REVIEW AND EXPERIMENT OF THE DIFFERENT METHODS FOR MOVING OBJECT EXTRACTION FROM A LIVE VIDEO

Sina Adham Khiabani and **Yun Zhang**
University of New Brunswick, Department of Geodesy and Geomatics Eng.
Fredericton, Canada
s.adham@unb.ca, YunZhang@unb.ca

ABSTRACT

Moving object extraction is a critical stage in a measurable 2D and 3D photogrammetric surveillance systems. In order to keep the frame rate of the live video, the chosen automatic method should be fast enough. Moreover, it is important to extract the moving objects precisely to be able to increase the accuracy of the measurements. Prior to the boundary extraction stage, the moving objects must be detected and separated from each other in each frame of the live video. In this paper, the various spatial domain (e.g. background subtraction) methods including intensity-based and feature-based strategies have been introduced for the purpose of moving object detection. Following the detection step, some utilized methods for object separation have been noted. Furthermore, some useful approaches of error reduction and accurate boundary extraction of detected moving object have also been noted. In addition to the spatial domain methods, some frequency domain methods (e.g. Wavelet decomposition) of moving object detection are also considered in this paper.

Keywords: Photogrammetry, Computer vision, Artificial intelligence, Moving object detection, extraction, Spatial domain, Frequency domain

INTRODUCTION

Automatic moving object extraction is an essential task in many surveillance projects. In many of these projects, it is necessary to detect moving objects on the scene in order to track or locate them through the interested scene. In some projects, it is needed to only detect the existence of moving objects or at most count the number of existing moving objects. While in some projects, it is additionally essential to extract borders of moving objects. Borders of moving objects could be used for different reasons like shape analysis, object matching or even accurate measurements. In summary, requirements of a project could impose the required steps and accuracy of the object detection and extraction stage. In a stereo system as ours, accurate boundary extraction is important due to the requirements of the system to match the extracted object for further measurements.

For a live-stream continuous capturing, it is important to extract moving objects completely automatic and in the fastest possible period. In order to keep the frame rate of a live video, the whole processing period on each frame must be less or at most equal to the frame refreshing interval. As an instance, for a 25 frame per second live video, each frame must be processed in less than 40 milliseconds. As result, in order to sustain the time requirements of live video streams such as rate, the utilized algorithms should not be time consuming and of course it must be completely automatic.

In our project, it is needed to monitor and extract moving objects in various security cameras in a same time. Moreover, further analysis on images of an object (or even multiple objects) from different views will be critical. Hence, in this project, it will be important to extract the shape of objects accurately. Furthermore, as security cameras are used, the live monitoring and analysis will be an issue in the project. In this paper, as needed, it will be attempted to review various automatic methods in which the extraction of multiple moving objects is possible with the required rate.

In object extraction stage, separating the signal (foreground) pixels from the background ones is the main duty. Technically, any interested moving object treated as foreground in the scene and other objects which remain unchanged during the capturing time, treated as background. As the first task, it is necessary to model or define the unchanged pixel as background. As foreground, it is possible to have unique or several moving objects in a single frame. In both cases, it is necessary to find the changing pixels or object in this step. Furthermore, for the more-than-one object cases, it is additionally essential to separate the moving objects from each other. In addition, to complete the extraction stage, error reduction strategies and accurate boundary extraction must be applied to the frames.

Like many other image processing projects, object detection and extraction could be implemented in both spatial or frequency domain. In the next two sections, studied methods in spatial domain and frequency domain are introduced separately. In the next section, utilized strategies for different stages of motion detection are noted. Moreover, in this section, some utilized methods to improve the detection accuracy are considered. Furthermore, in the last section, a brief discussion over different methods is mentioned and some suggestions for the future activities are also noted.

SPATIAL DOMAIN METHODS OF OBJECT EXTRACTION

As mentioned in the previous section, moving object extraction process could be separated into three steps: (1) Foreground (signal) detection, (2) object separation and (3) accurate boundary extraction. In the following section, different reviewed strategies on each step are noted. In this section, it has been tried to consider major noticed issues of each part.

Foreground Detection

Two types of intensity-based and feature-based methods have been proposed for the spatial domain object detection. Intensity-based methods are mostly based on the intensity difference of instantaneous frame with one or a set of reference image. The reference image is a scheme of background which could be acquired from a still scene (Li, 1997) or estimated from a set of information from unchanged pixels or objects in a scene (Haritaoglu et al, 2000). Background subtraction (Li, 1997; Spagnolo et al, 2006; Elhabian, et al, 2008; Li, et al, 2010; Li and Yang, 2011) and frame differencing (Cao et al, 2009) are two famous intensity based methods.

Feature-based methods are those in which the comparison or analysis on a certain feature(s) of the scene will result in the determination of being background or foreground of a pixel or object. Multi frame analysis (Stauffer and Grimson, 2000; Fang et al, 2006) and optical flow (Hu et al, 2004; Yokoyama and Poggio, 2005) are two common methods of feature-based foreground extraction.

Intensity-based methods. Back ground subtraction is the most common method of object detection. In this method, a reference image or a set of reference images are utilized as a base (Spagnolo et al, 2006). This reference image is accepted as a model of background and it is assumed that the reference image is a background, empty of any moving object. By subtracting the frames of the video from the reference image, the different areas could be extracted as the container of the moving objects. In this situation, usually, a simple threshold is defined (Elhabian, et al, 2008; Li, et al, 2010) as a criterion of changes. In this stage, the values over the threshold are assigned value of 1 and the ones below the threshold are assigned 0 in the subtracted image. This threshold has been mainly set to avoid noises and small illumination changes in the final product.

As illumination changes affect the pixel values of the image during the video surveillance, in practice, the simple background subtraction is not an ideal strategy of motion object detection. Updating reference image during the time (Peijiang, 2009) is a possible solution for this problem. However, updating the reference image during the time, especially in crowded places is a difficult task. Choi, et al, (2011) developed three kinds of illumination change, chromatically difference and brightness ratio models to remove the false detected pixels which have been chosen due to the illumination change. Chen, et al, (2011) have also suggested a method based on a bright and a dark reference images and made their background subtraction method illumination sensitive.

Cao, et al, (2009) have suggested a frame differencing method to avoid the illumination changes and temporary noise between the frames. In this method, two or a few respective frames are compared and unchanged pixels are removed completely and the remained pixels assumes as foreground. Incomplete detection of the shape and dependency on the size and texture of the moving object are the main problems of this method (Huang, 2011).

Feature-based methods. Multi frame analysis based on a specific feature of the frames is another way of motion detection in the surveillance videos. Colors (Wren, et al, 1997; Fang, et al, 2006), statistical features of the image (Stauffer and Grimson, 2000) or even texture (Choi, et al, 2011) could be used as a feature of interest in a multi frame analysis.

Stauffer and Grimson, (2000) suggested an elaborated statistical background subtraction model based on the local statistical features of an image. They used a mixture of Gaussians in neighborhood of each pixel instead of a single Gaussian density function for pixels. This strategy is more reliable for non-stationary backgrounds such as natural elements.

Wren, et al, (1997) and Fang, et al, (2006) suggested different methods using color features. Fang, et al, (2006) introduced a method using the color information differences to recognize the moving objects. They utilized the differences of the color properties of the moving objects and background in the YUV color system and extracted the

moving objects from the scene. Although color properties make the subtraction more reliable, the subtraction is also affected by the color noise of the frames.

Optical flow is a common feature-based method for motion detection. Yokoyama and Poggio, (2005) suggested a contour-based method in which by using the computed lines from gradient-based optical flow the moving objects will be extracted. Tian, et al, (2005) were also utilized optical flow to project the motion on the image plane. This method handled the complex background successfully. However, optical flow is difficult to implement and its computational complexity make it less applicable for the fast motion detection requirements (Huang, 2011).

Object Separation

The object separation is not a complicated issue for the objects which have been located far from each other. In this case, a simple thresholding could solve the problem. On the other hand, near objects, objects on a same line of sight and the objects which have been in the other objects' shadow are the main challenges in this area (Li, et. al, 2010). Hence, depending on the probable situations and required accuracy, an appropriate strategy could be used to separate the moving region to the extract objects. Stojkoska, et al, (2008) suggested a network to find the moving objects. They used N-queen network algorithm to label the moving objects. Vu, et al, (2010) have also introduced a grid-based method to find the correct interested object among the objects in the scene which has the potential for implementation in the high speed accurate moving object detection projects. Furthermore, it is possible to process multi-frames and use the tracking strategies to separate the objects by attention to their motion properties (like speed or path) (Augouris, et al, 2008). Huang, (2011) has also proposed a binary object detection mask to accurately separate and extract the moving objects.

Accurate Boundary Extraction

In order to extract boundaries with a high accuracy, the errors of detection should be reduced. There are some common errors like noise, false detections and missing objects which must be reduced in the detection image. Many research activities have been done to reduce the mentioned errors. Some of these errors could be reduced by use of the appropriate filters in the detection stage (e.g. Vargas et al, 2010). However, some of these errors need a further improvement on the detection stage results. In this section, it is tried to review some implemented strategies of noise reduction, shadow removal (as one of the most common issues of false detection) and aperture error removal (as a solution for a common missing object cause).

Illumination changes, minute changes of the camera, color noise, system noise and etc are the common noises which have unavoidable effect on the final detection product. To avoid the noise in the motion detection stage, filters (especially adaptive filters) are widely used. Vargas et al, (2010) suggested using the sigma-delta filter to the background subtraction stage to achieve more pure object map. This method is useful to reduce the instability of the background or the noises caused by small movement of the camera or background. Gaussian filter is also could be added to the background subtraction method to reduce the noise caused by instable background (Stauffer and Grimson, 2000). Moreover, multiple thresholds (Wang et al, 2010) could be used to reduce the noise (especially illumination noise). Hartatoglu, et al. (2000) also suggested a double thresholding method to reduce the noise.

One of the most common issues in this stage is the false detection of shadows (Hsieh, et. al, 2003; Huawei, et al, 2005; Leone and Distante, 2007). As many detection methods utilized the differentiating strategies, the shadow of moving objects or some local changes in the illumination of some parts may become problematic in this area. In order to separate the shadows from the moving objects, Gaussian modeling (Hsieh, et al, 2003), Color-space-based method (Huawei, et al, 2005) and texture analysis (Leone and Distante, 2007) have been utilized.

Uniform or low-gradient regions of an image could cause some misdetection in the final results (Elhabian, et al, 2008). This problem is known as aperture error in the motion detection research area. Antic et al. (2009) proposed a multi resolution temporal change detection to reduce the mentioned error. Low illumination area of an image is also caused such error of misdetection due to the small range of changes due to appearance of a moving object (Elhabian, et al, 2008).

FREQUENCY DOMAIN METHODS OF OBJECT EXTRACTION

Motion detection could be performed in frequency domain. Most of the proposed methods in this area have been implemented by using wavelet transforms. Antic, et al. (2009) suggested a wavelet-based multi resolution frame differencing technique. Toreyin, et al, (2006) used the wavelet transform to estimate the background from the past video frames and computationally extract the moving objects and their location. As the changes appears as

outliers in the high-low and low-high sub-images (Bagci, et al, 2002), fractional low-order statistic could be utilized to find and locate the moving objects. Changing the background illumination has less effect on the extraction stage in the frequency based methods (Toreyin et al, 2006). However, the misdetection of the shadows is still an issue in this method.

DISCUSSION AND CONCLUSION

As mentioned in the introduction section, our interested object detection point of view should meet the requirements of a photogrammetric system. In our system, automatic object detection and extraction must be performed to detect and separately extract the moving objects within a scene of interest. As 3D measurements should be possible in the further process, it is important to extract the boundaries and location of the moving objects with the highest possible accuracy. Moreover, as a video surveillance system is considered as a platform of the automated software, it is required to have the least possible noise to increase the reliability of the alarm system. Furthermore it is important to process the frames as fast as possible to meet the requirements of a live video.

In this paper, a review over the different methods of motion detection was performed. Two types of spatial domain and frequency domain method was mentioned in this research. Although frequency domain methods are less sensitive to the uniform illumination changes and uniform background objects, spatial domain methods are more popular in the video surveillance systems. It might be due to the relative complications of the frequency methods in comparison with the spatial domain methods. Moreover, it could be related to the suitable required processing time in the spatial domain methods including detection, separation and extraction steps in comparison with the required time to perform the frequency based method and preparation time for the visual purposes. However, frequency-based methods show their abilities in indoor locations (Antic, et al, 2009). Moreover, these kinds of methods have better compatibility with the frequency domain based methods of face detection.

Foreground detection, object separation and accurate boundary extraction were considered as different steps of a required solution. In each step, different methods were reviewed. Two types of intensity-based and feature-based methods were introduced for the foreground detection step. Intensity-based methods are mostly based on the single pixels' intensity. In these methods, detection step is based on a noticeable change intensity of a pixel. However, feature-based methods focused on a specific feature of the images or behavior of the intensity changes in neighborhood of each pixel. Selection of a method between a combination of an intensity-based strategy with an error reduction process and a more accurate feature-based method seems to be challenging. The complication of the method, processing tools and the required processing time could be taken into account in this selection. However, feature-based methods have shown better outputs in the crowded places and the scenes with a high illumination changes. On the other hand, intensity-based methods have shown their suitability and quickness at indoor situations and night surveillance.

Importance of the tracking stage makes the separation step more significant. For tracking purpose it is fundamental to separate and identify the moving objects in the scene. In some projects, it is needed to lock on specific types of moving objects or in some projects, it is needed to track the set off path of each object in the specific area. Accurate moving object separation could make the tracking stage more straightforward.

Identification of various pixel sets of detected moving objects is important for the shape analysis and measurement purposes. Noise reduced and illumination sensitive strategies make the extraction stage more accurate. Shadows are counted as common false detections in the motion detection systems. Different methods of shadow removal were listed in this paper. A fast texture-based or angel-based shadow removal tool is a great addition to the set of required process of accurate object extraction.

As a conclusion, based on the preformed review on different utilized methods of moving object extraction, it could be said that the rate of illumination changes, rate of movements, overall speed of moving objects and also the type of moving objects (cars or people) play a key role in determination of the suitable method for moving object detection. Moreover, requirements of further process on the extracted moving objects are important to design the different accuracy improvement strategies. Abilities of the processors, quality of the cameras and required process rate are also counted as effective parameters in choosing the best method.

REFERENCES

Agouris, P., P. Partsinevelos, A. Stefanidis. Differentiating and modeling multiple moving objects in motion imagery datasets. ISPRS congress proceeding, Commission V, WG V/5, Beijing, China. 2008.

Antić, B., V. Crnojević, and D. Ćulibrk, "Efficient wavelet based detection of moving objects," in Proc. 16th Int. Conf. Digital Signal Process. (DSP, pp. 1-6), Jul. 2009.

Bagci, M., Y. Yardimci, A.E. Cetin, Moving object detection using adaptive subband decomposition and fractional lower order statistics in video sequences, Signal Process. International Journal of Signal Processing. Elsevier. December 2002.

Cao, J., Wei Wang, and Yan Liang, "An Improved Detection Algorithm of Moving Vehicles Based on Computer Vision," Proc. IEEE ICIS 2009, pp. 20 -24, Nov. 2009.

Chen, Y.L, B.F. Wu, H.Y. Huang and C.J. Fan. "A Real-Time Vision System for Nighttime Vehicle Detection and Traffic Surveillance". IEEE Transactions on Industrial Electronics VOL. 58, NO. 5, MAY 2011.

Choi J.M., H. J. Chang, Y. J. Yoo, J. Y. Choi. "Robust moving object detection against fast illumination change". Elsevier Computer Vision and Image Understanding 116 (2012) 179–193. 2011.

Elhabian, S., K. El-Sayed, S. Ahmed, "Moving Object Detection in Spatial Domain using Background Removal Techniques – State-of-Art" (2008), Recent Patents on Computer Science, Volume 1, Number 1, pages 32-54, January 2008.

Fang, X., W. Xiong, B. Hu, L. Wang, "A Moving Object Detection Algorithm Based on Color Information," International Symposium on Instrumentation Science and Technology, vol. 48, pp. 384-387, 2006.

Haritaoglu, I., D. Harwood, and L. Davis, "W4 realtime surveillance of people and their activities," in IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, pp. 809-830, 2000.

Hsieh, J., H. Wenfong, C. Chiajung & C.,Yungsheng. Shadow elimination for effective moving object detection by Gaussian shadow modeling. Image and Vision Computing. 21, pp505-516. 2003.

Hu, W., T. Tan, L. Wang, and S. Maybank, "A survey on visual surveillance of object motion and behaviors," IEEE Trans. Syst., Man, Cybern. C, Appl. Rev., vol. 34, no. 3, pp. 334–352, Aug. 2004.

Huang, S., "An Advanced Motion Detection Algorithm with Video Quality Analysis for Video Surveillance Systems", IEEE Transactions on Circuits and Systems for Video Technology., Vol. 21, No. 1, JAN. 2011.

Huawei, W.H., B.J. Li., C.W. Quiying, A System for Moving Object Detection and Shadow Extermination. IEEE Int. Workshop VLSl Design & Video Tech. Suzhou, China, May 2005.

Leone, A. and C. Distante "Shadow detection for moving objects based on texture analysis", Pattern Recognit., vol. 40, pp. 1222. 2007.

Li, C.C., A New Approach to the Detection of Moving Objects, J. of Information sciences. Volume 103 Issue 2-4, Elsevier Science Inc. New York, NY, USA, doi>10.1016/S0020-0255(97)00056-X. 1997.

Li, J., C. Shao, C. Dong, D. Zhao and Y. Liu. A Video-based Algorithm for Moving Objects Detection at Signalized Intersection. World Academy of Science, Engineering and Technology 66. 2010.

Li, K. And Y. Yang, A method of background modeling and moving object detection in video surveillance, IEEE, 4th international congress on image and signal processing, 2011.

Peijiang, C. "Moving Object Detection Based on Background Extraction", Computer Network and Multimedia Technology, 2009. CNMT 2009. International Symposium on, On page(s): 1 - 4, Volume: Issue: , 18-20 Jan. 2009.

Spagnolo, P., T.D. Orazio, M. Leo and A. Distante, Moving object segmentation by background subtraction and temporal analysis, Image Vision Computing 24. pp. 411–423. 2006

Stojkoska, B., D. Davcev and V. Trajkovik, N-Queens-based Algorithm for Moving Object Detection in DistributedWireless Sensor Networks. Journal of Computing and Information Technology - CIT 16, 2008, 4, 325–332 doi:10.2498/cit.1001401. 2008.

Stauffer, C. and W. Grimson, "Learning patterns of activity using real-time tracking," in IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, pp. 747-757, 2000.

Tian, Y.L., M. Lu, and A. Hampapur, "Robust and efficient foreground analysis for real-time video surveillance," in Proc. IEEE Int. Conf. Comput. Vision Patt. Recog., vol. 1, pp. 1182–1187. Jun. 2005.

Töreyin, B.U., A. Enis Çetin, A. Aksay, M. B. Akhan, "Moving Object Detection in Wavelet Compressed Video", Signal Processing: Image Communication, EURASIP, Elsevier, vol. 20, pp. 255-264. Mar 2005.

Vargas, M., Milla, J.M., Toral, S.L., Barrero, F., "An Enhanced Background Estimation Algorithm for Vehicle Detection in Urban Traffic Scenes", Vehicular Technology, IEEE Transactions on, On page(s): 3694 - 3709, Volume: 59 Issue: 8, Oct. 2010.

Vu, T.D., J. Burlet, O. Aycard. Grid-based localization and local mapping with moving object detection and tracking. Journal of information Fusion. P. 58-62. February 2010.

Wang L. And N.H.C. Yung, Extraction of Moving objects from their Background based on multiple adaptive threshold and boundary evaluation, IEEE Trans. Intelligent transportation systems vol. 11, pp. 40-51, 2010.

Wren, C. R., A. Azarbayehani, T. Darrell, and A. P. Pentland, “Pfinder: Real-time tracking of the human body,” IEEE Trans. Patt. Anal. Mach. Intell., vol. 19, no. 7, pp. 780–785, Jul. 1997.

Yokoyama, M., T. Poggio, "A Contour-Based Moving Object Detection and Tracking", pp.271-276, 2005 IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance, 2005.