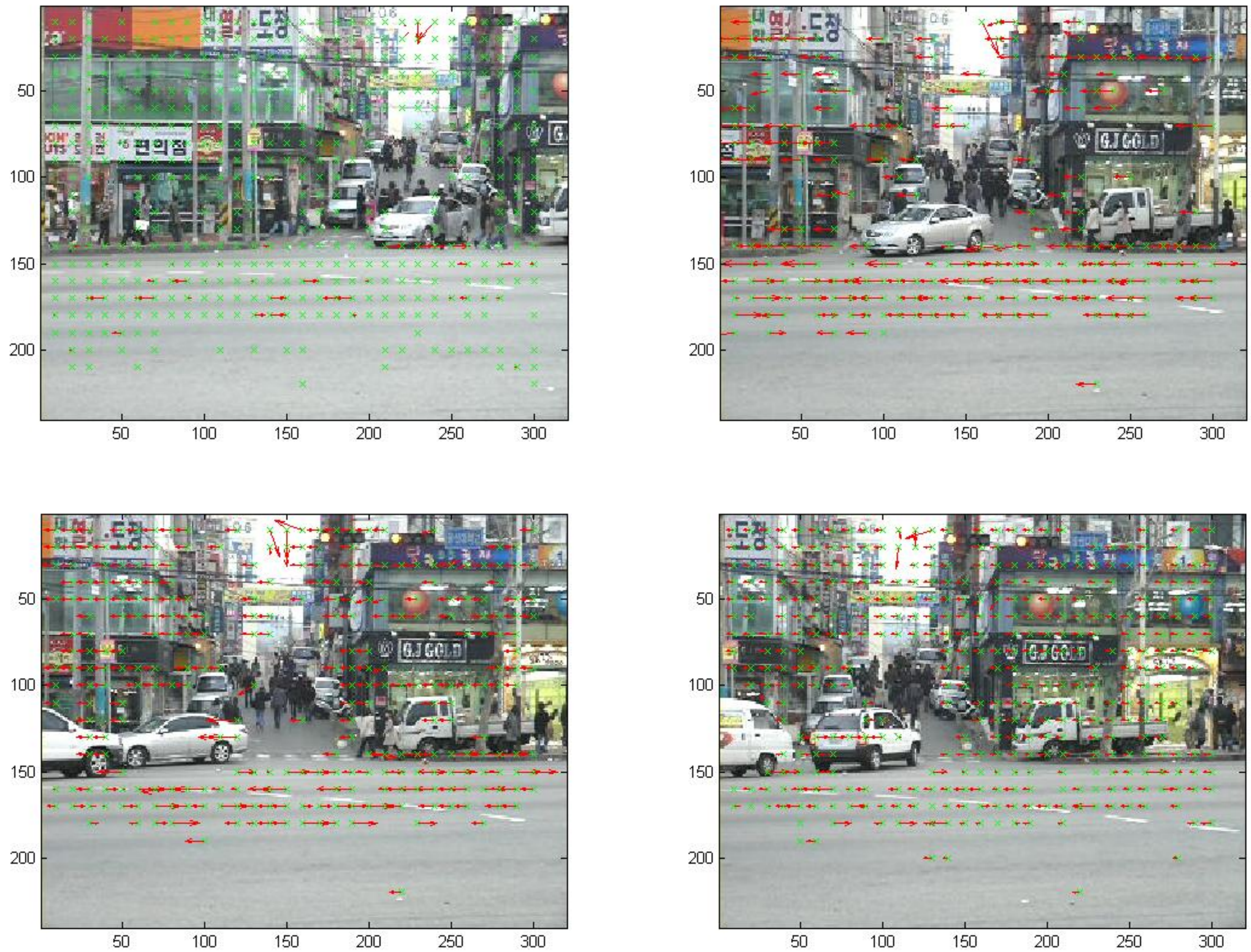


Fig.9. High resolution images for Fig. 5