Background subtraction algorithms with post processing
A review

Insaf Setitra
Research Center on Scientific
and Technical Information Cerist
University of Science and Technology USTHB
Algiers, Algeria
isetitra@cerist.dz

Slimane Larabi
University of Science and Technology USTHB
Algiers, Algeria
slarabi@usthb.dz

***Abstract*—Due to its several algorithms with their fast implementations, background subtraction becomes a very important step in many computer vision and video surveillance systems which assume static cameras. Literature counts a large number of robust background subtraction algorithms which try each to outperform the others in a quantitative and qualitative manner. This competition can sometimes confuse the user of this kind of process and make the choice of one of them difficult. To overcome this issue we review, in what follows, the background subtraction process by defining it and exploring most used algorithms of background subtraction. We then expose some post processing techniques used to remove superfluous content derived from background subtraction.**

***Keywords*-bakcground subtraction; video analysis; background modeling; post processing;state of the art;**

I. Introduction

Background subtraction is a widely used real-time method for identifying foreground objects in a video stream [1], [2], [3], [4]. In background subtraction, the main idea is to focus analysis of videos only on regions of interest which are mostly the moving objects of the scene. As input, background subtraction takes a sequence of images and produces as output, image masks where only moving objects are highlighted. The core problem of background subtraction is to identify the set of pixels of the image seen which are significantly different from the pixels seen in the previous images of the sequence; these pixels comprise what is called the change mask. The change mask is affected by several factors such as appearance, brightness and color, and shape changes of objects. Illumination changes are also a substantial issue involving the background subtraction process. Hence, in presence of all these conditions, the main idea behind background subtraction is to detect moving objects from the difference between the current frame and a reference frame, often called the background image, or background model. Background model must be a strong representation of the scene without any moving objects and must be kept regularly updated so as to adapt to the varying luminance conditions and geometry settings [5]. Literature counts a significant number of background subtraction methods and algorithms. However, only few of them are effectively applied in the computer vision tasks and are representative of the field. Relevant works directly related to the background subtraction task include surveys and comparative studies. [5] and [6] propose a comprehensive and systematic surveys on several image change algorithms. However, due to the fast growth of background subtraction algorithms, we believe it is crucial to survey as often as possible background subtraction evolution. We believe also that post processing techniques for background subtraction are necessary to describe when analyzing background subtraction techniques. In what follows, we first describe principles of evaluating and comparing background subtraction algorithms. Following paragraphs are dedicated to describe widespread algorithms of background subtraction. Description includes outline of the native algorithms with their strengths and weaknesses and some works done to improve them. Note that we avoid mathematical definitions in each description; we aim in this paper to give a simple yet comprehensive overview of background subtraction techniques. Readers can refer to original papers mentioned bellow to have a deeper mathematical modeling of each method. In our description, we follow the same fettering as the one presented in [5] i.e. we describe algorithms ranging from the most simple to the most sophisticated. We certainly contribute in this paper with more background subtraction methods than the ones described therein. We then describe some post-processing techniques used to overcome foreground misses derived for background subtraction process. We conclude with some critics of background subtraction algorithms and future works.

II. ACCURACY MEASUREMENT OF BACKGROUND SUBTRACTION

By considering background subtraction as the process of detecting important changes in images without considering unimportant or nuisance forms of change, Radke et. al. [6] consider that the concept of importance is application dependant. Examples include detecting swaying trees considered unimportant and detecting blood vessels movement which is important [7]. These issues make evaluation techniques divergent. Evaluation techniques can be divided into ground truth

[6] and non ground truth based techniques [8]. Ground truth is a mask generated most often manually by a human. Ground truth based evaluation techniques is prone to human error. It is also time consuming. Due to these difficulties, non ground truth based evaluation techniques are sometimes used. However, their drawback is that it is difficult to define superior criterion/criteria for such an evaluation.

A. Ground Truth based evaluation measures

Three methods for quantifying a classifiers performance can be used to measure background subtraction accuracy; Percentage Correct Classification PCC, Jaccard coefficient JC and Yule coefficient YC :

$$PCC = \frac{TP + TN}{TP + FP + TN + FN} \quad (1)$$

$$JC = \frac{TP}{TP + FP + FN} \quad (2)$$

$$YC = \mid \frac{TP}{TP + FP} + \frac{TN}{TN + FN} - 1 \quad (3)$$

TP (True positives) number of change pixels correctly detected; FP (False positives) number of no-change pixels incorrectly detected as change (also known as false alarms); TN (True negatives) number of no-change pixels correctly detected; FN (False negatives) number of change pixels incorrectly detected as no-change (also known as misses). The receiver-operating-characteristics (ROC) curve can also be used in comparing. ROC curve plots the detection probability versus the false alarm probability. Detection probability is the ratio between true positives and total actual positive. False alarm probability is the ratio between true negative and total actual negative. More sophisticated ROC curves can also be applied to background subtraction for a ground truth based evaluation [9].

B. Non Ground Truth based evaluation measures

In [7], measures are separated into three categories: boundary based, model based and assisted based methods. In boundary based methods, boundaries of detected foreground are used to extract internal and external regions in original images. Regions features such as color are then used to inspect their homogeneity and so background subtraction robustness [10]. In model based methods, further processing is used to improve false detections in background subtraction. In [11], detection of people is used to correct segmentation. Assisted based methods attempt to manage background subtraction errors by combining several background subtraction algorithms (defined bellow) [12]. An example of boundary based background subtraction measures are the ones described and used in [8]. The first one (Boundary Spatial Color Contrast BSCC) is based on color and the second one (Boundary Motion Contrast BMC) is based on motion.

BSCC measure first draws a virtual normal line of lengh $2L+1$ between each two opposite boundary pixels detected in the foreground mask contour. Let P_O^i and P_I^i be two pixels in two opposite directions traversing the contour of objet i at time t and let C_O^i and C_I^i the average colors calculated in the $M \times M$ neighborhood of the pixels P_O^i and P_I^i (using the RGB color space quantified into 256 levels). The formula used to compute the BSS is:

$$BSCC(t; i) = \frac{\parallel C_0^i(t) - C_I^i(t) \parallel}{\sqrt{3 \times 255^2}} \quad (4)$$

BMC measure draws similarly to BSS a $2L + 1$ normal line between each two opposite boundary pixels detected in the foreground mask contour. The motion is computed using the motion vector of the two regions C_O^i and C_I^i defined the same manner as for BSS. The formula used to caompute BMC is:

$$BMC(t; i) = \omega_t \left(1 - \exp\left(-\frac{\parallel v_0^i(t) - v_I^i(t) \parallel}{\sigma^2}\right)\right) \quad (5)$$

where $v_0^i(t)$ and $v_I^i(t)$ the average motion vectors calculated in the $M \times M$ region centered at the pixels P_O^i and P_I^i respectively. $\omega_t = R(v_0^i(t)).R(v_I^i(t))$ and $R(.)$ represents the reliability of the motion vectors [13].

III. BACKGROUND SUBTRACTION ALGORITHMS

In what follows, we present some background subtraction methods separated in two main categories: recursive and non-recursive methods. The difference between both is that the formers use a unique value to account for the background. The latters use an entire buffer to represent the background.

A. Non Recursive techniques for background subtraction

1) Temporal median filter: In temporal median filter [14], first, background model is constructed by computing the median value of a buffer of n last frames. Then, a new pixel is compared to the model, if it falls within a certain threshold, it is considered background, otherwise it is considered foreground. Update the background is done by adding the current value of the pixel to the buffer as long as the size of the buffer allows it. The method was improved by Cucchiara et al. in [15] and [16]. In [15], authors used medoid instead of median filtering. Although median and medoid filtering have proved their performance in many background subtraction works [14], [16], [15], [17], [1] and were considered as having competitive results in spite of their simplicity [17], they still suffer as most non recursive methods of high memory requirements because of the size of buffers needed for modeling.

2) Linear predictive filter: Linear predictive filter [18] method estimates the background by applying a linear predictive model in the buffer and at each frame the filter coefficients are estimated. Estimation is based on the sample covariance. This method has the disadvantage of being heavy in computation and memory consuming.

3) Non-parametric model: Non-parametric model [19], performs the background subtraction using the entire history of the pixel and estimating the background model using a Kernel function on it. In [19] the Kernel function used was a Gaussian one. For each observation, if the pixel is likely to come from this distribution i.e. difference is within a certain threshold then it is considered foreground, otherwise it is considered background. This method has the advantage to treat a multimodal background subtraction but is time and memory consuming.

4) Co-occurrence of image variations: This method [20], achieves background subtraction in two principal phases; training and classification phase. In training phase, an Eigen vector transformation is performed on co-occurrence matrix of the pixel with its N neighboring blocks and the mean of training pixels. In the classification phase, the current Eigen image variation of the observed pixel with its neighboring block is computed. New pixel is considered background if it is close to its estimated value. More details about the approach can be found in [5] and in [21]. Very time consuming, this technique requires high computations, however, it is robust to limited changes in the overall illumination level.

5) Eigenbackgrounds: This method [22] performs background subtraction using eigenvectors by applying principal component analysis to a set of N video frames that do not contain any foreground objects. This is done in order to capture spatial correlations between pixels history. The method follows two steps: the training and the classification phase. In the training phase, an eigenspace is constructed with covariance matrix of the L images in the buffer and their average. In the classification phase, the new image is projected into the eigenspace previously constructed and then back projected to the image space. If the difference between the original image and its backprojection falls within a threshold the current image is considered background, otherwise it is foreground. This technique is time and memory consuming and it seems not well updating the background. Moreover, it is not that obvious to have in each video a set of frames all not containing any moving objects. The comparative study performed in [1] reported that strong results were obtained by the Eigenbackground algorithm considering that it does not update the background model.

IV. Recursive techniques for background subtraction

6) Basic Motion Detection: The basic motion detection method is one of the easiest ways of performing background subtraction. Initially, the background of this method is the first frame. After each observation, pixels are compared to the background using some distance. If the pixel falls into a certain threshold it is considered background, otherwise, it is considered foreground. At each observation of time t, background is estimated by weighting the background of time $(t-1)$ with $(1-\alpha)$ and adding α times the current pixel. The strength of that method is that it is very easy to compute and is not memory consuming as no buffers are needed. However, many weaknesses can be observed. The method includes new foreground pixels to the background directly after each observation which may lose some foreground objects static in one or two frames, subtraction depends strongly on the threshold and the updating constant α which can be chosen inappropriately and we have no information about distribution of the background. [23] used that method for comparison and it was stated that its global threshold significantly penalizes the performance.

7) Frame differencing: In this method, current image pixels are compared only to pixels of the direct previous frame. No history of pixels is needed and there are no several computations. This is an advantage for the method since it is neither time nor memory consuming, however, since history of pixels is not known, multimodal background cannot be maintained and it is impossible to avoid detecting spurious objects, such as swaying tree branches. This is also known as the aperture problem described in [18] and [24]. Comparison in [24] showed that the method performs under best time; however, it has the worst results for waving trees and outdoor data set. A similar approach called first frame subtraction may suffer of the same problems; the latter is performed using only the first frame of the video considered as being background. Each new frame is compared to the first frame of the video, and is considered background if it falls within a certain threshold. While time and memory saving, this assumption is highly inadequate especially if the first frame contains moving objects.

8) Minimum, Maximum and Maximum Inter-Frame Difference: This method [25] computes 3 values, a minimum a maximum and a maximum of consecutive frames difference. For each new observation of the pixel, the latter is considered background if the difference between the pixel and one of the minimum, the maximum or the maximum of previous background pixels falls within a threshold weighted by a mean. This mean concerns the largest inter-frame difference over pixels. At each observation, the mean is updated. The method was used in [23] but still it is not efficient especially in presence of noise, and when used on multimodal videos.

9) Approximated median filter: The approach [26] is an alternative for median filtering, as the latter gave unsatisfying results in terms of memory and time consuming. Background is first initialized to the first image, then, for each new observation, the pixel is compared to the background. If it is either higher or lower it is considered foreground. Updating the background is done the following way; if the foreground pixel is higher than the background one, background is incremented, if it is lower, background is decremented, if theyre equals then background remains the same. Most advantages of this method are that it is simple, robust to noise and computationally efficient. In another hand, it has no history of pixels and does not model their variance.

10) Codebook Model: This method [27], considers background values over time and performs background subtraction as follow: For each pixel, get its codeword history, then each new pixel, is compared to its codeword, if it falls within a certain threshold it is considered background and the codeword is updated with its value. If not, it is considered foreground, and a new codeword is created. One advantage of the method is that it allows accounting for dynamic backgrounds since background is considered over a long period. It is also used as mentioned in [28] to bootstrap the system in the presence of foreground objects. Disadvantage of this method is that updating the background is done completely which means that background pixels are considered background during the whole process. Hence, a parked car which follows its same trajectory for a second time may be considered as background since its pixels are present in the codeword.

11) Running Gaussian average: The method [29] performs background subtraction as follows: First the running (or online cumulative) average of the background distribution assumed to be Gaussian is initialized to the first image weighted by , an empirical weight. For each new pixel, the latter is compared to the background by subtracting its value from the mean, if it is under k times the standard deviation it is considered background, otherwise, it is considered foreground. The background distribution is updated by including the observed pixel. An improvement of the method was proposed by Kolleref et. al. in [30]. This work aimed to avoid the excessiveness of background update. Authors proposed to add the current value to the background distribution only if it is considered background. The main advantages of this method are the speed and the low memory requirements. Instead of storing the previous n images to update the background image, only 2 values per pixel are stored. Disadvantage of that method is that it does not handle redundant movement of objects.

12) Single Gaussian: Single Gaussian [29], estimates the background first on a training phase where a Gaussian for each background pixel is constructed. New pixel is compared to the background model using a distance measure and mean and covariance of each background distribution pixel is updated with the current pixel and using an updating constant. Advantages and disadvantages of the method can be seen in [23] and [28]. In [23], the Gaussian weighted by a covariance matrix compensated well background instabilities. In [28] Single Gaussian was affected by the high variances of alternating pixels and the low variance of stable pixels.

13) Mixture of Gaussians: Stauffer and Grimson in [30] propose to model each background pixel as a mixture of Gaussian instead of a single Gaussian. The method creates k Gaussian distributions based on the first background frame; all distributions are ranked based on the ratio between their peak amplitude and standard deviation. Then, the same comparison is made. Once a pixel is classified, the best matching background distribution is updated with the value of the current pixel if the latter is considered background. Otherwise, the weaker distribution is replaced with a new Gaussian based on the observed pixel. Zivkovic and Heijden propose in [31] an adaptive GMM. The method adapts automatically the number of Gaussians being used to model a given pixel. This extension reduces the algorithms memory requirements, increases its computational efficiency, and can improve performance when the background is highly multi-modal.

14) Kernel Density Estimation: This method [19], models on a buffer of the last n backgrounds- the background distribution by a model based on Kernel Density Estimation values. More clearly, this approach estimates the background pixels using a Kernel function on the N most recent frames. For each new frame, pixels are considered background if they are likely to come from the distribution computed previously. When a pixel is considered background, it is added to the distribution by removing a value from the distribution in a FIFO order. In Kernel Density Estimation, Foreground pixels do not pollute the estimation. However, estimation of the Kernel parameters has to be done.

15) Bayes decision rule for classification: This method [32], uses a feature selection and a Bayes decision rule to decide whether the current pixel is considered background or foreground. Features in [32] are selected as vectors of color pixel values for background and color cooccurrences of inter-frame changes for foreground. In order to represent multiple sates of background pixels, the method keeps a table of feature statistics and updates the model in each classification. Bayesian decision rule based on features was compared to VeBe algorithm in [33] and gave competitive results compared to the method proposed therein. However, it was more than 20 times slower.

16) Kalman filter: This technique [34] is based on the intensity and its temporal derivative. Background in there is updated using two parameters: the background dynamics and measurement matrix. This method has been used in the comparative study of Cheung and Chandrika et. Al. in [17] and was easily affected by the foreground pixels and leaved a long trail after a moving object.

17) Multimodal Mean: [24] Background is modelled as k cells, each of which represents one mean. Mean of each cell is computed as a running sum of pixels averaged by number of appearances of the pixel. A new image pixel is a background pixel if a cell can be found whose mean for each color component x matches within the corresponding color component of the current pixel. Update of the model is done by adding the current pixel to the running sum and incrementing number of appearance to compute the mean. To deal with new stationary objects, a decimation rate is proposed and the model is updated using this rate. In [24], the method gave satisfying results, outperforming the Mixture of Gaussian. However, still this method does not deal with sudden luminance variations and strong background changes.

V. POST PROCESSING FOR BACKGROUND SUBTRACTION

In many computer vision applications, several issues may false the foreground detection. Examples include sudden changes of illumination which may result on several moving objects detected, similar pixels between moving objects and background, merge and split due to occlusions. In order to fix this kind of troubles, several techniques can be applied on foreground masks. In [1], a good insight of post processing techniques for background subtraction is given. In what follows, we describe briefly most used post processing techniques for background subtraction. We follow the same fettering as the one in [1], however, as the authors in there describe only techniques they used, we believe it is mandatory for us to discuss other post processing techniques not mentioned in their paper. We also add more details for techniques described in [1].

A. Remove noise and fill holes

Morphological operators and filtering are good ways to fill holes. Fig.1 shows an example of background subtraction which includes holes in the interest region and much clutter. Comparing the image mask with the original images shows that many regions detected are swaying trees, a swaying wire and some changes in illumination. They are considered definitely as clutter.

Fig. 1. Example of issues encountered in background subtraction. a. foreground mask. b. original image.

1) Image filtering for post processing: In [28] foreground mask was processed with a median filter and connected component analysis. Simpler methods consist of eliminating foreground regions that have a small size. The inconvenient of this kind of methods is the choice of the size. One can think of a dynamic threshold depending on the size of all regions of the foreground by keeping the history of all of them.

2) Morphological operations for post processing: Morphological transformations on images can be used to remove speckle noise, isolate individual elements, join disparate elements in an image, find intensity bumps or holes in an image, and so forth [35]. The two basic morphological operations used in post processing of background subtraction are dilatation which is described as the convolution of an image A with a Kernel B. Basic effect of the kernel on the image is to gradually enlarge the boundaries of bright regions and erosion which applies the converse processing to the foreground mask and causes darkness in the foreground mask. Fig.2 shows an example of using erosion and dilatation. Based on the two previous morphological operations, Opening consists of applying erosion followed by dilatation on the image and closing consists of applying erosion followed by dilatation on the image. For better results, connected components algorithms use an opening followed by a closing. Opening and closing in such algorithms are used instead of dilation and erosion since the latter operations preserve less the shape and area of regions than opening and closing. Other morphological operations less used include: morphological gradient, top hat, and black hat [35].

B. Remove shadows

Removing shadows in foreground masks is important in that shadows can alter region shapes and make further processing very difficult. Fig.3 shows an example of shadow considered as foreground. In [36], shadow can be self shadow or cast shadow. The self-shadow is a part of the object. The cast shadow is the area in the background projected by the object in the direction of light rays. Cast shadow and self-shadow regions have different colors if the background and object are of different colors. Based on this assumption, Javed and shah [36] use these cues with edges and gradients to remove shadows. In [37], color and texture are used to remove potential chromatic shadows from foreground masks. In [38], even though shadow is removed from static images and not from frames of a video, the concept is quite similar, where region and edge continuity are extracted from invariant images obtained by projecting the image log-chromaticities in the entropy minimizing direction [39], [40]. More simple methods remove shadows by simply converting the original image to an HSV coding, thus, different illumination of the same background will be ignored when taking only the hue and saturation of the image.

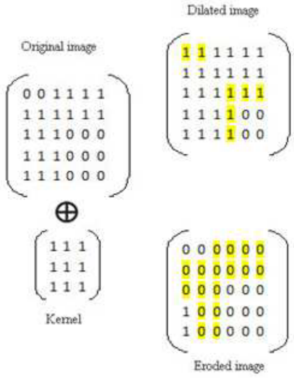

Fig. 2. Example of dilation and erosion in a matrix

Fig. 3. An example of shape distortion due to shadow.

VI. Conclusion

Methods cited previously are good methods to compare accuracy of background subtraction algorithms. However, visual inspection on results is often appreciated. Hence, human intervention including semantic knowledge about the scene and interpretation prevent of treating every pixel equally when pixels are of particular attention. We aim in further works to use background subtraction algorithms in subsequent processing such as tracking and classification.

References

[1] D. Parks and S. Fels, "Evaluation of background subtraction algorithms with post-processing," in *Advanced Video and Signal Based Surveillance, 2008. AVSS '08. IEEE Fifth International Conference on*, Sept 2008, pp. 192–199.

[2] T. B. Moeslund, A. Hilton, and V. Krger, "A survey of advances in vision-based human motion capture and analysis," *Computer Vision and Image Understanding*, vol. 104, no. 23, pp. 90 – 126, 2006, special Issue on Modeling People: Vision-based understanding of a persons shape, appearance, movement and behaviour.

[3] W. Hu, T. Tan, L. Wang, and S. Maybank, "A survey on visual surveillance of object motion and behaviors," *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, vol. 34, no. 3, pp. 334–352, Aug 2004.

[4] M. S. Lew, N. Sebe, C. Djeraba, and R. Jain, "Content-based multimedia information retrieval: State of the art and challenges," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 2, no. 1, pp. 1–19, Feb. 2006.

[5] M. Piccardi, "Background subtraction techniques: a review," in *Systems, Man and Cybernetics, 2004 IEEE International Conference on*, vol. 4, Oct 2004, pp. 3099–3104 vol.4.

[6] R. Radke, S. Andra, O. Al-Kofahi, and B. Roysam, "Image change detection algorithms: a systematic survey," *Image Processing, IEEE Transactions on*, vol. 14, no. 3, pp. 294–307, March 2005.

[7] T. R. Savarimuthu, A. Kjær-Nielsen, and A. S. Sørensen, "Real-time medical video processing, enabled by hardware accelerated correlations," *J. Real-Time Image Process.*, vol. 6, no. 3, pp. 187–197, Sep. 2011.

[8] J. SanMiguel and J. Martinez, "On the evaluation of background subtraction algorithms without ground-truth," in *Advanced Video and Signal Based Surveillance (AVSS), 2010 Seventh IEEE International Conference on*, Aug 2010, pp. 180–187.

[9] C. O'Reilly and T. Nielsen, "Revisiting the roc curve for diagnostic applications with an unbalanced class distribution," in *Systems, Signal Processing and their Applications (WoSSPA), 2013 8th International Workshop on*, May 2013, pp. 413–420.

[10] P. L. Correia and F. Pereira, "Stand-alone objective segmentation quality evaluation," *EURASIP J. Appl. Signal Process.*, vol. 2002, no. 1, pp. 389–400, Jan. 2002.

[11] M. Harville, "A framework for high-level feedback to adaptive, per-pixel, mixture-of-gaussian background models," 2002, pp. 543–560.

[12] L. Goldmann, T. Adamek, P. Vajda, M. Karaman, R. Mrzinger, E. Galmar, T. Sikora, N. OConnor, T. Ha-Minh, T. Ebrahimi, P. Schallauer, and B. Huet, "Towards fully automatic image segmentation evaluation," vol. 5259, pp. 566–577, 2008.

[13] C. Erdem, B. Sankur, and A. Tekalp, "Performance measures for video object segmentation and tracking," *Image Processing, IEEE Transactions on*, vol. 13, no. 7, pp. 937–951, July 2004.

[14] B. P. L. Lo and S. Velastin, "Automatic congestion detection system for underground platforms," in *Intelligent Multimedia, Video and Speech Processing, 2001. Proceedings of 2001 International Symposium on*, 2001, pp. 158–161.

[15] R. Cucchiara, C. Grana, M. Piccardi, and A. Prati, "Detecting objects, shadows and ghosts in video streams by exploiting color and motion information," in *Image Analysis and Processing, 2001. Proceedings. 11th International Conference on*, Sep 2001, pp. 360–365.

[16] S. Calderara, R. Melli, A. Prati, and R. Cucchiara, "Reliable background suppression for complex scenes," in *Proceedings of the 4th ACM International Workshop on Video Surveillance and Sensor Networks*, ser. VSSN '06. New York, NY, USA: ACM, 2006, pp. 211–214.

[17] S.-C. S. Cheung and C. Kamath, "Robust background subtraction with foreground validation for urban traffic video," *EURASIP J. Appl. Signal Process.*, vol. 2005, pp. 2330–2340, Jan. 2005.

[18] K. Toyama, J. Krumm, B. Brumitt, and B. Meyers, "Wallflower: principles and practice of background maintenance," in *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, vol. 1, 1999, pp. 255–261 vol.1.

[19] A. M. Elgammal, D. Harwood, and L. S. Davis, "Non-parametric model for background subtraction," in *Proceedings of the 6th European Conference on Computer Vision-Part II*, ser. ECCV '00. London, UK, UK: Springer-Verlag, 2000, pp. 751–767.

[20] M. Seki, T. Wada, H. Fujiwara, and K. Sumi, "Background subtraction based on cooccurrence of image variations," in *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, vol. 2, June 2003, pp. II–65–II–72 vol.2.

[21] B. Han, D. Comaniciu, and L. Davis, "Sequential kernel density approximation through mode propagation: applications to background modeling," in *Proc. ACCV*, vol. 4, 2004, pp. 818–823.

[22] N. Oliver, B. Rosario, and A. Pentland, "A bayesian computer vision system for modeling human interactions," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 22, no. 8, pp. 831–843, Aug 2000.

[23] Y. Benezeth, P.-M. Jodoin, B. Emile, H. Laurent, and C. Rosenberger, "Review and evaluation of commonly-implemented background subtraction algorithms," in *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*, Dec 2008, pp. 1–4.

[24] S. Apewokin, B. Valentine, L. Wills, S. Wills, and A. Gentile, "Multimodal mean adaptive backgrounding for embedded real-time video surveillance," in *Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on*, June 2007, pp. 1–6.

[25] I. Haritaoglu, D. Harwood, and L. Davis, "W4: A real time system for detecting and tracking people," in *Computer Vision and Pattern Recognition, 1998. Proceedings. 1998 IEEE Computer Society Conference on*, Jun 1998, pp. 962–962.

[26] N. McFarlane and C. Schofield, "Segmentation and tracking of piglets in images," *Machine Vision and Applications*, vol. 8, no. 3, pp. 187–193, 1995.

[27] K. Kim, T. H. Chalidabhongse, D. Harwood, and L. Davis, "Real-time foreground-background segmentation using codebook model," *Real-Time Imaging*, vol. 11, no. 3, pp. 172–185, Jun. 2005.

[28] K. Greff, A. Brando, S. Krau, D. Stricker, and E. Clua, "A comparison between background subtraction algorithms using a consumer depth camera." in *VISAPP (1)*, G. Csurka and J. Braz, Eds. SciTePress, 2012, pp. 431–436.

[29] C. Wren, A. Azarbayejani, T. Darrell, and A. Pentland, "Pfinder: real-time tracking of the human body," in *Automatic Face and Gesture Recognition, 1996., Proceedings of the Second International Conference on*, Oct 1996, pp. 51–56.

[30] D. Koller, J. Weber, T. Huang, J. Malik, G. Ogasawara, B. Rao, and S. Russell, "Towards robust automatic traffic scene analysis in real-time," in *Decision and Control, 1994., Proceedings of the 33rd IEEE Conference on*, vol. 4, Dec 1994, pp. 3776–3781 vol.4.

[31] Z. Zivkovic and F. van der Heijden, "Efficient adaptive density estimation per image pixel for the task of background subtraction," *Pattern Recogn. Lett.*, vol. 27, no. 7, pp. 773–780, May 2006.

[32] L. Li, W. Huang, I. Y. H. Gu, and Q. Tian, "Foreground object detection from videos containing complex background," in *Proceedings of the Eleventh ACM International Conference on Multimedia*, ser. MULTIMEDIA '03. New York, NY, USA: ACM, 2003, pp. 2–10.

[33] O. Barnich and M. Van Droogenbroeck, "Vibe: A universal background subtraction algorithm for video sequences," *Image Processing, IEEE Transactions on*, vol. 20, no. 6, pp. 1709–1724, June 2011.

[34] C. Ridder, O. Munkelt, and H. Kirchner, "Adaptive background estimation and foreground detection using kalman-filtering," 1995, pp. 193–199.

[35] D. G. R. Bradski and A. Kaehler, *Learning Opencv, 1st Edition*, 1st ed. O'Reilly Media, Inc., 2008.

[36] O. Javed and M. Shah, "Tracking and object classification for automated surveillance," in *Proceedings of the 7th European Conference on Computer Vision-Part IV*, ser. ECCV '02. London, UK, UK: Springer-Verlag, 2002, pp. 343–357.

[37] I. Huerta, M. Holte, T. Moeslund, and J. Gonzalez, "Detection and removal of chromatic moving shadows in surveillance scenarios," in *Computer Vision, 2009 IEEE 12th International Conference on*, Sept 2009, pp. 1499–1506.

[38] C. Fredembach and G. Finlayson, "Simple shadow removal," in *In ICPR 06: Proceedings of the 18th International Conference on Pattern Recognition (ICPR06*. IEEE Computer Society, 2006, pp. 832–835.

[39] G. D. Finlayson, M. S. Drew, and C. Lu, "Entropy minimization for shadow removal," *Int. J. Comput. Vision*, vol. 85, no. 1, pp. 35–57, Oct. 2009.

[40] C. Fredembach and G. D. Finlayson, "Hamiltonian path-based shadow removal," in *In: Proceedings of the 16th British Machine Vision Conference (BMVC*, 2005, pp. 502–511.