Pertanika J. Sci. & Technol. 25 (S): 159 - 168 (2017)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Parametric Tracking Across Multiple Cameras

Patrick Sebastian[1]*, Yap Vooi Voon[2] and Richard Comley[3]

[1]*Electrical and Electronic Engineering Department, Universiti Teknologi PETRONAS, 32610 Bandar Seri Iskandar, Perak, Malaysia*
[2]*Department of Electronic Engineering, Faculty of Engineering and Green Technology, Universiti Tunku Abdul Rahman, Kampar, Perak, Malaysia*
[3]*Department of Engineering and Information Science, Middlesex University, The Boroughs, Hendon, London NW4 4BT, United Kingdom*

ABSTRACT

This paper presents a tracking method based on parameters between colour blobs. The colour blobs are obtained from segmenting the overall target into multiple colour regions. The colour regions are segmented using EM method that determines the normal colour distributions from the overall colour pixel distribution. After segmenting into different regions on the different colour layers, parameters can be generated between colour regions of interest. In this instance, the colour regions of interest are the top and bottom colour regions. The parameters that are generated from these colour regions are the vector magnitude, vector angle and the value difference between colour regions. These parameters are used as a means for tracking targets of interest. These parameters are used for tracking the target of interest across an array of cameras which in this instance are three cameras. Three cameras have been set up with different background and foreground conditions. The summarised results of tracking targets across three cameras have shown that the consistency of colour regions across different cameras and different background settings provided sufficient parameters for targets to be tracked consistently. Example of tracking performance across three cameras were 0.88, 0.67 and 0.55. The remaining tracking performances across three cameras are shown in Table 2. The tracking performance indicate that the parameters between colour regions were able to be used for tracking a target across different cameras with different background scenarios. Based on results obtained, parameters between segmented colour regions have indicated robustness in tracking target of interest across three cameras.

Keywords: Blob tracking, parametric tracking, multiple camera, tracking

ARTICLE INFO

Article history:
Received: 24 August 2016
Accepted: 02 December 2016

E-mail addresses:
patrick_sebastian@petronas.com.my;
Patrick_sebastian@utp.edu.my (Patrick Sebastian),
yapvv@utar.edu.my (Yap Vooi Voon),
r.comley@mdx.ac.uk (Richard Comley)
*Corresponding Author

ISSN: 0128-7680

INTRODUCTION

Surveillance is an act of monitoring behaviour or activities of a person or a group of people in an area of interest. In this instance, video surveillance is the utilisation of video cameras to obtain video or visual imagery to observe areas of interest for the purposes of control, recognition and monitoring activities. Government, law enforcement and security organisations monitor human activities for the purpose of maintaining control, recognising and monitoring threats in a large crowd of people (Bredereck et al., 2012; Hommel et al., 2012; Wang et al., 2009). Video surveillance cameras are typically connected to a recording device or recording network. The video captured is usually observed or watched live or through a recording by an operator. A single operator would be able to observe all activities within the view field of a single camera. In surveillance of a large area, multiple video cameras provide complete coverage of the area of interest. In this instance, the operator observing the cameras would have a large number of monitors to observe and this would be difficult for a single operator to observe and track all activities from all cameras at the same time. It is difficult for a human operator to monitor multiple surveillance screens and thus, computer systems can intelligently track targets across multiple cameras in a large-scale surveillance system.

BACKGROUND

Tracking of targets in a surveillance system is based on a number of parameters such as area (Park & Aggarwal, 2002), colour (Zhu, 2009), trajectory and velocity. Apart from the general parameters mentioned earlier for tracking a target, targets can be tracked based on features such as edges, lines, corners (Trucco & Plakas, 2006), limbs and heads (Siebel & Maybank, 2002). With the wide availability of a number of parameters and features, tracking methods can vary from simple tracking methods to complex tracking methods.

Colour Space

The utilisation of different colour spaces which range from RGB, HSV and YCbCr colour spaces have different effects in tracking targets of interest. From available colour spaces, RGB colour space has been deemed the not preferred colour space for tracking as the brightness information is embedded in the respective layers of the RGB colour space. It was found from different studies that colour spaces with chrominance information on separate layers from brightness information such as HSV and YCbCr colour spaces have better tracking capabilities (Chippendale, 2006). Chrominance information such as HS and CbCr layers in the HSV and YCbCr colour space respectively could provide sufficient information to be used for tracking (Chippendale, 2006).

Blob Modelling

Foreground and background views are separated in a video scene by detecting blobs. From detected blobs, a number of blob parameters can be extracted such as blob size, motion, vector and centroid (Trucco & Plakas, 2006). Targets can be tracked by utilising the detected blob (Atsushi et al., 2002) or portions of the blob such as the head (Yan & Forsyth, 2005).

The target of interest in this scenario would be the tracking of a whole person which would lead tracking of the whole blob (Atsushi et al., 2002). For parametric tracking, a single blob can be segmented into separate blobs based on colours of the regions (Wang et al., 2014). Single blobs can be segmented into multiple blobs using methods such as watershed segmentation (Gonzalez et al., 2004) and mean shift segmentation (Yunji et al., 2014). When the single blob is segmented into separate blobs, the centroids (Ali et al., 2006) of each blob can be used to generate the parameters for tracking a multi blob target which is illustrated in Figure 6.

TRACKING METHODOLOGY

Tracking as used in this paper is based on the parameters that exist between colour regions of each target; colour region segmentation is the first step. Colour region segmentation is also dependent on the colour space that is used. The YCbCr colour space was selected based on better tracking performance compared with RGB and grayscale colour spaces that had been tested in different tracking methods (Sebastian et al., 2010).

Segmenting target of interest into multiple colour regions would begin with the extraction of colour information from the overall colour pixel distribution. Colour region information can be separated into different normal distributions using EM method (Sebastian et al., 2012; Yiming & Guirong, 2014). Each normal distribution parameters are used to generate different colour regions within the overall target blob region. A person can then be modelled or represented as a collection of coloured regions. Different coloured blobs can be used to denote the different portions of a person such as the head, face, torso and lower limbs (Park & Aggarwal, 2002). An example of segmentation is illustrated in the following figures. Figure 1 shows a snapshot input image while Figure 2 and Figure 3 illustrate the pixel distributions of the Cb and Cr layers for the person that is being tracked respectively.

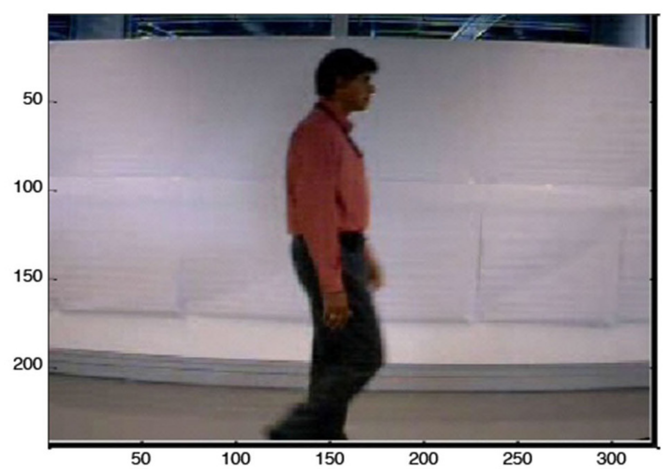

Figure 1. Sample Input Image

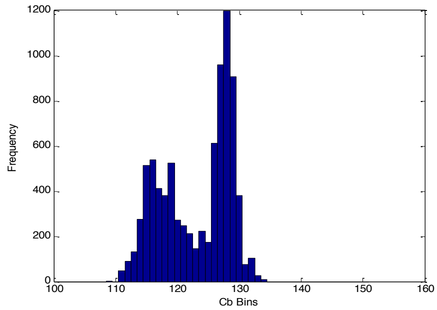

Figure 2. Cb Layer

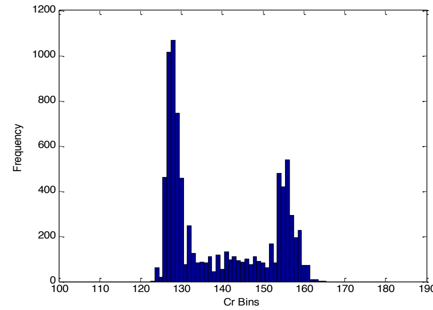

Figure 3. Cr Layer

Based on pixel distributions in Figure 2 and Figure 3 it can be seen that there are multiple peaks in the pixel distribution indicating different colour values and regions. Using the EM method to segment data into different normal distributions, as seen in Figure 4 and Figure 5, for layers Cb and Cr respectively, each detected normal distribution would have its own mean and standard deviation. Each detected distribution parameters would be used to segment the input image into different coloured blobs or regions. The result of the segmentation can be seen in Figure 6 where the centroids of each segmented regions are indicated as crosshairs. The red and green crosshairs are indicators of the centroids from regions detected on the Cb and Cr layers respectively.

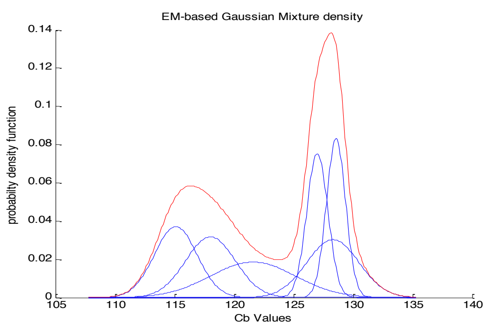

Figure 4. EM Gaussian Mixture on Cb Layer

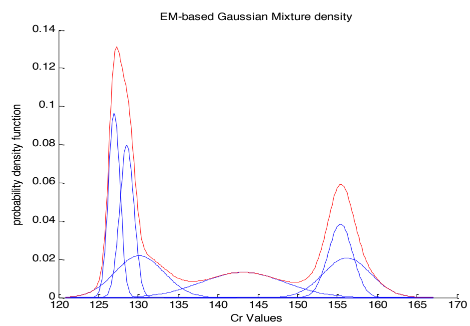

Figure 5. EM Gaussian Mixture on Cr Layer

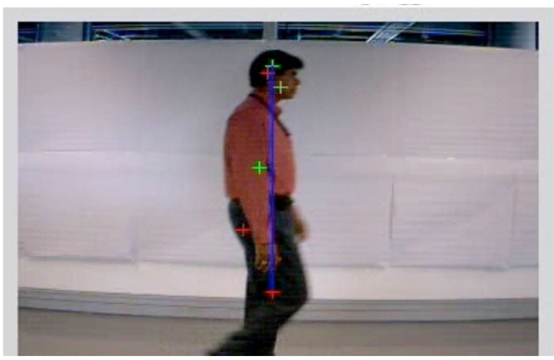

Figure 6. EM Sample Image Detected Centroids

The next step in tracking target of interest would be the extraction of parameters between colour regions of interest which in this case would be the top and bottom colour regions as reference parameters. The parameters to be used for tracking are the magnitude and angle of the vector between the colour regions. An additional parameter is the difference of the colour region mean values (Sebastian et al., 2012). Combination of data from the vector magnitude, vector angle and colour difference would generate a set of parameters that should be unique for each target of interest. An illustration of extracted data and clustering can be seen in Figures 7 and 8.

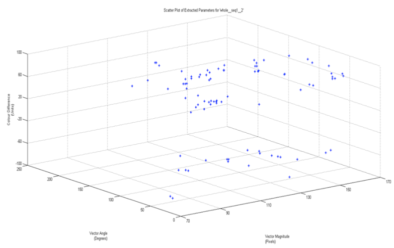

Figure 7. Scatter Plot of Extracted Data

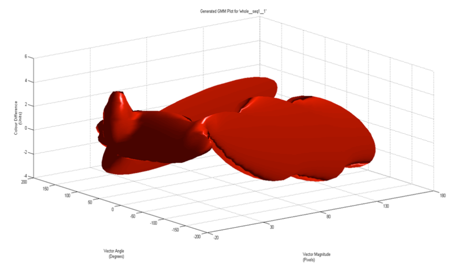

Figure 8. Data clusters of Extracted Data with Parameters

The tracking process is executed by comparing extracted or input parameters against reference parameters. A track is considered to be successful when the input parameters are statistically similar to the reference parameters. Table 1 shows the tracking performance of different tracking methodologies in the YCbCr colour space.

Table 1
Summarised Compared Tracking Methodology Tracking Data

Input Video	Histogram Similarity Comparison						
	Normalised Cross Correlation	Histogram Intersection	Euclidean Intersection	Chi-Squared 1	Chi-Squared 2	Kullbeck-Leibler	Parameteric Tracking
whole_seq1_1	0.00	0.29	0.19	0.31	0.25	0.27	0.92
whole_seq1_2	0.00	0.24	0.00	0.29	0.29	0.29	0.90
whole_seq1_3	0.00	0.20	0.00	0.15	0.16	0.12	0.95
whole_seq1_4	0.00	0.15	0.20	0.28	0.33	0.35	0.90
whole_seq1_5	0.00	0.20	0.00	0.19	0.28	0.19	0.94

Table 1 compares the tracking performances between different tracking methods. The same set of videos were used in determining the tracking performance for different methodologies investigated. Tracking performances from different tracking methods starting from normalised cross correlation and histogram comparison method had led to the development of parametric

Figure 9. Uneven Illumination

Figure 10. Uneven background

Figure 11. Noisy Environment

tracking method which were done on a single camera. The scores are determined by obtaining the ratio of the detection count of the target against the total count of the target appearance in the camera view field.

Multiple camera tracking setup was done with 3 web cameras with non-overlapping view fields. The overall camera setup was done in an indoor environment with different backgrounds. The cameras had a resolution of 320x240 pixels. Sample images of different background view fields captured by the different cameras are seen in the Figures 9, 10 and 11 where each figure illustrates the different background setup in each camera view field. The different background setup was used to evaluate the utilisation of same set of parameters between blobs as a method for tracking a target across the three cameras.

RESULTS

In measuring tracking capability or performance, a tracking metric known as track detection rate (TDR) (Ellis, 2002) was used. The tracking performance is determined by statistically comparing the input image parameters against the compiled parameters of the target of interest. A sample of raw tracking results can be seen in Table 2 which compiles the track detection rate in a confusion matrix. In this instance, the confusion matrix lists the tracking results of each input video against the tracking parameter of each target of interest

Table 2
Sample TDR of Camera 1

	Reference								
Input Video	whole _seq2_1	whole _seq2_2	whole _seq2_3	whole _seq2_4	whole _seq2_5	whole _seq2_6	whole _seq2_7	whol e_seq2_8	whole _seq2_9
whole_seq2_1	0.70	0.44	0.36	0.17	0.60	0.06	0.73	0.33	0.20
whole_seq2_2	0.87	0.97	0.87	0.38	0.69	0.11	0.90	0.84	0.34
whole_seq2_3	0.95	0.87	0.96	0.68	0.83	0.60	0.94	0.86	0.57
whole_seq2_4	1.00	1.00	1.00	0.92	0.91	0.29	0.96	0.96	0.53
whole_seq2_5	0.17	0.13	0.13	0.03	0.26	0.00	0.23	0.08	0.06
whole_seq2_6	0.78	0.92	0.77	0.56	0.62	0.55	0.83	0.84	0.23
whole_seq2_7	0.65	0.61	0.57	0.23	0.71	0.04	0.80	0.42	0.10
whole_seq2_8	0.89	0.97	0.92	0.44	0.73	0.21	0.92	0.97	0.48
whole_seq2_9	0.97	1.00	0.99	0.64	0.89	0.35	0.88	0.97	0.96

Comparing the input video against its own reference parameter in Table 2 should show the highest tracking result if the tracking parameters were the correct set of parameters which are highlighted. Compilation of the raw TDR data tracking performance can be seen in Table 3. The values that are recorded in the tables should range from 0 to 1.0 which indicates a correct tracking performance of 0% to 100% correct tracking. Table 3 shows the track results of targets that have successful track across three cameras. Empty cells in Table 3 indicate that correct tracking was not achieved across 3 cameras. The cells with tracking data rates indicate that the target was correctly tracked and had the highest tracking rate. Table 3 displays the number of input videos used for determining tracking performance, number of EM segmentations which indicate the number of colour regions that each tracked target is segmented into and the classification of data that is based on raw data tracking or outlier data removed tracking. Nine videos were used. The number of EM segments ranged from 3 to 6. Tracking was also done based on raw data and outlier removed data. Raw data indicates that all track data points are used in determining track performance whereas the outlier removed data set removes the outlier data track points from the calculation of track performance.

Table 3
Compiled TDR across 3 cameras

Ref Video	Number of Segmentation and Cameras											
	out6			out5			out4			out3		
Input Video	Cam1	Cam2	Cam3	Cam1	Cam2	Cam3	Cam1	Cam2	Cam3	Cam1	Cam2	Cam3
whole_seq2_1				0.76	0.7	0.82						
whole_seq2_2												
whole_seq2_3				0.89	0.61	0.57	0.94	0.62	0.62	0.64	0.61	0.37
whole_seq2_4	0.87	0.66	0.68	0.92	0.44	0.54				0.19	0.14	0.15
whole_seq2_5												
whole_seq2_6				0.71	0.59	0.26						
whole_seq2_7												
whole_seq2_8												
whole_seq2_9	0.88	0.67	0.55				0.84	0.77	0.57			
Ref Video	Number of Segmentation and Cameras											
	raw6			raw5			raw4			raw3		
	Cam1	Cam2	Cam3	Cam1	Cam2	Cam3	Cam1	Cam2	Cam3	Cam1	Cam2	Cam3
whole_seq2_1				0.76	0.7	0.78						
whole_seq2_2												
whole_seq2_3				0.92	0.69	0.59						
whole_seq2_4				0.93	0.58	0.57	0.84	0.16	0.23	0.20	0.17	0.10
whole_seq2_5												
whole_seq2_6												
whole_seq2_7				0.83	0.82	0.76						
whole_seq2_8												
whole_seq2_9	0.96	0.74	0.60				0.88	0.83	0.63	0.68	0.42	0.40

Table 3 shows random input videos labelled 'whole_seq2_1', 'whole_seq2_3', 'whole_seq2_4', whole_seq2_6', 'whole_seq2_7' and 'whole_seq2_9' had parameters that enabled the specific targets to be tracked across 3 cameras. Where the tracking parameters used in this instance are the combination vector magnitude, vector angle and colour difference that make each target have their own unique combination for tracking across the 3 cameras. The video sequences of 'whole_seq2_2', 'whole_seq2_5' and 'whole_seq2_8' did not generate the necessary tracking parameters for successful tracking across 3 cameras. Video sequence 'whole_seq2_5' could not be segmented into different colour regions as that particular target had only one colour for the whole target and thus this particular target could not generate the necessary parameters between colour regions for tracking. Video sequence 'whole_seq2_2' and 'whole_seq2_8' had the colour regions segmented and detected only on one colour layer. The utilization of parameters between colour regions detected on same colour layer did not give accurate tracking of targets (Sebastian et al., 2012). Apart from the video sequences that did not have any tracking results across three cameras, the other video sequences have tracking results that indicated that the targets are tracked across the three cameras. The tracking performance across three cameras are dependent on the number of colour region that the target of interest is segmented into.

CONCLUSION

This paper contributes in the area of target tracking in a video surveillance system where tracking the target of interest is based on the parameters between colour regions. The parameters used in tracking the target of interest were vector magnitude, vector angle and colour difference between colour regions. The tracking performance of the parametric tracking was initially compared against other tracking methods such as normalised cross correlation and histogram comparison methods. The results had shown that the parametric tracking had better tracking performance. In the initial evaluation, the targets of interest were tracked on a single camera and the results indicated that parametric track had significantly better tracking performance compared with other tracking methods. The parametric tracking methodology was then extended to tracking a target across an array of cameras. The tracking performance across an array of cameras had shown that the target of interest could be tracked based on the parameters between colour regions. The tracking performance had indicated that the consistent nature of the target colour was one of the primary reasons that the target of interest was able to be tracked across an array of cameras. The results obtained from this paper is an extension of a previous paper that determined that tracking a target was possible by utilising the parameters between the top and bottom colour regions of a target.

REFERENCES

Ali, M. A., Indupalli, S., & Boufama, B. (2006, June). *Tracking multiple people for video surveillance*. In First International Workshop on Video Processing for Security, Quebec City, Canada.

Atsushi, N., Hirokazu, K., Shinsaku, H., & Seiji, I. (2002). Tracking multiple people using distributed vision systems. In *Robotics and Automation, 2002. Proceedings. ICRA'02. IEEE International Conference on* (Vol. 3, pp. 2974-2981). IEEE.

Bredereck, M., Jiang, X., Körner, M., & Denzler, J. (2012, October). Data association for multi-object tracking-by-detection in multi-camera networks. In *Distributed Smart Cameras (ICDSC), 2012 Sixth International Conference on* (pp. 1-6). IEEE.

Chippendale, P. (2006). Towards Automatic Body Language Annotation. In *7th International Conference on Automatic Face and Gesture Recognition 2006 (FGR 2006)*.

Ellis, T. (2002). *Performance Metrics and Methods for Tracking in Surveillance*. Third IEEE International Workshop on Performance Evaluation of Tracking and Surveillance, Copenhagen, Denmark.

Gonzalez, W., & Woods, R. E. (2004). *Eddins, Digital Image Processing Using MATLAB*. Third New Jersey: Prentice Hall.

Hommel, S., Grimm, M. A., Voges, V., Handmann, U., & Weigmann, U. (2012, November). An intelligent system architecture for multi-camera human tracking at airports. In *Computational Intelligence and Informatics (CINTI), 2012 IEEE 13th International Symposium on* (pp. 175-180). IEEE.

Park, S. and J. K. Aggarwal (2002). S*egmentation and Tracking of interacting human body parts under occlusion and shadowing*. Workshop on Motion and Video Computing 2002.

Sebastian, P., Voon, Y. V., & Comley, R. (2010). Colour space effect on tracking in video surveillance. *International Journal on Electrical Engineering and Informatics, 2*(4), 298-312.

Sebastian, P., Voon, Y. V., & Comley, R. (2012, December). Parametric tracking of multiple segmented regions. In *Computer Applications and Industrial Electronics (ISCAIE), 2012 IEEE Symposium on* (pp. 185-189). IEEE.

Siebel, N. T. and S. Maybank (2002). Fusion of Multiple Tracking Algorithms for Robust People Tracking. In *Proceedings of the 7th European Conference on Computer Vision (ECCV 2002)*.

Trucco, E. and K. Plakas (2006). Video Tracking: A Concise Survey. *IEEE Journal of Oceanic Engineering, 3*1(2), 520-529.

Wang, H., Liu, C., Zhang, X., & Li, Q. (2009, October). Multi-camera tracking based on information fusion in video surveillance. In *Image and Signal Processing, 2009. CISP'09. 2nd International Congress on* (pp. 1-5). IEEE.

Wang, H., Sang, N., & Yan, Y. (2014, August). Real-time tracking combined with object segmentation. In *Pattern Recognition (ICPR), 2014 22nd International Conference on* (pp. 4098-4103). IEEE.

Yan, W., & Forsyth, D. A. (2005, January). Learning the behavior of users in a public space through video tracking. In *Application of Computer Vision, 2005. WACV/MOTIONS'05 Volume 1. Seventh IEEE Workshops on* (Vol. 1, pp. 370-377). IEEE.

Qian, Y., & Guirong, W. (2014, August). Lung nodule segmentation using EM algorithm. In *Intelligent Human-Machine Systems and Cybernetics (IHMSC), 2014 Sixth International Conference on* (Vol. 1, pp. 20-23). IEEE.

Zhao, Y., Zhang, B., & Zhang, X. (2014, July). Meanshift blob tracking with target model adaptive update. In *Control Conference (CCC), 2014 33rd Chinese* (pp. 4831-4835). IEEE.

Zhu, Y. F. (2009, December). Background subtraction and color clustering based moving objects detection. In *2009 International Conference on Information Engineering and Computer Science* (pp. 1-5). IEEE.