Improving Background Subtraction Based on a Casuistry of Colour-Motion Segmentation Problems

I. Huerta [1], D. Rowe [1], M. Mozerov [1], and J. Gonz`alez[2]

[1] Dept. d'Inform` atica, Computer Vision Centre, Edifici O. Campus UAB,
08193, Bellaterra, Spain
Ivan.Huerta@cvc.uab.es
[2] Institut de Rob` otica i Inform`atica Ind. UPC, Llorens i Artigas 4-6,
08028, Barcelona, Spain

Abstract. The basis for the high-level interpretation of observed patterns of human motion still relies on motion segmentation. Popular approaches based on background subtraction use colour information to model each pixel during a training period. Nevertheless, a deep analysis on colour segmentation problems demonstrates that colour segmentation is not enough to detect all foreground objects in the image, for instance when there is a lack of colour necessary to build the background model. In this paper, our segmentation procedure is based not only on colour, but also on intensity information. Consequently, the intensity model enhances segmentation when the use of colour is not feasible. Experimental results demonstrate the feasibility of our approach.

1 Introduction

The analysis of human motion in image sequences involves different tasks, such as motion segmentation and tracking, action recognition and behaviour reasoning [6]. However, the basis for high-level interpretation of observed patterns of human motion still relies on when and where motion is being detected in the image. This kind of information is critical for different applications such as smart video surveillance for intruder detection and suspicious behaviour detection.

To achieve robust detection, many researchers have proposed methods to address segmentation problems, such as illumination changes, shadows, camouflage, background in motion, or deposited and removed objects from the scene [7]. To overcome these difficulties, different techniques can be applied, such as temporal differencing, optical flow and background subtraction [9]. The latter is based on a background model used to compare the current image with such a model. Among these background subtraction methods, statistical approaches are very popular: W^4 [3] use a bimodal distribution; Pfinder uses a single Gaussian to model the background; Stauffer et al. [2] use a mixture of Gaussians; and Elgammal et al. [1] present a non-parametric background model.

On the other hand, several cues are used for segmentation in the literature: Horprasert et al. [4] use colour information to classify a pixel as foreground,

Case Analysis (Colour Model Casuistry)					
Cues	**Chromatic**	**Equal**			**Different**
	Brightness	**Lower**	**Equal**	**Higher**	**-**
Description	**Base case**	Global shadow Local shadow	Background	Global Highlight Local Highlight	Foreground
	Anomalies	Dark Camouflage Dark Foreground	Camouflage	Light Camouflage Light Foreground	Sensitivity of Sensor Change of illuminant Gleaming surface Saturation Minimum Intensity

Fig. 1. This table analyzes the differences between an input image and the background model

background, shadow or highlighted background, while Wallflower [8] uses a three-level categorization, namely pixel, region and frame level. Jabri et al. [5] use colour and edge information, and Shen [10] uses a RGB/HSL colour space plus fuzzy classification.

In this paper, a casuistry of colour-motion segmentation problems is first presented, since colour is not enough to detect all foreground objects. This allows to identify when a colour model can and can not be used. Thus, based upon this casuistry, different colour problems can be then addressed properly. As a result, a novel background subtraction technique is presented, which combines both colour and intensity cues in order to solve colour motion segmentation problems such as saturation or the lack of colour when building the background model.

This paper is organized as follows. Next section presents a casuistry of the problems when using colour information for motion segmentation. This leads to our approach to confront segmentation. Section 3 explains our approach to solve the above aforementioned problems using the colour and intensity cues. Experimental results are described in section 4. Lastly, section 5 concludes this contribution and discusses about future work.

2   Problems on Colour Models

Colour information obtained from the recording camera is based on three components which depend on the wavelength λ: the object reflectance R, the illuminant spectral potency distribution E and the sensor wavelength sensitivity S:

$$Colour = \int_{\lambda} R(\lambda)E(\lambda)S(\lambda)d\lambda. \tag{1}$$

Unfortunately, *sensitivity of the sensor* may depend on intensities for each channel which can cause chromaticity changes. In addition, if the *illuminant*
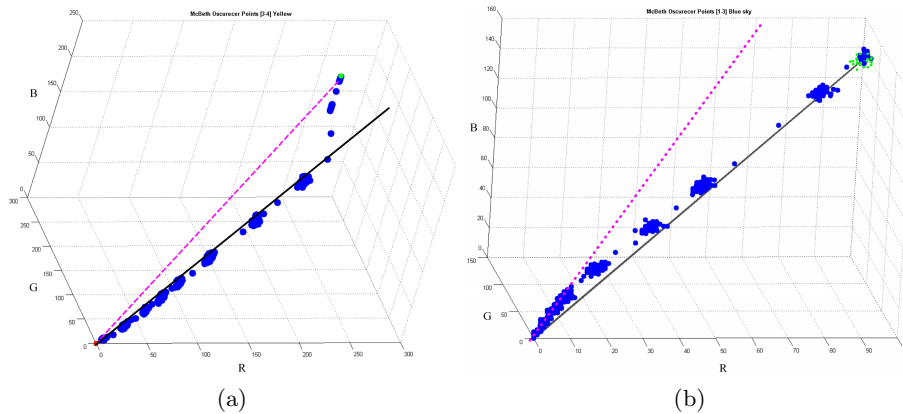
Fig. 2. Intensity RGB values of a) yellow and b) blue colours extracted from a Macbeth board which becomes lighter over time. The dashed line corresponds to a wrong chromaticity line and solid line corresponds to a correct chromaticity line due to: (a) at least one of the RGB channels is saturated, or (b) there is not enough intensity to build a colour model.

changes, the perceived chromaticity changes too, so the colour model can be wrongly built.

Fig. 1 shows a Colour Model Casuistry based on a background model which separates the chromaticity from the brightness component. The Base Case is the correct operation of the theoretical colour model, and the anomalies are problems that may appear. The theoretical base case solves some of the segmentation problems, as sudden or progressive global and local illumination changes, such as shadows and highlights. However, some problems remain. Foreground pixels with the same chromaticity component as the background model are not segmented. If the foreground pixel has the same brightness as the background model appears the *Camouflage* problem. A *Dark Camouflage* is considered when the pixel has less brightness and it cannot be distinguished from a shadow. Next, *Light Camouflage* happens when the pixel is brighter than the model, therefore the pixel cannot be distinguished from a highlight. *Dark Foreground* denotes pixels which do not have enough intensity to reliably compute the chromaticity. Therefore it can not be compared with the chromaticity background model. On the other hand is *Light Foreground* which happens when the present pixel is saturated and can not be compared with the chromaticity background model either.

Further, the perceived background chromaticity may change due to the *sensitivity of the sensor*, or *local* or *global illumination changes*. For instance, background pixels corresponding to shadows can be considered as foregrounds. The *Gleaming Surfaces* as mirrors cause that the reflect of the object is considered as foreground. On the other hand, due to saturation or minimum intensity problems the colour model can not be build correctly. Therefore, a background pixel can be considered foreground erroneously. *Saturation problem* happens when the

intensity value of a pixel for at least one channel is saturated or almost saturated. An example can be seen in the Fig. 2 where an experiment which consists on recording a Macbeth board while changing intensity was carried out. The Macbet board is broadly used because it contains a wide range of different colours. Fig. 2.(a) represents the RGB values of the yellow region extracted from a Macbeth board which becomes lighter over time. If the model is built when at least one of the RGB channels is saturated, the chromaticity line becomes erroneous. Therefore, the colour model would be build wrongly. In Fig. 2.(a), the dashed line is the erroneous chromaticity line and the solid line would be the correct chromaticity line. The second one is *minimum intensity problem* which happens when there is not enough chromaticity to build a colour model. This is mainly due to pixels do not have the minimum intensity value to built the chromaticity line, as shown in Fig. 2.(b) which represents the RGB values of Blue. The wrong chromaticity line is the dashed line which is built when there is not enough intensity. Then, a correct chromaticity line will be built when there is enough intensity. See the black solid in Fig. 2.(b).

3   Handling Colour-Based Segmentation Problems

The approach is based on background subtraction and uses either colour or intensity statistics, depending on the casuistry. First, the parameters of the background model are defined, next the colour and intensity models are explained in detail, and finally the segmentation procedure is presented. An sketch of the system can be seen in Fig. 3.
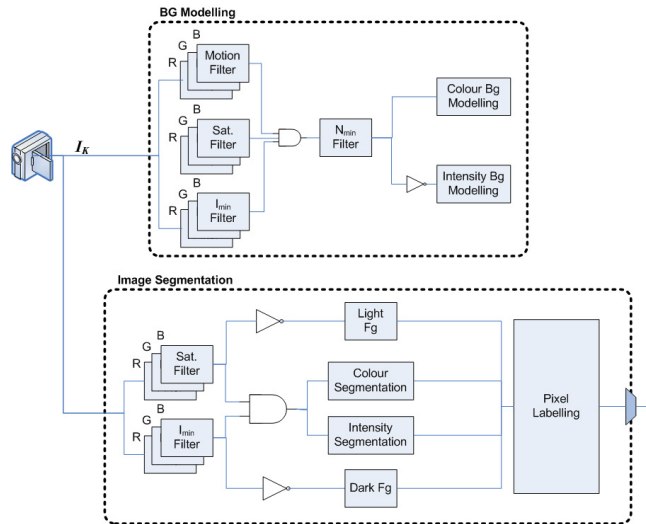

Fig. 3. Overview of the system

3.1   Background Modelling

The background model takes into account the problems depicted before. A motion filter is used to avoid moving foreground pixels of the images acquired during the training period, such as walking people or moving cars:

$$\text{Motion(x, t)} = \begin{cases} 1 & \text{if } |V(x,t) - \lambda(x)| < max(2*\sigma(x), T_m) \\ 0 & otherwise, \end{cases} \qquad (2)$$

where $V(x,t)$ is the intensity of a pixel location x in the t image of sequence V, $\lambda(x)$ is the median value, and $\sigma(x)$ is the standard deviation computed for all pixels in the image. The threshold T_m is a minimum constant value.

Pixels below a minimum intensity value (N_{Imin}) or pixels over a saturation intensity value (N_{Sat}) are not used to compute the colour model, but to build the intensity model:

$$\text{Imin(x, t)} = \begin{cases} 1 & \text{if } V(x,t) > N_{Imin} \\ 0 & otherwise \end{cases} \qquad (3)$$

$$\text{Sat(x, t)} = \begin{cases} 1 & \text{if } V(x,t) < N_{Sat} \\ 0 & otherwise \end{cases} \qquad (4)$$

Lastly, a Nmin filter is used to know if a pixel position x in the image has enough values to build colour statistics:

$$\text{Nmin(x)} = \begin{cases} 1 & \sum\limits_{t=1}^{N} Motion(x,t) \& Sat(x,t) \& Imin(x,t) > N_{Nmin} \\ 0 & otherwise \end{cases} \qquad (5)$$

As a result, the colour model is built using those pixels which have passed the Nmin filter, without saturation $Sat(x,t)$ neither minimum intensity $Imin(x,t)$ nor motion pixels $Motion(x,t)$ for each frame.

Next, the intensity model is built with those pixels which have not passed the Nmin filter but without considering motion pixels $Motion(x,t)$. Nevertheless, if almost all the pixels corresponding a place x are in motion, then there is not enough statistics to build the model correctly.

3.2   The Colour and Intensity Models

The colour model is based on the algorithm presented in [4], which computes the chromatic and brightness distortion components of each pixel. Furthermore, it can solve local and global shadows and highlights.

Each pixel is modelled by a 4-tuple $< E(x), s(x), a(x), b(x) >$, where $E(x) = [\mu_R(x), \mu_G(x), \mu_B(x)]$ is the expected colour value, μ is the arithmetic mean of the xth pixel's red, green, blue values computed over the training period of N frames, $s(x) = [\sigma_R(x), \sigma_G(x), \sigma_B(x)]$ is the standard deviation of colour value, $a(x)$ is the variation of the brightness distortion, and $b(x)$ is the variation of the chromaticity distortion of the xth pixel. See [4] for more details.

The intensity model is built based on a 2-tuple $< E(x), s(x) >$ for every pixel, where $E(x) = \mu[(R(x) + G(x) + B(x))/3]$ is the arithmetic mean of the xth pixel computed over the training period, and $s(x) = \sigma[(R(x) + G(x) + B(x))/3]$ is the standard deviation of colour value.

Foreground detection is thus achieved by using the normalised brightness and normalised chromaticity measures from colour model and the statistics from the intensity model for every new images, and then applying the pixel classification procedure explained next.

3.3   Image Segmentation

The combination of colour and intensity models allows to cope with different problems. In fact, our algorithm can detect different situations. If the background colour model of the current pixel is not available due to the lack of colour or saturation problems during the training period, then this pixel is segmented using intensity model. A pixel is considered foreground (FI) using a mean filter[1], in other case is considered background (BI):

$$\text{meanfilter}(x, t) = \begin{cases} 1 & \text{if } |V(x,t) - E(x)| < max(2 * s(x), T_i) then Bg. \\ 0 & otherwise foreground, \end{cases} \quad (6)$$

In other case, the colour model is used. When the current pixel has not chromaticity, it can be only segmented using the brightness component from the colour model. A threshold over brightness classifies if this pixel is a *Dark Foreground* (DF) or *Light Foreground* (LF). The thresholds $T_{CD}, T_{\alpha1}, T_{\alpha2}$ are given by the colour model defined in [4], and the thresholds $T_{\alpha lo} = k_1 * T_{\alpha1}$ and $T_{\alpha Hi} = k_1 * T_{\alpha2}$, where k_1 is a constant value. In other case this pixel is classified as *Background*(BB). Finally, if the pixel have chromaticity, then this will be compared with the background model chromaticity. The pixel is foreground (F) if it has different chromaticity. If the pixel have the same chromaticity, it is classified depending on the brightness as *Shaded Background or Shadow* (S), *Highlighted Background* (H), or *Original Background* (B).

Summarizing all the parameters of the background and colour model explained before, a pixel of the segmented image $M(i)$ is classified as:

$$M(i) = \begin{cases} FI & if & \neg meanfilter \;\; \& \;\; \neg Nmin & then & Fg. \\ BI & else\ if & \neg Nmin & then & Bg. \\ DF & else\ if & \hat{\alpha}_i < T_{\alpha lo} \;\; \& \;\; \neg Imin & \left.\vphantom{\begin{matrix}1\\1\end{matrix}}\right\} & Fg. \\ LF & else\ if & \hat{\alpha}_i > T_{\alpha Hi} \;\; \& \;\; \neg Sat & & \\ BB & else\ if & \neg Sat \;\; | \;\; \neg Imin & then & Bg. \\ F & else\ if & \hat{C}D_i > T_{CD} & then & Fg. \\ B & else\ if & \hat{\alpha}_i > T_{\alpha1} \;\; \& \;\; \hat{\alpha}_i < T_{\alpha2} & \left.\vphantom{\begin{matrix}1\\1\\1\end{matrix}}\right\} & \\ S & else\ if & \hat{\alpha}_i < 0 & & Bg. \\ H & & otherwise & & \end{cases} \quad (7)$$

---

[1] The threshold T_i is a minimum constant value.

4   Experimental Results

Our approach has been tested with multiple and different sequences with multiples segmentation problems. The first row of Fig. 4 shows the results obtained using colour model presented in [4]. In these images the saturated sky (blue sky colour) and saturated floor (yellow colour) are detected wrongly as foreground regions due to saturation problem. Furthermore, second image show as a black shadow is detected as foreground erroneously due to lack of colour problem. The second row shows that these problems are solved using our approach. The Fig. 5 shows that our approach works in different datasets, such as PETS and CAVIAR, among others.
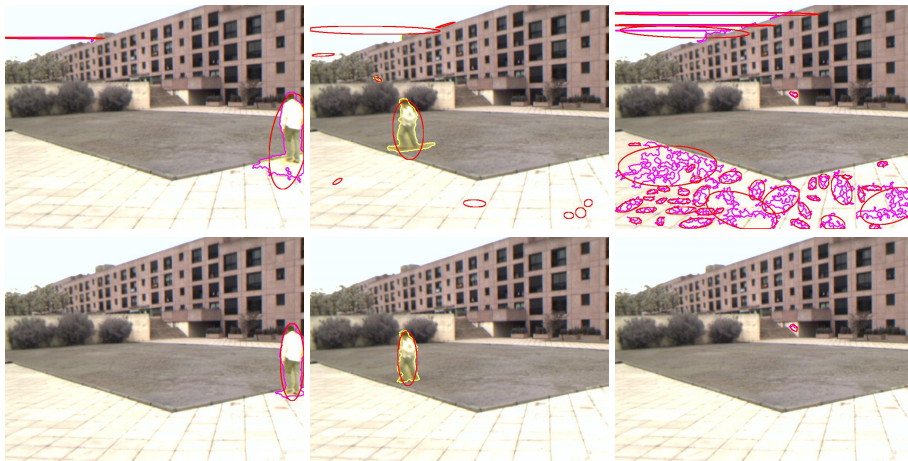

Fig. 4. First row shows erroneous blobs wrongly segmented by the colour model [4] due to the lack of colour and/or saturation. Second row shows that these blobs are removed by our approach, using both colour and intensity cues.

Fig. 5. Foreground region segmentation applying our approach to different datasets, such as PETS and CAVIAR, among others

5   Conclusion

In this paper, firstly a casuistry of the possible colour-motion segmentation problems is presented. This allows us to define when the colour model can be used. Then, an approach is proposed to cope with different colour problems as dark foreground and light foreground. Furthermore, it solves saturation problems and minimum intensity problems using intensity cue. The approach reduces the number of false negatives, false positives and increase the detected correct foreground regions as it can be seen in the experimental results. Future Work needs to address news cues, like edges or corners because the intensity model can not work with intense shadows and highlights. Furthermore, these kind of cues can help solving problems related to the sensitivity of sensor and changes of illumination which are not being tackled in this paper. Finally, an object based multilayer background model is required to face problems as ghosts [8].

References

1. Elgammal, A., Harwood, D., Davis, L.S.: Nonparametric background model for background subtraction. In: Proceedings European Conference Computer Vision (ECCV'00), pp. 751–767, Dublin (2000)
2. Grimson, W.E.L., Stauffer, C.: Adaptive background mixture models for real-time tracking. vol. 1, pp. 22–29 (1999)
3. Haritaoglu, I., Harwood, D., Davis, L.S.: W4: Real-time surveillance of people and their activities. IEEE Trans. Pattern Analysis and Machine Intelligence 22(8), 809–830 (2000)
4. Horprasert, T., Harwood, D., Davis, L.S.: A statistical approach for real-time robust background subtraction and shadow detection. IEEE Frame-Rate Applications Workshop (1999)
5. Jabri, H.W.S., Duric, Z., Rosenfeld, A.: Detection and location of people in video images using adaptive fusion of color and edge information. vol. 4, pp. 627–630 (September 2000)
6. Gonzàlez i Sabaté, J.: Human Sequence Evaluation: the Key-frame Approach. PhD thesis (May 2004)
7. Karaman, M., Goldmann, L., Yu, D., Sikora, T.: Comparison of static background segmentation methods. In: Visual Communications and Image Processing (VCIP '05) (July 2005)
8. Toyama, K., Krumm, J., Brumitt, B., Meyers, B.: Wallflower: Principles and practice of background maintenance, vol. 1, pp. 255–261 (1999)
9. Moeslund, T.B., Hilton, A., Krüger, V.: A survey of advances in vision-based human motion capture and analysis. Computer Vision and Image Understanding 104, 90–126 (2006)
10. Shen, J.: Motion detection in color image sequence and shadow elimination. Visual Communications and Image Processing 5308, 731–740 (2004)