Object Tracking under Illumination Variations using 2D-Cepstrum Characteristics of the Target

Fuat Cogun and A. Enis Cetin

Department of Electrical and Electronics Engineering, Bilkent University
Bilkent 06800, Ankara, TURKEY
{fuatc, cetin}@bilkent.edu.tr

Abstract—**Most video processing applications require object tracking as it is the base operation for real-time implementations such as surveillance, monitoring and video compression. Therefore, accurate tracking of an object under varying scene conditions is crucial for robustness. It is well known that illumination variations on the observed scene and target are an obstacle against robust object tracking causing the tracker lose the target. In this paper, a 2D-cepstrum based approach is proposed to overcome this problem. Cepstral domain features extracted from the target region are introduced into the covariance tracking algorithm and it is experimentally observed that 2D-cepstrum analysis of the target object provides robustness to varying illumination conditions. Another contribution of the paper is the development of the co-difference matrix based object tracking instead of the recently introduced covariance matrix based method.**

I. INTRODUCTION

Visual object tracking is defined as the process of estimating the location of the moving object in the current image frame given all previous frames of a video sequence. Tracking of moving objects is one of the most important tasks in computer vision as object tracking algorithms are used in many applications such as security surveillance, traffic flow analysis, and content-based video compression.

Robustness to varying illumination conditions is crucial for visual tracking algorithms because illumination variations are unavoidable in real-world environments. Globally or locally changing illumination conditions of the observed scene constitutes an important challenge for many video processing applications including object tracking algorithms. Illumination variations may cause misclassification of some background objects as moving objects or lost targets due to changes in the target model parameters.

There has been many algorithms proposed for object tracking in the past. Covariance matrix based object tracking method proposed by Porikli et al. [1] is used as the base line tracker in this study. Covariance tracking uses a covariance matrix based object description measure. Since the target model is represented as the covariance matrix of features covering the target region, the method is applicable to non-stationary camera sequences and less susceptible to noise than other methods mainly based on background subtraction. In

addition, computationally efficient co-difference matrix based object tracking method [2] introducing a new operator to replace the multiplication operator of the covariance matrix is also examined. Co-difference tracking method differs from the covariance tracking method in the formation of the matrix describing the object region. Co-difference method can be implemented without performing any multiplication leading to a computationally efficient tracking method.

Ideally, the color and texture information of objects are retained in images under changing light intensities. Therefore, for robustness to varying illumination conditions, measures independent of the light intensity that uses color and texture information to represent the target are required. In this study, two-dimensional (2D) cepstrum analysis of the target is used because the cepstrum is an amplitude invariant feature extraction method widely used in speech processing.

In this paper, a novel object tracking algorithm increasing the robustness of both covariance and co-difference tracking methods under varying illumination conditions is proposed. The proposed object tracking algorithm introduces the 2D-cepstrum analysis of the target region to the covariance and co-difference tracking methods. The light intensity-independent 2D-cepstrum coefficients of the target region are used to increase the robustness of the object tracking algorithms to varying illumination conditions.

The next section underlines some related work on object tracking under varying illumination conditions. Section III gives an overview of the covariance and co-difference based object tracking methods and presents the proposed 2D-cepstrum based object tracking method. Results of the proposed method implementation and comparisons with co-variance and co-difference tracking methods are presented in Section IV. The final section presents conclusions.

II. RELATED WORK

In recent years, many algorithms have been proposed to deal with robust object tracking under varying illumination conditions.

Background subtraction approaches are widely used to detect moving objects [3][4]. However, these approaches are susceptible to illumination variations of the background.

Some work in the literature discards the illumination-sensitive color information by using other features which are less-sensitive to illumination variations such as edges and

textures. In [6], a model-based method where edge information is used to capture hand articulation by learning hand natural constraints is implemented. In [7], the head is modeled as a texture mapped cylinder and tracking is formulated as an image registration problem for the cylinder's texture map image.

Other methods using only the color information to represent targets and their background are existing. In [9], a color modeling approach including intensity information in HSI color space using B-spline curves is used. Yang and Waibel [10] detected human faces by using a normalized RG plane and proposed a color-model adaptation algorithm based on the observation that shape of the histograms remain similar under illumination change. In [11], Gaussian mixture models are used to estimate probability densities of color for target and background objects. In their work, a technique for dynamically updating the models to accommodate changes in color due to varying illumination conditions is also introduced.

Methods combining both the color and texture information for robust object tracking are also available in the literature. The fusion of appearance and structural information in [12] is done using the condensation algorithm. In [5], illumination insensitive features are extracted from both target and background objects using the Bayesian framework for robust tracking. A method for moving object detection based on the background modeling and subtraction which uses both color and edge information is proposed in [8]. In this study, confidence maps are introduced to fuse intermediate results and represent the results of the background subtraction. Li and Leung [13] proposed a method which uses texture information by calculating the quotient between the cross-correlation and auto-correlation of a gradient vector to eliminate brightness variations.

III. Moving Object Tracking Method

The proposed tracking method introduces 2D-cepstral domain coefficients of the target region into the covariance and co-difference matrices that are used to describe the target characteristics. The tracking algorithms become more reliable under changing illumination conditions because cepstrum is an amplitude invariant feature extraction method.

In this section, first the covariance and co-difference tracking algorithms are introduced. In the second subsection, the 2D-cepstrum analysis of the target region and its incorporation to the tracking algorithms are presented.

A. Covariance and Co-difference Tracking Algorithms

Both covariance and co-difference tracking algorithms require the computation of a matrix representing the given target region by using the feature images formed for each frame to construct the target model. The covariance of feature vectors describing the target is called covariance matrix in the covariance tracking method. Similarly, the co-difference matrix is computed from feature vectors in the co-difference tracking method to model the moving target. In both covariance and co-difference methods, the aim is to find the region in a given

image frame having the minimum distance from the target model matrix and assign this region as the estimated location of the moving target at that frame. The first step of tracking algorithms is feature vector construction from a given image or image region.

1) Feature Images and Vectors: Let the observed m-dimensional image denoted as $\mathbf{I}$. Then, the corresponding m-dimensional feature image, $\mathbf{F}$, can be written as:

$$\mathbf{F}(x,y) = \gamma(\mathbf{I}, x, y)$$

where $\gamma(.)$ can be any mapping feature such as color, filter responses, image gradients $\mathbf{I}_x$, $\mathbf{I}_y$, $\mathbf{I}_{xx}$, ..., temporal frame differences, edge magnitudes, etc.

For a given window region $\mathbf{R} \subset \mathbf{F}$, let $\{\mathbf{f}_k\}_{k=1,...,n}$ be the d-dimensional feature vectors inside $\mathbf{R}$. Feature vector $\{\mathbf{f}_k\}$ is constructed using two types of mappings; spatial mappings using the pixel coordinates and appearance mappings using color and gradient values. The feature vector used in this work is:

$$\mathbf{f}_k = [\, x \;\; y \;\; \mathbf{I}(x,y) \;\; \mathbf{I}_x(x,y) \;\; \mathbf{I}_y(x,y) \;\; \mathbf{I}_{xx}(x,y) \;\; \mathbf{I}_{yy}(x,y) \,]^T$$

where x and y are the pixel coordinates, $\mathbf{I}(x,y)$ is the color value and $\mathbf{I}_x(x,y), \mathbf{I}_y(x,y), \mathbf{I}_{xx}(x,y), \mathbf{I}_{yy}(x,y)$ are the gradients along the x and y directions.

2) The Covariance Matrix: The second step of the tracking algorithms is the computation of the covariance matrix which is formed by using the feature vectors constructed in the previous step. To represent an $M \times N$ rectangular region $\mathbf{R}$, the covariance matrix of feature points of $\mathbf{R}$ is defined as:

$$\mathbf{C_R} = \frac{1}{MN} \sum_{k=1}^{MN} (\mathbf{f}_k - \mu_{\mathbf{R}})(\mathbf{f}_k - \mu_{\mathbf{R}})^T$$

where $\mu_{\mathbf{R}}$ is the mean of the feature vectors belonging to region $\mathbf{R}$.

3) The Co-difference Matrix: In the co-difference tracking method the co-difference matrix is computed from the feature vectors described in III-A1. To represent an $M \times N$ rectangular region $\mathbf{R}$, the co-difference matrix of feature points of $\mathbf{R}$ is defined as:

$$\mathbf{D_R} = \frac{1}{MN} \sum_{k=1}^{MN} (\mathbf{f}_k - \mu_{\mathbf{R}}) \odot (\mathbf{f}_k - \mu_{\mathbf{R}})^T$$

The operator $\odot$ acts like a matrix multiplication operator, however, the scalar multiplication is replaced by an additive operator $\oplus$ which is defined as follows

$$a \oplus b = \text{sign}(a \times b)(|a| + |b|)$$

Since $a \oplus b = b \oplus a$, the co-difference matrix is also symmetric. The new operator decreases the computational cost of tracking by replacing the multiplication operator by the addition operator in some processors.

For d-dimensional feature vector sets, the corresponding covariance and co-difference matrices of features has size $d \times d$. These matrices of the feature points inside $\mathbf{R}$ are used to represent region $\mathbf{R}$. Initially, the covariance and co-difference

matrices are used to model the target and in the following image frames of the video they are used to find the estimated target locations.

4) Distance Metric and Target Location Estimation: In both covariance and co-difference tracking methods, to obtain the most similar region to the given target window, distances between matrices corresponding to the target window and the candidate regions are calculated. Since the space of covariance and co-difference matrices are not vector spaces, subtraction of two matrices would not be a valid measure. Therefore, a distance metric needs to be introduced for finding the best candidate region match and estimating that region as the target location. The distance metric used in both tracking algorithms computes the dissimilarity between matrices as

$$\rho(\mathbf{C}_i, \mathbf{C}_j) = \sqrt{\sum_{k=1}^{d} \ln^2 \lambda_k(\mathbf{C}_i, \mathbf{C}_j)}$$

where $\{\lambda_k(\mathbf{C}_i, \mathbf{C}_j)\}$ is the set of generalized eigenvalues of matrices $\mathbf{C}_i$ and $\mathbf{C}_j$.

At each frame, neighboring regions of the previously estimated location of the target are defined as the candidate regions. The descriptive matrix of these candidate regions are computed and the region with the smallest distance to the matrix representing the target is assigned as the estimated target location in that image frame. This operation is repeated for each frame.

B. Two-dimensional (2D) Cepstrum Analysis of the Target

The proposed cepstrum analysis method includes the computation of the 2D-cepstrum of the initial target window and the candidate regions at each frame. The 2D-cepstrum analysis is used as cepstrum is an amplitude invariant feature extraction method, therefore, cepstral domain coefficients of a region remains unchanged under light intensity variations. This property of cepstrum provides robustness to illumination variations at the target region.

Cepstral domain feature extraction is widely used in speech processing applications [14]. The cepstrum operator was first introduced by Tukey [15]. The real cepstrum $\hat{x}[n]$ of a signal x is defined as the inverse Fourier transform of the log-magnitude Fourier spectrum of x.

Let $x[n]$ be a discrete signal. Its cepstrum $\hat{x}[n]$ is defined as follows:

$$\hat{x}[n] = F^{-1}\{\ln(|F\{x[n]\}|)\}$$

where $F\{.\}$ represents the discrete-time Fourier Transform, $|.|$ is the magnitude, $\ln(.)$ is the natural logarithm and $F^{-1}\{.\}$ is the inverse discrete-time Fourier Transform operator. In our approach, 2D-cepstrum is used. For a region $\mathbf{R}$, 2D-cepstrum of $\mathbf{R}$, $\hat{\mathbf{R}}$ is defined as follows:

$$\hat{\mathbf{R}} = F_{2D}^{-1}\{\ln(|F_{2D}\{\mathbf{R}\}|)\}$$

where $F_{2D}\{.\}$ is the 2D discrete-time Fourier Transform and $F_{2D}^{-1}\{.\}$ is the inverse 2D discrete-time Fourier Transform operator.

Let the initial $M \times N$ target window denoted as $\mathbf{W}$ and the shadowed version of the target window be represented as $\mathbf{W_s}$. According to [16], when the region of interest is shadowed, its intensity is scaled by a constant factor throughout that region. In our case, this statement corresponds to:

$$\mathbf{W_s} = \alpha \mathbf{W} \qquad (1)$$

where α is a positive real number less than 1 for the target window $\mathbf{W}$. When we compute the 2D-cepstrum of both sides in Eq. 1 we obtain:

$$\hat{\mathbf{W}}_\mathbf{s} = K_\alpha \delta(x,y) + \hat{\mathbf{W}} \qquad (2)$$

where δ is the dirac delta function, K_α is a constant, $\hat{\mathbf{W}}_\mathbf{s}$ and $\hat{\mathbf{W}}$ are the 2D-cepstrums of $\mathbf{W_s}$ and $\mathbf{W}$s, respectively. In the proposed 2D-cepstrum analysis, Eqn. 2 reveals the fact that the output 2D-cepstrum coefficients except the (0,0)-indexed (magnitude) coefficient remains unchanged under intensity variations for the analyzed region. That is:

$$\hat{\mathbf{W}}_\mathbf{s}(i,j) = \hat{\mathbf{W}}(i,j), \ \ \forall (i,j) \neq (0,0)$$

Therefore, to obtain additional target characteristics robust to illumination variations, the output 2D-cepstrum coefficients except the magnitude coefficient should be used as cepstral domain feature parameters.

Cepstral domain feature parameters of the target region are incorporated to the covariance and co-difference matrix as additional features. The approach introduced in this paper is to increase the size of the matrices and add the 2D-cepstrum coefficients except the magnitude coefficient found by analyzing the target region to the matrix.

For a d-dimensional feature vector set, the corresponding covariance or co-difference matrix of features has size $d \times d$. Let the matrix be denoted as $\mathbf{C}$. The modified covariance or co-difference matrix $\mathbf{C_m}$ including the output 2D-cepstrum coefficients is derived as

$$\mathbf{C_m} = \begin{bmatrix} \mathbf{C} & \mathbf{V} \\ \mathbf{V}^T & \mathbf{0} \end{bmatrix}$$

where $\mathbf{V}$ is $d \times z$ matrix containing output 2D-cepstrum coefficients and $\mathbf{V}^T$ is the transpose of $\mathbf{V}$. Therefore, additional $d.z$ values are included in the matrix for robustness. Notice that the modification is done in such a way that the symmetrical property of covariance and co-difference matrices is preserved.

IV. EXPERIMENTAL RESULTS

In this section, outcomes of the proposed algorithms are presented and comparisons with the covariance tracking approach are presented.

The experimental results are obtained for $\mathbf{V}$ selected as

$$\mathbf{V} = \begin{bmatrix} \hat{\mathbf{W}}(1,2)\hat{\mathbf{W}}(2,1)\hat{\mathbf{W}}(2,2)\hat{\mathbf{W}}(2,3)...\hat{\mathbf{W}}(3,4) \end{bmatrix}^T$$

where $\hat{\mathbf{W}}$ is the output 2D-cepstrum matrix of the target window $\mathbf{W}$. In this case, $d = 7$ and $z = 1$ and the corresponding modified covariance matrix has a size of 8×8.

The performance of the covariance, cepstrum-based covariance and cepstrum-based co-difference tracking methods

are measured using 10 video sequences adding up to more than 2000 frames. The video sequences are composed of moving and stationary camera recordings. In order to compare the performance of the tracking algorithms quantitatively, the approach taken in [17] is used. The detection rate is defined as the ratio of the number of frames the object location is accurately estimated to the total number of frames in the sequence. The estimated location is considered accurate if the estimate window center is within the 10×10 neighborhood of the tracked object center. Some of the resultant performances of the tracking methods are given in Table 1, 2 and 3. It is observed that the proposed algorithms increase the detection rates of the covariance tracking method and the cepstrum-based covariance tracking method performs slightly better than the cepstrum-based co-difference tracking method in most of the video sequences.

TABLE I: Performance of the Covariance Tracking Method

	Total Frames	Missed	Detection Rate
Shadowed Street	225	27	88.0
Woman walking	341	25	92.7
Dog running	156	11	92.9
Two friends	571	54	90.5
People group	125	7	94.4
Talking man	401	62	84.5
Pink shirt	364	22	93.9

TABLE II: Performance of the Cepstrum-based Covariance Tracking Method

	Total Frames	Missed	Detection Rate
Shadowed Street	225	12	94.6
Woman walking	341	7	97.9
Dog running	156	3	98.1
Two friends	571	21	96.3
People group	125	2	98.4
Talking man	401	17	95.8
Pink shirt	364	10	97.3

TABLE III: Performance of the Cepstrum-based Co-difference Tracking Method

	Total Frames	Missed	Detection Rate
Shadowed Street	225	16	92.9
Woman walking	341	12	96.5
Dog running	156	8	94.9
Two friends	571	16	97.2
People group	125	5	96.0
Talking man	401	22	94.5
Pink shirt	364	16	95.6

Some tracking algorithm outcomes are presented in figures. In Fig. 1, the man talking on the phone is walking into the building. When the man approaches to entrance of the building, the light intensity on him decreases. The covariance tracking method loses the target before he enters the building. However, the modified covariance and co-difference tracking methods track the man accurately until he enters the building. In Fig. 2, two men are walking into a shadowed area. Initially, the man wearing a pink shirt is introduced as target to all

tracking methods. It is observed in Fig. 2a that as illumination changes, the co-difference tracking method starts to lose the target and eventually the tracker completely loses the track of the target whereas in Fig. 2b and Fig. 2c, the modified covariance and co-difference tracking methods track the target successfully even there are abrupt light intensity changes in the scene. Two men walking into a darker region is tracked in Fig. 3 using the covariance tracking, the proposed modified covariance and co-difference tracking methods. The video sequence introduces a continuously decreasing target light intensity as the frames advances. The proposed object tracking algorithm manages to track the targets at each frame. However, the covariance tracking algorithm loses one of the targets at some point (Fig. 3a). In Fig. 4, a woman is walking into a darker region with her friends. In this case, Fig. 4a shows that although the co-difference tracker does not lose the target completely, the tracking is not robust due to the changes in the target model caused by illumination variations. It is observed from figures Fig. 4b and Fig. 4c that the proposed object tracking methods perform well by using the introduced additional cepstral domain target features. Fig. 5 can be seen in web-page [18]. The covariance tracking method differentiates the moving target from the clutter (other man) under sunlight successfully. However, when the target model changes due to illumination variation, the tracker fails and starts to track the clutter. In contrast, the proposed tracking methods are successful throughout the video sequence and even though the target gets into the shadowed region, they manage to track the target until it gets out of the video frame.

Fig. 1, 2, 3, 4 and 5 shows that the proposed object tracking methods perform better than the ordinary covariance and co-difference tracking methods under varying illumination conditions. The proposed methods are tested under abrupt illumination changes, continuously varying light intensity conditions and in the presence of a clutter. It is clear from our comparisons that there is need to adapt the changes in the target model under varying illumination conditions for robust object detection and accurate object recognition. It is observed that the introduction of the output 2D-cepstrum values of the target region to the covariance and co-difference matrices increases the robustness of the tracking algorithms to light intensity changes.

V. Conclusion

In this paper, an object tracking method based on cepstrum analysis is proposed for robust object tracking under varying illumination conditions. The proposed object tracking method combines the covariance tracking method and the 2D-cepstral features of the target region. The 2D-cepstrum is used because the cepstrum retains the underlying color and texture information under light-intensity variations. The method is applied to video sequences in which the intensity of the target region varies and it is experimentally observed that the proposed method produces better results than the ordinary covariance tracking method. Co-difference method provides a computationally efficient trade-off compared to the covariance
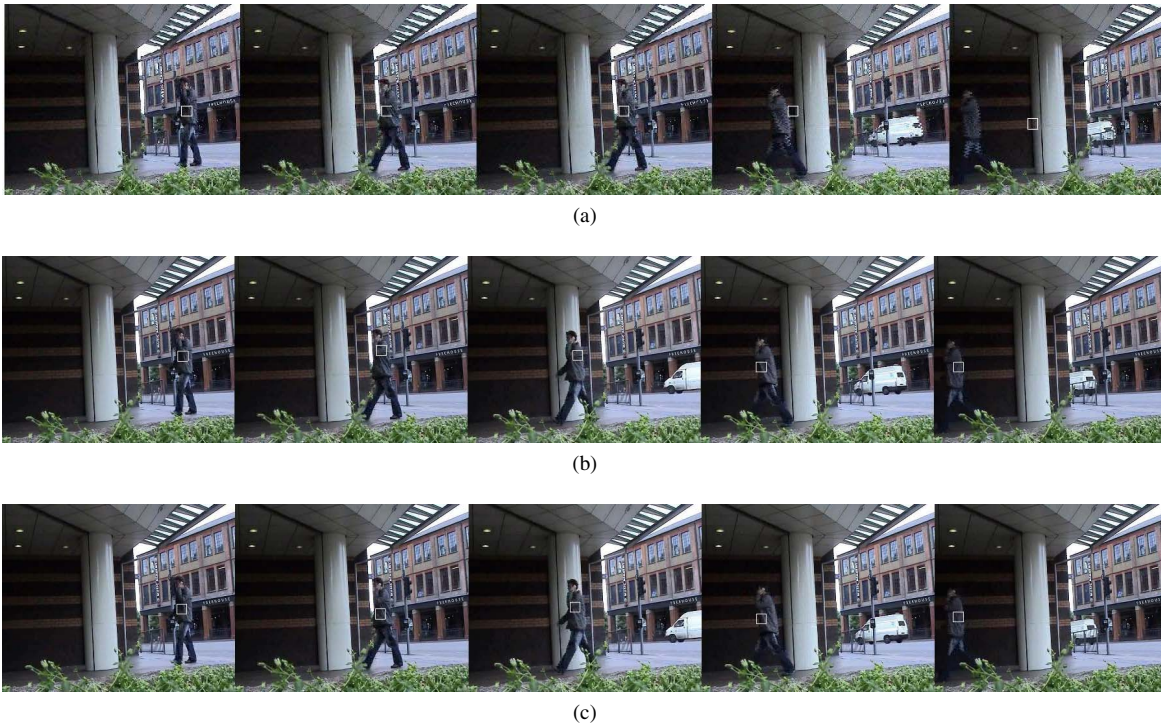
Fig. 1: A man walking into a building. (a) Covariance Tracker (b) Cepstrum-based Covariance Tracker (c) Cepstrum-based Co-difference Tracker

method because it does not require any multiplications during tracking.

REFERENCES

[1] F. Porikli, O. Tuzel and P. Meer, "Covariance Tracking using Model Update Based on Means on Riemannian Manifolds," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, 2006, vol. 1, pp. 728-735.

[2] H. Tuna, I. Onaran and A.E. Cetin, "Image Description Using a Multiplier-Less Operator," *IEEE Signal Processing Lett.*, vol. 16, no. 9, pp. 751-753, Sep. 2009.

[3] C. Stauffer and W.E.L. Grimson, "Learning Patterns of Activity using Real-time Tracking," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 747-757, Aug. 2000.

[4] A. Elgammal, D. Harwood and L. Davis, "Non-parametric Model for Background Subtraction," in *Proc. 6th European Conf. on Computer Vision*, 2000, pp. 751-767.

[5] C. Shen, X. Lin and Y. Shi, "Moving Object Tracking under Varying Illumination Conditions," *Pattern Recognition Lett.*, vol. 27, no. 14, pp. 1632-1643, Oct. 2006.

[6] Y. Wu, J.Y. Lin and T.S. Huang, "Capturing Natural Hand Articulation," in *Proc. IEEE Int. Conf. on Computer Vision*, 2001, vol. II, pp. 426-432.

[7] M.L. Cascia, S. Sclaroff and V. Athitsos, "Fast, Reliable Head Tracking under Varying Illumination: An Approach Based on Registration of Texture-mapped 3D Models," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 4, pp. 322-336, April 2000.

[8] S. Jabri, Z. Duric, H. Wechsler and A. Rosenfeld, "Detection and Location of People in Video Images using Adaptive Fusion of Color and Edge Information," in *Proc. IEEE Int. Conf. on Pattern Recognition*, 2000, vol.4, pp. 627-630.

[9] Y.-B. Lee, B.-J. You and S.-W. Lee, "A Real-time Color-based Object Tracking Robust to Irregular Illumination Variations," in *Proc. IEEE Int. Conf. on Robotics and Automation*, 2001, vol. 2, pp. 1659-1664.

[10] J. Yang and A. Waibel, "A Real-time Face Tracker," in *Proc. IEEE Workshop on Applications of Computer Vision*, 1996, pp. 142-147.

[11] Y. Raja, S.J. McKenna and S. Gong, "Tracking and Segmenting People in Varying Lightning Conditions using Colour," in *Proc. 3rd. Int. Conf. on Automatic Face and Gesture Recognition*, 1998, pp. 228-233.

[12] F. Moreno-Noguer, J. Andrade-Cetto and A. Sanfeliu, "Fusion of Color and Shape for Object Tracking under Varying Illumination," in *Proc. IEEE Iberian Conf. on Pattern Recognition and Image Analysis*, 2003, pp. 580-588.

[13] L. Li and M.K.H. Leung, "Integrating Intensity and Texture Differences for Robust Change Detection," *IEEE Trans. on Image Processing*, vol. 11, no. 2, pp. 105-112, Feb. 2002.

[14] S. Furui, "Cepstral Analysis Technique for Automatic Speaker Verification," *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. 29, no. 2, pp. 254-272, April 1981.

[15] A.V. Oppenheim and R.W. Schafer. (2004, Sep.). From Frequency to Quefrency: A History of the Cepstrum. *IEEE Signal Processing Mag.* pp. 95-106.

[16] T. Horprasert, D. Harwood and L. Davis, "A Statistical Approach for Real-time Robust Background Subtraction and Shadow Detection," in *Proc. 7th IEEE Int. Conf. on Computer Vision Frame-Rate Workshop*, 1999, pp. 1-19.

[17] F. Porikli, O. Tuzel and P. Meer, "Covariance Tracking Using Model Update Based on Lie Algebra," in *Proc. IEEE Int. Conf. on Computer Vision and Pattern Recognition*, 2006, pp. 728-735.

[18] "Tinypic" [Online]. Available: *http://i42.tinypic.com/xrr0z.png* [Accessed: Sept. 25, 2010].
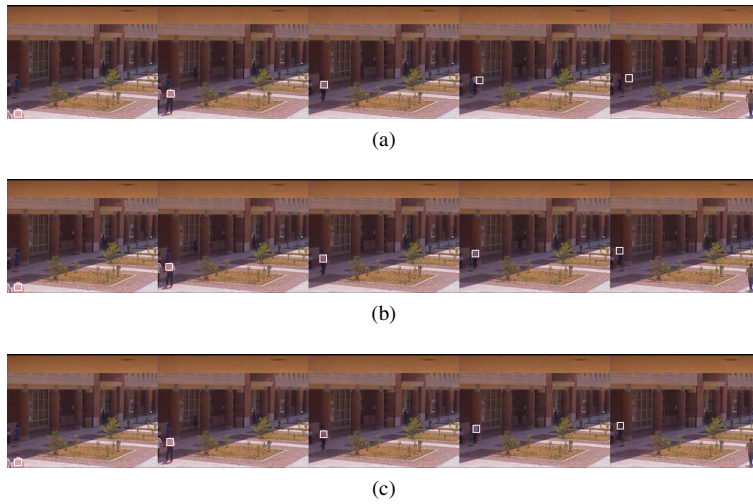
Fig. 2: A man walking to a shadowed area. (a) Co-difference Tracker (b) Cepstrum-based Covariance Tracker (c) Cepstrum-based Co-difference Tracker
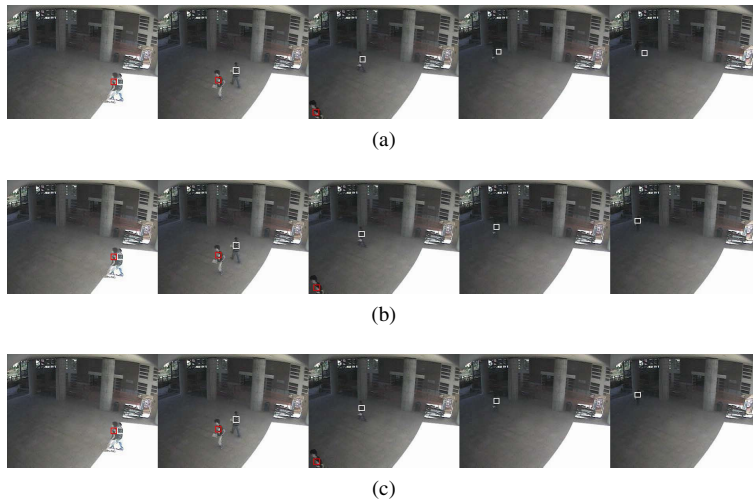

Fig. 3: A man walking into a covered area. (a) Covariance Tracker (b) Cepstrum-based Covariance Tracker (c) Cepstrum-based Co-difference Tracker
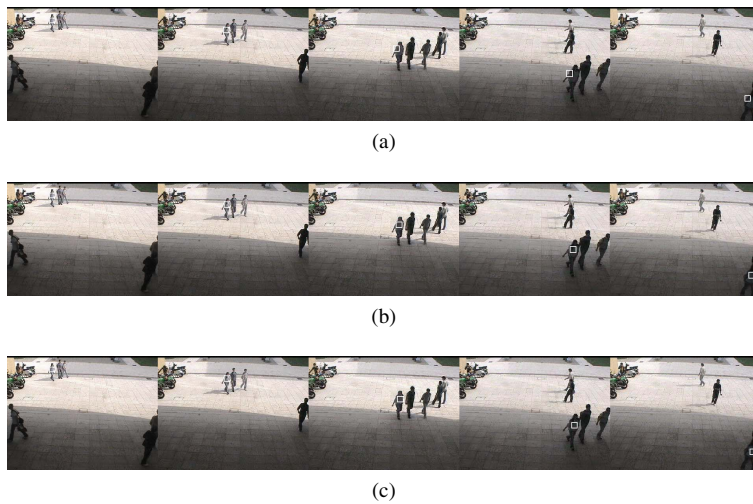

Fig. 4: A woman walking into a darker region. (a) Co-difference Tracker (b) Cepstrum-based Covariance Tracker (c) Cepstrum-based Co-difference Tracker