ELSEVIER

CrossMark

Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 59 (2015) 445 – 453

Procedia Computer Science

International Conference on Computer Science and Computational Intelligence (ICCSCI 2015)

Evaluation of Recursive Background Subtraction Algorithms for Real-Time Passenger Counting at Bus Rapid Transit System

Jonathan Samuel Lumentut[a*], Fergyanto E. Gunawan[b], Diana[a]

[a]*School of Computer Science, Bina Nusantara University, Jl. K. H. Syahdan 9, Jakarta 11480, Indonesia*

[b]*Binus Graduate Program, Bina Nusantara University, Jl. Kebon Jeruk Raya No. 27, Jakarta 11480, Indonesia*

Abstract

This work discusses a method to count the number of passengers waiting in Bus Rapid Transit station. The proposed system relies on computer vision technique to monitor the movement of passengers crossing doors on the station. In this work, three background subtraction techniques, namely, Running Gaussian Average, Gaussian Mixture Model, and Adaptive Gaussian Mixture Model, were used to count the passengers crossing an entrance on a BRT station from a pre-recorded motion picture. The results indicates that the tree algorithms are able to identify the passenger crossing with a reasonable high level of recall and but low level of precision. These results indicates that many false positives are identified by the three algorithms. In addition, the empirical data indicate that the three algorithms tend to have better performance with higher value of the learning rate.

Peer-review under responsibility of organizing committee of the International Conference on Computer Science and Computational Intelligence (ICCSCI 2015)

Keywords: Computer vision, Passenger counting system, Background subtraction, Bus rapid transit

* Corresponding author. Tel.: +62-21-5345830 (ext. 2188); fax: +62-21-5300244

E-mail address: jlumentut@binus.edu

Peer-review under responsibility of organizing committee of the International Conference on Computer Science and Computational Intelligence (ICCSCI 2015)

doi:10.1016/j.procs.2015.07.565

Nomenclature

t	Time unit
I_t	Pixel's intensity at time t
α	Size of temporal window to fit the Gaussian probability density function (learning rate)
μ_t	Mean value of the Gaussian probability density function
σ_t^2	Variance value of the Gaussian probability density function; recommended $\sigma_t^2 = 36$
k	Threshold weight value; recommended $k = 0.75$
X_t	Pixel history value at time t
K	A Constant for Gaussian distribution function; recommended $K = 3$
N	Gaussian probability density function
B	Estimated median
FG	Pixel belongs to foreground
BG	Pixel belongs to background
M	Component of each pixel
o	Ownership value
δ_m	$x_t - \pi_m$

1. Introduction

Traffic congestion is one of the important issues faced by many large cities in the globe including Jakarta, the capital of Republic of Indonesia. The problem is extremely complex and requires comprehensive solution involving collaboration of many sectors. Despite of this fact, Morichi[1] has advised a structure of the transportation strategy and city planning that suitable for megacities. At the heart of this proposal is a well-structured public transportation with various performance requirements.

According to Morichi[1], the mass transit system within the mega-city is essential to serve large volume of passengers. This system can be developed on the basis of the bus-based or train-based system. The bus-based system is often called as the bus rapid transit (BRT) and is required significantly lower cost and time to develop in comparison to the train-based system. However, the BRT system tends to have lower performance and higher variation in the travel time reliability[2,3], which is important to measure performance of a transportation system. Meanwhile, the train-based system has nearly zero variation in the travel time[4,5].

However, during the last thirty years, BRT system has received a large rate of adoption. The number of BRT and rail-based system that operate around the world described by Campo [6], see Fig. 1. The city of Jakarta also adopts this BRT system since 2004 and it is called TransJakarta BRT. Currently, TransJakarta has 13 corridors with a total busway length around 180 km. The first corridor was developed within three years from project initiation until its operation.

Similar to any other systems, BRT system can also be divided into two sub-systems: the supply side system and the demand side system. The trade-off between the two systems determines the level of service received by passengers. Maintaining high level of service quality is necessary in order to attract more passengers.

Many previous studies reported that the performance of TransJakarta BRT is relatively low. The main complaint is that passengers have to wait in uncertainty for long hours[7]; see Fig. 2. The other complaint is the limited number of available buses.

Monitoring systems for the supply-side and demand-side of the BRT system is imperative for various purposes such as for measuring the level of services, fleet management, and cost-effectiveness in the bus operation. The bus fleet monitoring can clearly be done using using floating-car data technique[8–11]. However, monitoring the number of passengers in a long and narrow BRT station is still a challenging issue. For this case, we had previously proposed a computer-vision-based approach[12]. However, the system performance in term of the passenger counting accuracy is relatively low and further study is of importance.

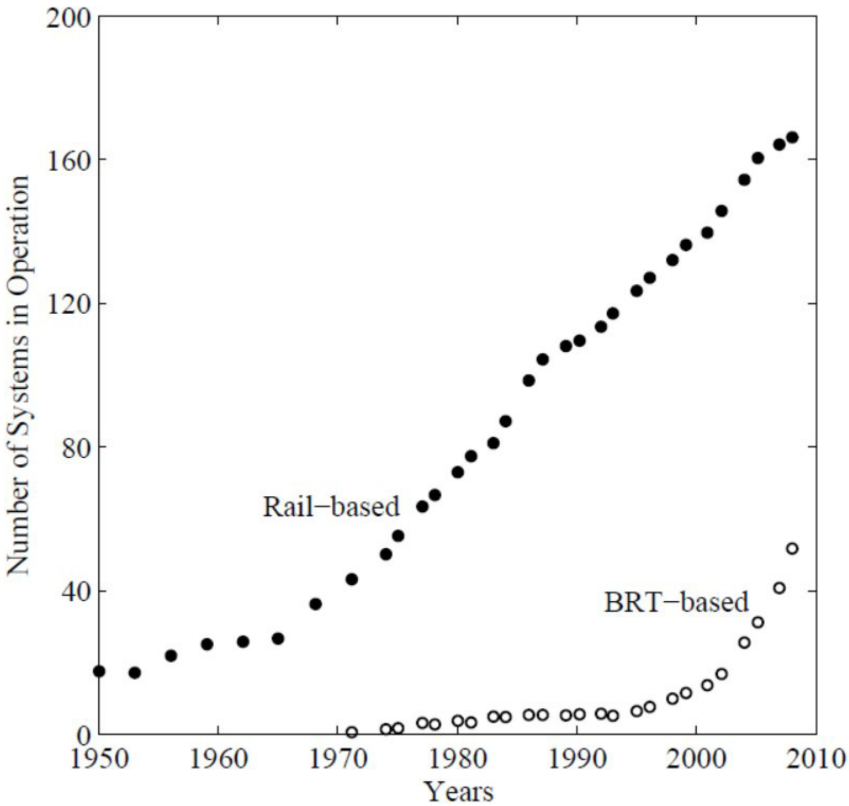

Fig. 1. The number of development of the train-based and BRT-based public transportation system across the globe[6]

Fig. 2. A long queue of TransJakarta BRT

2. Research Method

The passenger monitoring system discussed in this work is proposed in the context of the bus rapid transit system. In this system, each station is usually designed to serve the movement of buses in two directions and is placed in the road median. The station is usually long and narrow. On this basis of the station design, we proposed the passenger monitoring system as depicted in Fig. 3.

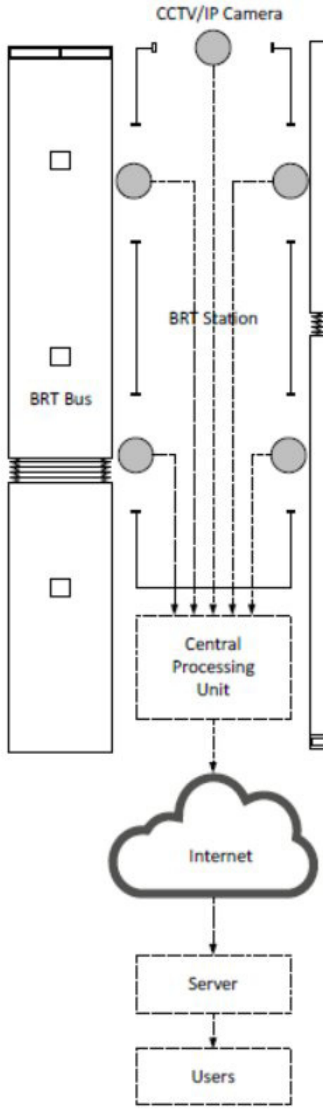

Fig. 3. The proposed passenger monitoring system for TransJakarta BRT stations[12]

The monitoring system works as the following. Each door will be monitored with an over-the-top camera. The camera shall record the movement of passengers crossing the door. The counting is done by implementing a computer vision technique. Information regarding the passenger movement across the all passages in the station should be aggregated in order to determine the remaining passenger in the station.

From the over-the-top camera, a station passage should look like that in Fig. 4. We then establish a red virtual-line to separate the passage into two region: Region A and Region B. Region A is located to the left on the line. The bus door will be in region A. In our experiment, we record the activity in this access point for 30 minutes duration. In practice, passengers often wait for the bus along the red line, resulted in the counting error. Our test will use the F1-score method to track the recall, defined by Eq. (1), and precision, defined by Eq. (2), of these recursive algorithms that implemented in real-time video of passenger movements on the access point of BRT station.

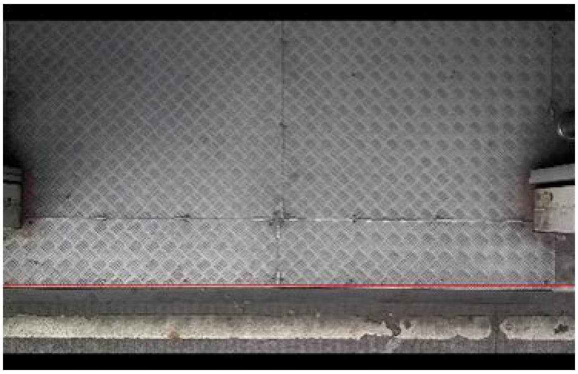

Fig. 4. An access point to a typical BRT station. The image recording system is located on the top. The red virtual line which is constructed to separate Region A, to the left of the line, and Region B, to the right of the line[12].

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \quad (1)$$

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \quad (2)$$

The false positive values are counted by the number of incorrect objects captured by the algorithm. The division between number of correct object captured by algorithm and total objects (correct and incorrect) captured by the algorithm creates the precision values. Meanwhile the recall values are counted by the division between number of correct object captured by the algorithm and total number of actual objects

3. Supporting Theories

According to Pavlidis, many passenger counting is collected manually using human counter (example: turnstile) that are expensive and can disrupt traffic [22]. In 2010, Yahiaoui with his research team conclude that there are two most reliable approaches to count passenger in BRT, which are: firstly,the use of infrared directional sensors and secondly, using video sensing and image processing. The infrared approach has several disadvantages in crowded situations. This approach is less reliable in crowded situations because of its high sensitivity to noise, variations in temperature, dust and smokes. Also, it cannot distinguish between one passenger and group of passengers. Therefore,they suggest that video-based system are very promising for this task [23]. Their suggestion is taken to develop our evaluation with computer vision techniques.

Many previous studies have been performed to develop methodology to track moving objects using computer vision. Kang and Kim [13] develops method to track the movement of many human objects in real-time utilizing the Conditional Density Propagation algorithm. Three improvement were made in his work: the use of an effective template for the human form using self-organizing map; the use of the hidden Markov model for modeling the dynamic of the human shape; and the use of a competition rule to separate a person from others. In addition,Wang et al.[14] study passenger detection using the characteristics of the head area. They intend to track the pedestrian movement. The tracking is achieved in two steps: applying the background difference algorithm, dynamic threshold algorithm, and the method of morphological processing to filter the image noise; and applying head matching algorithm using a mask template. Our previous research proposed a real-time passenger counting system using three algorithm: adaptive median filtering (AMF)[15]; pixel-based adaptive segmenter (PBAS)[16]; and background-subtraction by Godbehere-Matsukawa-Goldberg (GMG)[17].

Background Subtraction. The use of background subtraction algorithms are the common way to track moving objects. The movement of passengers in and out at the BRT station can be tracked using this algorithm. Since our previous study resulted that AMF is the best algorithm to used, we will compare it with other algorithms that are categorized as one type of the AMF. AMF is classified as recursive algorithm. Recursive background subtraction techniques maintain a single background model thta is updated with each new video frame. According to Parks and Fels [18] , there techniques are generally computationally efficient and have minimal memory requirements. As stated previously, there are four most-widely used recursive background subctraction algorithms[18] to track moving objects; they are RGA, GMM, AGMM, and AMF.

Running Gaussian Average (RGA). The Running Gaussian Average algorithm is proposed by Wren et al. [19] and identifies the background pixels with the following procedure. We assume that μ_t denote the image mean at time t and the related variance is σ_t^2. The mean will be updated following:

$$\mu_t = \alpha I_t + (1-\alpha)\mu_{t-1}, \tag{3}$$

where the variance update is

$$\sigma_t{}^2 = d^2 I_t + (1-\rho)\sigma_{t-1}{}^2 \,. \tag{4}$$

A pixel is cosidered to be a background if its intensity lies between within a confidence interval with a threshold k such that

$$\frac{|(I_t-\mu_t)|}{\sigma_t} \geq k. \tag{5}$$

Gaussian Mixture Model (GMM). The adopted algorithm is that proposed by Stauffer and Grimson [20] and was designed to model the multi-modal backgrounds. The algorithm assumes that every pixel's in a video frame can be modeled by using Gaussian mixture model. Each channel of pixel represented as a mixture of K Gaussian. The pixels that do not match the probability of the background decision are called foreground pixels. In a time series t of pixel values, a particular pixel (x_0,y_0) can be modeled as the following:

$$\{X_1, \ldots, X_t\} = \{I(x_o, y_o, i) \colon 1 \leq i \leq t\} \tag{6}$$

The recent history of each pixel $\{X_1, \ldots, X_t\}$ is modeled by a mixture of K Gaussian distribution. Mathematically, the probability of observing the current pixel value is written as of the following:

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

where $\omega_{i,t}$ is the weight associated with the ith Gaussian at time t with mean $\mu_{i,t}$ and standard deviation $\Sigma_{i,t}$ and N denotes the Gaussian probability density which can be written as the following:

$$N\big(X_t, \mu_{it}, \Sigma_{i,t}\big) = \frac{1}{(2\pi)^{\frac{D}{2}} |\Sigma_{i,t}|^{\frac{1}{2}}} e^{\left(-\frac{1}{2}(X_t-\mu_t)^T \Sigma_{i,t}^{-1}(X_t-\mu_t)\right)}.$$

Gaussian Mixture Model with Adaptive Number of Gaussian (AGMM). Stauffer-Grimson algorithm uses a fixed number of Gaussian to model each pixel[20]. Zivkovic and van der Heijden[21] proposed an interesting extension to this mode which shows how to automatically adapt the number of Gaussian being used to mode a given pixel. Their previous studies resulted in reducing the required memory, increasing computational efficiency, and improvement in the performance when the background is highly multi-model. They choosed a reasonable time adaptation period T. At time t, we have $K_T = \{x_t, \ldots, x_{t-T}\}$. For each new sample, we update the training set K_T and re-estimate the density. The estimated density $P(x|K_T, BG + FG)$ in M component can be denoted as the following:

$$P(x|K_T, BG + FG) = \sum_{m=1}^{M} \pi_m N(x;\, \mu_m, \sigma_m{}^2).$$

The variable μ_m are the estimate of the means while ${\sigma_m}^2$ are the estimate of the variances. Variable π_m is the mixing weight which is non-negative and add up to one. The adaptive number of Gaussian is determined by updating process of the parameters. The update formula is as the following:

$$\pi_m \leftarrow \pi_m + \alpha(o_m^{(t)} - \pi_m),$$

$$\mu_m \leftarrow \mu_m + o_m^{(t)}\left(\frac{\alpha}{\pi_m}\right)\delta_m\ ,$$

and

$$\sigma_m{}^2 \leftarrow \sigma_m{}^2 + o_m^{(t)}\left(\frac{\alpha}{\pi_m}\right)(\delta_m^T\delta_m - \sigma_m{}^2),$$

where

$$\delta_m = x_t - \pi_m\ .$$

The reader is advised to consult Zivkovic and van der Heijden[21] for detail explanation.

Adaptive Median Filtering (AMF). Adaptive median filtering is a non-recursive median filtering technique. In this algorithm, the estimated median is added by 1 if the input is larger than the previously estimated median. Inversely, the median is subtracted. Mathematically, it is written as:

$$B_{t+1}^c = \begin{cases} B_t^c + 1 & if\ \ I_t^c > B_t^c \\ B_t^c - 1 & if\ \ I_t^c < B_t^c. \\ B_t^c & if\ \ I_t^c = B_t^c \end{cases}$$

The four algorithms above are the most-widely recursive techniques used for background subtraction. The previous study by Parks and Fels noted that parameters: learning rate, gaussian value, initial variance, and weight threshold affected the recall and precision of those recursive algorithms for object tracking. Their studies showed that Gaussian with value of 3, initial variance with value of 36 and weight threshold with value of 0.75 gives the best result while our previous study[3] resulted that AMF will work with the best recall and precision with the sampling rate value of 13. Our study in this paper aims to test the learning rate parameters for its influence on the result of recall and precision in real-time passenger counting in BRT station. The values we used in the learning rate parameter are the recommendation done by Parks and Fels on their study of evaluating the background subtraction algorithms with post processing. Those recommended values of the learning rate parameter are: 5.0×10^{-4}, 1.0×10^{-3}, 5.0×10^{-3}, 1.0×10^{-2}, 2.0×10^{-2}, and 1.0×10^{-1}.

4. Results

The previous study[12] concluded that, for this case, the AMF algorithm was able to count the passengers crossing the door with the level of recall and precision of 71% and 36%, respectively. The results of the current studied algorithms are presented in Tables 1 and 2.

In term of the level of recall, the three considered algorithms, RGA, GMM, and AGMM, have a tendency of improvement in the level of recall and precision with increasing the learning rate. In the case of RGA algorithm, a significant improvement of the level of recall is found when the learning rate is increased from 2.0×10^{-2} to 1.0×10^{-1}. The level of precision has also been increased significantly.

With respect to the global performance, one can clearly see that AGMM outperforms the other algorithms for more cases. The best observed level of recall is only 75%, which is slightly better than the previous approach. However, in term of precision, the three current algorithms show a rather low performance. At their best, it can only achieve the level of precision of 19%. This indicates that there are too many false positives existed in the counting process.

Table 1. Evaluation of the recall of the three algorithms

	Learning Rate, α					
Algorithm	5.0×10^{-4}	1.0×10^{-3}	5.0×10^{-3}	1.0×10^{-2}	2.0×10^{-2}	1.0×10^{-1}
RGA	0.46	0.46	0.58	0.58	0.58	0.88
GMM	0.42	0.29	0.58	0.88	0.54	0.71
AGMM	0.58	0.71	0.71	0.71	0.46	0.75

Table 2. Evaluation of the precision of the three algorithms

	Learning Rate, α					
Algorithm	5.0×10^{-4}	1.0×10^{-3}	5.0×10^{-3}	1.0×10^{-2}	2.0×10^{-2}	1.0×10^{-1}
RGA	0.08	0.04	0.06	0.05	0.07	0.15
GMM	0.07	0.04	0.09	0.13	0.07	0.12
AGMM	0.15	0.12	0.08	0.1	0.08	0.19

5. Conclusion

The problem of counting the number of passengers existed in a BRT station is an important issue to be solved. The solution will have significant impact to the operational of the system and society in general. However, the attempt to solve this problem is still facing problem particularly in term of the accuracy of the counted number of passengers. The results of this study and the previous one show that the deployed recognition algorithms produced a rather low accuracy with respect to the precision, indicating that the algorithms tend to produce false positive counting. However, the all evaluated algorithms have result in reasonable level of recall. At this point, with a proper adjustment of algorithm parameters, the recall level can achieve the level of 88%. However, if the all the recalls and precision in each learning parameter are summed and get the means of it, the AGMM has better recall and precision values than RGA and GMM. The AGMM technique resulted mean recall value of 0.65 and mean precision value of 0.12. In this study, the AGMM technique is better than RGA and GMM. But when these 3 algorithms are to be compared with our previous study[12], AMF with the level of recall and precision of 0.71 and 0.36 outperforms these 3 recursive techniques.

References

1. Morichi, S.. Long-term strategy for transport system in asian megacities. *Journal of the eastern asia society for transportation studies* 2005;**6**:1-22.
2. Gunawan, F.E., Koesnandar, E., Pranolo, F.L., Hartono, T., Pangestu, V.. Empirical travel time reliability of transjakarta bus rapid transit. In: *Proceeding international logistics seminar and workshop 2012: Competitive logistics operation as key to success in the global economy*. Jakarta; 2012, p. 7-11.
3. Gunawan, F.E.. Empirical assessment on factors affecting travel time of bus rapid transit. *International Journal of Engineering and Technology* 2015;**7**(1):327-334.
4. Gunawan, F.E., Suharjito, , Gunawan, A.A.S.. Simulation model for bus rapid transit. *EPJ Web of Conferences (Scopus indexed)* 2014;**68**:00021.
5. Gunawan, F.E.. Design and implementation of discrete-event simulation framework for modelling bus rapid transit system. *Elsevier Journal of Transportation Systems Engineering and Information Technology (scopus indexed) 14* 2014;**14**(4):97-110.
6. Campo, C.. Bus rapid transit: *Theory and practice in the united states and abroad*. Master's thesis; School of civil and environmental engineering, Georgia Institute of Technology; 2010. URL: http://hdl.handle.net/1853/37089
7. Yunita, R.. Busway for jakarta: A pressing need. AsiaViews; 2008. AsiaViews.
8. Gunawan, F.E., Chandra, F.Y.. Optimal averaging time for predicting tra_c velocity using floating car data technique for advanced traveler information system. Procedia - Social and Behavioral Sciences 2014; **138**(0):566-575. URL: http://www.sciencedirect.com/science/article/pii/S1877042814041603. doi:http://dx.doi.org/10.1016/j.sbspro.2014.07.240; the 9th International Conference on Traffic and Transportation Studies (ICTTS 2014).
9. Gunawan, F.E., Chandra, F.Y.. *Real-time Traffc Monitoring System Using Floating Car Method*. LAP Lambert Academic Publishing; 2014.

10. Gunawan, F.E., Chandra, F.Y.. Optimal averaging time for predicting traffic velocity using floating car data technique for advanced traveler information system. *Procedia - Social and Behavioural Sciences* 2014;**138**:566-575. URL: http://www.sciencedirect.com/science/article/pii/S1877042814041603. doi: http://dx.doi.org/10.1016/j.sbspro.2014.07.240.
11. Gunawan, F.E., Abbas, B.S., Atmadja, W., Chandra, F.Y., Agung, A.A., Kusnandar, E.. Ergodicity of traffc flow with constant penetration rate for traffic monitoring via floating vehicle technique. *EPJ Web of Conferences* 2014;**68**:00020. URL: http://www.scopus.com/inward/record.url?eid=2-s2.0-84901428304&partnerID=40&md5=4ec9707fb351b9126eb1b2826b15cb5f.
12. Lumentut, J., Gunawan, F.E., Atmadja, W., Abbas, B.S.. A system for real-time passenger monitoring system for bus rapid transit system.In: *The 7th Asian Conference on Intelligent Information and Database Systems (ACIIDS)*. Bali, Indonesia; 2015, .
13. Kang, H., Kim, D.. Real-time multiple tracking using competitive condensation. *Pattern Recognition* 2005;**38**:1045–1058.
14. Wang, B., Chen, Z., Wang, J., Zhang, L.. A fast passenger detection algorithm based on head area characteristics. In: *International workshop on information and electronic engineering (IWIEE)*. Harbin, China: Elsevier; 2012, p. 184–188.
15. McFarlane, N., Schofield, C.. Segmentation and tracking of piglets in images. *Machine Vision and Applications* 1995;**8**(3):187-193. URL: http://dx.doi.org/10.1007/BF01215814. doi:10.1007/BF01215814
16. Hofmann, M., Tiefenbacher, P., Rigoll, G.. Background segmentation with feedback: The pixel-based adaptive segmenter. In: *Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on*. 2012, p. 38-43. doi:10.1109/CVPRW.2012.6238925.
17. Godbehere, A., Matsukawa, A., Goldberg, K.. Visual tracking of human visitors under variable-lighting conditions for a responsive audio art installation. In: *American Control Conference (ACC), 2012*. 2012, p. 4305–4312. doi:10.1109/ACC.2012.6315174.
18. Parks, D., Fels, S.. Evaluation of background subtraction algorithms with post-processing. In: *Advanced Video and Signal Based Surveillance, 2008. AVSS '08. IEEE Fifth International Conference on.* 2008, p. 192–199. doi:10.1109/AVSS.2008.19.
19. Wren, C., Azarbayejani, A., Darrell, T., Pentland, A.. Pfinder: Real-time tracking of the human body. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 1997;**19**(7):780–785.
20. Stauffer, C., Grimson, W.. Adaptive background mixture models for real-time tracking. In: *IEEE Computer Society Conference on ComputerVision and Pattern Recognition*. 1999, .
21. Zivkovic, Z., van der Heijden, F.. Efficient adaptive density estimation per image pixel for the task of background subtraction. *Pattern recognition letters* 2006;**27**:773–780.
22. Pavlidis, I., Symosek, P., Morellas, V., Fritz, B., Papanikolopoulos, N.P., Sfarzo, R. *Automatic passenger counting in the HOV lane*. Department of Computer Science, University of Minnesota; 1999.
23. Yahiaoui, T., Khoudour, L., Meurie, C. Real-time passenger counting in buses using dense stereovision. Journal of Electronic Imaging 2010; **19**(3): 031202 -031202