Subsequent Processing of Background Modeling for Traffic Surveillance System

Duong Nguyen-Ngoc Tran, Tien Phuoc Nguyen, Tai Nhu Do, and Synh Viet-Uyen Ha

Abstract—**The background model finds difficult to extract the good foreground if it is implemented in real-time, the change of light and weather that make the detected-foreground unclear, has countless noise. Therefore, to avoid that problem, the Mixture of Gaussian need subsequent processing step to make the foreground better. The paper proposes a method, degrades several drawbacks of Background Subtraction, which intended to be used in urban traffic scenes.**

Index Terms—**Background subtraction, mixture of gaussian, contour, morphology, quantitative performance, surveillance system.**

I. Introduction

The traffic monitoring system is used to collect, process, analysis and disseminate the crowded traffic. It has couple of parts which have the relationship with each other. The better previous part works, the better the following part does. The paper shows the way to improve background subtraction.

To consider the challenges in video surveillance [1] and monitoring, a robust background subtraction method should be capable of dealing with critical situations: noise image, camera jitter, camera automatic adjustment, placement of camera, light switch, time of the day, bootstrapping, shadows, slow-moving objects [2]. Traditional approaches based on backgrounds methods typically fail in these general situations.

Much work has been done since the introduction of the Mixture of Gaussian (MOG) model [3] by Stauffer and Grimson [4]. In their approach, the mixture of K Gaussian representing the statistics of one pixel over the time can cope with multi-modal background distribution. This model can lessen the effect of small repetitive motions; for example, lightly swaying trees and bushes or small displacement of camera. KaewTraKulPong and Bowden [5] as well as Zoran Zivkovic [6] presented methods which improve the computation performance by reinvestigating the update equations. Fuzzy is also implemented to the background subtraction, each step must be designed and the features in relation has to be chosen in relation to the critical situation [7], [8]. Moreover, the ViBe [9] is a technique for motion detection that incorporates several innovative mechanisms. To conclude, Surveys and comparisons of different algorithms for background subtraction can be found in the literature [10]-[12].

At the present time, the method or algorithm which are complicated need more time to compute, the better result the more spend in time. For that reason, the development of hardware that help to speed up the complex model. For example, CUDA (NVidia's parallel computing architecture) or OpenGL, which was implement to support from specific GPU hardware [13].

In the paper, the Mixture of Gaussian is consider to for the reason that it keep the best quality in traffic surveillance system.

The rest of this paper is organized as follows: In the next section, we overview the statistic method of background subtraction, which is Mixture of Gaussian. The proposed improvement is showed on the following section which has the illustration and algorithm. In addition, the algorithm is shown in step by step to get the best result. In conclusion, the quantitative evolution is shown in last section, which illustrates a qualitative comparison of the improvement with the other method in background subtraction field.

II. Related Work

By the idea of Stauffer and Grimson [1], Mixture of Gaussian (MOG) algorithm aims to formulate and maintain a statistical representation of each pixel in the scene. The feature of given pixel at time is represented by its intensity in RGB color and observed under a PDF:

$$P(X_t) = \sum_{i=1}^{K(X_t)} \omega_{i,t}\eta(X_t, \mu_{i,t}, \Sigma_{i,t}) \tag{1}$$

K_{X_t} is the number of distributions in each mixture. To make it more flexible we do not set K_{X_t} as constant; it can vary, and the maximum value is 5, depending on computation performance of resource. $\omega_{i,t}$ is the weight associated to the i^{th} component at time *t*, $\mu_{i,t}$ is mean, and $\sum_{i,t}$ denotes as standard deviation. The Gaussian density function η is:

$$\eta(X_t, \mu, \Sigma) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} e^{-\frac{1}{2}(X_t-\mu)^T\Sigma^{-1}(X_t-\mu)} \tag{2}$$

In order to find the matching distribution, they used a statistical measure Mahalanobis distance:

$$\sqrt{\left(X_{t+1}-\mu_{i,t}\right)^T . \Sigma_{i,t}^{-1} . \left(X_{t+1}-\mu_{i,t}\right)} < k.\sigma_{i,t} \tag{3}$$

where k is a constant threshold equal to 2.5 [1]. Zivkovic [2] applied moving average filter to update the weight — $\omega_{i,t}$, $\mu_{i,t}$, and $\sigma_{i,t}$ for matching distribution:

Manuscript received October 21, 2014; revised March 8, 2015.

The authors are with School of Computer Science Engineering, International University, Ho Chi Minh City, Vietnam (e-mail: hvusynh@hcmiu.edu.vn).

DOI: 10.7763/IJCTE.2016.V8.1050

$$\omega_{i,t+1} = (1-\alpha)\omega_{i,t} + \alpha \quad (4)$$

$$\mu_{i,t+1} = (1-\alpha)\mu_{i,t} + \alpha.X_{t+1} \quad (5)$$

$$\sigma_{i,t}^2 = (1-\alpha)\sigma_{i,t}^2 + \alpha\left(X_{t+1} - \mu_{i,t+1}\right).\left(X_{t+1} - \mu_{i,t+1}\right)^T \quad (6)$$

To simplify the updating equations, we assume α as the same learning rate for mean, weight and variation. For other unmatched distributions, only weight is updated by:

$$\omega_{j,t+1} = (1-\alpha)\omega_{j,t} \quad (7)$$

The new distribution is added to the mixture or replaced to the lowest weight distribution whenever the mixture is full. The highest weight distribution is selected as background.

III. The Subsequent Processing

At the moment, the surveillance is widespread used in the street, and the most them are set up at the same heights and angles. To be clearly, cameras are installed on the 4-5 meter traffic lights and the angel is 15-20 degree with the horizontal. The observation zone do not change significantly in several consecutive frames. Therefore, the number of vehicles in one frame, conclude truck, car and motorcycle, is not numerous in any condition. On the other hand, the size of vehicles is increase from the top down to bottom, that lead to the size of the pattern further small enough to be determined the noise or not important to detected.

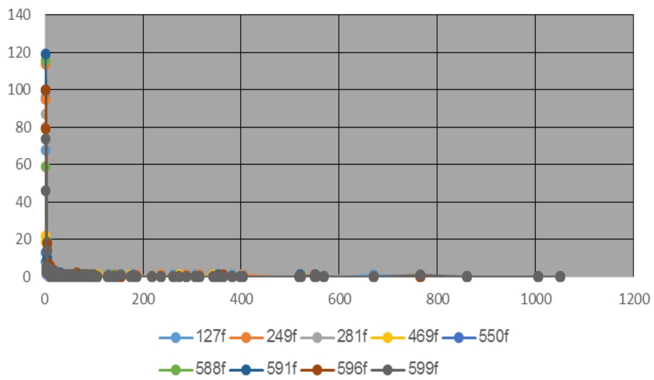


Fig. 1. The collection of sizes of contours.

TABLE I: The Gaussian Value of Foregrounds

	Mean	Variance	Standard deviation	Threshold
127f	14,37	3909,24	62,52	76,89
249f	11,13	1663,53	40,79	51,92
281f	12,72	3581,85	59,85	72,57
469f	20,20	2677,43	51,74	71,95
550f	18,54	1208,32	34,76	53,30
588f	11,49	4100,56	64,04	75,53
591f	11,01	4701,56	68,57	79,57
596f	13,20	5994,52	77,42	90,62
599f	15,13	4958,63	70,42	85,55
		Mean		73,10

The method that accurately defines the smallest contour size which is count for the vehicle, which guide to our method. Let x denote the size of contour gotten in foreground frame, and $f(x)$ be the number of each size, x is horizontal axis and $f(x)$ is vertical axis in Fig. 1. We run the Mixture of Gaussian on some video sand get the typical foregrounds on them. The result of counting contour is showed in the Fig. 1. Afterwards, we can use the value to calculate the mean and standard deviation then the standard distribution of the normal (Gaussian) distribution. To get the smallest size of the contour which is measured the pattern (vehicle) in the frame, we can plus the mean with the standard deviation to be the threshold. All of calculation is presented in Table I.

The Fig. 1 represent the number of size contour and the quantity of them in foreground. The Table I shows all the values from which mean and the standard normal distribution is deliberate, the totality of mean and standard deviation is called threshold. As can be seen on the Fig. 1, the most size of contour is between 1 and 50. In addition, on the table, the mean is from 10-20. To conclude, the number 73 is the threshold for that the size of noise or inessential pattern for detecting are smaller. The result to find the number is that we make the foreground clearer just rely on the few number of weighty pattern.

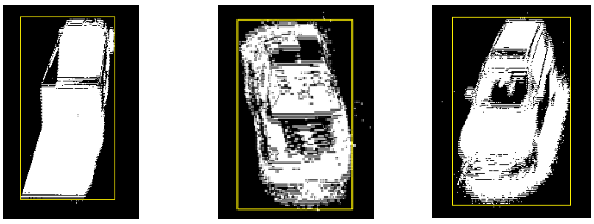


Fig. 2. The vehicle is bounded in rectangle contour.

The Fig. 2 shows the vehicle which is bounded in rectangle contour, the first is the back of the truck, the second is the back of the car and the last is the front of the car. As can be seen, the vehicles detected in foreground are not clear, there are a lot of holes in there.

After collecting the suitable vehicle, which are not clear and still have holes. After Background method, in normal way, erode and dilate are use sequentially to discard the noise and fill the hole. However, the size of hole is fill depend on the parameter of erode and dilate, that lead to the shape of each pattern is not the same as original. We define white is the foreground color and black is the background color.

To solve the problem, we find all the contour in the given foreground, for each the contour, only sizes more than 73 are kept and others is erase by draw it with background color. As the result, only few contour are stored.

For each stored-contour, we consider each pixel by spread out 8 direction from them. The condition that the pixel is recognized the foreground pixel is either has the foreground color in advanced or 5 of 8 direction meets the white pixel. Thus, we draw all foreground pixel with white color.

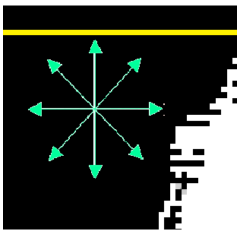


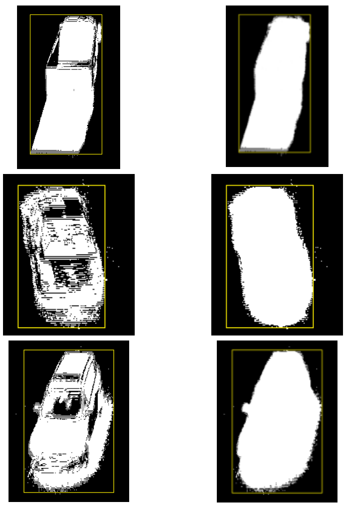
Fig. 3. The direction when we spread out.

The Fig. 3 illustrates the way to spread out from the current pixel. The yellow line is the part of the rectangle contour, the white pixel is the available foreground pixel and the green arrows are the direction from the pixel at the middle. As can be seen, only 2 routes could meet the foreground pixel, 3 directions reach the contour and an others cannot meet the white pixels. Therefore, based on the algorithm, that pixel just have value in count equal 2, less than 5, and still keep it a background pixel with black color.

TABLE II: DRAW THE PATTERN ALGORITHM

Algorithm: Draw the partern

```
for each contour C
  if size of C more than 73
    draw all pixel with black color
  else
    for each black pixel P in contour C
      count = 0
      for 1 to 8 direction from pixel P
        if direction reach the White point
          count = count + 1
      if count >= 5
        change into white color
      else
        keep the black color
```



Source Result

Fig. 4. The result after drawing.

The Fig. 4 shows the results after subsequence processing, the source is consider in Fig. 2, the vehicles are clear and have not any holes.

IV. RESULTS AND ANALYSIS

A. Comparative Experiments

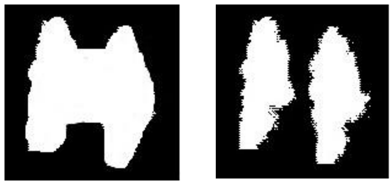
The section shows the comparison of the improvement with some currently method. The chosen algorithms for the simulation are: Fuzzy Gaussian [14] (FG), Advanced MOG [6] (MOG), Gaussian Mixture Model [5] (GMM), Multi-Layer Background subtraction [15] (ML). The parameters of each algorithm are setup according to author's propositions.

Datasets such as CDnet [16] and BMC 2012 [17] are often used in recent publications. However, in order to evaluate the performance in traffic context of crowed urban in Viet Nam [18], all experiments will demonstrate with the dataset collected by IU DIP team.

The Fig. 7 shows the result after run some of current methods and our improvement. In the first input, we can see the Fuzzy Gaussian has problem and needs time for training to get the normal background, others run well but some of them get the vehicles with the hole on it, which are the same with the input 2, 3, and 4. Specially, the Gaussian Mixture Model has only bounded line on the truck in the fourth experiment. Moreover, except out improvement, others tend to keep the small pattern which is not need to be detected in the surveillance system. In our improvement, we just keep the big one, which can be detected what kind of lorry, .e.g., motorbike, car, truck…

B. Comparison with Morphology

There is already method that can do the same thing with our method, it is Morphology, but it has some disadvantages. The Morphology depends on the size and the shape of the structure element, and in some case, it works worse and makes some patterns become one.



Morphology Improvement

Fig. 5. Comparative results of methods with respect to ground-truth segmentation.

In Fig. 5 shows 2 motorbike which run nearly, but in the morphology, they become one pattern, therefore that can make the traffic system detect them become the 1 car. In our improvement, two motorbikes still keep on the foreground.

C. Quantitative Performance

The method to find the accuracy of performance in study is conducted. In this section, we concern the evaluations of object detectors, which compare the performance under particular method which is mentioned in [19] such as Precision (PR), Versus Recall (RC), False Negative Rate (FNR), False Positive Rate (FPR), the accuracy (AC) is the ratio of correct decisions compared to all decisions and PWC is the wrong classifications.

Another measure for fitness quantification in the context of background subtraction is F-measure (Fm) [20]-[22] which combines precision and recall to provide more representative than PR and RC themselves.

If foreground elements are present in small part of image, there is not much difference in the achieved high ratings of this coefficient with respect to the case of simple classifying everything as background. Therefore, the Jaccard coefficient [22] is used to solve the problem.

Table III compares the results of quantitative test in the dataset in order to evaluate the performance of each algorithms adequately. FP-rate, AC and Fm are important and described in Fig. 8. The lower the PWC the better; AC and Fm otherwise. As can be seen in the bar chart, the gap of PWC between algorithms is not satisfactory, but AC and Fm from new method are higher than others. These evaluations consolidate that the proposed method gives better result over other algorithms as comparative test. In addition, the J is good for our improvement.

D. Speed Performance

In this section, the chosen-algorithms were implemented in the same hardware system with 3.4GHz AMD Quad core CPU (equivalent of I5-CPU) and DDR3 4GB. The library we used is OpenCV [23] [24] with dataset of 800×480 pixels. The number is computed by the average number of millisecond per frame.

In Fig. 6, the table illustrated that our improvement is slower lightly then MOG method, as 92% speed of MOG, which show that the subsequence process take it small time for running.

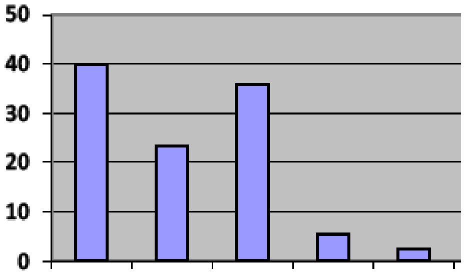


Fig. 6. Speed performance (frame/s) on the dataset recorded.

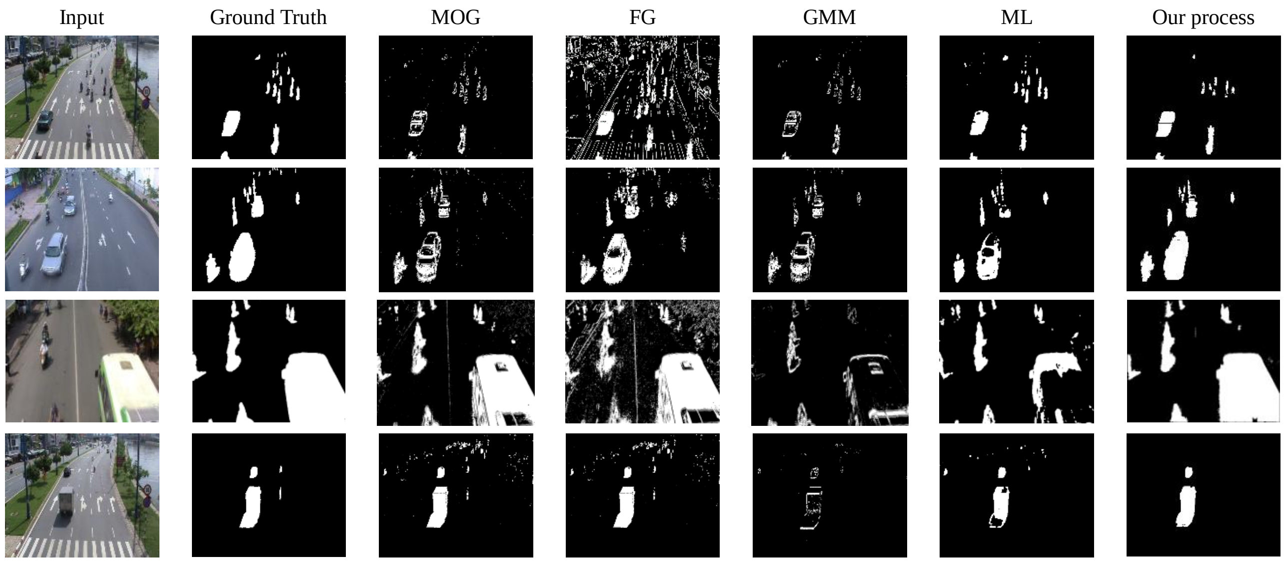


Fig. 7. Comparative results of methods with respect to ground-truth segmentation.

TABLE III: QUANTITATIVE PERFORMANCE OF THE FOREGROUND SEGMENTATION

	TP	FP	TN	FN	PR	RC	FNR	FPR	AC	PWC	Fm	J
MOG	52317	18877	180054	55952	0,73	0,48	0,52	0,09	0,76	0,24	0,58	0,41
FG	52300	18894	180054	55952	0,73	0,48	0,52	0,09	0,76	0,24	0,58	0,41
GMM	43000	21194	187054	55952	0,67	0,43	0,57	0,10	0,75	0,25	0,53	0,36
ML	49329	25865	151911	80095	0,66	0,38	0,62	0,15	0,66	0,34	0,48	0,32
New	47400	24824	227600	7325	0,66	0,87	0,13	0,10	0,90	0,10	0,75	0,60

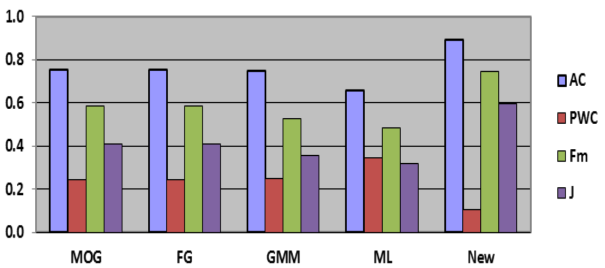


Fig. 8. The quantitative performance of the foreground segmentation on AC, PWC, FM, J.

V. CONCLUSION

In the traffic monitoring system, background subtraction is the important part which makes the following part run well. The paper gives the information about the improvement of background subtraction. After considering the challenges in video surveillance and monitoring, a robust background subtraction method should be capable of dealing with critical situations: Noise image, Camera jitter, Camera automatic adjustment, Placement of camera, Light switch, Time of the day, Bootstrapping, Shadows, slow-moving objects. Traditional approaches based on background methods typically fail in these general situations, there are some challenges which could be deal by improvement.

In the improvement, the foreground is handling. After processing the foreground, the noise and not valuable pattern are removed and keep the worth vehicle.

In conclusion, at the current time, the surveillance traffic system runs the current method which is used in the western country and run as well. However, Vietnamese traffic has some particular features which make them run worse. The improvement, mentioned in the paper, proposed the method can run in real time and get the good result.

REFERENCES

[1] G. H. S. Brutzer and B. Hoferlin, "Evaluation of background subtraction techniques for video surveillance," in *Proc. CVPR*, 2011, vol. 4, pp. 1937–1944.

[2] H. T. J. L. J. Yang and Y. Tan, "Accurate dynamic scene model for moving object detection," in *Proc. Int. Conf. on Image Processing*, 2007.

[3] S. P. N. Atev and O. Masoud, "Practical mixtures of gaussians with brightness monitoring," in *Proc. IEEE Conf. on Intt. Transportation Systems,* 2004.

[4] C. Stauffer and W. Grimson, "Adaptive background mixture models for real-time tracking," in *Proc. CVPR*, 1999, pp. 431–447.

[5] P. K. T. K. Pong and R. Bowden, "An improved adaptive background mixture model for realtime tracking with shadow detection," Tech. Rep. AVBS01, Department of Systems Engineering, Brunel University, Middlesex, UB8 3PH, UK, 2001.

[6] Z. Zivkovic, "Improved adaptive gaussian mixture model for background subtraction," in *Proc. ICPR*, 2004.

[7] F. E. Baf, T. Bouwmans, and B. Vachon, "Type-2 fuzzy mixture of Gaussians model: Application to background modeling," in *Proc. International Symposium on Visual Computing*, Las Vegas, USA, December 2008, pp. 772-781.

[8] F. E. Baf, T. Bouwmans, and B. Vachon, "Foreground detection using the choquet integral," in *Proc. International Workshop on Image Analysis for Multimedia Interactive Services*, Klagenfurt, Austria, May 2008, pp. 187-190.

[9] O. Barnich and M. V. Droogenbroeck, "ViBe: A universal background subtraction algorithm for video sequences," *IEEE Transactions on Image Processing*, vol. 20, issue 6.

[10] M. Piccardi, "Background subtraction techniques: A review," in *Proc. IEEE Int. Conf. Syst.*, Man, Cybern, 2004, vol. 4, pp. 3099–3104.

[11] T. H. Chalidabhongse, K. Kim, D. Harwood, and L. Davis, "A perturbation method for evaluating background subtraction algorithms," in *Proc. Joint IEEE Int. Workshop Vis. Surveillance Perform. Eval. Tracking Surveillance*, Nice, France, 2003, pp. 1–7.

[12] S.-C. Cheung and C. Kamath, "Robust techniques for background subtraction in urban traffic video," in *Proc. SPIE Electron. Imaging VideoCommun. Image Process*, San Jose, CA, Jan. 2004, pp. 1–12.

[13] V. Pham, P. Vo, H. V. Thanh, and B. L. Hoai, "GPU implementation of extended gaussian mixture model for background subtraction," in *Proc. IEEE International Conference on Computing and Telecommunication Technologies*, Vietnam National University, November 2010.

[14] H. S. Mohamad, M. Naser, and R. P. Hamid, "Fuzzy running average and fuzzy background subtraction: concepts and application," *International Journal of Computer Science and Network Security*, vol. 8, no. 2, February 2008.

[15] J. Yao and J. M. Odobez, "Multi-layer background subtraction based on color and texture," in *Proc. IEEE the CVPR Visual Surveillance Workshop (CVPR-VS)*, Minneapolis, June 2007.

[16] 1st IEEE Change Detection Workshop in Conjunction with CVPR. (2014). [Online]. Available: http://www.changedetection.net

[17] BMC 2012 background models challenge dataset. [Online]. Available: http: //www.bmc.univ-bpclermont.fr

[18] N. Tang, C. Do, T. B. Dinh, and T. B. Dinh, "Urban traffic monitoring system," *Lecture Notes in Computer Science*, vol. 6839, pp. 573-580, 2011.

[19] P. L. Rosin and E. Ioannidis, "Evaluation of global image thresholding for change detection," *Pattern Recognition*, vol. 24, no. 14, pp. 2345–2356, Oct. 2003.

[20] B. White and M. Shah, "Automatically tuning background subtraction parameters using particle swarm optimization," in *Proc. IEEE ICME*, Beijing, China, 2007, pp. 1826–1829.

[21] A. Ilyas, M. Scuturici, and S. Miguet, "Real time foreground-background segmentation using a modified Codebook model," in *Proc. 6th IEEE Int. Conf. AVSS*, 2009, pp. 454–459.

[22] P. Sneath and R. Sokal, *Numerical Taxonomy — The Principle and Practice of Numerical Classification*, San Francisco, CA: Freeman, 1973.

[23] Open Computer Vision Library (OpenCV 2.4.9). [Online]. Available: http://www.sourceforge.net/projects/opencvlibrary/

[24] G. Bradski and A. Kaehler, *Learning Open CV: Computer Vision With the Open CV Library, Wheeling*, IL: O'Reilly, 2008.

Duong Nguyen–Ngoc Tran is with the Vision Research Group at International University of Ho Chi Minh National University, Ho Chi Minh, Vietnam, as a student in the School of Computer Science Engineering under the supervision of Dr. Synh Viet-Uyen Ha. His research interests include computer machines, computer vision, and image processing.

Tien Phuoc Nguyen is with the Vision Research Group at International University of Ho Chi Minh National University, Ho Chi Minh, Vietnam, as a student in the School of Computer Science Engineering under the supervision of Dr. Synh Viet-Uyen Ha. His research interests include computer machines, computer vision, and image processing.

Tai Nhu Do received the bachelor degree in information technology from Ho Chi Minh City University of Foreign Language–Information Technology, Ho Chi Minh City, Vietnam, in 2005.

In 2005, he joined the Faculty of Information Technology, Ho Chi Minh City University of Foreign Languages and Information Technology, Ho Chi Minh City, Vietnam, where he is currently a lecturer. His research interests include computer vision, image processing, and parallel programming.

Synh Viet-Uyen Ha received the engineer degree from Ho Chi Minh City University of Technology (HCMUT), Vietnam in 1996; and a M.Sc. degree from Ho Chi Minh University of Natural Science (HCMUNS), Vietnam in 1999, both in computer science and engineering, and the Ph.D. degree from School of Information and Communication Engineering, Sungkyunkwan University, Suwon, Korea in 2010. In 2011, he joined the School of Computer Science and Engineering, International University-Vietnam National University, as a lecturer. His research interests include computer machine, computer vision and image processing.