Bayesian Illumination-Invariant Motion Detection

Til Aach and Lutz Dümbgen and Rudolf Mester and Daniel Toth

Institute of Imaging and Computer Vision
RWTH Aachen University, 52056 Aachen, Germany
tel: +49 241 80 27860, fax: +49 241 80 22200
web: www.lfb.rwth-aachen.de

in: Proceedings IEEE International Conference on Image Processing (ICIP), Vol. III. See also BIBTEX entry below.

RWTHAACHEN
RHEINISCH-WESTFÄLISCHE TECHNISCHE HOCHSCHULE AACHEN

BAYESIAN ILLUMINATION-INVARIANT MOTION DETECTION

Til Aach [1], Lutz Dümbgen [2]

[1]Institute for Signal Processing,
[2]Institute for Mathematics,
University of Lübeck,
Ratzeburger Allee 160, 23538 Germany
aach@isip.mu-luebeck.de

Rudolf Mester [3], Daniel Toth [1]

[3]Institute for Applied Physics,
University of Frankfurt,
Robert-Mayer-Str. 2-4,
60054 Frankfurt, Germany
mester@iap.uni-frankfurt.de

ABSTRACT

We describe an algorithm for change detection which is insensitive to both slow and fast temporal variations of scene illumination. Our algorithm is based on statistical decision theory by using a Bayesian approach. The goal is to detect only temporal changes which are induced by true scene changes, like motion, but not changes due to noise or varying illumination. To this end, our algorithm uses a simple illumination model which is invariant to common camera nonlinearities like gamma-nonlinearity. This is combined with a model for the influence of noise as well as an a priori model for the expected properties of the sought change masks. Key ingredients of the resulting algorithm are a suitable test statistic and an adaptive threshold mechanism. As the algorithm can be applied in a noniterative manner, it is also computationally attractive.

1. INTRODUCTION

In scene analysis for object oriented coding schemes or surveillance applications, it is often required to detect and segment moving objects in video data which are acquired by a static camera [1, 2, 3, 4]. Such changes include the projections of moving objects and their potential shadows onto the image plane together with any uncovered background. Statistical methods to detect such changes in the presence of noise exist, both with [5, 6, 7] and without [1, 8] the use of a priori models for the sought change masks. First, these methods employ statistically optimal test statistics. Secondly, they couple the detection procedure to error probabilities, usually the false alarm rate via a significance test. Thirdly, Bayesian approaches using a priori knowledge about the change masks strongly reduce error probabilities by making both noise-like false alarms and holes in changed areas unlikely to appear. Indeed, the benefits of the Bayesian approach in [5, 6, 7] have led to its use in a variety of applications, e.g. [4, 9, 10].

Since these algorithms rely on the comparison of the grey levels at the same locations in subsequent images of a video sequence, they are insensitive with respect to slow temporal variations of scene illumination. The reason for this is that subtraction of subsequent frames, as done in [5, 6, 7], is a temporal highpass operation. Fast changes of scene illumination, however, are not sufficiently suppressed and may therefore lead to detection errors in the sense of motion detection. For applications where fast illumination changes cannot be ruled out, so-called illumination-invariant methods have been developed [11, 12, 13]. Not all of these are based on statistical decision theory, and none of these use a priori models for the change masks. In [14, 15], we proposed two-step approaches to illumination invariant change detection that employ spatial [14] or spatiotemporal [15] Markov Random Fields (MRFs) as a priori models. In a first step, the influence of illumination is suppressed by a homomorphic filter (cf. [16, 17]) before the Bayesian algorithm of [7, 6] is applied. This method thus uses both statistical decision theory and an a priori model.

Unlike separating suppression of illumination and change detection in two steps, [13] uses an ad hoc test statistic which is, within certain limits, insensitive to changes of illumination in successive image blocks. However, neither a statistical framework nor a priori models are invoked. In this paper, we derive a new approach to illumination-insensitive change detection that combines key ideas from [6, 7, 13, 14, 15]. The resulting one-step method is computationally similarly appealing as our earlier algorithms in [6, 7].

2. THE BAYESIAN APPROACH

2.1. The illumination model

Homomorphic filtering [16] models the recorded grey levels $g(m, n)$ as the product of scene illumination $i(m, n)$ and surface reflectance $r(m, n)$. Clearly, structural scene changes are captured by $r(m, n)$. Scene illumination is assumed to vary slowly over the spatial coordinates (m, n), and can hence be suppressed by applying a linear highpass

filter to $\log(g(m,n))$. A similar, but slightly more stringent illumination model is used in [13], where illumination is modelled as a constant factor within small image blocks. We note here that this relationship between illumination and surface reflectance may be altered by potential camera non-linearities. Commonly, however, the camera gain is described by a so-called γ-curve, for which it can be shown that the multiplicative relationship is preserved [14, 15].

2.2. Change detection as hypothesis testing

To decide whether or not a change did occur between the successive frames G_t and G_{t-1} at pixel (j,l), we compare the grey levels from G_t and G_{t-1} which lie within a small sliding window, which is centered around (j,l). These grey levels are ordered into column vectors $\mathbf{x}$ and $\mathbf{y}$, respectively. If the window contains N pixels, each of these vectors contains N components $x(n)$ and $y(n)$, respectively, where $n = 1, 2, \ldots, N$. If no scene change occurs within the window, and neglecting the effects of noise, both vectors would be identical if illumination remained constant between times $t-1$ and t. Under the same circumstances, a change in illumination would change the norms of these vectors, but not their directions. Consequently, in [13] an ad-hoc test function is employed which detects whether or not $\mathbf{x}$ and $\mathbf{y}$ are linearly dependent. If they are, a structural change did not occur. Here, we now express this reasoning as a hypothesis test.

Let $\mathbf{s} = (s(1), \ldots, s(N))$ denote the noisefree signal in $\mathbf{x}$. Assuming additive white Gaussian camera noise, we observe

$$\mathbf{x} = \mathbf{s} + \epsilon_1 \tag{1}$$

where ϵ_1 is a noise vector obeying $\mathcal{N}(0, \sigma^2 \mathbf{I})$. The parameter σ^2 is the variance of the camera noise, and $\mathbf{I}$ the identity matrix. If no scene change occurs, $\mathbf{y}$ contains the same signal as $\mathbf{x}$, which may be scaled by a factor k. Hence,

$$\mathbf{y} = k \cdot \mathbf{s} + \epsilon_2 \tag{2}$$

where ϵ_2 is another realization of the camera noise which is independent of ϵ_1. If the illumination remains constant, we have $k = 1$, otherwise $0 < k < 1$. (This can always be achieved by assigning $\mathbf{x}$ and $\mathbf{y}$ accordingly to the frames G_t and G_{t-1}, and implies no loss of generality. When deriving the distribution of our test statistic, we will at one point approximate the norm of $\mathbf{x}$ — but not its direction — by the norm of $\mathbf{s}$. The relative approximation error is smaller when $\mathbf{x}$ is chosen as done above.) We call the hypothesis that $\mathbf{x}$ and $\mathbf{y}$ are given by (1) and (2) the null hypothesis H_0.

Let $|\mathbf{x}|$ denote the Euclidean norm of $\mathbf{x}$, and $\mathbf{x}_U$ the unit vector pointing in the direction of $\mathbf{x}$, given by $\mathbf{x}_U = \mathbf{x}/|\mathbf{x}|$. The transpose of $\mathbf{x}$ is denoted by $\mathbf{x}^T$. To find a suitable test statistic which allows us to decide whether or not to accept H_0 based on the noisy observations $\mathbf{x}$ and $\mathbf{y}$, let us consider

the diagram in Fig. 1. We project $\mathbf{y}$ onto $\mathbf{x}$, yielding a new vector that is parallel to $\mathbf{x}$, and whose norm is given by the absolute inner product $|\mathbf{x}_U^T \mathbf{y}|$. Clearly, this vector can be expressed as $\mathbf{x}_U^T \mathbf{y} \cdot \mathbf{x}_U$. Subtracting this vector from $\mathbf{y}$ yields the difference vector

$$\mathbf{d} = \mathbf{y} - \mathbf{x}_U^T \mathbf{y} \cdot \mathbf{x}_U \tag{3}$$

which is zero if and only if $\mathbf{x}$ and $\mathbf{y}$ are collinear. The squared norm of $\mathbf{d}$ can be calculated as

$$|\mathbf{d}|^2 = |\mathbf{y} - \mathbf{x}_U^T \mathbf{y} \cdot \mathbf{x}_U|^2 = |\mathbf{y}|^2 - (\mathbf{x}_U^T \mathbf{y})^2 \tag{4}$$

since $\mathbf{x}_U^T \mathbf{y} \mathbf{x}_U$ and $\mathbf{d}$ are orthogonal (Fig. 1). Testing H_0 is then equivalent to testing if $|\mathbf{d}|^2$ can be explained by camera noise. Like in [5, 7], we assume that the variance σ^2 of the camera noise is known, or can be estimated recursively from unchanged areas in the images.
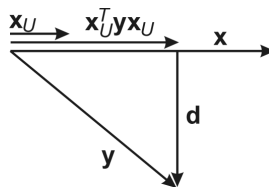

Fig. 1. Geometrical interpretation of the projection of $\mathbf{y}$ onto $\mathbf{x}$, given by $\mathbf{x}_U^T \mathbf{y} \cdot \mathbf{x}_U$. Also depicted is the difference vector $\mathbf{d}$.

As we will show in the final paper, $\mathbf{d}$ can be expressed in a suitable basis (which depends on $\mathbf{s}$) such that its components $d(1), d(2), \ldots d(N-1)$ are approximately independent, while $d(N)$ is negligible. Each component $d(n)$ except $d(N)$ is then normally distributed with variance

$$\sigma_d^2 = \sigma^2 \cdot \left(1 + k^2\right) \tag{5}$$

if the null hypothesis H_0 holds, yielding the conditional probability density function (pdf)

$$p(\mathbf{d}|H_0) = \left(\frac{1}{\sqrt{2\pi}\sigma_d}\right)^{N-1} \cdot \exp\left\{-\frac{|\mathbf{d}|^2}{2 \cdot \sigma_d^2}\right\} \tag{6}$$

Hence, the squared norm of $\mathbf{d}$ normalized by the variance σ_d^2 is χ^2-distributed with $N-1$ degrees of freedom. Since the parameter k is not known, we estimate its value from the noisy observations to $\hat{k} = |\mathbf{y}|/|\mathbf{x}|$. Inserting this estimate into σ_d^2, we finally obtain the test statistic

$$T = \frac{|\mathbf{y}|^2 - (\mathbf{x}_U^T \mathbf{y})^2}{\sigma^2 \cdot \left(1 + \frac{|\mathbf{y}|^2}{|\mathbf{x}|^2}\right)} = \frac{|\mathbf{x}|^2 |\mathbf{y}|^2 - \left(\mathbf{x}^T \cdot \mathbf{y}\right)^2}{\sigma^2 \cdot (|\mathbf{x}|^2 + |\mathbf{y}|^2)} \tag{7}$$

The denominator of T is just a normalization factor, while the numerator is easily interpreted in terms of the Cauchy-Schwarz inequality: it is zero if and only if $\mathbf{x}$ and $\mathbf{y}$ are collinear, otherwise it is larger than zero. The pdf $p(T|H_0)$

of T conditioned on H_0 is approximately a χ^2-pdf with $N - 1$ degrees of freedom. A significance test can now be carried out as follows [18, 5]: we specify an acceptable false alarm rate α, from which a decision threshold t is derived such that

$$\alpha = \text{Prob}(T > t|H_0) = \int_t^\infty p(T|H_0)dT \qquad (8)$$

using a χ^2-table for $N - 1$ degrees of freedom. If T exceeds t, we reject H_0, and assign the label c for changed to the window centre, otherwise, we accept H_0 and assign the label u for unchanged. We express this decision by

$$T \overset{c}{\underset{u}{\gtrless}} t \qquad (9)$$

Eq. (9) is a valid decision procedure. It does, however, not yet allow the use of an a priori model. In the following, we therefore briefly sketch how (9) is developed into a maximum a posteriori (MAP) decision by extending the derivations in [7], where the details can be found.

2.3. MAP change detection

Given the alternative hypothesis H_1, we model the conditional pdf $p(\mathbf{d}|H_1)$ by [5, 7]

$$p(\mathbf{d}|H_1) = \left(\frac{1}{\sqrt{2\pi}\sigma_c}\right)^{N-1} \cdot \exp\left\{-\frac{|\mathbf{d}|^2}{2 \cdot \sigma_c^2}\right\} \qquad (10)$$

with $\sigma_c^2 \gg \sigma^2$. Furthermore, we model the sought change masks by an MRF such that the detected changed regions tend to be compact and smoothly shaped. From this model, a priori probabilities $\text{Prob}(c)$ and $\text{Prob}(u)$ for the labels c and u can be obtained. The MAP decision rule then is

$$\frac{p(\mathbf{d}|H_1)}{p(\mathbf{d}|H_0)} \overset{c}{\underset{u}{\gtrless}} \frac{\text{Prob}(u)}{\text{Prob}(c)} \qquad (11)$$

This can be manipulated into the context adaptive decision [7]

$$T \overset{c}{\underset{u}{\gtrless}} t + (4 - \nu_c) \cdot B \qquad (12)$$

where T is the test statistic of (7), and t the threshold according to (8). The parameter ν_c denotes the number of pixels that carry the label c and lie in the 3×3-neighbourhood of the pixel to be processed (Fig. 2). These labels are known for those neighbouring pixels which have already been processed while scanning the image raster (causal neighbourhood). For the pixels which are not yet processed we simply take the labels from the previous change mask (anticausal neighbourhood). Clearly, the adaptive threshold on the right hand side of (12) can only take the nine different values $\nu_c = 0, 1, \ldots, 8$. The parameter B is a positive cost.

The adaptive threshold hence is the lower, the higher the number ν_c of adjacent pixels with label c. It is obvious that this behaviour favours the emergence of smoothly shaped changed regions, and discourages noise-like decision errors. The nine different possible values for the adaptive threshold can be precomputed and stored in a look-up table.
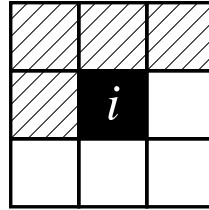

Fig. 2. 3×3-neighbourhood of a pixel i, with its causal neighbours shown shaded.

3. RESULTS

Fig. 3 shows one frame of a sequence where a beam of light moves quickly across the scene. Application of the illumination sensitive Bayesian algorithm of [6, 7] generates the expected compact change masks, but detects both the moving engines and the changes caused by varying illumination. The significance test of (9) using the illumination insensitive test statistic of (7) results in a change mask which almost completely ignores the effects of light, but contains noise-like false alarms. These are eliminated by the MRF-based MAP-algorithm.
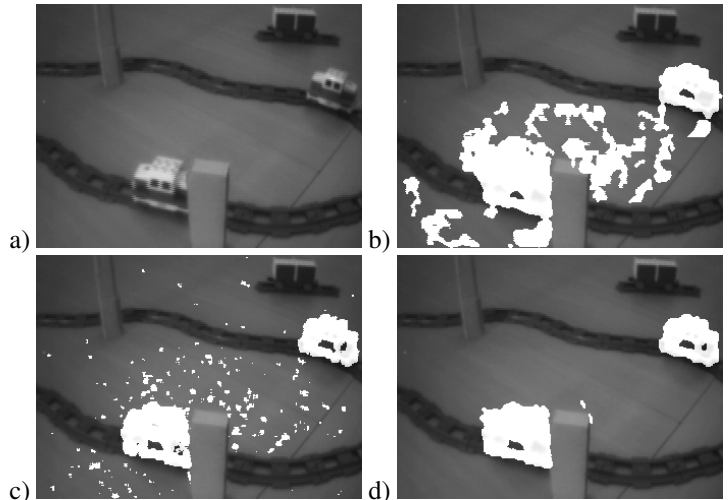

Fig. 3. a) Original frame from a sequence with moving toy engines. A beam of light crosses this scene quickly from left to right. b) Result of the illumination sensitive change detection algorithm in [7]. c) Result of illumination invariant change detection according to Eqs. (7), (8) and (9). d) Result of the MAP-algorithm according to Eqs. (7), (8) and (12).

4. CONCLUSIONS

We have described a new algorithm for illumination-invariant change detection by combining an illumination model from [13] with decision theoretic approaches to change detection from [5, 6, 7]. The core of our algorithm is a new statistical test for linear dependence of noisy observations arranged into vectors, which can be carried out as a significance test or be integrated into a Bayesian framework. In the latter approach, an MRF-based prior model for the sought change masks can be brought bear in a similar manner as in our earlier illumination-sensitive change detection methods. Noise-like decision errors are then almost fully eliminated.

5. REFERENCES

[1] Y. Z. Hsu, H.-H. Nagel, and G. Rekers, "New likelihood test methods for change detection in image sequences," *Computer Vision, Graphics, and Image Processing* vol. 26, pp. 73–106, 1984.

[2] M. Hötter, R. Mester, and F. Müller, "Detection and description of moving objects by stochastic modelling and analysis of complex scenes," *Signal Processing: Image Communication* vol. 8, pp. 281–293, 1996.

[3] A. Mitiche and P. Bouthemy, "Computation and analysis of image motion: A synopsis of current problems and methods," *International Journal of Computer Vision*, vol. 19, no. 1, pp. 29–55, 1996.

[4] ISO/IEC JTC 1/SC 29/WG 11 N2688, "Information technology - generic coding of audio-visual objects - part 2: Visual," Proposed Draft Amendment, Annex F, Seoul, 1999.

[5] T. Aach, A. Kaup, and R. Mester, "Statistical model-based change detection in moving video," *Signal Processing*, vol. 31, pp. 165–180, 1993.

[6] T. Aach, A. Kaup, and R. Mester, "Change detection in image sequences using Gibbs random fields," in *Proc. Intl. Works. Intell. Sig. Proc. Comm. Syst.*, Sendai, Japan, October 1993, pp. 56–61.

[7] T. Aach and A. Kaup, "Bayesian algorithms for change detection in image sequences using Markov random fields," *Signal Processing: Image Communication*, vol. 7, no. 2, pp. 147–160, 1995.

[8] J. F. Y. Cheung, M. C. Wicks, G. J. Genello, and L. Kurz, "A statistical theory for optimal detection of moving objects in variable corruptive noise," *IEEE Trans. Image Proc.*, vol. 8, no. 12, pp. 1772–1787, 1999.

[9] R. Mech and M. Wollborn, "A noise robust method for segmentation of moving objects in video sequences," in *Proc. ICASSP*, Munich, Germany, April 1997, pp. 2657–2660.

[10] H. Luo, A. Eleftheriadis, and J. Kouloheris, "Statistical model-based video segmentation and its application to very low bit-rate video coding," *Signal Processing: Image Communication*, vol. 16, pp. 333–352, 2000.

[11] K. Skifstad and R. Jain, "Illumination independent change detection for real world image sequences," *Computer Vision, Graphics, and Image Processing*, vol. 46, pp. 387–399, 1989.

[12] S.-Z. Liu, C.-W. Fu, and S. Chang, "Statistical change detection with moments under time-varying illumination," *IEEE Trans. Image Proc.*, vol. 7, no. 9, pp. 1258–1268, 1998.

[13] E. Durucan and T. Ebrahimi, "Robust and illumination invariant change detection based on linear dependence for surveillance application," in *Europ. Sig. Proc. Conf. (EUSIPCO*, M. Gabbouj and P. Kuosmanen, Eds., Tampere, Finland, Sept. 3-8 2000, pp. 1041–1044.

[14] D. Toth, T. Aach, and V. Metzler, "Illumination-invariant change detection," in *4th IEEE Southwest Symp. Image Anal. Interpret. (SSIAI)*, Austin, TX, April 2 - 4 2000, pp. 3–7.

[15] D. Toth, T. Aach, and V. Metzler, "Bayesian spatio-temporal motion detection under varying illumination," in *Europ. Sig. Proc. Conf. (EUSIPCO*, M. Gabbouj and P. Kuosmanen, Eds., Tampere, Finland, Sept. 3-8 2000, pp. 2081–2084.

[16] A. V. Oppenheim, R. W. Schafer, and T. G. Stockham Jr., "Nonlinear filtering of multiplied and convolved signals," *Proc. IEEE*, vol. 56, no. 8, pp. 1264–1291, 1968.

[17] C. R. Moloney, "Methods for illumination-independent processing of digital images," in *IEEE Pacific Rim Conf. Comm. Comp. Sig. Proc.*. May 9–10 1991, pp. 811–814.

[18] C. W. Therrien, T. F. Quatieri, and D. E. Dudgeon, "Statistical model-based algorithms for image analysis," *Proc. IEEE*, vol. 74, no. 4, pp. 532–551, 1986.