Improving video segmentation algorithms by detection of and adaption to altered illumination

Lindström, Johan; Lindgren, Finn; Holst, Ulla; Åström, Karl

Published in:
Preprints in Mathematical Sciences

2008

Improving Video Segmentation Algorithms by Detection of and Adaption to Altered Illumination

Johan Lindström
Finn Lindgren
Ulla Holst
Kalle Åström

Preprints in Mathematical Sciences
2008:9

Lund Institute of Technology
Lund University

Centre for Mathematical Sciences
Mathematical Statistics

CENTRUM SCIENTIARUM MATHEMATICARUM

Improving Video Segmentation Algorithms by Detection of and Adaption to Altered Illumination

Johan Lindström Finn Lindgren Ulla Holst
Kalle Åström

16th April 2008

Abstract

Changing illumination constitutes a serious challenge for video segmentation algorithms, especially in outdoor scenes under cloudy conditions. Rapid illumination changes, e.g. caused by varying cloud cover, often cause existing segmentation algorithms to erroneously classify large parts of the image as foreground.

Here a method that extends existing segmentation algorithms by detecting illumination changes using a CUSUM detector and adjusting the background model to conform with the new illumination is presented. The method is shown to work for two segmentation algorithms, and it is indicated how the method could be extended to other algorithms.

1 Introduction

The segmentation of image sequences in foreground and background is a fundamental low-level task in computer vision. The result is used in different high level operations such as object tracking, video surveillance, and monitoring.

To handle dynamically changing backgrounds most segmentation algorithms today use adaptive estimation methods. Early work includes the use of Kalman (Karman & von Brandt, 1990) and Wiener filters (Toyama *et al.*, 1999).

Allowing for only one background description per pixel these models encounter problems with multi-modal backgrounds. Multi-modal backgrounds can be handled by modelling the background using Gaussian mixtures (Stauffer & Grimson, 1999; Tuzel *et al.*, 2005; Lindström *et al.*, 2006). Using similar background models these methods differ primarily in the recursive updates of the model parameters. Further work in this rapidly growing field includes post processing to remove shadows (KaewTraKulPong & Bowden, 2001), use of depth data (Harville *et al.*, 2001) and compensation for camera movement (Hayman & Eklundh, 2003), to mention a few. An alternative to the Gaussian mixtures is use use kernel density estimators (Elgammal *et al.*, 2000).

The recursive parameter estimation methods of the algorithms mentioned above handle gradual illumination changes. But, rapid illumination changes often cause the algorithms to, erroneously, classify large parts of the image as foreground (Lindström *et al.*, 2006). In outdoor applications illumination changes are mainly caused by varying cloud cover. The adaptive nature of the algorithms ensures that they will readjust to the new background after a limited period, during which a substantial part of the image may be misclassified as foreground. Unfortunately, this adaptation can take 10-20 seconds (Lindström *et al.*, 2006), severely hampering high level algorithms that depend on the foreground segmentation.

Previous attempts to alleviate this drawback include temporarily increasing the forgetting factor (Schindler & Wang, 2006), switching between parallel background models when the amount of foreground exceeds a fixed threshold (Toyama *et al.*, 1999) or using edge detector data (Jabri *et al.*, 2000; Lindström *et al.*, 2006). However, these approaches do not completely resolve the issue, and potential complications include: 1) Increasing the learning rate still requires some time for the model to adapt; 2) Using a fixed threshold might not detect cases when only parts of the image is affected by cloud cover; 3) Switching between parallel models works better for indoor applications, where the illumination changes often are caused by lights turning on or off, than for outdoor applications; 4) Under constant illumination, edge detector data causes more false positives than RGB data.

Other alternatives include the use of depth information from multiple cameras (Ivanov *et al.*, 2000; Lim *et al.*, 2005). However this increases the hardware requirements and places some restrictions on the camera placement (Lim *et al.*, 2005).

Here a method based on detection of the rapid illumination changes using a CUSUM detector (Page, 1954; Gustafsson, 2000) followed by adjustments of the background model is proposed. The method allows a straightforward extension of existing segmentation algorithms that improves their ability to handle rapid illumination changes. The method is shown to work with the algorithms in Stauffer & Grimson (1999) and Lindström *et al.* (2006), and should be adaptable to other algorithms that model the background using (Gaussian) mixtures.

The proposed method is described in Section 2, with 2.1 describing the change detection and 2.2 describing how the background model is modified to account for illumination changes. Results are presented in Section 3 and Section 4 gives conclusions and a short summary.

2 Theory and method

Handling rapid illumination changes can conceptually be divided into two parts. Firstly the illumination change has to be detected, and secondly the model should be adjusted to better reflect the new background.

2.1 Detection of illumination changes

Illumination changes caused by rapidly changing cloud cover differ from illumination changes caused by lights turning on or off. In the former case the change is gradual with an increasing part of the image being affected as the clouds drift over the scene; in the latter case the change is momentaneous. This difference strongly influences the formulation of the change detection.

Illumination changes are very difficult to detect at a pixel level, implying that use of frame level information would be advantageous (Toyama *et al.*, 1999). Here the total amount of foreground, P_t, in the image is used

$$P_t = \frac{\sum_i \mathsf{P}\big(x_{it} \in \text{foreground}|\text{model at time t}\big)}{\text{nbr. of pixels}}, \tag{1}$$

where x_{it} is pixel i at time t. For models that do not give foreground probabilities, e.g. Stauffer & Grimson (1999), the probabilities in (1) are replaced by indicators for the pixel belongings.

If clouds start to affect the segmentation at time T_0, a simple model for $\Delta P_t = P_t - P_{t-1}$ is

$$\Delta P_t \in \begin{cases} \mathsf{N}\big(0, \sigma^2\big), & t \leq T_0, \\ \mathsf{N}\big(k, \sigma^2\big), & t > T_0, \end{cases} \tag{2}$$

i.e. the total amount of foreground follows a random walk with mean zero before the clouds cause an (approximately linear) increase in the expected amount of foreground, differentiation reduces the problem to detection of a shift in expectation. Here σ^2 relates to the normal variation in P_t caused by objects entering and leaving the scene, and k depends on how fast clouds affect the illumination.

As a comparison an appropriate model for momentaneous illumination change would be obtained by replacing ΔP_t with P_t in (2). A detection scheme similar to that presented here should then be applicable. Since we focus on the rapid but not instantaneous illumination changes that are prevalent in outdoor applications, this is outside the scope of this article

A complication is that variations in ΔP_t can be caused both by objects entering or leaving the scene and by illumination changes. Thus the detection becomes a question of separating natural variations from increases caused by illumination changes, see Figure 1.

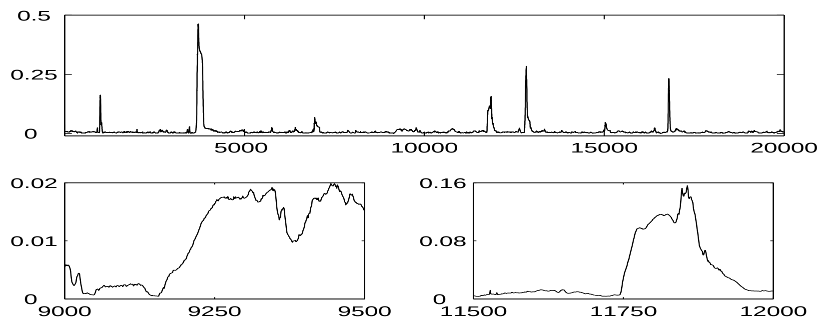

Figure 1: The total amount of foreground, P_t, in the video sequence as a function of the frame number (15 fps). The two lower panels show enlargements of what happens when a car enters the scene (left) and when a cloud alters the illumination in a section of the scene (right).

If $\hat{s}_t^2$ is a recursive estimate of σ^2 at time t and using ΔP_t defined in (2), then a CUSUM detector (Gustafsson, 2000, Ch. 3.4) can be formulated as

$$y_t = \max\left(y_{t-1} + \Delta P_t/\sqrt{\hat{s}_t^2} - C, 0\right), \tag{3}$$

with $y_0 = 0$, and a change is detected when y_t passes a threshold, h. The drift, C, and threshold, h, are design parameters.

If σ^2 is estimated using a delayed, windowed estimate

$$\hat{s}_t^2 = \frac{\sum_{i=T+1}^{L+T}\left(\Delta P_{t-i} - \overline{\Delta P}_t\right)^2}{L-1}, \tag{4}$$

where $\overline{\Delta P}_t = \sum_{i=T+1}^{L+T} \Delta P_{t-i}/L$, and assuming that the covariance

$$\mathsf{C}(\Delta P_t, \Delta P_\tau) = 0 \text{ if } t \neq \tau$$

(in practise for this application the auto-correlation is very small), then $\Delta P_t/\sqrt{\hat{s}_t^2}$ follows a doubly non-central Student's t distribution (Krishnan, 1967) (see Appendix, page 16, for details). The approximate expectation of $\Delta P_t/\sqrt{\hat{s}_t^2}$ for large values of L is

$$\mathsf{E}\left(\frac{\Delta P_t}{\sqrt{\hat{s}_t^2}}\right) \approx \begin{cases} 0, & t \leq T_0, \\ \frac{k}{\sigma}\frac{4L-1}{4(L-1)}, & T_0 < t \leq \widehat{T}, \\ \frac{k}{\sigma}\frac{4L-1}{4(L-1)} - \frac{k^3}{\sigma^3}\frac{(t-\widehat{T})\left(1-\frac{t-\widehat{T}}{L}\right)}{2(L-1)}, & \widehat{T} < t < L + \widehat{T}, \\ \frac{k}{\sigma}\frac{4L-1}{4(L-1)}, & L + \widehat{T} \geq t, \end{cases} \tag{5}$$

where $\widehat{T} = T + T_0 + 1$.

Note that the expectation is large for $T_0 < t \leq \widehat{T}$ where after it decreases. This decrease is caused by (4) overestimating σ^2 since it uses ΔP_t:s both before and after T_0. The decrease is a motivation for the use of a suitably chosen lag, T, in (4).

When selecting C it should be noted that $\Delta P_t/\sqrt{\hat{s}_t^2}$ follows a standard Student's t distribution with $L-1$ degrees of freedom if $t \leq T_0$ (see Appendix, page 16). If L is large $\Delta P_t/\sqrt{\hat{s}_t^2}$ will be approximately $\mathsf{N}(0,1)$ and C could be taken as a Gaussian quantile, e.g. $C = \lambda_{0.99} = 2.3263$, which should ensure that $y_t \approx 0$ for $t \leq T_0$.

After the illumination change at T_0, y_t will increase approximately as

$$\mathsf{E}(y_t) \approx \begin{cases} 0, & t \leq T_0, \\ \left(\frac{k}{\sigma} - C\right)\frac{4L-1}{4(L-1)}(t - T_0), & t > T_0, \end{cases} \tag{6}$$

with a somewhat slower increase for $t > \widehat{T}$ due to the overestimation of σ^2. The expression is approximate since the true expression for $\mathsf{E}(y_t)$ is intractable, due to the maximum operation in (3). Given some knowledge about the normal variations, σ^2, in ΔP_t and the speed, k, at which the illumination changes affect P_t, (6) can be used to select a suitable threshold, h.

2.2 Modifying the background model

Having detected a potential illumination change the background model should be adapted to represent the new illumination. In this section a simple illumination model is used to motivate a functional relationship between the old and new background models, estimation of parameters in the relationship is described, and details for altering the background model for the two segmentation algorithms in Stauffer & Grimson (1999) and Lindström *et al.* (2006), are given.

Since both the algorithms studied here operate in RGB colour space the following discussion is focused on that space. However for algorithms working in different colour spaces, such as normalised RGB or L$\alpha\beta$ the change detection presented earlier and the principles of altering the background model are still valid. But a slightly different method for calculating the new background is needed.

2.2.1 A simple illumination model

If ρ_{ic} is the measured intensity of channel c ($c \in \{1,2,3\}$ for RGB images) for pixel i on a Lambertian surface, then a simple illumination model (Tsin *et al.*, 2001) is

$$\rho_{ic} = \int f_c(\lambda)S_i(\lambda)l(\lambda)d\lambda.$$

Here $f_c(\lambda)$ is the sensitivity of the channel, $S_i(\lambda)$ is the surface reflectance and $l(\lambda)$ is the light spectrum. The simplifying assumption of a Lambertian surface is made since specular highlights only occur in few, if any, pixels so ignoring specular highlights will have limited effects.

In case of altered illumination only $l(\lambda)$ is affected, and the new pixel values become

$$\hat{\rho}_{ic} = \alpha \int f_c(\lambda) S_i(\lambda) \hat{l}(\lambda) d\lambda, \tag{7}$$

where α is a possible attenuation due to the automatic gain control (AGC) of the camera. If the illumination change is roughly homogeneous over the support of $f_c(\lambda)$ then (7) can be reformulated as

$$\hat{\rho}_{ic} = \alpha \int f_c(\lambda) S_i(\lambda) \hat{l}(\lambda) d\lambda \approx \alpha k_c \int f_c(\lambda) S_i(\lambda) l(\lambda) d\lambda = \kappa_c \rho_{ic}, \tag{8}$$

where $\kappa_c = \alpha k_c$ are channel specific constants. The approximately linear relationship is illustrated in Figure 2.

The assumption of homogeneous illumination change is fulfilled if, either the illumination change mainly affects the intensity and not the colour of the light or, the sensor has a narrow bandwidth compared to the structures in the illumination change, e.g. if the illumination change consists of one homogeneous change in the light affecting the red sensor and, another, homogeneous change in the light affecting the green sensor. The latter is usually true for most illumination changes (Finlayson *et al.*, 2000), including those caused by clouds.

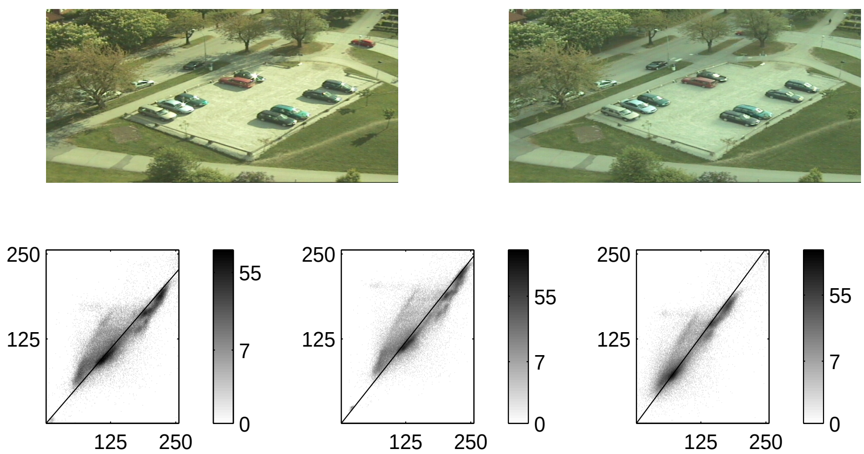

Figure 2: The top row depicts the scene before and after an illumination change. The bottom row contains 2-dimensional histograms of pixel values (logarithmic scale) before and after the change in intensities for the three colours (from left to right: red, green and blue), together with a regression line estimated using (9).

2.2.2 Estimating κ_c

Denote the current estimate of the pixel values of the background for channel c, pixel i, as μ_{ic}, and denote the current pixel value as x_{ic}. It should now be noted that μ_{ic} and x_{ic} are noisy observations of ρ_{ic} and $\hat{\rho}_{ic}$ respectively. This implies that κ_c (8) can be estimated using regression with errors in both variables (Fuller, 1987). Further, only those pixels actually affected by the illumination change should be used to estimate κ_c. Since the illumination change has been detected as an abnormal amount of foreground, the foreground probabilities (indicators), p_i, of each pixel can be used as regression weights. Thus we take

$$\kappa_c = \frac{S_{xx} - S_{\mu\mu} + \sqrt{(S_{xx} - S_{\mu\mu})^2 + 4S_{\mu x}^2}}{2S_{\mu x}}, \tag{9}$$

where $S_{\mu\mu} = \sum_i \mu_{ic}^2 p_i$, and similarly for S_{xx} and $S_{\mu x}$. The new estimates of the background become $\hat{\mu}_{ic} = \kappa_c \mu_{ic}$, see Figure 2.

A potential issue here is that actual foreground objects in the scene will act as outliers, leading to strange estimates of κ_c. However, experiments indicate that using a robust estimator of κ_c makes very little difference, and preserving the old background drastically reduces the problems caused by incorrect κ_c estimates.

2.2.3 Applying the changes

The adaptation of the background described above now has to be incorporated into existing background segmentation algorithms. The principles for doing this are very similar for all algorithms but details may vary.

In principle the most likely and least likely background components for each pixel are located, κ_c is estimated and a new possible background is determined. The least likely background is then replaced with the new, altered background and the two background components (most likely and altered) are given equal prior probabilities. The motivation for this strategy is twofold: 1) Parts not yet affected by the illumination change still need the old background model; 2) In case of an erroneous change detection retention of the old background models allows for graceful failure. Note that this only alters the expectation of the background components and not their variances. We have found that estimating appropriate new in variances is a far harder task than estimating the new expectations. Further a small increase in the forgetting factor for the first few frames after the change detection seems to be sufficient for the recursive algorithm to adjust the variance of the background components.

Below details for the algorithms proposed by Stauffer & Grimson (1999) and Lindström *et al.* (2006) are given. These should be fairly easy to extend to other segmentation algorithms.

The algorithm proposed by Stauffer & Grimson (1999) models each pixel value, x_i, as a Gaussian mixture

$$p(x_i) = \sum_{k=1}^{K} \pi_{ik} f(x_i | \mu_{ik}, \sigma_{ik}^2),$$

where π_{ik} are mixture probabilities and $f(x_i|\mu_{ik}, \sigma_{ik}^2)$ are Gaussian densities with expectation μ_{ik} and covariance matrix $\sigma_{ik}^2 \mathbf{I}$, where $\mathbf{I}$ is an identity matrix of suitable size. Here k indices the different Gaussian components used to describe the value taken by each pixel. The components are then ranked in decreasing order by their weight to standard deviation ratio, π_{ik}/σ_{ik}, and the first few components are considered to be background. The most likely background model and least likely model can therefore be found by considering the ratio π_{ik}/σ_{ik}.

Introducing $I_i = \operatorname{argmax}_k \frac{\pi_{ik}}{\sigma_{ik}}$ and $J_i = \operatorname{argmin}_k \frac{\pi_{ik}}{\sigma_{ik}}$, the most likely pixel values of the background are $\mu_{i,I_i,c}$, where the additional subscript c indices the colour channel. Using $\mu_{i,I_i,c}$ together with the pixels from the current frame, x_{ic}, and the foreground probabilities, p_i, in (9) gives estimates of κ_c which allows calculation of new background pixel values that accounts for the altered illumination. The least likely component is replaced with the new background, i.e.

$$\mu_{i,J_i,c} := \kappa_c \mu_{i,I_i,c}, \quad \text{and} \quad \sigma_{i,J_i}^2 := \sigma_{i,I_i}^2.$$

And the probability of the most likely background component is divided evenly between the old and new background components

$$\pi_{i,J_i} := \pi_{i,I_i}/2, \quad \text{and} \quad \pi_{i,I_i} := \pi_{i,I_i}/2,$$

renormalising so that $\sum_k \pi_{i,k} = 1$ if necessary. This replaces the old, most likely, background model with two, equally probable, models relating to the image with and without illumination changes. Note that as much as possible of the old background model is retained by letting the new component replace the component with the lowest mixture weight.

For the algorithm proposed by Lindström *et al.* (2006), the procedure is very similar. Here the pixel values are modelled using a common, global Gaussian mixture for the foreground and local Gaussian mixtures for the background. Let π_{il}^B and μ_{il}^B be the mixture probabilities and expectations of pixel i in the l^{th}, local, background class. We once again introduce, $I_i = \operatorname{argmax}_l \pi_{il}^B$ and $J_i = \operatorname{argmin}_l \pi_{il}^B$, to index the most likely and least likely background. As previously we use (9) to obtain estimates of κ_c, and replace the least likely background with the altered background, i.e. $\mu_{i,J_i,c}^B := \kappa_c \mu_{i,I_i,c}^B$ and $\Sigma_{i,J_i}^B := \Sigma_{i,I_i}^B$. The mixture probabilities, π_{il}^B, are divided evenly between the old and new background components by splitting the cumulative sums (see Lindström *et al.*, 2006, for details), $S^B_{i,J_i} := S^B_{i,I_i}/2$ and $S^B_{i,I_i} := S^B_{i,I_i}/2$.

Another aspect remains to be considered. If the illumination change is detected rapidly, the κ_c used to calculate the new background component (8) will be estimated using only a small portion of the image, giving a reasonable but far from perfect estimate. This will in turn cause a less than optimal new background component. To mitigate the effects of slightly incorrect new background component, the forgetting factor should be increased for a short time period, allowing faster changes to the background model. Increasing the forgetting factor also gives the segmentation algorithms more margin to adapt the covariance matrices of the new background component, avoiding the cumbersome issue of estimating the covariance matrices.

3 Results

The algorithms from (Stauffer & Grimson, 1999) and (Lindström *et al.*, 2006) have been used to segment a video sequence of 20 000 frames. The video is from a traffic surveillance camera under cloudy conditions and contains several illumination changes caused by clouds obscuring the sun.

Parameter values used for the algorithm in (Stauffer & Grimson, 1999) were $\alpha = 0.001$, $T = 0.7$, $K = 5$, $\omega_{\text{init}} = 0.05$ and $\sigma^2_{\text{init}} = \sum_i \left(x_{1i} - \overline{x}_{1i}\right)^2 / (10N)$, i.e. one tenth of the variance in the first frame, with α temporarily increased to 0.01 after a detection. Parameters for the algorithm in (Lindström *et al.*, 2006) were $\alpha = 0.99$, $K_{\max} = 8$, $L_{\max} = 3$, $C = 0.1$, $U_{\text{add}} = 135$, $U_{\text{mem}} = 900$, $\pi^F_{\text{init}} = 0.2$, $\pi^F_{\min} = 0.01$, $v = 3$ and Σ_{init} was taken as one fifth of the covariance matrix from the first frame with the off diagonal elements set to zero. The parameter α was temporarily decreased to 0.90 after a detection. For further details about the parameters and their interpretation, see Stauffer & Grimson (1999) and Lindström *et al.* (2006).

The change detector from Section 2.1 was used to detect anomalous events and three different approaches for handling the illumination changes were applied: 1) do nothing, 2) increase the forgetting factor for a limited time period (150 frames or 10 s), 3) modify the background according to Section 2.2 and increase the forgetting factor for a limited time period (150 frames or 10 s). The parameters of the change detector (3) and (4) were set to $L = 150$ frames (10 seconds), $T = 30$ frames (2 seconds), $C = 2.3263$ and $h = 50$.

Both approaches 2) and 3) reduce the amount of erroneously detected foreground after illumination changes, see Figure 3, with the largest reduction taking place for approach 3). Studying a few frames before and after a detected illumination change (see Figures 4, 5 and 6) shows that only increasing the forgetting factor still leads to a short period with erroneously detected foreground. It also shows that modifying the background

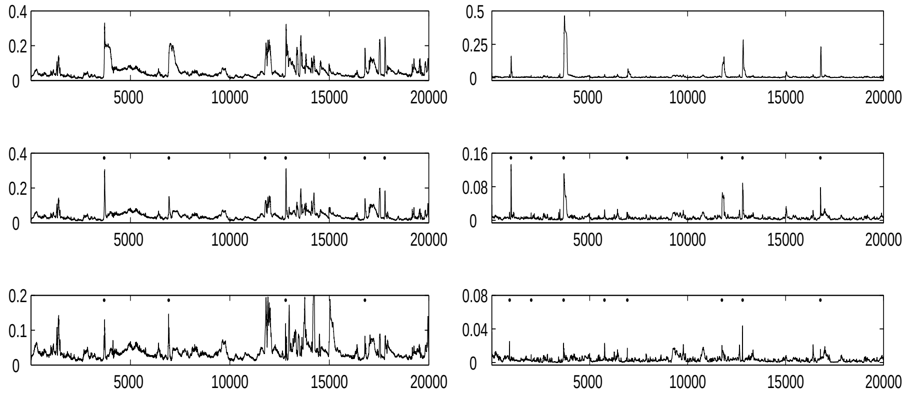

Figure 3: The graphs depict the amount of foreground in the video as a function of the frame number (15 fps) using the algorithm in (Stauffer & Grimson, 1999) (left) and (Lindström *et al.*, 2006) (right) for the three scenarios: 1) no intervention (top), 2) increased forgetting factor (middle) and 3) altered background model and increased forgetting factor (bottom). Intervention points are marked with dots in the two lower figures. Note the difference in vertical scale for the three graphs.

improves the performance for both segmentation algorithms. It should be noted that the change is detected eleven frames later when using the algorithm in (Stauffer & Grimson, 1999). This is due to the slightly higher number of false foreground detections in this algorithm, which reduces the signal to noise ratio of the change detector.

Maintaining the old background introduces robustness against incorrectly detected changes, i.e. false positives, see Figure 7. Here a bus causes an incorrect change detection and the background is modified. However, since the new and old background models are used in parallel after a change the incorrect detection does not degrade the segmentation.

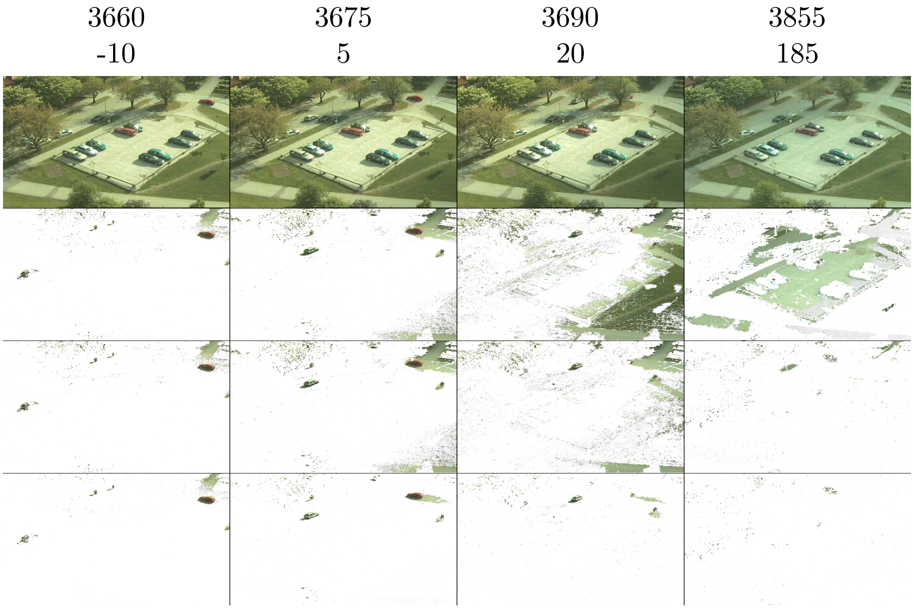

Figure 4: Results for a correctly detected illumination change using the segmentation algorithm in (Lindström *et al.*, 2006). From top to bottom, frame number and frame number relative to the frame at which the change was detected, original data, detected foreground with no intervention, with increased forgetting factor, and with altered background model and increased forgetting factor.

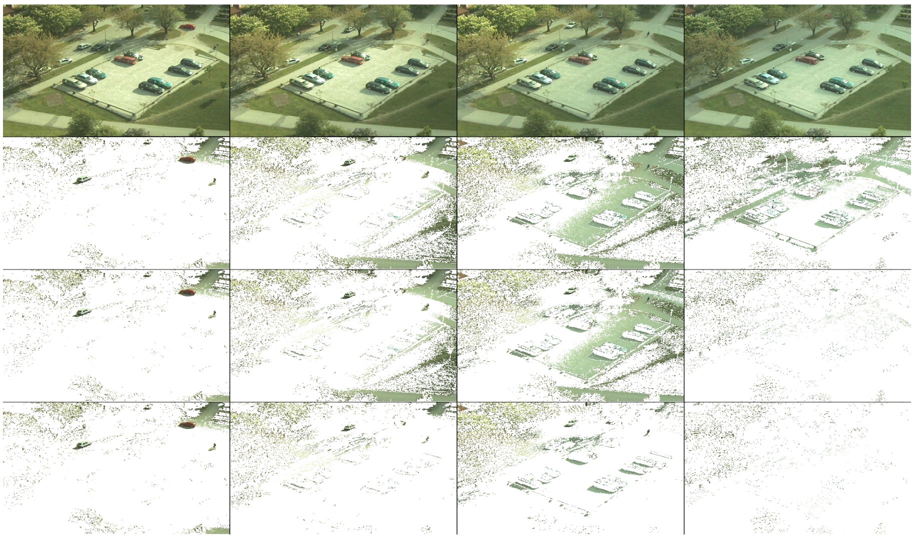

Figure 5: Results for a correctly detected illumination change using the segmentation algorithm in (Stauffer & Grimson, 1999). Note that the change is detected eleven frames later than when using the algorithm in (Lindström *et al.*, 2006). From top to bottom, frame number and frame number relative to the frame at which the change was detected, original data, detected foreground with no intervention, with increased forgetting factor, and with altered background model and increased forgetting factor.

16788 16790

Figure 6: Results for a correctly detected illumination change using the segmentation algorithm in (Stauffer & Grimson, 1999). From top to bottom, frame number, original data, detected foreground with no intervention, and with altered background model and increased forgetting factor. The difference in segmentation prior to the detection is due to previous compensations for illumination changes. Note that the two pedestrians remain after altering the background model.

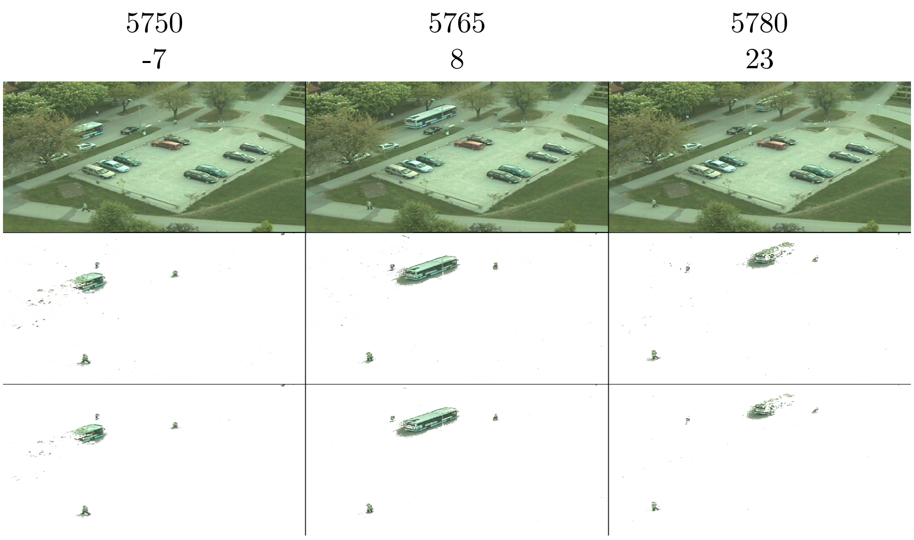

Figure 7: Results for a incorrectly detected illumination change using the segmentation algorithm in (Lindström *et al.*, 2006). From top to bottom, frame number relative to the frame at which the change was detected, original data, detected foreground with no intervention, and with altered background model and increased forgetting factor. Note that the incorrect detection does not affect the segmentation.

4 Conclusions

A method that extends existing segmentation algorithms by detecting illumination changes and adjusting the background model to conform with the new illumination has been presented. Other segmentation algorithms that attempt to handle illumination changes exist but commonly either require more hardware (Ivanov *et al.*, 2000; Lim *et al.*, 2005) or are relatively complex to implement (Javed *et al.*, 2002). The aim of the method presented in this paper is to give a relatively simple extension that allows existing mixture based segmentation algorithms to handle illumination changes.

The method uses a CUSUM detector to find illumination changes followed by weighted regression with errors in both variables to estimate a new background that reflects the altered illumination. Details for the CUSUM detector were given, and possible detection speeds and choices of detector parameters were discussed. Regarding the change detection two things should be noted: 1) The increase in foreground caused by the illumination change has to be notably larger than the variation ordinarily caused by foreground objects, for the change to be detected. 2) A better, i.e. less false positives, segmentation algorithm gives less noise in the total amount of foreground P_t thus requiring smaller changes before detection of an illumination change.

Results for two segmentation algorithms were given and it was indicated how the method can be adapted for use with other segmentation algorithms, allowing a simple extension of existing algorithms that improve their ability to handle rapid illumination changes. The results show that the proposed method significantly reduces the effects of rapid illumination changes in an outdoor surveillance application. Further it was shown that the method is robust against false positives, i.e. incorrect detection of changes.

Acknowledgement

The project has been funded by the Swedish Foundation for Strategic Research (SSF) under grant A3 02:125, Spatial statistics and image analysis for environment and medicine and through the programme Vision in Cognitive Systems II (VISCOS II). The project has also received support from the European Commission's Sixth Framework Program under grant no. 011838 as part of the Integrated Project SMErobot, from the Swedish Road Administration (Vägverket) and from the Swedish Governmental Agency for Innovation Systems (Vinnova).

A Appendix

The change detector (3) contains terms of the kind $\Delta P_t/\sqrt{\hat{s}_t^2}$. To analyse the properties of the change detector we need the expectation and variance this term where ΔP_t and $\hat{s}_t^2$ are defined according to (2) and (4).

In the following superscript arrows, $\vec{x}$, denotes column vectors, with $\vec{1}$ being a vector of ones, bold face, $\mathbf{A}$, denotes matrices with $\mathbf{I}$ being the identity matrix, and $^\top$ denotes transpose. To simplify notation the ΔP_t:s used in the estimate of $\hat{s}_t^2$ are stacked in a vector, $\vec{x}_t$, giving

$$\vec{x}_t = \big(\Delta P_{t-T-1},\ \ldots,\ \Delta P_{t-T-L}\big)^\top.$$

Under the assumptions outlined in Section 2.1 the $\vec{x}$-vector follows a multivariate Gaussian distribution, $\vec{x}_t \in \mathsf{N}\big(\vec{\mu}_t, \sigma^2\mathbf{I}\big)$. The expectation, μ_t, varies with t depending on the position of the change point T_0 relative to $\vec{x}_t$ and consists of: 1) a zero vector, 2) a set of $t-\widehat{T}$ k:s followed by $L+\widehat{T}-t$ 0:s, or 3) a vector containing only k:s, i.e.

$$\vec{\mu}_t = \begin{cases} \vec{0}, & t \le \widehat{T} \\ (\underbrace{k,\ldots,k}_{t-\widehat{T}}, \underbrace{0,\ldots,0}_{L+\widehat{T}-t})^\top, & \widehat{T} < t < L+\widehat{T} \\ k\vec{1}, & L+\widehat{T} \le t \end{cases}.$$

The estimate of σ^2 (4) can now be rewritten as

$$\frac{\hat{s}_t^2(L-1)}{\sigma^2} = \frac{\sum_{i=T+1}^{L+T}\big(\Delta P_{t-i} - \overline{\Delta P}_t\big)^2}{\sigma^2} = \vec{x}^\top \mathbf{A}\vec{x},$$

where $\mathbf{A} = \big(\mathbf{I} - (\vec{1}\vec{1}^\top)/L\big)/\sigma^2$. Now $\big(\hat{s}_t^2(L-1)\big)/\sigma^2$ follows a non-central χ^2 distribution with $L-1$ degrees of freedom and non-centrality parameter δ_t^2, denoted $\chi^2_{L-1}(\delta_t^2)$, iff the following conditions are fulfilled (Styan, 1970, Theorem 4)

$$(\sigma^2\mathbf{I})\big(\mathbf{A}(\sigma^2\mathbf{I})\big)^2 = (\sigma^2\mathbf{I})\mathbf{A}(\sigma^2\mathbf{I}), \tag{10}$$

$$\operatorname{rank}\big((\sigma^2\mathbf{I})\mathbf{A}(\sigma^2\mathbf{I})\big) = \operatorname{tr}\big(\mathbf{A}(\sigma^2\mathbf{I})\big) = L-1, \tag{11}$$

$$\vec{\mu}_t^\top(\mathbf{A}(\sigma^2\mathbf{I}))^2 = \vec{\mu}_t^\top\mathbf{A}(\sigma^2\mathbf{I}), \tag{12}$$

$$\vec{\mu}_t^\top\mathbf{A}(\sigma^2\mathbf{I})\mathbf{A}\vec{\mu}_t = \vec{\mu}_t^\top\mathbf{A}\vec{\mu}_t = \delta_t^2. \tag{13}$$

The matrix $\mathbf{A}\sigma^2$ is a projection matrix and thus idempotent, which implies that (10), (12) and (13) are satisfied. It is obvious that

$$\operatorname{rank}\big(\sigma^2\mathbf{I}\mathbf{A}\sigma^2\mathbf{I}\big) = \operatorname{rank}\big(\mathbf{A}\sigma^2\big) = L-1,$$

with the last equality following since $\mathbf{A}\sigma^2$ is a projection matrix. And since $\operatorname{tr}\big(\mathbf{A}\sigma^2\big) = L - 1$, (11) is fulfilled. Finally δ_t^2 in (13) becomes

$$\delta_t^2 = \begin{cases} \frac{k^2}{\sigma^2}(t - \widehat{T})\big(1 - \frac{t-\widehat{T}}{L}\big), & \widehat{T} < t < L + \widehat{T}, \\ 0, & \text{otherwise.} \end{cases}$$

Having obtained the density for $\hat{s}_t^2$, we rewrite $\Delta P_t / \sqrt{\hat{s}_t^2}$ as

$$\frac{\Delta P_t}{\sqrt{\hat{s}_t^2}} = \frac{\frac{\Delta P_t}{\sigma}}{\sqrt{\frac{\hat{s}_t^2 (L-1)}{\sigma^2}}} \sqrt{L-1}, \tag{14}$$

which gives a random variable on the form of a uni-variate Gaussian divided by the square root of a non-central χ^2 scaled by the square root of the degrees of freedom in the χ^2 distribution. If ΔP_t and $\hat{s}_t^2$ are independent (true if $T \geq 1$) (14) follows a doubly non-central Student's t distribution, see Krishnan (1967), denoted t''. Note that the doubly non-central Student's t distribution in (14) is reduced to a standard Student's t distribution with $L - 1$ degrees of freedom for $t \leq T_0$, since $\delta_t = 0$ and $\mathsf{E}(\Delta P_t) = 0$ if $t \leq T_0$.

The density of a t'' distribution is intractable but the raw moments are given in Krishnan (1967), e.g. the first two moments are

$$\mathsf{E}\Big(\frac{\Delta P_t}{\sqrt{\hat{s}_t^2}}\Big) = \begin{cases} 0, & t \leq T_0, \\ \frac{k}{\sigma}\frac{\Gamma\left(\frac{L-2}{2}\right)}{\Gamma\left(\frac{L-1}{2}\right)}\sqrt{\frac{L-1}{2}}, & T_0 < t \leq \widehat{T}, \\ \frac{k}{\sigma}\frac{\Gamma\left(\frac{L-2}{2}\right)}{\Gamma\left(\frac{L-1}{2}\right)}\sqrt{\frac{L-1}{2}} H\big(\frac{1}{2}, \frac{L-1}{2}; \frac{-\delta_t^2}{2}\big), & \widehat{T} < t < L + \widehat{T}, \\ \frac{k}{\sigma}\frac{\Gamma\left(\frac{L-2}{2}\right)}{\Gamma\left(\frac{L-1}{2}\right)}\sqrt{\frac{L-1}{2}}, & L + \widehat{T} \leq t, \end{cases}$$

and

$$\mathsf{E}\Big(\big(\frac{\Delta P_t}{\sqrt{\hat{s}_t^2}}\big)^2\Big) = \begin{cases} \frac{L-1}{L-3}, & t \leq T_0, \\ \big(1 + \frac{k^2}{\sigma^2}\big)\frac{L-1}{L-3}, & T_0 < t \leq \widehat{T}, \\ \big(1 + \frac{k^2}{\sigma^2}\big)\frac{L-1}{L-3} H\big(1, \frac{L-1}{2}; \frac{-\delta_t^2}{2}\big), & \widehat{T} < t < L + \widehat{T}, \\ \big(1 + \frac{k^2}{\sigma^2}\big)\frac{L-1}{L-3}, & L + \widehat{T} \leq t, \end{cases}$$

where $H\big(1/2, (L-1)/2; -\delta_t^2/2\big)$ is the hyper-geometric function.

Now if the window used to estimate σ^2 is large the degrees of freedom, $L - 1$, will be large and the series expansion for the moments in terms of $1/(L-1)$ given by Krishnan (1967) can be used to obtain approximate expressions for the expectation

$$\mathsf{E}\left(\frac{\Delta P_t}{\sqrt{\hat{s}_t^2}}\right) \approx \begin{cases} 0, & t \leq T_0, \\ \frac{k}{\sigma}\frac{4L-1}{4(L-1)}, & T_0 < t \leq \widehat{T}, \\ \frac{k}{\sigma}\frac{4L-1}{4(L-1)} - \frac{k^3}{\sigma^3}\frac{(t-\widehat{T})\left(1-\frac{t-\widehat{T}}{L}\right)}{2(L-1)}, & \widehat{T} < t < L+\widehat{T} \\ \frac{k}{\sigma}\frac{4L-1}{4(L-1)}, & L+\widehat{T} \leq t. \end{cases}$$

and variance

$$\mathsf{V}\left(\frac{\Delta P_t}{\sqrt{\hat{s}_t^2}}\right) \approx \begin{cases} 1+\frac{2}{L-1}, & t \leq T_0, \\ 1+\frac{2}{L-1}+\frac{k^2}{\sigma^2}\frac{1}{2(L-1)}, & T_0 < t \leq \widehat{T}, \\ 1+\frac{2}{L-1}+\frac{k^2}{\sigma^2}\frac{1-2(t-\widehat{T})\left(1-\frac{t-\widehat{T}}{L}\right)}{2(L-1)}, & \widehat{T} < t < L+\widehat{T}, \\ 1+\frac{2}{L-1}+\frac{k^2}{\sigma^2}\frac{1}{2(L-1)}, & L+\widehat{T} \leq t. \end{cases}$$

References

Elgammal, A. M., Harwood, D. & Davis, L. S. (2000). Non-parametric model for background subtraction. In *Computer vision - ECCV 2000, 6th european conference on computer vision*, vol. 2. Springer, pp. 751–767.

Finlayson, G. D., Hordley, S. D., Marchant, J. A. & Onyango, C. M. (2000). Colour invariance at a pixel. In *Proc. british machine vision conference*. pp. 13–22.

Fuller, W. A. (1987). *Measurement error models*. John Wiley & Sons Ltd.

Gustafsson, F. (2000). *Adaptive filtering and change detection*. John Wiley & Sons Ltd.

Harville, M., Gordon, G. & Woodfill, J. (2001). Foreground segmentation using adaptive mixture models in color and depth. In *IEEE workshop on detection and recognition of events in video*. pp. 3–11.

Hayman, E. & Eklundh, J.-O. (2003). Statistical background subtraction for a mobile observer. In *Ninth IEEE international conference on computer vision*, vol. 1. IEEE, pp. 67–74.

Ivanov, Y., Bobick, A. & Liu, J. (2000). Fast lighting independent background subtraction. *Internat. J. Comput. Vision* **37**, 199–207.

Jabri, S., Duric, Z., Wechsler, H. & Rosenfeld, A. (2000). Detection and location of people in video images using adaptive fusion of color and edge information. In *15th international conference on pattern recognition*, vol. 4. IEEE, pp. 627–630.

Javed, O., Shafique, K. & Shah, M. (2002). A hierarchical approach to robust background subtraction using color and gradient information. In *Workshop on motion and video computing*. IEEE, pp. 22–27.

KaewTraKulPong, P. & Bowden, R. (2001). An improved adaptive background mixture model for real-time tracking with shadow detection. In *2nd european workshop on advanced video-based surveillance systems*. pp. 149–158.

Karman, K.-P. & von Brandt, A. (1990). Moving object recognition using an adaptive background memory. In V. Cappellini, ed., *Time-varying image processing and moving object recognition*, vol. 2. Elsevier, pp. 297–307.

Krishnan, M. (1967). The moments of a doubly noncentral t-distribution. *J. Amer. Statist. Assoc.* **62**, 278–287.

Lim, S.-N., Mittal, A., Davis, L. S. & Paragios, N. (2005). Fast illumination-invariant background subtraction using two views: Error analysis, sensor placement and applications. In *IEEE computer society conference on computer vision and pattern recognition*, vol. 1. IEEE, pp. 1071–1078.

Lindström, J., Lindgren, F., Åström, K., Holst, J. & Holst, U. (2006). Background and foreground modelling using an online EM algorithm. In *The sixth IEEE international workshop on visual surveillance (VS 2006)*. IEEE, pp. 9–16.

Page, E. (1954). Continuous inspection schemes. *Biometrika* **41**, 100–115.

Schindler, K. & Wang, H. (2006). Smooth foreground-background segmentation for video processing. In *Computer vision – ACCV 2006*, vol. 3852 of *Lecture Notes in Computer Science*. Springer-Verlag, pp. 581–590.

Stauffer, C. & Grimson, W. E. L. (1999). Adaptive background mixture models for real-time tracking. In *IEEE computer society conference on computer vision and pattern recognition*, vol. 2. IEEE, pp. 246–252.

Styan, G. P. (1970). Notes on the distribution of quadratic forms in singular normal variables. *Biometrika* **57**, 567–572.

Toyama, K., Krumm, J., Brumitt, B. & Meyers, B. (1999). Wallflower: Principles and practice of background maintenance. In *Seventh IEEE international conference on computer vision*, vol. 1. IEEE, pp. 255–261.

Tsin, Y., Collins, R. T., Ramesh, V. & Kanade, T. (2001). Bayesian color constancy for outdoor object recognition. In *IEEE computer society conference on computer vision and pattern recognition*, vol. 1. IEEE, pp. 1132–1139.

Tuzel, O., Porikli, F. & Meer, P. (2005). A Bayesian approach to background modeling. In *IEEE computer society conference on computer vision and pattern recognition*, vol. 3. IEEE, pp. 58–58.

Preprints in Mathematical Sciences 2008:9
ISSN 1403-9338
LUTFMS-5075-2008

Mathematical Statistics
Centre for Mathematical Sciences
Lund University
Box 118, SE-221 00 Lund, Sweden
http://www.maths.lth.se/