A Kalman Filter Based Background Updating Algorithm Robust to Sharp Illumination Changes

Stefano Messelodi[2], Carla Maria Modena[2], Nicola Segata[1], and Michele Zanin[2]

[1] University of Trento, Italy,
[2] ITC-irst, Via Sommarive, 18 - 38050 Povo (Trento), Italy

Abstract. A novel algorithm, based on Kalman filtering is presented for updating the background image within video sequences. Unlike existing implementations of the Kalman filter for this task, our algorithm is able to deal with both gradual and sudden global illumination changes. The basic idea is to measure global illumination change and to use it as an external control of the filter. This allows the system to better fit the assumptions about the process to be modeled. Moreover, we propose methods to estimate measurement noise variance and to deal with the problem of saturated pixels, to improve the accuracy and robustness of the algorithm. The algorithm has been successfully tested in a traffic surveillance task by comparing it to a background updating algorithm, based on Kalman filtering, taken from literature.

1 Introduction

The most popular techniques for the detection of moving objects in complex environments observed by a static camera, are based on the background differencing method. It consists of maintaining a background image of the scene and detecting foreground objects by subtracting it from the current frame. Background updating is a critical task in outdoor scenes which undergo significant changes caused both by natural events, e.g. the sun suddenly disappearing behind clouds, and artificial events, like the change of the exposure time of the acquisition device, or the switching-on of artificial lights. A background updating module should be able to detect static background pixels, dynamic background pixels and to deal with gradual and sharp illumination changes. Many different methods have been proposed in literature that vary in their adopted features. A common classification distinguishes the techniques depending on the features they use: a) pixel level [8,2,4,3,9,10]: only the temporal distribution of intensities is used, b) region level [12]: a small neighbourhood of each pixel is considered to take into account local structure; c) spatial-temporal level [13,11]: both spatial and temporal features are used mainly to detect non-static background pixels (moving leaves, rippling water,...). Several techniques take into account the result of foreground object detection as a feedback in order to apply different updating criteria to the background and foreground regions. As background updating is

a preliminary step within more complex surveillance systems, a typical requirement is computational efficiency. Kalman filter techniques are characterized by low computational cost and, being based on a solid statistical model, by a good robustness level. Their use for background updating was firstly proposed in [2]. Some modifications have been recently presented in order to better manage slow illumination variations [3]. This method can adapt to gradual changes but fails in cases of sudden change. In this paper we propose a different use of the Kalman filter paradigm and present a background updating algorithm able to deal with both gradual and sharp global illumination changes. Moreover, we address two problems that can complicate the functioning of predictive filters: the estimation of the measurement noise variance and the management of saturated pixels. Section 2 describes the use of the Kalman filter schema for background updating, Section 3 motivates and illustrates our method. Experimental results are described in Section 4 and Section 5 concludes the paper.

2   Kalman Filter for Background Updating

The Kalman filter [1,5] is an optimal estimator of the state of processes which satisfies: (a) they can be modeled by a linear system, (b) the measurement and the process noise are *white*, and have zero mean gaussian distributions. Under these conditions, knowing the input (external controls u_t) and the output (measurements z_t) of the system, the Kalman filter provides an optimal estimate of the state of the process (x_t), by minimizing the variance of the estimation error and constraining the average of the estimated outputs and the average of the measures to be the same. It is characterized by two main equations: the state equation (1) and the measurement equation (2):

$$x_t = Ax_{t-1} + Bu_{t-1} + w_{t-1} \tag{1}$$
$$z_t = Cx_t + v_t \tag{2}$$

A is the state transition matrix, B is the external control transition matrix, w represents the process noise, C is the transition matrix that maps the process state to the measurement, and v represents the measurement noise. The Kalman filter works in two steps: prediction and correction steps. The former uses the state of the system and the external control at time $t-1$ to predict the current state $(\hat{x}_t^-)$, the latter uses the current measure z_t to correct the state estimation $(\hat{x}_t)$. The working schema of the Kalman filter is illustrated in Figure 1. The factor K_t, the *gain of the filter*, is chosen in order to minimize the variance of the estimate error (P_t). The difference between the measure and the state predicted value $(z_t - C\hat{x}_t^-)$ is called *innovation*.

The application of Kalman theory to the background updating task, e.g. [2] [4], typically considers the temporal distribution of the intensity levels of each pixel p and models it with a Kalman filter: the state of the system is the background value of pixel $x_t(p)$, and the measurement at time t is the value $I_t(p)$ of the pixel in the current image. The system input term (u_t) is set to zero, and the
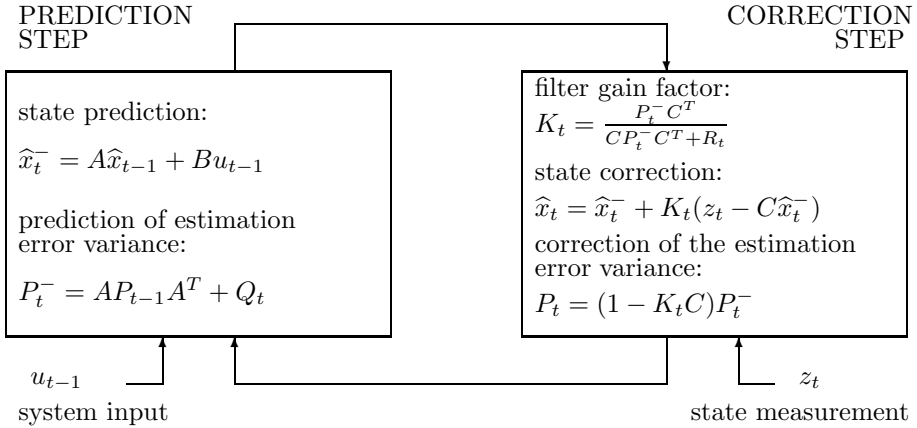
PREDICTION
STEP

CORRECTION
STEP

state prediction:

$$\widehat{x_t^-} = A\widehat{x}_{t-1} + Bu_{t-1}$$

prediction of estimation
error variance:

$$P_t^- = AP_{t-1}A^T + Q_t$$

filter gain factor:
$$K_t = \frac{P_t^- C^T}{CP_t^- C^T + R_t}$$

state correction:
$$\widehat{x}_t = \widehat{x_t^-} + K_t(z_t - C\widehat{x_t^-})$$

correction of the estimation
error variance:
$$P_t = (1 - K_t C)P_t^-$$

u_{t-1}

system input

z_t

state measurement

Fig. 1. The general schema of a Kalman filter. R_t and Q_t represent the variances of the gaussian random variables describing, respectively, the measurment noise and the process noise: $v_t = N(0, R_t)$ and $w_t = N(0, Q_t)$.

temporal distributions of the background intensities are considered constant. All unexpected changes are described by the random noise w_t, which by hypothesis is a zero mean gaussian variable. In order to prevent foreground pixels modifying the background image, a different gain factor is introduced if the innovation overcomes a given threshold. In this approach, gradual illumination changes can be captured only by the random noise term (w_t) which has to vary in time according to them. In this way the zero mean hypothesis could be, in principle, violated. Moreover, sudden illumination changes cause intensity variations that are considered as foreground pixels and cannot be correctly managed. In [3] a new term is introduced to model the intensity variations caused by gradual illumination changes. The random noise term is split into a slow varying component (β_t) and a zero mean gaussian component (μ_t). In this way the state of the system consists of two components: the background value (x_t) and its changing rate (β_t) whose role is to model gradual illumination variations. In order to deal with foreground objects the authors introduce a statistical test which takes into consideration the innovation value. If the innovation is below a given threshold they say that the measure accords with the prediction and the filter works as usual (the filter is open), otherwise a different updating criterion is adopted (the filter is closed). The threshold, computed on a statistical base, is related to the sum of the estimate error variance (P) and the measurement noise variance (R).

3   The Proposed Background Updating Module

We argue that it is incorrect to model all variations in the background intensities as noise added to the process. Variations caused by global illumination changes are external events that are logically different from variations due to the presence

of foreground objects. We develop a module that measures such global changes, and use this information as an external input to the system. The module exploits the fact that global illumination changes affect in a consistent way all the pixels in the image while variations due to foreground objects involve only a subset of them. In the subsequent paragraphs, the three main contributions of our paper are detailed, followed by the final schema of the proposed Kalman filter.

Estimation of global illumination changes. Following the light reflection model introduced by [6], the intensity at a given point on an object is the product of the illumination and a shading coefficient which is a characteristic of the object. A global illumination variation causes a modification of the pixel values that is approximately proportional to the preceding ones. Although this is an oversimplification, the introduced errors are generally acceptable and they can be modeled inside the Kalman filtering framework. To detect global illumination changes we analyze the distribution of the ratios $I_t(p)/x_{t-1}(p)$, for each pixel p. If the illumination is stable the distribution has a main peak around the value 1 and only foreground pixels contribute to the distribution tails. When a change takes place the main peak shifts towards values greater than 1 if the illumination increases, lower than 1 otherwise. Under the hypothesis that the majority of pixels belong to the background, the median of the distribution of ratios provides a global estimation of the changing factor. We verify experimentally that a single factor is not adequate for all tbackground pixels, but slight local variations permit the system to accommodate for the non uniform reflection characteristics of different image locations. These local variations with respect to the global changing factor have to be small enough to prevent foreground pixels to be detected as background. This is controlled by a threshold ($K_{thr} = 1.02$ in the experiments). The estimation procedure is reported in Algorithm 1.

Estimation of the measurement noise variance. The parameter R plays a crucial role in the functioning of the filter. Generally, it is assumed independent of time (t) and pixel position (p) and is estimated by analysing a portion of the input sequence. After running the background updating module on several sequences, acquired by devices (camera/lenses) with different characteristics and encoded in different ways, we noted a non homogeneous behaviour in correspondence of dark and bright pixels. We verified that one cause was the inaccurate estimation of the measurement error variance. To reduce this effect, we estimate the parameter R considering the intensity level ℓ of the pixel p. The estimation procedure is reported in Algorithm 2 ($N = 400$ and $L = 256$ the experiments).

Management of saturated pixels. Saturated pixels have the maximum intensity level L. Their actual intensity is unknown as all intensities above the maximum are mapped to L. Let us consider a background image where some pixels have the value L and after an increase of illumination a number of new pixels assume the value L. When the illumination comes back to the previous level, all the pixels in the image reduce their intensity levels by approximatively the same factor. As a result all of the saturated pixels receive incorrectly the

---

Algorithm 1 Estimate of the global illumination change

---

divide image I_t into a rectangular grid r_{ij} $(i = 1 \ldots N, j = 1 \ldots M)$
for each rectangle r_{ij} **do**
  $D_{ij} \leftarrow$ distribution of ratios $I_t(p)/x_{t-1}(p)$ for all pixel $p \in r_{ij}$ such that:
  $(x_{t-1}(p) \neq 0) \wedge (x_{t-1}(p) < L) \wedge (I(p) < L)$ {deal with saturated pixel}
  $k_t(ij) \leftarrow$ median of D_{ij}
end for
$K \leftarrow$ median of $k_t(ij)$ over all rectangles r_{ij}
for each rectangle r_{ij} **do**
  $k_t(ij) \leftarrow unknown$ if $(k_t(ij)/K > K_{thr}) \vee (k_t(ij)/K < 1/K_{thr})$
end for
for each rectangle r_{ij} such that $k_t(ij) = unknown$ **do**
  $k_t(ij) \leftarrow$ average of K and $k_t(uv)$, with r_{uv} adjacent to r_{ij} and $k_t(uv) \neq unknown$
end for
for all pixels p **do**
  let r_1, r_2, r_3, r_4 be the 4 rectangles closest to p
  let w_1, w_2, w_3, w_4 be the distance from p to the rectangles centers
  $k_t(p) \leftarrow (\sum_{h=1}^{4} k_t(h)w_h)/\sum_{h=1}^{4} w_h$
end for
return the map $k_t(p)$

Algorithm 2 Estimate of the measurement noise variance

---

consider the image sequence $I_1(p) \ldots I_N(p)$
let L be the maximum quantization level of the illumination intensity
for each level $\ell \in [0, L]$ **do**
  build an histogram H_ℓ and initialize it to 0
end for
for all image pair (I_k, I_{k+1}) with $k = 1 \ldots N - 1$ **do**
  for all pixels p **do**
    $\ell \leftarrow I_k(p)$
    increment $H_\ell[I_{k+1}]$
  end for
end for
for each level $\ell \in [0, L]$ **do**
  $R_\ell \leftarrow$ median of the H_ℓ portion between its 10th and 90th percentile
  $R_\ell \leftarrow R_\ell/\sqrt{2}$
end for
smooth R_ℓ with a mean filter of length $w \simeq 0.1 * L$
$m \leftarrow \arg\max(R_\ell)$
for each level from $m + 1$ to L **do**
  $R_\ell \leftarrow R_m$
end for
return the vector R

Table 1. $\hat{x}_t$ estimate of the background value; $\hat{x}_t^-$ *a priori* estimate of the background value; $P_{g,t}$ estimation error variance; $P_{g,t}^-$ *a priori* estimation error variance; $K_{g,t}$ gain factor of the filter when the measure is compatible with the prediction (the filter is open); $K'_{g,t}$ gain of the filter when there is incompatibility between prediction and measure (the filter is closed); z_t is the current measure; α controls the noise of the k_t estimation process; γ controls the test for establishing if the prediction and the measure are compatible (if to open or to close the filter); ϱ controls the gain of the closed filter.

Initialization
Estimate initial background image $x_0(p)$
Estimate the measurement noise variance R_ℓ (Algorithm 2)
Initialize a map of labels: $L(p) \leftarrow \begin{cases} unknown \text{ if } x_0(p) = L \\ known \quad \text{otherwise} \end{cases}$
Initialize a map of counters: $C(p) \leftarrow 0$

For each frame t
Estimate the illumination change factors map $k_t(p)$ (Algorithm 1)

For each pixel p		
$\hat{x}_t^- \leftarrow k_t x_{t-1}$	$\Delta \leftarrow z_t - \hat{x}_t^-$	Prediction
$\ell \leftarrow Range_0^L(\hat{x}_t^-)$	$\eta_{\ell,t} \leftarrow \gamma\sqrt{P_{\ell,t}^- + R_\ell}$	step
$P_{\ell,t}^- \leftarrow (\alpha|1-k_t|+1)P_{\ell,t-1}^-$		

$L(p)$						
unknown	$C(p) \le 0$	true	$z_t < L - \eta_{\ell,t}$	true	$\hat{x}_t = z_t$; increment $C(p)$	Correction step
				false		
		false	$\Delta^2 < \eta_{\ell,t}$	true	increment $C(p)$; if $(C(p) > C_{thr})$ $L(p) = known$; $K_{\ell,t} = \frac{P_{\ell,t}^-}{P_{\ell,t}^- + R_\ell}$; $\hat{x}_t = \hat{x}_t^- + K_{\ell,t}\Delta$; $P_{\ell,t} = (1 - K_{\ell,t})P_{\ell,t}^-$	
				false	decrement $C(p)$; if $(C(p) = 0)$ $\hat{x}_t = L$	
known	$\hat{x}_t^- < L$	true	$\Delta^2 < \eta_{\ell,t}$	true	$K_{\ell,t} = \frac{P_{\ell,t}^-}{P_{\ell,t}^- + R_\ell}$; $\hat{x}_t = \hat{x}_t^- + K_{\ell,t}\Delta$; $P_{\ell,t} = (1 - K_{\ell,t})P_{\ell,t}^-$	
				false	$K'_{\ell,t} = \varrho\frac{P_{\ell,t}}{P_{\ell,t}+\Delta^2}$; $\hat{x}_t = \hat{x}_t^- + K'_{\ell,t}\Delta$	
		false	$z_t < L - \eta_{\ell,t}$	true	$K'_{\ell,t} = \varrho\frac{P_{\ell,t}}{P_{\ell,t}+\Delta^2}$; $\hat{x}_t = \hat{x}_t^- + K'_{\ell,t}\Delta$	
				false	$\hat{x}_t = \hat{x}_t^-$	

same value, and the information about the different values before the illumination variations are lost. We introduce a modification to the background updating schema in order to distinguish between pixels with a *known* or *unknown* intensity level. Initially, all of the unsaturated pixels are labelled *known* while the saturated ones are labelled *unknown*. They can move to the *known* state if their background value, after an illumination reduction, remains unsaturated for a sufficiently long period (controlled by a threshold C_{thr}). The resulting background updating schema is reported in Table 1.

4   Experimental Results

The proposed algorithm has been tested within a vison-based traffic control system named SCOCA [7]. Its purpose is the counting and classification of vehicles crossing a road intersection observed by a camera. The detection of vehicles relies on a background differencing technique. Hence the performance of our algorithm compared to that of [3] was measured by observing the output of SCOCA. We used three image sequences coming from different intersections. Sequences $S1$ and $S2$ are composed of frames taken from a standard color surveillance camera, endowed with autoiris lens, and compressed into *mpeg* files. $S3$ is composed of gray scale frames taken from a progressive scan camera with electronic shutter, each one compressed in a *jpg* file. $S1$ and $S2$ contain gradual illumination changes both due to real lighting variations and the activation of the autoiris lenses. $S3$ contains some sharp illumination changes due to the electronic variation of the shutter, and a significant presence of saturated pixels. The table in Figure 2 reports the performance of SCOCA in terms of correct detection rate and correct classification rate. For the sequences $S1$ and $S2$, the use of the proposed algorithm provides better performances showing its superiority in detecting foreground objects and in localizing their boundaries, as suggested by the improvements in the classification scores. Considering the third sequence a

	Detection rate		Classification rate	
	alg. [3]	our	alg. [3]	our
S1	90.7%	91.4%	85.4%	87.6%
S2	75.5%	77.1%	70.6%	74.8%
S3	–	78.5%	–	73.8%

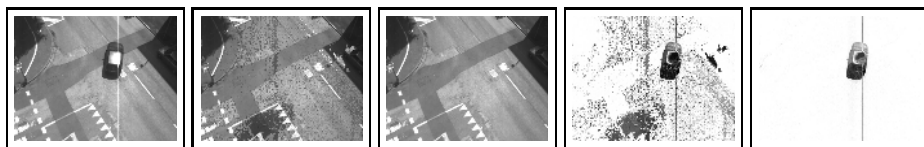**Fig. 2.** Comparison of SCOCA performances using the two algorithms

Fig. 3. From left to right: a frame in the middle of sequence S3; the corresponding backgrounds computed by alg. [3] and ours; the foreground pixels

direct comparison is not possible because only the proposed algorithm provides a reasonable result, the other one went into an inconsistent state after the first sharp illumination change (see Figure 3).

5   Conclusions

We have presented a new algorithm for background updating based on Kalman filtering technique, which is robust to gradual and sharp illumination changes. The most significant novelty we have introduced, which make the algorithm robust to gradual and sharp illumination changes, is the estimation of the global illumination variations and its use as an external control of the Kalman filter. The experiments have shown that the algorithm extends the range of sequences where it can work successfully. Its major limitation, i.e. inability to manage dynamic background pixels, is the subject of our current research.

References

1. Kalman, R.E.: A New Approach to Linear Filtering and Prediction Problems. Transactions of the ASME Journal of Basic Engineering, 82 (Series D):35–45, 1960.
2. Karman, K., Brandt, A., Gerl, R.: Moving object segmentation based on adaptive reference images. In Signal Processing V: Theories and Applications, pp 951–954, Barcelona, Spain, September 1990.
3. Boninsegna, M., Bozzoli, A.: A Tunable Algorithm to Update a Reference Image. Signal Processing: Image Communication, 16(4) pp 353–365, November 2000.
4. Koller, D., Weber, J., Huang, T., Malik, J., Ogasawara, G., Rao, B., Russell, S.: Toward robust automatic traffic scene analyis in real-time. Proc. Int. Conf. Pattern Recognition, pp 126–131, 1994
5. Catlin. D.E.: Estimation, control, and the discrete Kalman filter. Springer-Verlag, 1989
6. Oppenheim, A.V., Schafer, R.W., Stockham Jr, T.G.: Nonlinear filtering of multiplied and convolved signals. Proc. IEEE, vol. 56, pp 1264–1291, August 1968
7. Messelodi, S., Modena, C.M., Zanin, M.: A computer vision system for the detection and classification of vehicles at urban road intersections. ITC-irst Technical Report T04-02-07, February 2004
8. Wren, C.R., Azarbayejani, A., Darrell, T., Pentland, P.: Pfinder: Real-time tracking of human body. IEEE Transaction PAMI, vol.19, pp 780–785, July 1997
9. Grimson, W.E.L., Stauffer, C.: Adaptive background mixture models for real-time tracking. Proc. IEEE Computer Vision Pattern Recognition, vol.1, pp 22–29, 1999
10. Elgammal, A., Duraiswami, R., Harwood, D., Davis, L.S.: Background and foreground modeling using nonparametric kernel density estimation for video surveillance. Proc. of the IEEE, vol.90, pp 1151–1163, N.7, July 2002
11. Li, L., Huang. W., Gu, I., Tian, Q.: Statistical modeling of complex backgrounds for foreground object detection. IEEE Trans. Image Processing, vol.13, Nov. 2004
12. Li, L., Leung, M.: Integrating intensity and texture differences for robust change detection. IEEE Trans. Image Processing, vol.11, pp 105–112, February 2002
13. Wixson, L.: Detecting salient motion by accumulating directionary-consistent flow. IEEE Trans. PAMI, vol.22, pp 269–274, August 2000