# CHANGE DETECTION AND BACKGROUND UPDATE THROUGH STATISTIC SEGMENTATION FOR TRAFFIC MONITORING

T. Alexandropoulos, V. Loumos and E. Kayafas
*Multimedia Technology Laboratory, National Technical University of Athens, 9 Heroon Polytechneiou, Athens, Greece*

Keywords: Highway Surveillance, Traffic Monitoring, Background Update, Change Detection.

Abstract: Recent advances in computer imaging have led to the emergence of video-based surveillance as a monitoring solution in Intelligent Transportation Systems (ITS). The deployment of CCTV infrastructure in highway scenes facilitates the evaluation of traffic conditions. However, the majority of video-based ITS are restricted to manual assessment and lack the ability to support automatic event notification. This is due to the fact that, the effective operation of intelligent traffic management relies strongly on the performance of an image processing front end, which performs change detection and background update. Each one of these tasks needs to cope with specific challenges. Change detection is required to perform the effective isolation of content changes from noise-level fluctuations, while background update needs to adapt to time-varying lighting variations, without incorporating stationary occlusions to the background. This paper presents the operation principle of a video-based ITS front end. A block-based statistic segmentation method for feature extraction in highway scenes is analyzed. The presented segmentation algorithm focuses on the estimation of the noise model. The extracted noise model is utilized in change detection in order to separate content changes from noise fluctuations. Additionally, a statistic background estimation method, which adapts to gradual illumination variations, is presented.

## 1 INTRODUCTION

The continuous development in computer imaging technology has resulted in the adoption of video-based surveillance as a monitoring solution in Intelligent Transportation Systems. The video-based approach to traffic monitoring suggests the utilization of CCTV infrastructure for the surveillance of highway scenes by human operators. Video monitoring, as a non-intrusive approach, allows the easy and secure installation of monitoring infrastructure and facilitates the repositioning of the equipment when needed.

A video-based ITS provides visual information to a human operator, who is responsible for traffic condition assessment. Depending on the observed conditions, the operator selects an appropriate course of action. Some decision examples include the adjustment of top speed limits, depending on the present traffic congestion, or the notification of emergency medical services, in the case on an accident.

Most installed video-based highway surveillance systems restrict to visualization purposes, that is, traffic conditions and events are assessed manually by the operator. However, the ongoing advances in computer processing speed have increased the interest for the expansion of video surveillance capabilities to the level where intelligent event detection and notification features are supported.

The operation of an ITS as an autonomous decision support system relies strongly on the effective operation of a front end which executes machine vision algorithms, in order to perform feature extraction. In video-based ITS, the term “feature extraction” refers to the detection of the regions, which are likely to generate traffic events. Apparently, in a traffic surveillance scene, these regions consist of virtually every present target, including vehicles and pedestrians.

In the present problem, it is convenient to assume a static background and compare it against successive frames, in order to track the emergence of targets of interest. Therefore, feature extraction is treated as a change detection problem. The



Alexandropoulos T., Loumos V. and Kayafas E. (2007).
CHANGE DETECTION AND BACKGROUND UPDATE THROUGH STATISTIC SEGMENTATION FOR TRAFFIC MONITORING.
In *Proceedings of the Second International Conference on Signal Processing and Multimedia Applications*, pages 257-264
DOI: 10.5220/0002134802570264

performance of a video-based ITS front end depends on the joint operation of two operations: change detection and background update.

All imaging devices introduce inter-frame variations due to sensor noise. Thus, direct frame differencing fails to provide accurate results as a change detection method. In general, change detection demands the utilization of methods which separate optimally content changes from noise-level fluctuations.

In addition, although a static background is assumed, its appearance exhibits intensity variations due to gradual changes in lighting conditions. In highway scenes, an apparent factor that causes alterations in the background scene is the influence of daylight variations. This fact dictates the use of a background update method, which adapts to the present lighting conditions, without incorporating occlusions to the background model.

The present paper analyzes the operation of a video-based ITS front end which performs feature extraction in the image domain. In detail, the paper is structured as follows. In Section 2, a block-based clustering procedure, which performs noise model estimation, is presented. In Section 3, a statistic change detection method which takes into account the noise model information is analyzed. In Section 4, an algorithm which performs background model adaptation to gradual illumination variations is presented. In the end, a short discussion analyzes further development plans on the architecture of the proposed system.

## 2 NOISE MODEL ESTIMATION

Feature extraction in traffic surveillance is feasible through the application of change detection. A highway frame, which is free of occlusions, can be selected as reference. The change detection method is responsible for the detection of occlusions on the surveyed scene. This is performed through comparison of subsequent frames against the reference frame. The operator can restrict the feature extraction process by specifying a region of interest or even multiple regions of interest. The latter case applies when the operator is interested in the inspection of each highway lane individually.

Since change detection requires the existence of a reference frame, there is the need to ensure its availability. The background frame may either be selected manually or be created by applying temporal median to a training video sequence (Massey, 1996). In the latter case, manual verification is suggested, in order to ensure that no static occlusions have been erroneously incorporated to the background model.

An algorithm which performs change detection needs to cope with the noise effect. Noise is an inherent characteristic which is present in every image acquisition sensor and introduces inter-frame variations even in “unchanged” scenes. This fact precludes direct differencing and requires the use of a technique that separates content changes from noise-level fluctuations. Obviously, content changes and noise variations are not always separable. Thus, change detection algorithms focus on the estimation of the optimum “threshold of perception” and the formulation of an optimized classification rule.

Change detection methods encountered in the literature (Radke, 2005) are based on statistic criteria, in order to decide whether a pixel or a block of pixels corresponds to a changed or an unchanged region. A statistic approach proposed in (Aach, 1993), (Cavallaro, 2001) applies a statistic significance test over a rectangular window which is centred in the pixel of interest. The approach assumes a Gaussian noise model and introduces a $\chi^2$ probability distribution function in the formulation of the classification rule. However, the utilized significance test relies on prior knowledge of the noise model and therefore requires the availability of statistic noise information in order to operate reliably.

The method which is analyzed in the present paper aims to estimate the noise model - that is, the noise mean value $m_n$ and standard deviation $\sigma_n$ - from the image blocks of the absolute difference. Let $m_{ij}$ and $\sigma_{ij}$ denote the observed mean value and standard deviation in each block $B_{ij}$ of the absolute difference. In order to achieve an accurate noise model approximation, it is sufficient to restrict the noise-model calculations to the block subset which exhibits noise-level fluctuations. Therefore, the goal of the proposed procedure is to group blocks of the absolute difference to clusters, according to their statistic similarity (Alexandropoulos, 2005).

In order to perform the clustering operation, each cluster $C_k$ is described by two statistic parameters: the averaged mean value $m_{C_k}$ and the averaged standard deviation of its member blocks. Thus, for a cluster with $N_k$ member blocks, we define:

$$m_{C_k} = \sum_{i=1}^{N_k} \frac{m_i}{N_k} \qquad (1)$$

$$\sigma_{C_k} = \sum_{i=1}^{N_k} \frac{\sigma_i}{N_k} \qquad (2)$$

Obviously, equations (1) and (2) define the *centroid* of cluster $C_k$.

The proposed block clustering procedure is performed in the following steps.

i) Cluster $C_1$ is initialized:

$$m_{C_1} = m_{11} \qquad (3)$$

$$\sigma_{C_1} = \sigma_{11} \qquad (4)$$

ii) For each block $B_{ij}$ of the absolute difference and for each existing cluster $C_k$ $(k = 1,2,...,n)$, their *mean value distance* $d_{ij,k}$ is estimated.

$$d_{ij,k} = \left| m_{ij} - m_{C_k} \right| \qquad (5)$$

If $d_{ij,k} \le \sigma_{C_k}$, cluster $C_k$ is added to the list of candidates.

iii) If candidate clusters have been found, block $B_{ij}$ is classified to the cluster which yields the minimum mean value distance and the centroid of the "winner cluster" is updated, as indicated by equations (1) and (2). If no candidate clusters have been found, a new cluster $C_{n+1}$ is initialized and block $B_{ij}$ is added to it. Therefore:

$$m_{C_{n+1}} = m_{ij} \qquad (6)$$

$$\sigma_{C_{n+1}} = \sigma_{ij} \qquad (7)$$

The flowchart of the clustering procedure is presented in Figure 1.

Upon completion of the clustering algorithm, a *segmentation map* of the absolute difference is obtained. In this representation, blocks which have been grouped into the same cluster are displayed with the same intensity value. When the algorithm is applied on colour images, a segmentation map is extracted for each colour component. In the present work, the segmentation technique is used in the HSL colour space. Specifically, it is applied on the hue and saturation components of the absolute difference.

The segmentation maps are used as a qualitative criterion and indicate the effectiveness of the procedure in the noise model estimation. In an effective segmentation, the majority of unchanged regions – and only unchanged regions - is grouped in a single cluster.

An example which shows the results of the segmentation method is displayed in Figure 2. Figure 2a presents a background frame which is used as reference. The inspected region of interest is shown in Figure 2b. Figure 2c shows an occluded frame. The absolute difference of the reference and the occluded frame is displayed in Figure 2d. Figures 2e and 2f present the segmentation maps which are produced when the statistic segmentation algorithm is applied on the hue and saturation components of the absolute difference. It is evident that the majority of background blocks have been grouped in a single cluster. Thus, the segmentation procedure allows accurate noise model estimation, by taking into account the properties of the block subset which is grouped in the "background cluster".

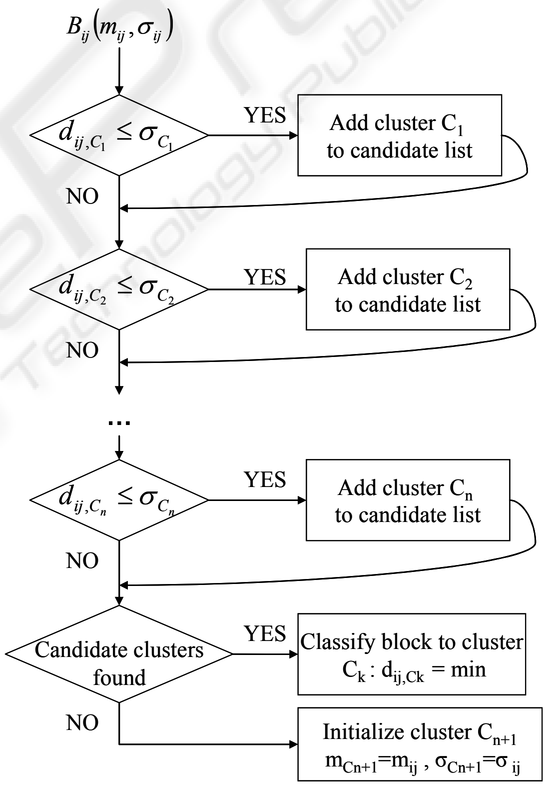


Figure 1: Flowchart of the proposed block-based statistic segmentation algorithm.

The noise model estimation is based on the assumption that *the largest cluster carries noise model information*. This assumption is considered safe when the occlusions occupy up to half of the surveyed scene. This fact allows the application of the noise model estimation method without the strict necessity to ensure complete absence of occlusions in the training video sequence.

Let $m_n$ and $\sigma_n$ denote noise mean value and standard deviation respectively. In order to utilize the statistic properties of the background cluster, we define:

$$m_n = m_{C_{\max}} \tag{8}$$

$$\sigma_n = \max\{\sigma_{ij} \in C_{\max}\} \tag{9}$$

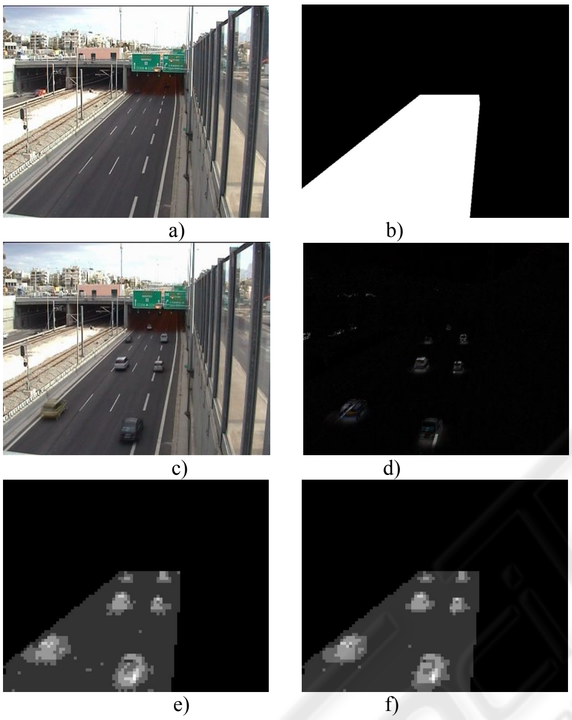


Figure 2: Block-based segmentation results: a) Reference background, b) The inspected region of interest c) A frame with occlusions d) Absolute difference, e) Segmentation map extracted fron the the hue component of the absolute difference f) Segmentation map extracted from the saturation component of the absolute difference.

# 3 CHANGE MASK EXTRACTION

In change detection, the noise model, which is obtained by the segmentation procedure, is used in order to establish a noise-dependent decision rule and separate content changes from noise-level fluctuations. For each block $B_{ij}$ of the absolute difference, the mean value distance $d_{ij,n}$ is estimated.

$$d_{ij,n} = \left| m_{ij} - m_n \right| \tag{10}$$

The mean value distance $d_{ij,n}$ is then compared against the noise standard deviation $\sigma_n$.

i) If $d_{ij,n} > \sigma_n$ then block $B_{ij}$ is classified as changed.

ii) If $d_{ij,n} \leq \sigma_n$ block $B_{ij}$ is unchanged.

This decision rule produces a binary change mask which is applied on the present frame, in order to isolate content changes. When change detection is performed on colour images, a binary change mask is produced for each colour component. The partial masks are then merged into a single change mask through the application of an OR operator. In the last stage, block-level median filtering is applied on the change mask for the suppression of inconsistencies.

In the present work the algorithm is applied on the HSL colour space and produces change masks which correspond to the hue and saturation components of the absolute difference. The respective algorithm flowchart is shown in Figure 3.

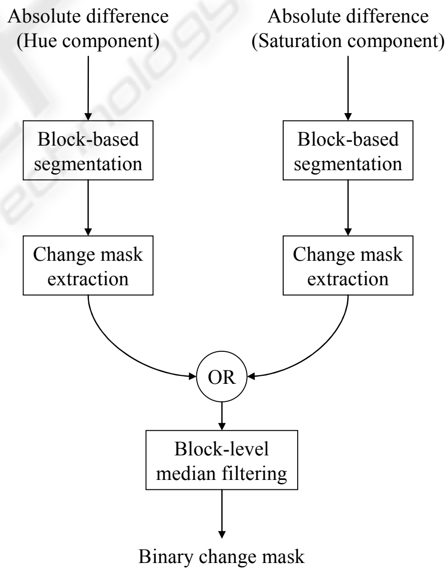


Figure 3: Flowchart of binary change mask extraction in colour images.

The results of the change detection algorithm are presented in Fig 4. Figure 4a shows the reference background and Figure 4b presents a frame with occlusions. Figure 4c displays the changes which are detected through the employment of the proposed statistic segmentation and change detection criteria. It is clear that the inclusion of the extracted noise model to the change detection decision rule achieves the detection of content alterations and the concurrent suppression of noise-level fluctuations.

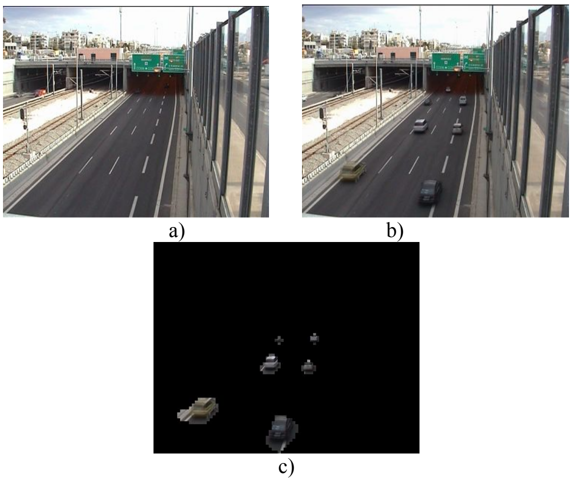

Figure 4: a) Reference background, b) A highway frame with occlusions, c) Changes detected through the application of the proposed block-based change detection method.

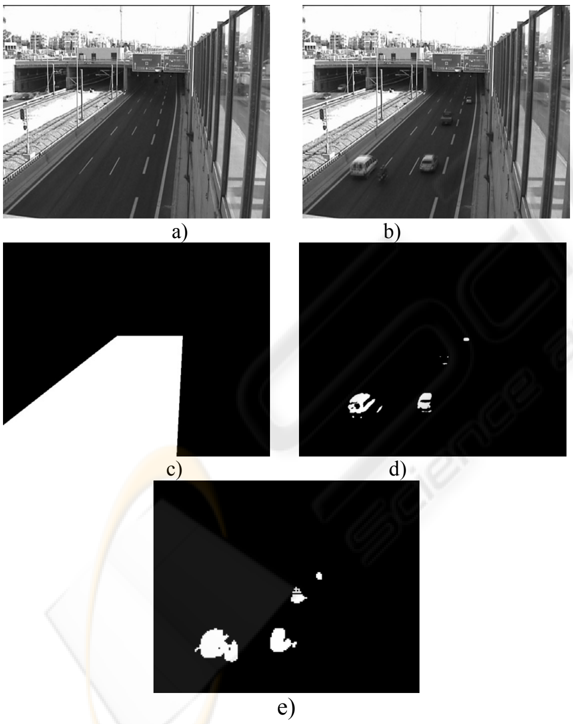

Figure 5: a) Reference frame b) Occluded frame c) Region of interest d) Binary mask extracted by the significance test presented in (Aach,1993) e) Binary change mask extracted by the proposed segmentation procedure.

In the example shown in Figure 5, the results of the proposed technique are compared to the respective results extracted by the algorithm proposed in (Aach, 1993). Both algorithms have been tested on grayscale images. Figures 5a and 5b display the reference frame and an occluded frame respectively. The change detection algorithms are applied on the region of interest displayed in Figure 6c. Figure 6d presents the binary change mask which is obtained through the application of the significance test proposed in (Aach, 1993). The change mask displayed in Figure 5e is obtained through application of the proposed algorithm. In the latter case, one can notice that the utilization of the statistic segmentation criterion has resulted in an increased detection rate of occluded regions.

# 4 BACKGROUND UPDATE

Video-based ITS operate continuously under time varying lighting condition. Thus, the surveyed scenes inevitably exhibit gradual illumination variations. These are caused by changes in daylight intensity and colour temperature. Moreover, the time-varying positioning of the light sources results in the shifting of cast shadows. These factors affect the appearance of the background and introduce the necessity to update the reference frame when performing change detection.

Background estimation methods which are found in the literature attempt to separate foreground from background regions based on pixel-wise temporal measurements. Linear prediction methods, such as Weiner and Kalman filtering, (Toyama, 1999) (Ridder, 1995) take into account the history of the observed intensity values at each pixel. Haritaoglu et al. (2000) track the intensity fluctuations which are observed between the frames of a background video sequence and use them as a training parameter in a thresholding criterion. Statistical methods track the temporal intensity variations at each pixel, in order to approximate a noise probability function distribution. Early approaches model background pixel intensities with a Gaussian pdf (Cavallaro, 2001). In most recent approaches, a mixture of Gaussians is estimated for the expression of both background and foreground models (Lee, 2005), (Harville, 2001), (Stauffer, 1999).

The aforementioned techniques share a common characteristic: Persistent static occlusions are eventually incorporated in the background model. Consequently, these methods are oriented towards motion detection rather than change detection. However, in the present case, the addition of still targets to the updated background is considered as an undesirable effect. In traffic monitoring, the existence of a static occlusion is likely to indicate significant events. For example, a static target may indicate the presence of an immobilized vehicle or an accident scene.

Toyama et al. (1999) assume discrete illumination states and select the appropriate background model based on k-means clustering. This approach addresses background update in indoor scenes with finite illumination states. However, outdoor scenes are characterized by complex illumination variation. Thus, the utilization of this algorithm would require a very large number of illumination profiles.

In the present work, a block-based statistic background update method, which adapts to gradual illumination alterations, is implemented. The algorithm is based on successive comparisons of the background against each observed frame. It utilizes the noise model which is extracted by the statistic segmentation method and produces a "foreground mask".

The decision rule which is used for the classification of image blocks into foreground and background blocks is similar to the one used for change detection. In this case, the averaged mean value $m_{G_{\max}}$ and averaged standard deviation $\sigma_{C_{\max}}$ of the dominant cluster $C_{\max}$ are introduced in the criterion.

Initially, the mean value of each block of the absolute difference is compared against the averaged mean value of the dominant cluster.

$$d_{ij,C_{\max}} = \left| m_{ij} - m_{C_{\max}} \right| \qquad (11)$$

i) If $d_{ij,C_{\max}} > \sigma_{C_{\max}}$ block $B_{ij}$ is added to the foreground mask.

ii) If $d_{ij,C_{\max}} \leq \sigma_{C_{\max}}$ block $B_{ij}$ denotes a background region

In colour images, this decision rule extracts a foreground mask for the absolute difference of each colour component.

In the next stage of the procedure, the foreground mask of each colour component is subjected to block-level median filtering, in order to suppress inconsistencies. The foreground masks are then merged into a single mask through an OR operation.

Image blocks which lie in occlusion boundaries frequently exhibit statistic characteristics which are similar to the ones of unoccluded regions, because a percentage of their member pixels display background regions. The incorporation of these blocks to the background model would cause the appearance of edge traces and the overall degradation of the background scene. In order to cope with this issue, the foreground mask is subjected to block-level dilation and the occlusion boundaries are added into it.

The flowchart of the foreground mask extraction is presented on Figure 6.

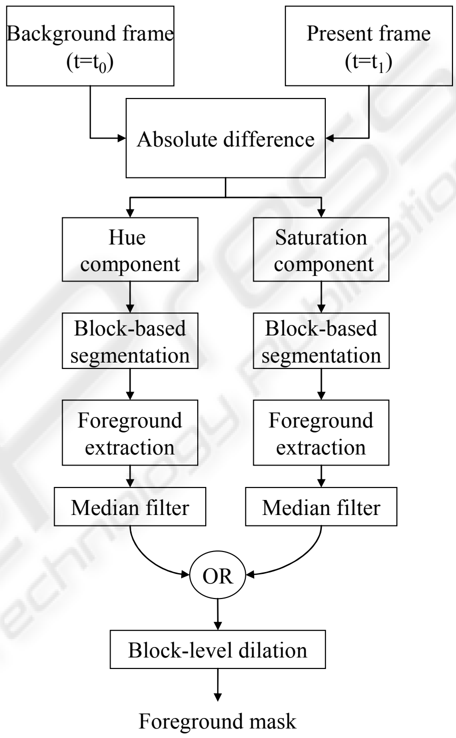


Figure 6: Flowchart of the foreground mask extraction algorithm.

After the extraction of each blocks $F_{ij}$ of the binary foreground mask, the background model is updated as follows:

i) If $F_{ij} = 0$, the background frame is updated in the respective block.

ii) If $F_{ij} = 1$, the background model maintains its previous state in the specific block.

The proposed algorithm has been applied on highway video sequences, in order to demonstrate its capability to adapt to time-varying illumination conditions. An example is presented in Figure 7. Figure 7a shows the background frame at the initialization of the algorithm. Figure 7b shows an occluded frame, in which ambient light variations have been introduced. Figure 7c displays the differences between frames 7a and 7b. One can observe the chrominance variations which have emerged on the highway surface. If the noise model estimation is performed offline - during an initial training stage - a direct comparison between frames 7a and 7b would result in the misclassification of all image blocks as changed. If inter-frame noise is estimated continuously, direct differencing of these frames would result in the estimation of an exaggerated noise model and the formulation of a conservative decision rule.

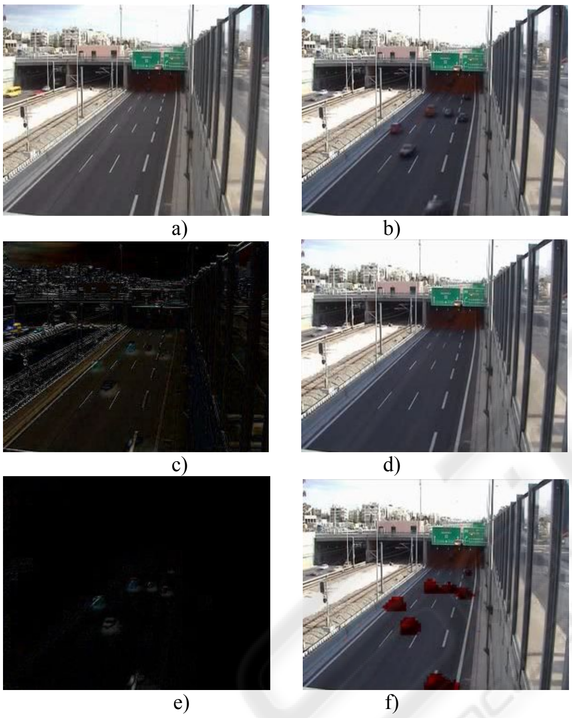


Figure 7: Background update example a) Reference frame b) Changed frame c) Differences between the initial background model and the changed frame, d) Updated background e) Differences between the updated background and the changed frame f) Detected features.

The utilization of the proposed background update procedure produces the background model which is shown in Figure 7d. Figure 7e displays the differences between the present frame and the updated model. It is evident that the update process managed the incorporation of gradual illumination changes to the background model without the addition of occlusions. The change detection results are highlighted in Figure 7f.

# 5 FURTHER DEVELOPMENT

This paper has presented a combination of image processing algorithms, which can be included in the front end of a video-based ITS for target tracking in highway scenes. The presented approach addresses the problem through the combination of change detection and background update algorithms.

The analysis in the present work has focused exclusively on the description of early vision algorithms for feature extraction. Further development on the implementation of an integrated ITS solution demands the cooperation of machine vision algorithms with content interpretation algorithms. This integration is necessary for the assessment of the temporal behaviour of each detected feature and the extraction of high-level attributes. In traffic surveillance, high-level attributes are mostly related to motion characteristics and are considered useful for the estimation of vehicle properties, such as vehicle speed and trajectory patterns, or the detection of significant events, such as traffic jam conditions or accidents.

The estimation of high-level attributes requires the inter-frame correlation of the extracted features and can be addressed through the employment of MPEG-7 visual descriptors. The use of visual description schemas establishes a framework which offers the capability to formulate event representations. This approach attempts to enable the ITS to interpret the behaviour of each feature by matching it with the appropriate event profile. Therefore, it is expected that the ITS decision support capabilities will be enhanced.

# REFERENCES

Massey, M., and Bender, W., 1996. Salient Stills: Process and Practice, In *IBM Systems Journal*, Vol. 35, No 3 & 4, pp. 557-573.

Radke, R. J., Andra, S., Al-Kofahi, O., and Roysam, B., 2005. Image Change detection Algorithms: a systematic survey, In *IEEE Transactions on Image Processing,* Vol. 14, Issue 3, pp. 294-307.

Aach, T., Kaup, A., and Mester, R., 1993. Statistical model based change detection in moving video, In *Signal Processing*, Vol. 31, Issue 2, pp. 165-180.

Cavallaro, A., and Ebrahimi, T., 2001. Video object extraction based on adaptive background and statistical change detection, In *Proc. SPIE Visual Communications and Image Processing*, pp. 465-475.

Alexandropoulos, T., Boutas, S., Loumos, V., and Kayafas, E., 2005. Real-time change detection for surveillance in public transportation, In *IEEE*

*International Conference on Advanced Video and Signal-Based Surveillance*, pp. 58-63.

Toyama, K., Krumm, J., Brummit, B., and Meyers, B., 1999. Wallflower: Principles and Practice of Background Maintenance, In *7th IEEE International Conference on Computer Vision*, pp. 255-261.

Ridder, C., Munkelt, O., and Kirchner, H., 1995. Adaptive background estimation and foreground detection using Kalman filtering, In *International Conference on Recent Advances in Mechatronics*, pp.193-199.

Haritaoglu, I., Harwood, D., and Davis, L. S., 2000. W4: Real-time surveillance of people and their activities, In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 8, pp. 809-830.

Lee, D. S., 2005. Effective Gaussian Mixture Learning for Video Background Subtraction, In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 27, No. 5, pp. 827-832.

Harville, M., Gordon, G., and Woodfill, J., 2001. Foreground Segmentation Using Adaptive Mixture Models in Color and Depth, In *Proc. IEEE Workshop Detection and Recognition of Events in Video*, pp. 3-11.

Stauffer, C., and Grimson, W.E.L., 1999. Adaptive Background Mixture Models for Real-Time Tracking, *Proc. Conf. Computer Vision and Pattern Recognition*, Vol. 2, pp. 246-252.