Foreground segmentation using luminance contrast

LUIS M. FUENTES, SERGIO A. VELASTIN
Department of Electronic Engineering
King's College London
Strand, London WC2R 2LS
UNITED KINGDOM

Abstract: - The authors suggest the convenience of using luminance contrast for a normalised threshold in the foreground segmentation process for video-based surveillance applications. An individual pixel is part of the foreground when its contrast is greater than a given threshold. This method presents two major advantages: it simplifies the background model and increases the processing speed. The background model requires a single image, which may be updated in outdoors applications. No statistical information about the variation of each pixel value is needed. The algorithm runs faster in colour images because it operates with just one coordinate, luminance in YUV colour space, keeping colour information available for further use. The method has been tested in images from four different scenarios, indoors and outdoors and the results in two particular contexts, negative and colour contrast are discussed.

Key-Words: - **Segmentation, contrast, luminance contrast, foreground detection**

1 Introduction

Luminance (L) is a photometric quantity related with the visual sensation of brightness. Although it is the basic magnitude in light measurement, its definition is derived from psychophysical measurements in humans. Primarily it is the luminous flux by area and solid angle units integrated over wavelength using V(λ) as weighting function. V(λ) is the human luminous efficiency curve, as defined by the Commission Internationale de l'Eclairage (CIE), and represents the relative efficiency of light of different wavelength in exciting the visual system. As a result of being an integrated value, two stimuli with different spectral distributions can lead to the same luminance value (isoluminance).

Most of the information acquired by the visual system – motion, form, shading, detail, stereo - comes mainly through the luminance channel. Colour channels are used mainly for surface detailing and object labelling [1,2]. In both cases, the human visual system is very accurate at computing relative differences but is largely unable to compute absolute values. The brightness or colour of a particular element in the field of view depends on the brightness and colour of the environment. The visual system has, however, a kind of compensating attribute called colour constancy. Basically, it is the ability of the visual system through which objects tend to be recognised as having nearly the same colour under many different lighting conditions (different amounts of light spectral composition).

These properties of the visual system are clearly reflected in the way colour images are coded. The RGB system is related to three types of colour detectors, sensitive to long (R), medium (G) and short (B) wavelengths respectively and video signal are coded using formats, with one luminance and two colour components, YUV for PAL and NTSC signals. This separation between the 'luminance' information in an image from the 'colour' information allows the resolution of the latter to be reduced, exploiting the lack of colour acuity inherent in the human visual system.

Thus, the greater importance of the luminance parameter is reflected in the video standards. In this paper we propose to take advantage of this luminance-chrominance separation in video signals, digitising images into YUV colour space and using just luminance information for foreground detection. Relative differences between current and background images are used in the segmentation of foreground. Thresholds based on luminance contrast do not need a complex statistical description of background variations and requires a single image.

After the definition of contrast and its application to foreground detection, a discussion on three mayor issues of the method follows. The effect of changes in the lighting condition, the sign of the contrast and the relative importance of colour information are discussed. The last two issues are analysed in more than 1700 different images from three different indoor scenarios –standard footage of CCTV system in a London Underground station- and an additional outdoor scenario [14].

2 Previous work

Foreground detection algorithms are normally based on background subtraction algorithms (BSAs) [4,5,6,7,8,9], although some approaches combine this method with a temporal difference [6]. These methods are based on extracting motion information by thresholding the differences between the current image and a reference image (background) or the previous image respectively. BSAs are widely used because they detect not only moving objects but also stationary objects not belonging to the scene. The reference image is defined by assuming a Gaussian model for each pixel. BSAs are normally improved by means of updating their statistical description so as to deal with changing lighting conditions [6,7,8.9], normally linked with outdoor environments. Some authors present a different model of background, using pixels' maximum and minimum values and the maximum difference between two consecutive frames [5], a model that can clearly take advantage of the updating process. Pixels of each new frame are then classified belonging to the background or the foreground using a statistical description to define a threshold. When dealing with colour images, the background method is normally extrapolated to three coordinates, either using RGB [8] or YUV space [9], without differentiating between luminance and colour processing. However, alternative approaches for motion vectors estimation using mainly colour information can be found in the literature [10].

3 Method

We assume fixed cameras (as commonly used in surveillance scenarios) and generally small global changes in lighting conditions. However, the foreground detection algorithm is also valid using continuously updated background images.

3.1 Colour space

Image sequences to be analysed came through standard PAL video signals. Faster processing benefits from digitalisation into YUV colour co-ordinates. A couple of remarks have to be made about this colour format. Firstly the common use of the term luminance for Y is a misrepresentation and a distinction should be drawn between this term and true CIE luminance. Secondly YUV is a shorthand commonly but incorrectly used to refer to two different aspects: the YCbCr format used for component digital video (JPEG, MPEG) and the YPbPr format used for component analogue video (PAL). Both systems are based on the luminance and colour difference signals – Y, B-Y and R-Y - but the coefficients to compute analogue or digital RGB values are different [11]. A typical output format for capture boards, named YUV16, has a (4:2:2) ratio between luminance and colour components, which are halved along the horizontal axis or line.

As pointed before, the former Y value does not correspond with the CIE luminance value but the RGB triplet stored in a digital image neither represents the proportion of three monochromatic lights used in the definition of luminance. Therefore when a digitiser uses RGB system or the image sequence is stored in any format using the RGB system, a transformation to YUV is required. This is actually the inverse of the transformation usually required in a digital video encoder [12] and can be defined as follows:

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 0.256 & 0.504 & 0.098 \\ -0.148 & -0.291 & 0.439 \\ 0.439 & -0.368 & -0.072 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \qquad (1)$$

where the RGB coordinates are in the range (0,255), Y values in a compressed range (16,235) and colour difference coordinates UV in (15,240), varying from yellow to blue and from cyan to red respectively.

3.2 Background

Using luminance contrast in foreground detection, just a single background image is needed for background modelling. A stationary background without significant changes in the illumination levels is normally modelled with a normal distribution [6] and therefore the background image obtained by simply averaging N different images. Although statistical information about deviations from this averaged image can be used to determine contrast thresholds for each pixel, it makes the segmentation process more complicated without a significant improvement in its performance.

3.3 Definition of luminance contrast

Luminance contrast is an important magnitude in psychophysics and the central point in the definition of the visibility of a particular object. Typically, luminance contrast is defined as the relative difference between luminances of the object, L_O, and the surrounding background, L_B, equation 2. As can be seen, positive and negative values are possible, negative contrast meaning an object darker than the background.

$$C_L = \frac{L_O - L_B}{L_B} \tag{2}$$

To apply this concept in foreground detection we propose an alternative contrast definition comparing the luminance 'y' of a pixel P(i,j) in both the current and the background images:

$$C(i,j) = \frac{y(i,j) - y_B(i,j)}{y_B(i,j)} \tag{3}$$

Although null values for background 'y' coordinate are not likely to be found, due to camera and digitalisation noise, they may be changed to one because the infinite contrast value they produce has no physical meaning.

Unlike real luminance contrast, contrast values for pixels are not symmetrically distributed around zero. Due to the integer quantisation, actual values for pixels' contrast varies in the range (-1,254). An example of luminance contrast values of pixels is shown below, figure 1.

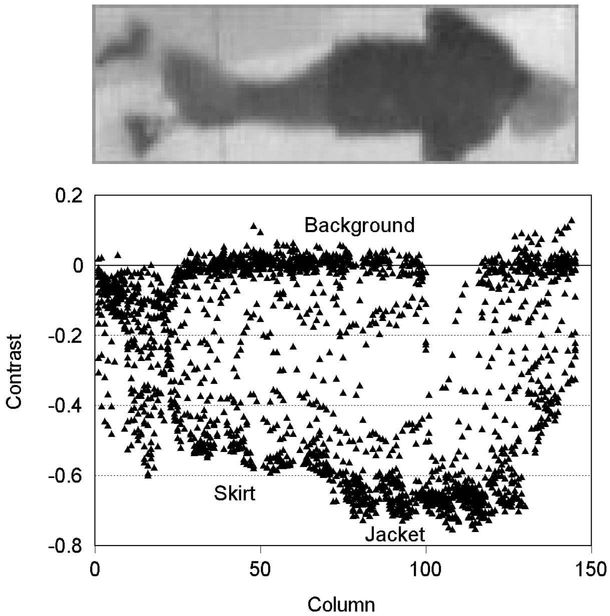

Fig.1 Values of luminance contrast for individual pixels. Points around null contrast correspond to background pixels.

3.4 Foreground detection

According to the non-symmetrical distribution of contrast around zero, foreground detection algorithm should use two different thresholds for positive C_P and negative C_N values of contrast, depending on the nature of both the background and objects to be segmented. To simplify the notation, we assume from now onwards a single contrast threshold C, that is $C_P = - C_N = C$. So, a pixel P(i,j) is set to foreground when the absolute value of its contrast is bigger than the chosen threshold C. Otherwise is set to background.

4 Discussion

After analysing the contrast of individual pixels within the boundary boxes of many blobs in different frames from sequences of different cameras, three main facts are significant. There is sometimes a small shift in background contrast value, foreground pixel's contrast is normally negative and the use of luminance contrast in segmentation is justified by the low significance of colour contrast. Images from three different CCTV cameras in London Underground Liverpool Street station were analysed. An additional set of outdoor images has been taken from the PETS2000 data set [14] to provide a very different background and object for both, colour and negative contrast and different objects (cars).

4.1 Illumination

There are always variations in the illumination parameters between two images of the same scene taken at different days. However, indoor backgrounds provide a stable lighting configuration whereby variation is normally due to a lamp replacement or momentary lamp failure. There are many other factors, such as changes in voltage or obstruction of reflected light, that can lead to minor illumination changes, but their effect on the general illumination level is relatively small. These minor modifications produce a global shift in the contrast plot, with the "background contrast" moving from zero to positive or negative values depending on whether the new scenario is darker or lighter than the stored background image. Any background-updating algorithm [5,6,7,8,9] should be able to deal with these minor changes, although as the observed shifts are not bigger than 15%, figure 2, the selection of an appropriate contrast threshold can deal with these small illumination changes.

4.2 Negative contrast

Negative contrast values were found to occur for most of the foreground pixels. In this particular application of people tracking in London Underground Stations, that means that usually

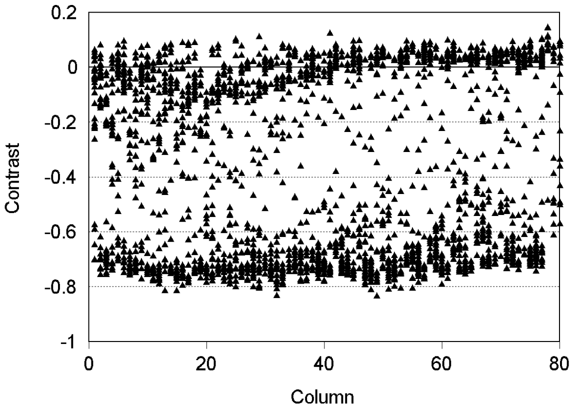

Fig.2. Effect of changes in illumination conditions: luminance contrast of background pixels shifts to positive values. The cloud of pixels with low negative contrast (columns 0 to 40) is associated with the shadow of the person.

people are dressed in dark colours. As interior design tends to use light colours to improve lighting efficiency, foreground object will be normally darker than the background and will exhibit negative contrast. Individual foreground pixel contrast was analysed from blobs on images from three indoor scenarios resulting in an average percentage above 95% of foreground pixels with negative contrast, table 1, and a high value in the outdoors example, lowered by the fact that one of the three cars in the sequence is white. It is important to note that the sign in the contrast, figure 3, does not affect the performance of the segmentation algorithm.

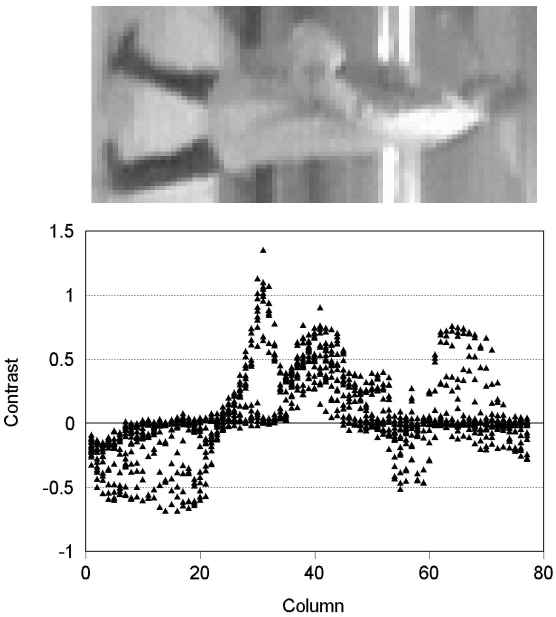

Fig.3 Example of positive luminance (Y) contrast.

4.3 Colour contrast

Using just the luminance co-ordinate in a people detection algorithm can be seen to lose additional information. Although this may be true in general, it is not the case in the kind of images we are analysing here. Indoor environments are normally badly illuminated in terms of the sensitivity requirements of standard colour cameras used in CCTV surveillance systems. Together with the fact that people in north-western countries wear, mostly, dark and colourless clothes, colour information is not significant in most of the cases. This can be verified using luminance and colour information to foreground detection and calculating the percentage, over the total foreground pixels, of pixels labelled as foreground due just to colour contrast. Table 1 shows a colour analysis of blobs in three different indoor scenarios (I1, I2, I3) and an outdoors (O) example.

Although in some particular instances, colour information can lead to a better segmentation in bright coloured areas, figure 4, it will fail in the rest. Ideally, a combination of both methods could lead to an optimum result. This solution, however, is more expensive in computation time, especially if a co-ordinate transformation is needed.

	I1	I2	I3	O
Number of Images	300	300	530	620
Number of Blobs	434	564	1010	735
Negative contrast (%)	99.7	98.6	96.4	83.4
Colour contrast (%)	1.38	3.19	2.18	7.76

Table 1. Percentage of foreground pixel with negative contrast and detected by means of colour contrast.

5 Conclusion

In this paper we have proposed a new foreground detection method relying on luminance contrast for segmentation and simplifying therefore the background model. Just a single background image, that may be updated, is needed. In addition, working on one and not three co-ordinates for segmentation purposes reduces computation time. The YUV colour scheme is recommended for digital images, allowing fast access to pixel luminance and keeping colour information for further use. Having indoors people tracking applications in mind, experimental evidence of lack of colour and darkness of the clothes people wear is provided adding weight to the selection of luminance contrast for segmentation.

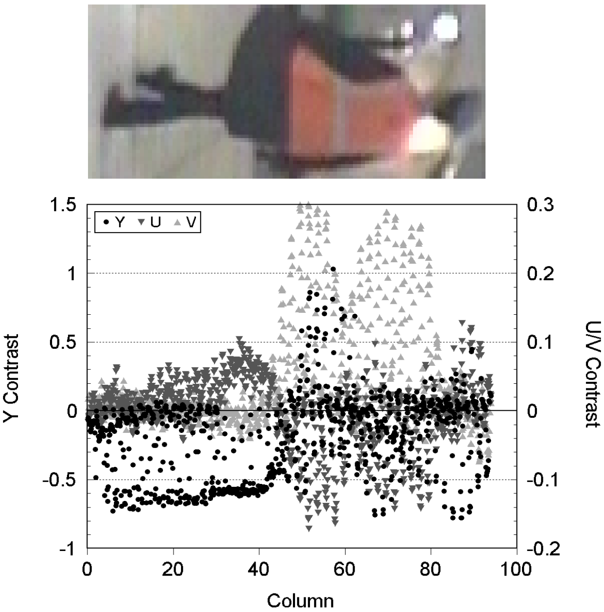

Fig.4 Contrast plot of pixel within a blob comprising a person wearing a bright orange waistcoat –negative U contrast for yellow and positive V contrast for red- with a grey band in the middle.

6 Acknowledgements

The work reported here has been carried out as part of EPSRC project "PerSec: Assessment of image processing techniques as a means of improving personal security in public transport" in collaboration with the Centre for Transport Studies, University College London. The authors are also grateful to London Underground Limited for access to their sites and advice.

References:

[1] T. Cornsweet, Visual perception, Academic Press, New York, 1970.

[2] D. Marr, Vision, W.H. Freeman, New Yortk, 1982.

[3] S.S. Intille, J.W. Davis and A. F. Bobick, "Real-time Closed-World Tracking", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'97), 1997, pp. 697-703.

[4] F. De la Torre, E. Martinez, M. E. Santamaria and J.A.Moran, "Moving Object Detection and Tracking System: a Real-time Implementation", Proceedings of the Symposium on Signal and Image Processing GRETSI 97, Grenoble, 1997.

[5] I. Haritaoglu, D. Harwood and L.S. Davis, "W4: Real-Time Surveillance of People and Their Activities", IEEE Trans. Pattern Analysis and Machine Intelligence, 22(8), 2000, pp. 809-822.

[6] S. Huwer and H. Niemann, "Adaptive Change Detection for Real-time Surveillance Applications", Proceedings of The IEEE Workshop on Visual Surveillance, Dublin, 2000, pp. 37-43.

[7] N. Rota and M. Thonnat, "Video Sequence Interpretation for Visual Surveillance", Proceedings of The IEEE Workshop on Visual Surveillance, Dublin, 2000, pp. 59-68.

[8] S. McKenna, S. Jabri, Z. Duric, A. Rosenfeld and H. Wechsler, "Tracking Groups of People", Computer Vision and Image Understanding 80, 2000, pp. 42-56.

[9] C. R. Wren, A. Azarbayejani, T. Darrel and P. Pentland, "Pfinder: Real-Time Tracking of the Human Body", Trans. Pattern Analysis and Machine Intelligence, 17(6), 1997, pp. 780-785.

[10] J. Magarey, A. Kokaram and N. Kingsbury, "Robust motion estimation using chrominance information in colour image sequences", Proceedings, Int. Conf. On Image Analysis and Processing, Florence, September 1997.

[11] K.N. Plataniotis and A.N. Venetsanopoulos, Color Image Processing and Applications, Springer, 2000.

[12] Philips, Product specification : SAA7110 One Chip Front-end 1 and SAA7126H Digital video encoder; 1999

[13] M. Bollmann, T. Hempel and B. Mertsching, "Improved Edge detection by evaluation of color Contrast Information", Proceedings of 2nd Workshop Farbbildverarbeitung. Schriftenreihe des Zentrums für Bild- und Signalverarbeitung, Ilmenau, Report 1/96, 1996, pp. 1-6

[14] First IEEE International Workshop on Performance Evaluation of Tracking and Surveillance (PETS2000), March 31st 2000; ftp://ftp.cs.rdg.ac.uk/pub/PETS2000