SNRFAI'2001. Vol II. Pág. 145-150. Benicásim, Mayo 2001.

Face tracking based on colour constancy*

J.M. Buenaposada, D. Sopeña, L. Baumela
Departamento de Inteligencia Artificial, Universidad Politécnica de Madrid
Campus de Montegancedo s/n, 28660, Madrid, SPAIN
{jmbuena,dsgalindo}@dia.fi.upm.es, lbaumela@fi.upm.es

Abstract

In this paper we present a colour constancy algorithm for real-time face tracking using skin colour segmentation. It is based on a modification of the well known Grey World algorithm in order to use the redundant information available in an image sequence. In the experiments conducted it is clearly more robust to sudden illuminant colour changes than the rg-normalised algorithm.

Keywords: Colour constacy, grey-world, face tracking.

1 Introduction

Automatic analysis and synthesis of facial expressions is a popular research topic in computer vision. With the advent of multimodal computer interfaces [12], face gesture analysis has become a key element for visual interaction with the computer and for image compresion and gesture-based image coding [3, 8]. This interest on gesture analysis has also fostered research in other traditional areas such as tracking [11] (3D face location), structure from motion (face structure computation), computer graphics (clone rendering), and many others. In this paper we will study the problem of face tracking using colour.

Skin colour is the most frequently used feature for face detection and tracking [11]. Trackers based on this feature are used as an initial estimate or follow-up verification of face location in the image plane or as a recovery process when more accurate and computer demanding trackers can not cope with face motion.

The primary problem in automatic skin detection is colour constancy. The RGB colour of an image pixel depends not only in the imaged object colour, but also on the lighting geometry, illuminant colour and camera response. For example [4], if the light intensity is scaled by a factor s, each captured pixel becomes (sr, sg, sb). The rg-normalisation algorithm provides a colour constancy solution which is independent of the illuminant intensity: $(sr, sg, sb) \mapsto (\frac{sr}{s(r+g+b)}, \frac{sg}{s(r+g+b)})$. On the other hand, a

*Work funded by CICyT under project number TIC1999-1021

change in illumination colour can be modelled as a scaling α, β and γ in the R, G and B channels. The Grey World (GW) algorithm [4] provides a constancy solution independent of the illuminant colour by dividing each channel by its average value: $(\alpha r, \beta g, \gamma b) \mapsto (\frac{\alpha r}{\frac{\alpha}{N}\sum_i r_i}, \frac{\beta g}{\frac{\beta}{N}\sum_i g_i}, \frac{\gamma b}{\frac{\gamma}{N}\sum_i b_i})$.

The most widely used colour constancy algorithm in face tracking is the rg-normalised space, as the skin colour distribution in the rg-normalised chromaticity space is gaussian [14], a bayesian classifier can be used to track fast and reliably a moving face [2]. Variuous improvements to this constancy model have been proposed in order to deal with small changes in the colour of the illumination. These consist on tracking the motion of the skin colour cluster in rg-space by using stochastic prediction models [9] or by clustering [13]. These algoritms have a problem in common: they can not deal with sudden changes in lighting colour. Other algorithms that also work in rg-normalised space deal with sudden changes in lighting by matching the colour distributions by a shift in illuminant colour [1]. Unfortunatelly this solution is unfeasible for real-time tracking as it incurs in a substantial computational overhead. Other colour constancy algorithms are more difficult to use in the analysis of a real-time image sequence. Either because they need too much information or are much computationally too complex, as is the case in physics-based models [5, 10].

The GW algorithm is fast and robust to illuminant colour variations, but it only works for scenes with constant colour distribution, e.g. it fails when a new object appears in the scene. In this paper we propose an extension to the GW algorithm that, using the redundant information available in an image sequence, overcomes the static colour distribution limitation.

2 Face tracking using GW

In this section we present a colour-based face tracking algorithm. First we will briefly describe how to track a coloured patch using simple statistics, afterwards the modified GW algorithm is presented.

Given a sequence of RGB images, building a face traker is straight forward if we know the probability that a pixel with colour (r, g, b) be *skin*, $P(skin|r, g, b)$. Then tracking can be performed with a mode seeking algorithm by computing the position and orientation of the face color cluster in each frame [2].

Let $p_s(r, g, b)$ and $p_b(r, g, b)$ be the conditional colour distributions of the skin and background respectively (we assume that backgroud is anything that is not skin). The problem now is computing these distributions in a reliable and robust way. In real-time systems this has been done using a rg-normalised space. As we previously mentioned, these methods fail when there is a sudden change of the illuminant colour. On the other hand, the GW model is invariant to these changes, but it was conceived for static images. Here we propose an dynamic extension to GW that computes colour distributions in image sequences, which will be invariant to big changes in the illuminant colour. We will use one image of the sequece, whose colour distribution is known, as a reference for GW-based segmentation. Let I^r be the reference image, $p^r_{sgw}(r,g,b)$ and $p^r_{bgw}(r,g,b)$ be respectively the *skin* and *background* colour distributions of I^r in GW space, h^r_r, h^r_g and h^r_b be the R, G and B histograms of I^r and let $\overline{\mu}^r_{gw}$ be the average RGB value of I^r. The dynamic extension to GW is based on the analysis of two different situations that arise in an image sequence:

1. If there is no significant change in the scene illumination between the actual image, I, and the reference image, then I can be segmented using $p^r_{sgw}(r,g,b)$, $p^r_{bgw}(r,g,b)$ and $\overline{\mu}^r_{gw}$. This situation can be detected when the minimum value for the SSD correlations of h_r, h_g and h_b (I histograms) with h^r_r, h^r_g and h^r_b, respectively is obtained with a common zero displacement.

2. If there is a change in the scene illumination, then I must be segmented using $p^r_{sgw}(r,g,b)$, $p^r_{bgw}(r,g,b)$ and $\overline{\mu}_{gw}$, the average RGB value of I. This situation can be detected when the minimum value for the SSD correlations of h_r, h_g and h_b with h^r_r, h^r_g and h^r_b, respectively is different from zero.

I^r must be updated whenever there is a significant change in the scene. This situation occurs when the previously mentioned correlations show a high minimum value. Based on these ideas we propose the Dynamic-GW (DGW) algorithm:

```
Initialisation
  /*Initialise the reference image model using motion
  segmentation and a precalculated colour model*/
While (true)   /* select model */
  If MinSSDPos(h_r^r, h_r) & MinSSDPos(h_g^r, h_g) & MinSSDPos(h_b^r, h_b) ≠ 0
    then \overline{μ}_{gw}^r = Mean(I)
    /* segment image */
  I_{gw} = I / \overline{μ}_{gw}^r
  Do ∀(x,y) ∈ I_{gw}
     If p_{sgw}^r[I_{gw}(x,y)] > p_{bgw}^r[I_{gw}(x,y)], then I(x,y) ∈ skin
                                                  else I(x,y) ∈ background
    /* update reference model */
  If MinSSDVal(h_r^r, h_r) & MinSSDVal(h_r^g, h_g) & MinSSDVal(h_r^b, h_b) > U_h
    then Update(p_{sgw}^r, p_{bgw}^r, I_{gw}),
         h_r^r = h_r,  h_g^r = h_g,  h_b^r = h_b,
         \overline{μ}_{gw}^r = Mean(I)
end /* while */
```

In our implementation $p_{sgw}^r \sim N(\mu_{sgw}^r, \Sigma_{sgw}^r)$ and the background model, p_{bgw}^r, is a non informative uniform distribution.

3 Experiments

In our experiments we used a VL500 Sony colour digital camera at 320×240 resolution, iris open, no gain, no gamma correction. Images were taken with regular roof fluorescent lights and variations in illumination colour were obtained using a controled tungsten light.

In the first experiment we validate the hipothesis used to detect variations in the color distribution and illumination of the scene. We acquired a sequence of images with a change in illumination, at one point, and a red object appearing in the scene at a different moment. Results are presented in Fig.1. In the first, second and third columns are displayed the h_r histograms and SSD correlation results of pairs of images with no significant environmental change, appearance of a red object and illumination variation, respectively. Results coincide with our hypothesis: a) with no change, the minimum SSD correlation value is the smallest and this is achieved with a cero displacement; b) when a new object appear, the minimum SSD value is higher and it is also achieved with a cero displacement; c) illumination variations manifest themselves with a high minimum SSD value and non cero displacements.

In the following experiment we validate the performance of the DGW algorithm.

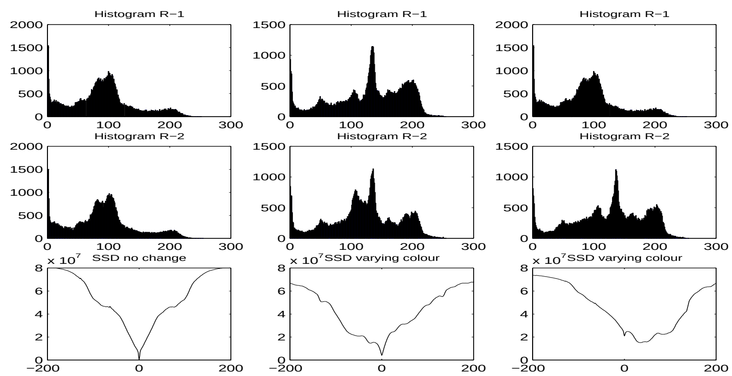

Figure 1: Histogram SSD correlation

We have taken a sequence with "difficult" background (i.e. brownish door and shelves to distract the segmentation). In Fig.2 four frames of the sequence are shown in each colum, representing each one of the following situations: initial image, red object appears, tungsten frontal light turns on, green lights turn on. DWG segmentation results are shown in the second row and rg-normalised results in the third one. Visual inspection of these results show that both algorithms have similar results in the least favourable cases for the DGW algorithm (second and third columns) and a clear success of the DGW compared to the rg-normalisation when the illuminant colour abruptly changes (fourth column).

4 Conclusions

We have introduced a new colour constacy algorithm based on an extension of the well known GW algorithm. It was designed to work with sequences of images with varying environmental conditions. In the experiments conducted it performed better than the rg-normalised algorithm, when sudden changes in the illuminant colour take place. The least favorable case for our algorithm occurs when changes in the illuminant geometry or intensity take place. In this case the performance is simmilar to the rg-normalised one. On the other hand, the detection of an illumination change based on RGB SSD histogram correlations fail in some cases so we are working on a different method to robustly choose between $\overline{\mu}_{gw}$ and $\overline{\mu}^{r}_{gw}$ in the algorithm.

Figure 2: Segmentation results

References

[1] D. Berwick and S.W. Lee. A chromaticity space for specularity-, illumination color- and illumination pose invariant 3-d object recognition. *Proc. of the Int. Conf. on Computer Vision*. Bombay, India. 1998.

[2] J.M. Buenaposada and L. Baumela. Seguimiento robusto del rostro mediante visión computacional. *Actas Conf. de la Asociación Española para la Inteligencia Artificial. Vol I.* 48–53, Murcia. Spain. 1999.

[3] J. L. Crowley and J. Schwerdt. Robust tracking and compression for video communication. *Proc. of the Int. Workshop on Recognition, Analysis and Tracking of Faces and Gestures in Real-Time* (RATFG'99), 2–9, Corfu. Greece. 1999.

[4] G.D. Finlayson, B. Shiele and J.L. Crowley. Comprehensive colour normalization. *Proc. European Conf. on Computer Vison (ECCV). Vol. I*, 475–490, Freiburg, Germany. 1998.

[5] G.D. Finlayson and G. Shaefer. Constrained dichromatic colour constancy. *Proc. ECCV. Vol. II*, 342–358, Dublin. Ireland. 2000.

[6] G.J. Klinker, S.A. Shafer and T. Kanade. A physical approach to color image understanding. *International Journal of Computer Vision*, 4: 7–38, 1990.

[7] H. Lee. Method for computing the scene illuminant chromaticity from specular highlights. *Journal of the Optical Society of America A*, 3: 1694–1699, 1986.

[8] MPEG-4 Video Group, "Coding of audio-visual objects: video", *ISO/IEC JTC1/SC29/WG11 N2202*, March 1998.

[9] Y. Raja, S.J. McKenna, S. Gong. Colour model selection and adaptation in dynamic scenes. *Proc. ECCV. Vol. I*, 460–474. 1998.

[10] M. Störring, H.J. Andersen and E. Granum. Estimation of the illuminant colour from human skin colour. *Proceedings of the Int. Conference on Automatic Face and Gesture Recognition* (FG'00), 64–69, Grenoble. France. 2000.

[11] K. Toyama. Prolegomena for robust face tracking. MSR-TR-98-65. Microsoft Research, November 1998.

[12] A. Waibel, M.T. Vo, P. Duchnowski, S. Manke. Multimodal interfaces. *Artificial Intelligence Review*, 10: 299-319. 1996.

[13] Y. Wu, Q. Liu and T.S. Huang. Robust real-time hand localization by self-organizing color segmentation. *Proceedings RATFG'99*, 161–166. 1999.

[14] J. Yang, W. Lu, A. Waibel. Skin-color modeling and adaptation. *Proceedings Third Asian Conference on Computer Vision Vol. II*, 142-147. 1998.

[15] M. D'Zmura and P. Lennie. Mechanisms of colour constancy. *Journal of the Optical Society of America A*, 3: 1662–1672, 1986.