A Modified Codebook-based Background Subtraction Technique to improve Activity Classification in Highly Variable Environments

Yusuf A. Syed, Shetty S., Wilkinson N., and Brown D. J.

Abstract**—** Background subtraction techniques are commonly used for moving object detection in videos to detect foreground changes under highly variable environments such as moving trees, traffic, shadows or cloud cover. The article proposes an extension to the existing background subtraction state-of-the-art by means of a modified codebook algorithm (MCB-HED). The technique exploits the ability of Euclidean distance-based RGB color representation, Hue and intensity of pixels over consecutive frames to identify foreground objects in the presence of intensity variations due to shadows and highlights.

The methodology constructs a codebook of entries that are organized into intensity-based groups with each containing a number of codewords whose number depends upon the degree of variability of that pixel. Each pixel's behavior overtime is thus recorded in the codebook based upon the degree of change it exhibit via a statistical probability descriptor called kurtosis. When evaluated against a range of benchmarking video sequences, the outcome presents efficient background subtraction. The proposed technique is found to efficiently eliminate shadows and near-color foreground-background pixel similarities.

Keywords**—Background subtraction, video surveillance, image processing, artificial intelligence**

I. Introduction

The ability to identify non-static objects from a sequence of video frames via a static camera is the first step in a computer vision-based video surveillance system. Particularly, processing complex backgrounds that are cluttered with moving background such as trees, water, or vehicles is a challenging problem. Due to the presence of these objects, the classification of various image regions in video frames in order to differentiate background from foreground pixels requires careful modeling in order to minimize false alarms. The problem is further complicated if the frames contain abrupt intensity variations, object shadows, occlusions or objects moving at variable speeds. The field has attracted a lot of attention lately with a substantial level of research focusing on background estimation, modeling and subtraction.

Yusuf A. Syed is with the STS Defence Ltd, Gosport, PO12 1AF UK (phone: 02392 584222; fax: 02392 529598; e-mail: adnan.yusuf@sts-defence.com).

Shetty S. is with the STS Defence Ltd, Gosport, PO12 1AF UK (phone: 02392 584222; fax: 02392 529598; e-mail: siddharth.shetty@sts-defence.com).

Wilkinson N. is with the STS Defence Ltd, Gosport, PO12 1AF UK (phone: 02392 584222; fax: 02392 529598; e-mail: nigel.wilkinson@sts-defence.com).

Brown D. J. is with the Institute of Industrial Research, University of Portsmouth, Portsmouth, PO1 2EG UK (phone: 02392 844446; e-mail: david.j.brown@port.ac.uk).

Background subtraction is a technique that compares each new frame in a video sequence with a model of the scene background which is trained from image frames taken from a static camera. A well-known and robust technique to background subtraction is to construct a statistic model representing each pixel's distribution of intensity and color distortion. To address the problems of intensity or color variations, [1] presented a probabilistic method based on Gaussian Mixture Models (GMM) to treat each pixel as a mixture model trained over time. Moreover, a method utilizing statistical Bayesian framework for background segmentation via mixture of Gaussians (MOGs) was developed by [2]. [3] used time-adaptive pixel-based Gaussian modeling where pixel-based background models were updated as new changes to the scene were detected leading to older observations being discarded over time. The methodology was able to detect repetitive scenes such as object insertions and removals or intensity variations due to events such as cloud cover change or moving dials of a clock. [4] employed color and gradient information over pixel, region and frame level that used high gradients on object boundaries to eliminate false objects or abrupt intensity variation-induced foregrounds.

Despite promising results, these methodologies presented a number of significant disadvantages due to the underlying principles of MOGs. The method is sensitive to abrupt lighting changes and can only model slow intensity changes. Moreover, it cannot model fast background variations with fewer Gaussians. Moreover, if the learning rate is low, the technique cannot detect sudden changes in the background. On the other hand, a high learning rate can adapt quickly but slow moving foreground is absorbed by the background.

In order to eliminate shortcomings of Gaussians [5] proposed a memory and speed efficient technique for background model via a modified codebook model. The technique uses maximum negative runtime length (MNRL) based on a color distortion and intensity variation measure compared between pixels of consecutive video frames in order to identify foreground objects in a scene. The technique presents good results against abrupt shadow and intensity variations with a substantial improvement against Gaussian-based models that are more prone to errors due to object shadows. In order to address a number of these issues, building upon the work by [5], an advanced codebook model based upon weighted Euclidean distance was used by [6]. In addition to MNRL, the methodology introduced frequency parameters in the codebook to access, delete, match or add code words from the codebook. Though the technique claim to produce better results when compared to [5] and MOG, it still presents an additional memory overhead of keeping two frequency variables in memory.

Based on the work presented above, this work proposes a novel modified codebook model based upon reduced number of codewords and parameters. In the technique, a measure to identify abrupt or long-duration intensity changes is proposed. The measure acts as a probability distribution shape descriptor of the real-valued variable of intensity to identify the co-called "peakedness" present in the temporal values. Moreover, the use of an RGB Euclidean distance (ED)-based operator addresses the issues of sudden intensity variations in identically colored backgrounds in addition to shadows. Based on the above-mention criteria, a modified codebook model (MCB-HED) is proposed with an objective to reduce the number of codebook entries to improve memory utilization. The work is organized as follows: Section II presents an overview to existing background subtraction methods along with the underlying shortcomings. Section III presents the methodology based on a modified Euclidean distance (ED) and vector angle (VA) measurement in addition to time-based Intensity-kurtosis to add, update or remove codebook entries. Section IV performs a comparative analysis of the results obtained via the proposed methodology in comparison to an existing methodology. Section V finally concludes with a discussion on the possible future directions of this research.

II. Overview

In background subtraction based modeling, pixels of non-static objects are detected by subtracting current frame from the trained background model. The well-known codebook methodology identifies moving objects on the basis of each underlying pixel with its behavior saved in a codebook. For instance, for a pixel $p_{i,j}$ with its incoming RGB value $p_{i,j}(RGB)$ and codebook $C_{i,j}$, the image subtraction process will typically comprise of three steps as follows:

A. Overview of MOG-based Background Subtraction

Using conventional MOGs has its own disadvantages. It must be noted that proper tuning of MOG is paramount in accurate detection for foreground pixels in frames as, if the model updates to fast, the system takes more time to learn from light intensity changes [6].

Consequently, a background subtraction process on the MOG method tends to suffer from a difficulty to classify object shadows. Shadows attached to actual objects, particularly in the detection of human bodies, pose a big challenge in the surveillance research as it makes the next phase of gait, gesture or limb recognition more problematic. Cast shadows tend to consist of a dark part where the light source is totally absent and a "boundary region", or penumbra, where light source is partly blocked.

For an image frame defined in RGB color space, the intensity of a pixel $p_{i,j}$ is defined as a linear blend of RGB color space donated by $I_{i,j}$.

$$I_{i,j} = \omega_r R_+ \omega_g G + \omega_b B \tag{1}$$

When a pixel has a shadow cast on it, it assumes a darker intensity due to lesser amount of light reaching it which effectively changes its brightness as follows:

$$\tilde{I}_{i,j} = \alpha I_{i,j} \tag{2}$$

Where $\tilde{I}_{i,j}$ is the new brightness level for pixel $I_{i,j}$. According to [7], the effect is similar to specific illumination changes in a frame, for instance, when the lights a turned on/off. The value of α is < 1 for shadows and > 1 for highlights. However, as a shadow is regarded as an area that is only partially illuminated due to the obstruction of the light source due to the moving object's opacity, if it is assumed that the illumination source only consists of white light, the chromaticity tend to stay similar in the shadowed as well as the neighboring illuminated regions. Hence, a normalized color space r = R + G + B, is immune to shadows though with a loss of color information. For instance, if the foreground target is a white vehicle moving against a grey background, there is no color information involved and hence, the foreground detection algorithm will lose track of it due to white and grey pixels having same chromaticity coordinates. In fact, in the r, g space, all the gray $(R = G = B)$ line projects to is a point $(\frac{1}{3}, \frac{1}{3})$ in the space. Therefore, usage of a brightness variable is of a third variable is mandatory to cater for chromaticity-matching pixels. However, the chromaticity variables do not change under shadows where the variable $s = R + G + B$ is expected to change under limits in the case of object shadows. Moreover, another approach that can be considered for shadows is the fact that texture information of background pixels remains consistent under shadow. Therefore, usage of an image gradient feature to cast shadows via a background edge or gradient model along the chromaticity model is widely used in the literature for shadow modeling [7].

A wide range of pixel-level shadow removal methods have been used that employ brightness or color distortions. Yet, the challenging aspect of detecting shadow pixels remains due to two main challenges in the classification of foreground results from background: the normal color variation over a set of frames may cause some foreground regions to behave as part of the non-required shadow foreground pixels resulting in false negatives. Also, most of the proposed methodologies tend to tune their foreground detection models via manually configured shadow pixels.

A. Background subtraction via intensity and chromatic color information

Selection of an appropriate color space to accurately model the underlying color distortion measure is paramount in accurate realization of foreground and background models in a sequence of frames. A number of different color spaces have been used to model foreground pixels along with shadows and highlights. The tri-chromatic color space is regarded as a model closest to human perception that utilizes three variables Y, Cb and Cr to describe a point in color space though the RGB model corresponds more closely with the sensor cones in the human eye or most colored CCD devices. However, the actual perception of color quantifies more closely to the Hue, Saturation, and Intensity (HSI) model. As discussed earlier, the intensity value in this space represents the overall brightness of a pixel whereas the Hue value represents the core color value separated at 120^0 angles. The saturation value depicts the purity of a color where, lesser shades have low saturation values compared to pure spectral colors [8]. [8] presented an edge detection approach based on the vector angle (VA) and Euclidean distance (ED) between adjacent RGB pixel values.

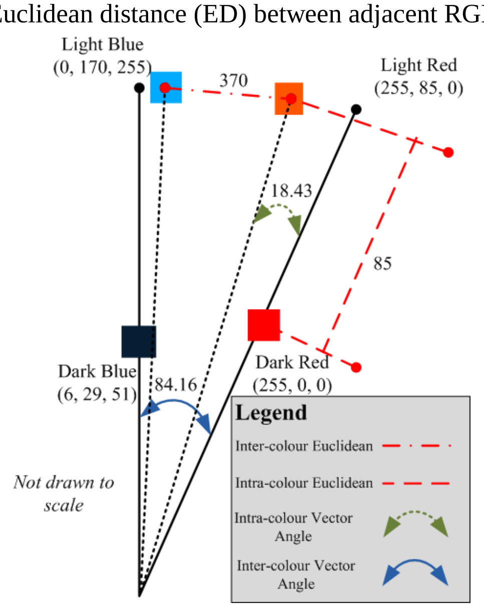

Fig. 1 Vector and Euclidean geometry measurement between four sample colors at variable chromatic (light blue/light red) and luminance (light red, dark red) levels

The edge detection methodology by [8] used VA to detect pixel chromaticity variation independent of luminance or intensity for edge detection. As shown in Fig. 1, the approach showed promising scope to be used to identify color-variations between background and foreground pixels. Based on this attribute, VA measurement was initially used to detect pixels under the influence of shadows or highlights. ED measurement was another chromatic identifier used by [8] to find detect both across luminance and chromaticity boundaries. Based on this attribute, ED was also cross-analyzed for its usefulness against VA to identify color and intensity-based pixel variations (Fig. 2). The figure shows ED measurement to represent color variation both at chromatic (370 between light blue and light red) and luminosity (85 between light red to dark red). The VA methodology, on other hand, is shown in represent chromaticity difference only via an angle measured as 84.16^0 between different colors and 18.43^0 between similar colors with different intensity levels.

The main challenges faced by conventional background subtraction algorithms are due to abrupt intensity variations. Changes in background light intensities in an image may be due to a number of factors including cloud cover, object shadows or highlights. These factors are generally induced when background pixels change intermittently due to dynamic objects such as tree leaves, waves, reflections, etc. Based on the concept of color and intensity modeling presented above, these variations were originally envisaged to be modeled via a VA/ED-based codebook model. However, Fig. 2 (b) shows a shortcoming in the use of VA technique to detect identical colors with changing intensities. This indicates a potential inability of VA to detect false foregrounds such as shadows which eventually lead to its removal from the proposed algorithm.

The issue of intensity variation due to changing background attributes such as cloud cover, time-variant shadows or highlights can be handled via a technique introduced by [5]. The technique alone however cannot be sufficiently used to filter a wide range of code-words in codebooks. Moreover, moving a codeword into a codebook merely on the basis of minimum and maximum intensity comparison is not sufficient to identify foreground pixels from background pixels especially when both assume similar color information in addition to similar intensity values. Additionally, [5] implemented a variable-length codebook based upon the degree-of-variability of pixels. Despite its promising results, the approach is likely to induce unwanted codebook entries or updated especially due to multiple pixel changes with multiple overlapping backgrounds as in the case of pedestrian walkways or intermittently flowing traffic passage which is likely to create a massive codeword length for each pixel over extended periods of time.

Based on the issue discussed above, this work presents a new technique of Intensity-kurtosis to update a codebook based on the detection of abnormal background update due to intensity changes. Moreover, the work eliminates the issue of unrestricted codebook size by limiting intensity changes into a set of groups or containers. Therefore, 10 containers are assigned for to pixel, covering roughly a 25-level intensity change over a 256 value intensity variable. The proposed approach is further explained in the next section.

III. Methodology

The ED models are commonly used to create intensity invariant color detection mainly for the purpose of segmentation or edge detection. Based on the idea adapted from [8], the proposed algorithm utilizes a straight forward approach to incorporate both chromaticity (color) and luminance (intensity) information in a single model. The principle used here is to utilize RGB Euclidean distance between the modeled and evaluated pixels to calculate a measure of intensity change in the background either due to an alien object or a dynamically background (See Fig. 1). An efficient modeling of this measure was envisaged to robustly differentiate between changes induced due to changing backgrounds or foreground objects. The distance for intensity changes between two temporally consecutive pixels is given in (3):

$$\mathrm{ED}(p_{i,j}(t)) = \|\hat{c} - \check{c}\| \tag{3}$$

Where $\|.\|$ is the L_2 vector norm, $c = [\![r\ g\ b]\!]^T$ for RGB color space and $\hat{c}\ and\ \check{c}$ are the maximum and minimum color intensity values observed over a pixel $p_{ij} \in N$ where N is the total frames being observed. Based on this condition, the Euclidean distance is calculated as given in (4):

$$\|\hat{c} - \check{c}\| = [\![(R_1 - R_2)^2 + (G_1 - G_2)^2 + (B_1 - B_2)^2]\!] \tag{4}$$

As the object shadows are merely an obstruction of light where, based on the pixel location over the shadow region, the pixel color remains same with a loss or gain in the intensity values due to presence of shadows or highlights respectively.

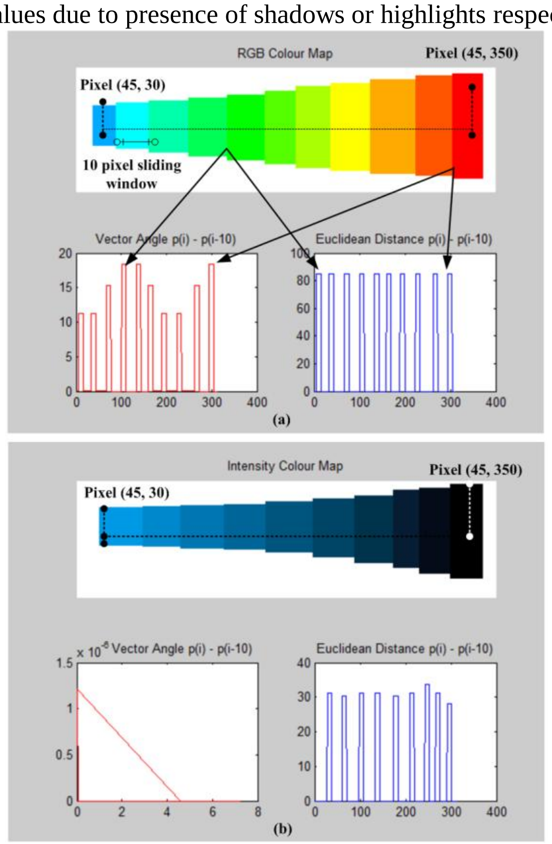

Fig. 2 Geometrical variation based on ED and VA-based variation measured via a 10-pixel-wide sliding window-based assessment of multicolor and single-color-multiple-intensity palettes shown in (a) and (b) respectively

For the case of a sequence of pixels undergoing changes due to intensity variations, the measure of change can be detected by statistically modeling the "measure of peakedness" of the intensity variable along a temporal scale. Kurtosis is a statistical technique used as a shape descriptor for the probability distribution of variables changing overtime. The attribute represents the so-called "degree of peakedness" of a dataset which is represented as a ratio of the fourth central moment of the data to its squared variance. Considering the case of a sinusoidal signal, it can be seen that kurtosis can be used to detect minor changes compared to the central moments [9]. In the current case, the kurtosis of intensity variable of pixels changing on a time scale was evaluated to measure deviation from a normal Gaussian behavior. Such a change is zero for a process with Gaussian distribution and infinite for a linear distribution. For a pixel, it was noted that a change in intensity regardless of its duration was reflected in the corresponding kurtosis value as shown in Fig. 3 for pixel changes shown in Fig. 4 and Fig. 5.

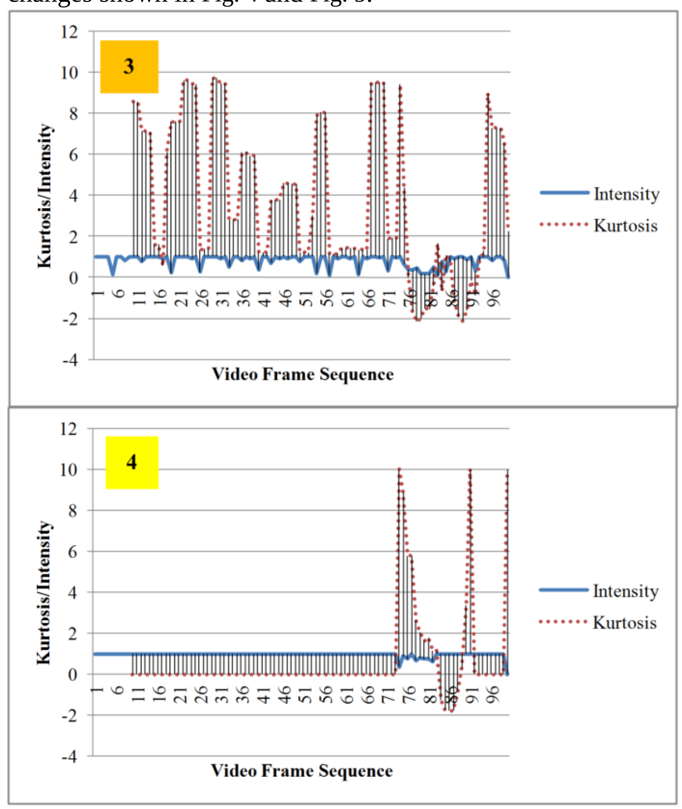

Fig. 3: Kurtosis profile of pixel locations (3) and (4) shown in Fig. 4 and Fig. 5 showing substantial changes for even slight intensity variations over a duration of a 10 frame sliding window

A. Modified codebook model (MCB-HED) based on ED, Hue and Intensity-kurtosis

The two ED/VA measures were analyzed and plotted against a time-based sliding window. The respective graphs show the 10-value intensity and color variations in Fig. 2. It must be noted that for color variations shown in Fig. 2(a), both VA and ED performed well. However, for various single-color intensities of blue or any other color so-to-speak, VA failed to properly represent the underlying intensity changes. Yet, the measure efficiently modeled color variations for various color groups such as different shades of blue, green and red as shown in Fig. 2(a). Based upon the inability of VA-based approach to detect incorporate intensity changes in images with little or no chromatic variation, ED-based technique was selected as one of the parameters in the proposed codebook model. Moreover, another issue addressed in this work is that of the presence of background objects for a short-to-extended period of time such as leaves or clouds. Presence of such background pixels imparts a substantial challenge especially when a color similar to the background is present.

The kurtosis K_i^4 for temporal intensity values I can be defined as a fourth-order stationary zero-mean, stochastic process. K_i^4 is defined by [10] as the normalized zero-order cumulant for zero time lag which is defined as given in (5):

$$K_i^4 = \frac{E\{I^4\}}{(E\{I^2\})^2} - 3 \tag{5}$$

Where $E\{\cdot\}$ represents the expected value operator and the '-3' is often regarded as a correction made to make a normal distribution kurtosis equal to zero. According to [11], kurtosis of any process exist in a confidence interval that is based upon the probability properties of the estimator. For the case of intensity variation modeling, given a desired confidence interval, the estimator can be frames between two values depending upon the first statistics of the estimator.

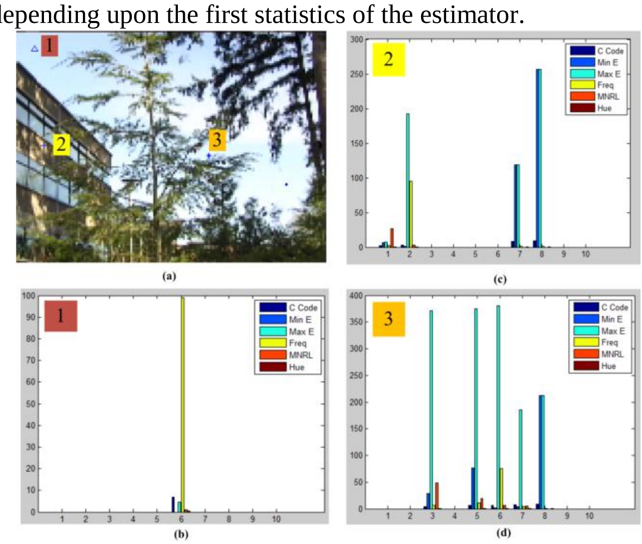

Fig. 4 Frame 1 containing four unique pixel locations analyzed for various intensity and ED-based color distortions with (1) representing uniform sky location, (2) moving tree shadow, (3) moving leaves against sky and changing cloud cover and (4) changing cloud cover against variable cloud cover

Therefore, based on [12], for an $N-sample$ intensity signal $\{I(m,n)\}, n = 1, \ldots N$ where N is the total image intensity matrix frames, its kurtosis K_i^4 can be formulated as follows:

$$K_i^4 = (N-1)\frac{\sum_{n=1}^{N} I^4(n)}{\left(\sum_{n=1}^{N} I^2(n)\right)^2} - 3 \tag{6}$$

Based on (6) for the image shown in Fig. 4, four unique pixel locations were analyzed to assess the behavior of kurtosis in highly dynamic images containing a diverse set of intensity changes and distributions. The image is taken from a change detection benchmarking dataset located at [13] covered under the work done by [14].

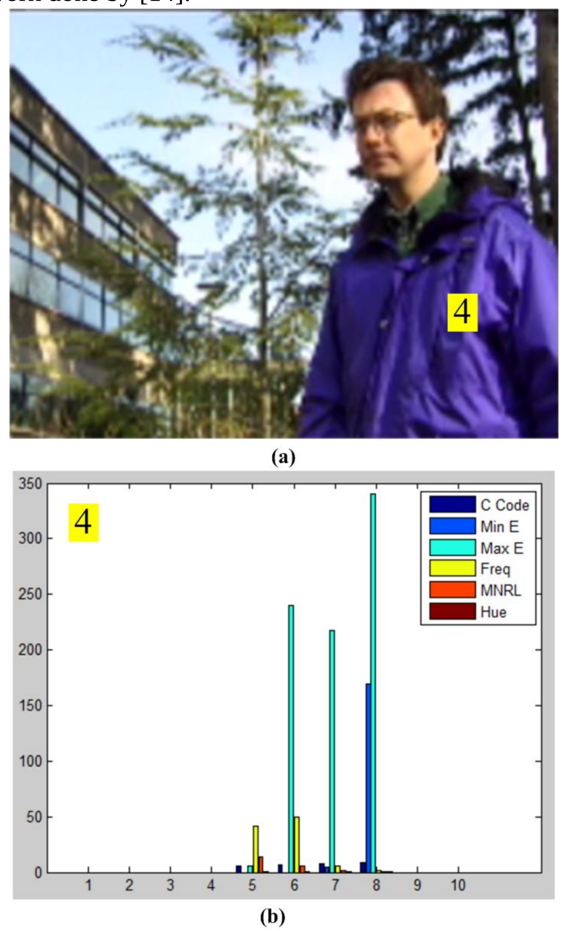

Fig. 5 Pixel '4' with a person moving-in at the 245th frame of the dataset depicting substantial change in pixel intensity compared to other pixels shown in Fig. 4

B. Modified Codebook construction

The MCB-HED model proposed in this work builds a background model based on continued, pixel-based scene observation. For each pixel in the sequence of temporally distributed frames, codewords are saved on the basis of 10 intensity groups. Each of these groups represent the underlying intensity distribution ranging from 0 – 255. The number of codewords for each of these groups or "codeword containers" depends upon the measure of background variation present at that pixel location. For instance, pixel (1) shown in Fig. 4 only contains a single codeword in container 6 (See Fig. 4 (a)) due to a low intensity variation area. On the other hand, pixel (3) is the most diverse pixel location in the codebook which is exposed to a wide range of intensity and chromaticity changes due to moving leaves, cloud cover and eventually a foreground body of a person moving-in (See Fig. 5 (a) and (d)).

Each codeword is represented by a RGB vector $\vartheta_i = (\bar{R} + \bar{G} + \bar{B})$ and a tuple $X \Leftarrow < \hat{\varepsilon}_i, \breve{\varepsilon}_i, \varphi_i, \lambda_i, \tau_i^o, \tau_i^n, K_i^4 >$ where $\hat{\varepsilon}_i$ and $\breve{\varepsilon}_i$ are the maximum and minimum Euclidean distances between frames are k and $k-1$, φ_i is the frequency with which the codeword has repeated, λ_i is time span in which the codeword has not repeated itself [5, 6], τ_i^o and τ_i^n are the first and last time the codeword occurred and K_i^4 is the kurtosis based upon previous $k-10$ frames.

For a frame sequence $\bar{I} := (i_1, i_2, \ldots, i_N)$, a codebook $\mathfrak{B} = (\beta_1, \beta_2, \ldots, \beta_M)$ where N is the total number of frames and M are the total codeword entries for container κ_i of the pixel $p_{k,l}$. The size of each container's codebook varies, depending upon the degree of variability present for that pixel location. The algorithm for background codebook model construction is given in Table I with the actual background subtraction algorithm explained in Table II:

TABLE I
MCB-HED ALGORITHM

For each pixel $p_{i,j} \in M(r,c)_{RGB}$, at frame n
$\forall\, p_{k,l} \in M(r,c)$ and $\forall\, n \in N$ compute the following:

$$ED\left(p_{i,j}^{N} / p_{i,j}^{N-1}\right)$$
$$K_i^4\left(p_{i,j}^{N} / p_{i,j}^{N-10}\right)$$

If $\mathfrak{B} = \emptyset$ *or* no match in codebook then $\beta = \beta + 1$ and generate a new codebook entry β_{k+1}

$$\vartheta_i \Leftarrow (\bar{R} + \bar{G} + \bar{B})$$
$$X_i \Leftarrow < \varepsilon, \varepsilon, 1, t-1, t, t, \infty >$$

Else, update the codebook β_m

TABLE II
BACKGROUND SUBTRACTION ALGORITHM

$$\vartheta_i = (\bar{R} + \bar{G} + \bar{B}), \varepsilon_i$$

$\forall\, \beta_k \in codebook\ \mathfrak{B}$ look for codeword β_k matching ϑ_i based on the following two conditions:

$$\varepsilon_i \leq T_1$$
$$K_i^4 \leq T_2$$

Update the matched codeword as shown in Table 1, Step 4

$$p_{i,j} = \begin{Bmatrix} 1\ if\ matched \\ 0\ if\ no\ match\ found \end{Bmatrix}$$

IV. RESULTS AND OUTCOMES

The constructed codebook model was evaluated against a benchmarking dataset containing numerous moving individuals with shadows cast on the ground and walls. The training or codebook construction phase was performed over the first 100 frames of the "bus station" dataset [13].

A. Evaluation against benchmarking datasets

The proposed approach was evaluated against two benchmarking datasets. The preliminary analysis were performed over the waving tress dataset containing a total of 286 frames and contained variable background movement including clouds, leaves and shadows cast over the building on the left. The dataset was primarily used to act as a proof-of-concept for the MCB-HED algorithm for its susceptibility against background variability in the presence of shadows, branches and cloud cover. For each dataset, the system was trained with the initial 100 frames and evaluated against a selected set of frame sequences with variable level of activity and illumination change present. The outcome shown in Fig. 7 was taken from the "bus station dataset" with 1250 frames. Fig. 7 contains two individuals moving in opposite directions with the individual on the right having his shadow cast on the brick wall and his reflection present on the glass panels. The system was trained over the first 100 frames with little background variation. The constructed codebook was finally evaluated against frames 1020 – 1110 containing a number of foreground bodies of individuals moving in different directions with their shadows present on the ground and walls.

B. Background Modeling via MCB-HED algorithm

Table III represents the codebook entries of various hue/intensity-based variations present at pixel 'P' shown in Fig. 6. The pixel was chosen as it contained a person moving through it with his shadow following shortly thereafter. An analysis of the table shows two groups of data shown as light and dark grey rows. The Hue-based CCode containers shown at the bottom of the table are recorded in the codebook when the person's body moves over the pixel. The dark-grey "CCode 2" column represents the overall code for the background which is apparent from its high 'Freq' value of 55. The shortest frequency entry in the codebook is that of "CCode 9" which appeared when the person's right arm crossed the pixel at 1056th frame. The remaining 2 pixels represent the person himself.

Most importantly, it must be noted that the person's shadow was successfully integrated into the "CCode 2" codebook entry based upon the Hue model. This can further be seen in Fig.

TABLE III
MCB-HED-BASED BACKGROUND SUBTRACTION

CCode	ED	Min(E)	Max(E)	Freq	MNRL	Ftime	Ltime	Hue	Code insertion
2	0	0	171.72	55	15	1031	1099	0.05	Due to normal background
9	5.83	5.8309	5.83	1	0	1056	1056	0.78	Individual's right arm
10	2.23	1.7320	11.18	9	2	1050	1060	0.86	Individual's body
11	5.09	4.4721	33.54	4	7	1049	1062	0.96	Individual's body

6(a) and (d) which represents a selected frame once the individual shown on the right has passed pixel ‘P’. The intensity profile of these sequences shows the shadow from local frame count 52 to 65 which justifies the incorporation of the additional “shadow removal parameter” Hue which successfully removes the shadow. The Euclidean MNRL parameter shows the four unique transitions that resulted into codebook entries (CCode 2, 9, 10, 11) in Fig. 6(c). The first peak appears when the person’s body (arm) moves over the pixel, the second when his body moves away from the pixel, third when the shadow moves-in and fourth when the shadow moves out. It must be noted here that the two shadow entries are integrated into CCode 2 background entry whereas the arm, person’s body moving-over and moving-out are recorded as three codebook entries (CCode 9, 10 and 11).

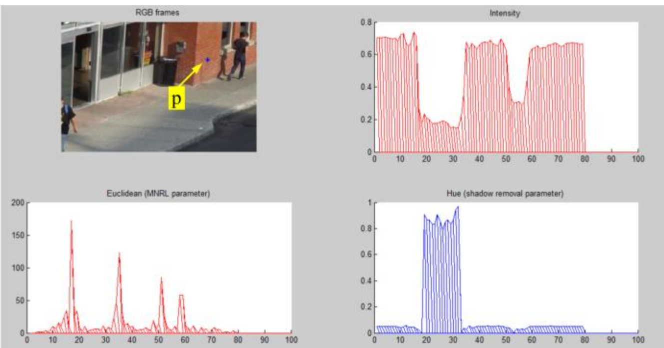

Fig. 6: Outcome of MCB-HED algorithm over pixel P with remaining frames showing codebook parameter outcomes for intensity, ED and Hue variables

The actual background subtraction algorithm given in Table II was implemented as a standard Matlab routine run against the “bus station” sequence shown in Fig. 7(a). The initial test outcome shown in Fig. 7 (b) shows a MOG-based outcome failing at shadow detection. Both (b) and (c) were evaluated against Fig. 7 (b) shows a substantial disadvantage in using MOG-based techniques for scenes with abrupt intensity changes as was the case in these frames due to the presence of shadows. Since the shadows presented a spontaneous change of intensity at pixels, their presence was not only detected as unwanted regions, the underlying change detection model was tuned for long enough a duration to miss fast moving objects present in this scene. Moreover, the technique also detected the reflection cast on glass panel by the individual on the right. Fig. 7 (c) shows the outcome based on the proposed MCB-HED algorithm. The segmented foreground images of two individuals walking in opposite directions were detected while the shadow cast by the person on the right was effectively removed by the model. The overall outcome shows promising prospects of continuing with this investigation to further challenging issues as discussed in the later section.

(a)

(b)

(c)

Fig. 7: (a) Background subtraction outcome for the bus station benchmarking dataset evaluated with (b) MOG and (c) proposed methodology

V. CONCLUSIONS

The paper presents a novel background subtraction methodology based upon a Euclidean distance-based intensity and color distortion modeling. These two measures were integrated into a proposed codebook model and analyzed and tested against a standard benchmarking database. The methodology utilized kurtosis assessment of time-based intensity profiles of pixels to add further pixel variations in the codebook occurring due to unusual intensity variations. The implementation outcomes with two benchmarking datasets present promising results where the reflections and shadows are successfully removed. The study was primarily performed in order to further continue the work against more complex issues such as multiple, overlapping backgrounds, behavioral profiling, temporal scene analysis and most importantly, behavioral analysis of individuals such as abandoned objects, improperly parked cars or trespassing in security sensitive environments. Moreover, via a preliminary analysis it is understood that the methodology can be further improved and utilized with the integration of advanced AI methodologies such as neural-fuzzy systems, Markov models with their performance further improved via various evolutionary optimization techniques.

ACKNOWLEDGMENT

The authors would like to offer their thanks to the Knowledge Transfer Partnerships/Technology Strategy Board for their support. The project is a joint venture between STS Defence Ltd and University of Portsmouth.

REFERENCES

[1] Stauffer, C. and W.E.L. Grimson, Learning patterns of activity using real-time tracking. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000. 22(8): p. 747-757.

[2] Lee, D.S., et al., A Bayesian framework for Gaussian mixture background modeling. 2003 International Conference on Image Processing, Vol 3, Proceedings, 2003: p. 973-976.

[3] Harville, M., A framework for high-level feedback to adaptive, Per-Pixel, Mixture-Of-Gaussian background models. Computer Vision - Eccv 2002 Pt Iii, 2002. 2352: p. 543-560.

[4] Javed, O., K. Shafique, and M. Shah, A hierarchical approach to robust background subtraction using color and gradient information. Ieee Workshop on Motion and Video Computing (Motion 2002), Proceedings, 2002: p. 22-27.

[5] Kim, K., et al., Real-time foreground-background segmentation using codebook model. Real-Time Imaging, 2005. 11(3).

[6] Ilyas, A., et al., Real Time Foreground-Background Segmentation Using a Modified Codebook Model. Avss: 2009 6th Ieee International Conference on Advanced Video and Signal Based Surveillance, 2009: p. 454-459.

[7] Moeslund, T.B. and SpringerLink, Visual analysis of humans [electronic resource] : looking at people. 2011, London ; New York: Springer-Verlag London Limited. 1 online resource (xxi, 632 p.).

[8] Dony, R.D. and S. Wesolkowski, Edge detection on color images using RGB vector angles, in Electrical and Computer Engineering, 1999 IEEE Canadian Conference on1999. p. 687-692.

[9] Pakrashi, V., B. Basu, and A. O' Connor, Structural damage detection and calibration using a wavelet-kurtosis technique. Engineering Structures, 2007. 29(9).

[10] Nikias, C.L. and A.P. Petropulu, Higher-order spectra analysis : a nonlinear signal processing framework. Prentice-Hall signal processing series. 1993, Englewood Cliffs, N.J: PTR Prentice Hall. xxii, 537p.

[11] Ravier, P. and P.O. Amblard, Wavelet packets and de-noising based on higher-order-statistics for transient detection. Signal Processing, 2001. 81(9).

[12] Bickel, P.J. and K.A. Doksum, Mathematical statistics : basic ideas and selected topics. 2nd ed. 2007, Upper Saddle River, N.J.: Pearson Prentice Hall. xviii, 556 p.

[13] ChangeDetection. ChangeDetection Video Database. 2012 [cited 2012 19th August, 2012]; Available from: http://www.changedetection.net/.

[14] Goyette, N., et al. changedetection.net: A new change detection benchmark dataset. in Proc. IEEE Workshop on Change Detection (CDW'12). 2012. Providence, RI.