Variations of Grey World for face tracking

José M. Buenaposada, Luis Baumela
Departamento de Inteligencia Artificial
Universidad Politécnica de Madrid
Campus de Montegancedo s/n, 28660, Madrid, SPAIN
jmbuena@dia.fi.upm.es, lbaumela@fi.upm.es

Abstract

This paper deals with the problem of colour constancy for real-time face tracking. We present two extensions of the Grey World (GW) algorithm in order to deal with dynamic image sequences with varying illumination intensity. The first extension consists in a new colour normalisation, the *Projective Grey World Normalisation*, which makes GW more robust to illumination intensity variations. The second extension, the *Dynamic Grey World Algorithm*, permits the use of GW with dynamic image sequences. We have conducted several experiments with these extensions and compared the results obtained with the rg-normalisation algorithm, possibly the most popular colour constancy procedure among the face tracking community.

Keywords: Face Tracking, Grey World, Colour Constancy

1 Introduction

Automatic analysis and synthesis of facial expressions is a popular research topic in computer vision. With the advent of multimodal computer interfaces [16], face gesture analysis has become a key element for visual interaction with the computer and for image compression and gesture-based image coding [5, 11]. This interest on gesture analysis has also fostered research in other traditional areas such as tracking [15] (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 [15]. Trackers based on this feature are fast and reliable, but not very accurate. They 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 colour of an image pixel depends not only on the imaged object colour, but also on the lighting geometry, illuminant colour and camera response. This means that the RGB value of a patch of skin can be very different depending on the camera used to capture the image, the colour and intensity of the illumination, the relative orientation between camera, skin surface and light source, or the existence of shadows or highlights in the image. Colour constancy algorithms model these effects and try to obtain colour invariants that facilitate the identification of a given colour under varying environmental conditions. For example [6], if the scene light intensity is scaled by a factor s, then each perceived pixel colour 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 change in illuminant colour can be modelled as a scaling α, β and γ in the R, G and B image colour channels. In this case the previous normalisation fails. The Grey World (GW) algorithm [6] provides a constancy solution independent of the illuminant colour by dividing each colour channel by its average value:

$$[\alpha R, \beta G, \gamma B] \mapsto [\frac{\alpha R}{\frac{\alpha}{n}\sum_i R}, \frac{\alpha G}{\frac{\beta}{n}\sum_i G}, \frac{\alpha B}{\frac{\gamma}{n}\sum_i B}].$$

The most widely used colour constancy algorithm in face tracking is rg-normalisation. As the skin colour distribution in the rg-normalised chromaticity space is Gaussian, a Bayesian classifier can be used to track fast and reliably a moving face [18]. Various 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 [12] or by clustering [17]. These algorithms 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]. Unfortunately this solution is infeasible 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 because they are computationally too complex. Physics-based colour constancy is based on an understanding of how physical processes such as light reflectance and camera sensitivity manifest themselves in the image. Many of these algorithms are based on the dichromatic reflectance model [9]. It postulates that the light reflected from a surface comprises a body reflection component, which reflects the object colour, and a surface reflection component, which models highlights and has the same spectral power distribution as the illuminant. A well known method of colour constancy is to find the illuminant colour as the intersection of at least two dichromatic planes [10, 19]. This method assumes the existence of several surfaces on the image, something which is not always true. More recent methods work in the chrominance space by modelling the range of possible illuminants and locating their chrominance as the intersection of the dichromatic line of a single object with the locus of illuminants [14, 7]. The method proposed by Störring [14] is impractical for a real-time application, whereas the results presented by Finlayson [7] are not very satisfactory for skin colour, as human skin is near achromatic and errors in the orientation of the dichromatic line become important.

In this paper we introduce two colour constancy algorithms that can be used for real-time colour-based image segmentation. The algorithms are both based on GW and exploit the redundancy of the image sequence in order to compute the relative change in illumination between the images of the sequence. As shown in the experiments conducted, these constancy algorithms are clearly more robust to big sudden illuminant colour changes than the popular rg-normalised algorithm.

2 Grey World-based colour constancy

Colour constancy is the perceptual ability to assign the same colour to objects under different lighting conditions. The goal of any colour constancy algorithm is to transform the original $[RGB]$ values of the image into constant colour descriptors. In the case of Lambertian surfaces, the colour of an image pixel $I(ij)$ can be modelled by a lighting geometry component s_{ij}, which scales the $[rgb]$ surface reflectance of every pixel independently, and three colour illuminant components (α, β, γ), which scale respectively the red, green and blue colour channels of the image as a whole [6]. The lighting geometry component accounts for surface geometry and illuminant intensity variations, while the colour illuminant components account for variations in the illuminant colour. According to this model, two pixels $I(ij)$ and $I(kl)$ of an image would have the following $[RGB]$ values: $[s_{ij}\alpha r_{ij}, s_{ij}\beta g_{ij}, s_{ij}\gamma b_{ij}]$, $[s_{kl}\alpha r_{kl}, s_{kl}\beta g_{kl}, s_{kl}\gamma b_{kl}]$, where $[r_{ij}, g_{ij}, b_{ij}]$ and $[r_{kl}, g_{kl}, b_{kl}]$ represent surface reflectance; i.e. real object colour, independent of the illuminant.

The GW algorithm proposed by Buchsbaum [2] assumes that the average surface reflectance in an image with enough different surfaces is grey. So, the average reflected intensity corresponds to the illuminant colour, which can be used to compute the colour descriptors. This algorithm was refined in [8] by actually obtaining an average model of surface reflectance and proposing a procedure to compute the average image reflectance. On the basis of this, the colour normalisation proposed by GW consists on dividing each colour channel by its average value:

$$[s_{ij}\alpha r_{ij}, s_{ij}\beta g_{ij}, s_{ij}\gamma b_{ij}] \mapsto \left[\frac{\alpha s_{ij} r_{ij}}{\frac{\alpha}{n}\sum_{\forall ij\in I} s_{ij} r_{ij}}, \frac{\beta s_{ij} g_{ij}}{\frac{\beta}{n}\sum_{\forall ij\in I} s_{ij} g_{ij}}, \frac{\gamma s_{ij} b_{ij}}{\frac{\gamma}{n}\sum_{\forall ij\in I} s_{ij} b_{ij}}\right] \tag{1}$$

Equation (1) is what we call *Basic GW Normalisation*. It is invariant to illuminant colour variations (α, β, γ), but it has one important drawback: it assumes that s_{ij} is constant. This means that:

- It does not account for all illuminant intensity variations (some of these variations will cancel with the colour coefficients). In the sequel we will present a new normalisation procedure that will account for most of the illuminant intensity variations.

- It fails when the surface reflectances of the scene (r_{ij},g_{ij},b_{ij}) vary, i.e. when new objects appear or disappear in the scene. This means that GW is only valid for static scenes. In the next section we will introduce a dynamic extension to the GW algorithm that solves this problem using the redundant information available in an image sequence.

Let us define the *image average geometrical reflectance*, $\bar{\mu}$, as

$$\bar{\mu} = [\mu_r, \mu_g, \mu_b] = \left[\frac{1}{n} \sum_{ij \in I} s_{ij} r_{ij}, \frac{1}{n} \sum_{ij \in I} s_{ij} g_{ij}, \frac{1}{n} \sum_{ij \in I} s_{ij} b_{ij} \right],$$

where n is the number of image pixels. It represents the average $[RGB]$ image values, once the colour illuminant component has been removed.

If we assume that the average geometrical reflectance is constant over the image sequence, then the following normalisation, called *Projective GW Normalisation*, removes the illuminant colour and intensity changes:

$$\begin{aligned} [s_{ij}\alpha r_{ij}, s_{ij}\beta g_{ij}, s_{ij}\gamma b_{ij}] &\mapsto \left[\frac{\alpha s_{ij} r_{ij}}{\alpha \mu_r}, \frac{\beta s_{ij} g_{ij}}{\beta \mu_g}, \frac{\gamma s_{ij} b_{ij}}{\gamma \mu_b} \right] \mapsto \\ &\mapsto \left[\frac{s_{ij} r_{ij}}{s_{ij}(\frac{r_{ij}}{\mu_r} + \frac{g_{ij}}{\mu_g} + \frac{b_{ij}}{\mu_b})}, \frac{s_{ij} g_{ij}}{s_{ij}(\frac{r_{ij}}{\mu_r} + \frac{g_{ij}}{\mu_g} + \frac{b_{ij}}{\mu_b})} \right] \end{aligned} \quad (2)$$

This new normalisation is still valid just for static scenes. In the next section we will present a dynamic extension to GW in order to overcome this problem.

3 Face tracking using Dynamic Grey World

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 Dynamic GW (DGW) algorithm is presented.

3.1 Face segmentation and tracking using a skin colour model

Given a sequence of colour images, building a face tracker is straight forward if we have a reliable model of the image colour distributions. Let I_{rgb} be the $[RGB]$ channels of image I, and let $p(I_{rgb}|skin)$ and $p(I_{rgb}|back)$ be the conditional colour probability density functions (pfds) of the skin and background respectively (we assume that background is anything that is not skin). Using the Bayes formula, the probability that a pixel with colour I_{rgb} be *skin*, $P(skin|I_{rgb})$, can be computed as follows:

$$P(skin|I_{rgb}) = \frac{p(I_{rgb}|skin)P_s}{p(I_{rgb}|skin)P_s + p(I_{rgb}|back)P_b},$$

where P_s and P_b are the a priori probabilities of *skin* and *background*. The transformation $\mathcal{T}(I_{rgb}) = 255 \times P(skin|I_{rgb})$ returns an image whose grey values represent the probability of being skin (see Fig. 1). Face tracking on this image can be performed with a mode seeking algorithm, like [4], by computing the position and orientation of the face colour cluster in each frame [3].

The problem now is to make the previous statistical model robust to variations in the scene illumination. This will be achieved by working in one of the normalised colour spaces presented in the previous section.

Figure 1: Face segmentation based on skin colour. On the left is shown the colour image, on the right the probability image.

3.2 Face colour models

Various authors have indicated different preferences for modelling the colour distributions. In [12] Gaussian mixture models are used, whereas in [5] and [13] pure histogram-based representations are chosen. In our experiments we found that pure histogram-based models are faster and represent more accurately the skin colour distribution than continuous ones, if enough samples are available. On the other hand, continuous models are adequate when the number of samples is low.

As the Basic GW normalisation defines a colour space of dimension three, $I_{\hat{r}\hat{g}\hat{b}}$, given the small number of samples available in an image, we propose modelling the *skin* Basic GW colour distribution with a continuous Gaussian model. As can be seen in Fig. 2, $p(I_{\hat{r}\hat{g}\hat{b}}|skin)$ is approximately Gaussian. On the left are shown the Chi-square and Gaussian plots of the $I_{\hat{r}}$, $I_{\hat{g}}$ and $I_{\hat{b}}$ marginals and the $I_{\hat{r}\hat{g}\hat{b}}$ multivariate distribution. From the analysis of these plots we can verify that the assumption $p(I_{\hat{r}\hat{g}\hat{b}}|skin) \sim N(\bar{m}_s, \Sigma_s, I_{\hat{r}\hat{g}\hat{b}})$ can not be rejected. On the other hand, it is not possible to find an analytic model for the *background*, so we will model it with a uniform pdf, $u_b(I_{\hat{r}\hat{g}\hat{b}})$.

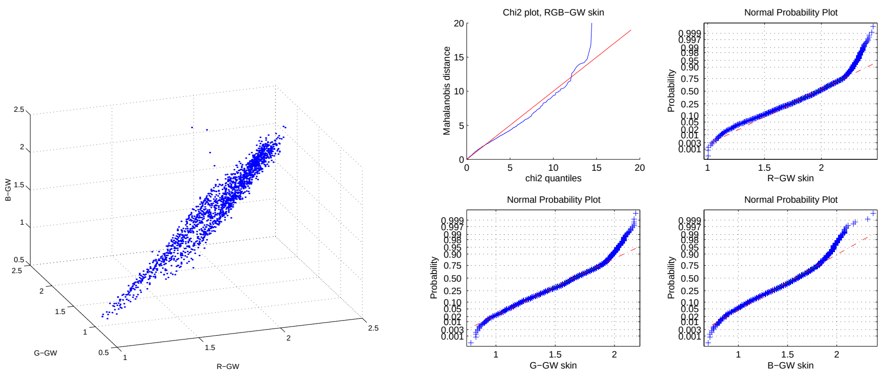

Figure 2: Skin colour pdf in Basic GW space. On the left is shown the skin colour cluster. On the right are shown the Chi-square plot for the multivariate distribution and the Normal plots for the marginals.

If we approximate the priors $P_s \approx n_s/n$ and $P_b \approx n_b/n$, where n_s and n_b

are respectively the number of the *skin* and *background* pixels, then

$$P(skin|I_{rgb}) = \frac{n_s\, N(\bar{m},\, \Sigma, I_{\hat{r}\hat{g}\hat{b}})}{n_s\, N(\bar{m},\, \Sigma, I_{\hat{r}\hat{g}\hat{b}}) + n_b\, u_b(I_{\hat{r}\hat{g}\hat{b}})}.$$

On the other hand, the Projective GW Normalisation defines a two dimensional colour space, $I_{\hat{r}\hat{g}}$, in which the colour distributions can be modeled with the skin, $h_s(I_{\hat{r}\hat{g}})$, and background, $h_b(I_{\hat{r}\hat{g}})$, colour histograms. Consequently, in our implementation we model the posterior probability as:

$$P(skin|I_{rgb}) = \frac{h_s(I_{\hat{r}\hat{g}})}{h_s(I_{\hat{r}\hat{g}}) + h_b(I_{\hat{r}\hat{g}})}.$$

3.3 The Dynamic Grey World algorithm

The main problem of GW is that it was conceived for static images; i.e. it fails when there is a big change in the image average geometrical reflectance. In this section we propose a dynamic extension to GW (DGW) which will detect this situation and update the GW model.

In the following we assume that there exists a partition of the image sequence into a set of image subsequences such that the image average geometrical reflectance is constant over each subsequence; i.e. the basic GW algorithm can be used as a colour constancy criterion over each subsequence. We will use the first image of each subsequence as a *reference image.* The other images of the subsequence will be segmented using the colour descriptors of the reference image.

Let I^r_{rgb}, I^t_{rgb} and I^{t-1}_{rgb} be respectively the reference image, the present and the previous image, $F^r_{\hat{r}\hat{g}\hat{b}}$ be the face pixels in GW space, $\bar{\mu}^{I^t}_{rgb}$ be the average value for each colour channel in I^t_{rgb}, $\bar{\mu}^{F^r}_{\hat{r}\hat{g}\hat{b}}$ and $\bar{\mu}^{F^t}_{\hat{r}\hat{g}\hat{b}}$ be the average GW descriptors for the face pixels in the reference and present image respectively, and $\mathcal{E}$ be the statistical distribution of the GW space colour descriptors for the reference image.

The dynamic extension to the basic GW is based on the fact that when a change in the average geometrical reflectance ($\bar{\mu}$) is detected, the GW colour descriptors for the present image $I^t_{\hat{r}\hat{g}\hat{b}}$ can still be computed with the the average pixel values of the previous image, $\bar{\mu}^{I^{t-1}}_{rgb}$. In this situation we segment the present image with the the average pixel values of the previous one, and let the present image be the reference image.

The problem now is how to detect a change of subsequence. We do this just by searching for a change in the average geometrical reflectance. This can not be accomplished on the basis of analysing $\bar{\mu}^I_{rgb}$, as it also changes with the illuminant colour. We solve this problem by monitoring the average GW descriptors of the face pixels. As they are invariant to illuminant colour changes, a change in these descriptors is necessarily caused by a change in average geometrical reflectance.

Based on these ideas in Fig. 3 we propose the Dynamic-GW (DGW) algorithm.

```
Initialisation
  /*Initialise the reference image model using motion
  segmentation and a precalculated colour model*/
  [ℰ, μ̄^{F^r}_{r̂ĝb̂}] = InitTracking();
While (true)   /* tracker main loop */
  μ̄^{I^t}_{rgb} = Mean(I^t_{rgb}); /* image mean rgb values */
  I^t_{r̂ĝb̂} = I^t_{rgb} / μ̄^{I^t}_{rgb}   /* GW normalisation */
  F^t_{r̂ĝb̂} = ProbabilisticSegment(I^t_{r̂ĝb̂}, ℰ); /* segment img */
  μ̄^{F^t}_{r̂ĝb̂} = ComputeAvgFaceGW(F^t_{r̂ĝb̂}); /* face avg GW descriptors */
  If ||μ̄^{F^r}_{r̂ĝb̂} − μ̄^{F^t}_{r̂ĝb̂}|| > Δ then /* change of subsequence */
    I^t_{r̂ĝb̂} = I^t_{rgb} / μ̄^{I^{t-1}}_{rgb}   /* GW normalise with previous mean */
    F^t_{r̂ĝb̂} = ProbabilisticSegment(I^t_{r̂ĝb̂}, ℰ); /* segment image */
    I^r_{r̂ĝb̂} = I^t_{r̂ĝb̂} /* update reference image */
    μ̄^{F^r}_{r̂ĝb̂} = ComputeAvgFaceGW(F^t_{r̂ĝb̂}); /* face GW descriptors */
    [ℰ] = ColourDistrib(F^t_{r̂ĝb̂}); /* ref. colour distrib */
  end /* if */
end /* while */
```

Figure 3: Dymanic Grey World Algorithm

4 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 controlled tungsten light.

In the first experiment we validate the hypothesis on which the DGW algorithms rests, namely: variations in the average geometrical reflectance can be detected, and the reference image of each subsequence can be segmented. We acquired a sequence of 200 images with a green object appearing at one point and illuminant geometrical variations taking place at a different moments. The result of this experiment is shown in Fig. 4. Four images of the sequence are shown, each in one column of the figure. Each of them represents respectively the first image of the sequence (image 1), a change in the illuminant (roof lights turned off) (image 26), and the appearance and disappearance of an object (images 88 and 139). In this experiment the system detects three subsequences (1 to 87, 88 to 138, and 139 to 200). This is clearly visible in the plot at the bottom of Fig. 4. In image 26 the roof fluorescent lights are turned off. This geometrical illumination variation can be perceived again in the face GW descriptors plot. In this case the segmentation is good. This is an example of "worst case" test. In similar situations with stronger variations in the illuminant geometry, the system may not be able to segment the image and eventually may loose the target. Images 88 and 139 show the first segmented image in the two last subsequences, that coincide with the appearance and disappearance of an object in the image. Here we can see how the system detects a change of subsequence and correctly segments the images.

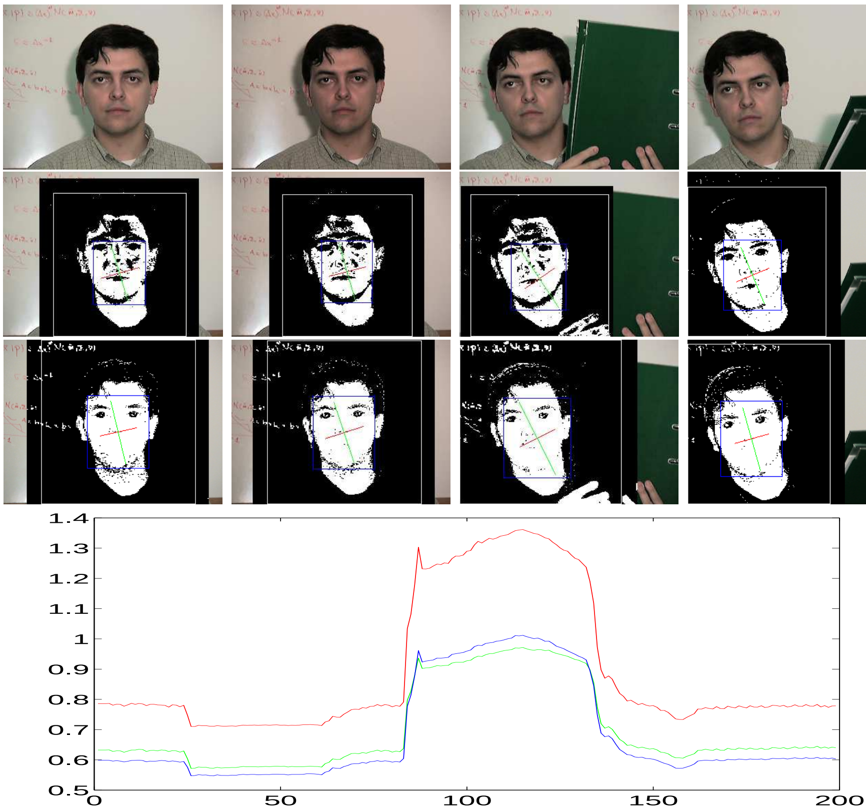

Figure 4: Hypothesis validation experiment. On the first row four images of a sequence are shown. Their segmentation with the DGW algorithm using respectively the Basic GW and the Projective GW normalisation are presented on the second and third row. The average r,g and b face GW descriptors (in red, green and blue color respectively) are shown on the fourth row.

The goal of the next experiment is to check that the dynamic extension to GW is necessary; i.e. we want to check what would happen if we segment the previous sequence with the Basic GW Normalisation without the dynamic extension. In Fig. 5 the same sequence as in Fig. 4 is used and the same images are shown. We can clearly perceive that without the dynamic extension, the initial colour model is invalid when a change in the image average geometrical reflectance (caused by the appearance of an object) takes place. The initial model gradually becomes valid again as the object disappears (see last column).

In the following experiment we compare the performance of the DGW algorithm with the rg-normalised colour constancy algorithm. We use a sequence with a set of images with "difficult" background (i.e. brownish books and shelves to distract the segmentation). In Fig. 6 four frames of the sequence are shown in each column, representing each one of them following situations: initial image, tungsten frontal light turned off, tungsten frontal light partially turned on,

Figure 5: DGW algorithm versus Basic GW without dynamic extension. Basic GW Normalisation with DGW algorithm results are shown in first row. Basic GW without dynamic extension is shown in the second row.

green object is introduced. Raw images are shown in the first row, rg-normalised results in the second row, and DGW segmentation results with Basic GW and Projective GW Normalisation are shown in the third and fourth row respectively. Visual inspection of these results show that Basic and Projective GW Normalisation with DGW have similar results, the later being a slightly more robust to light intensity variations. The results of the second row show a clear success of the DGW compared to the rg-normalisation when the illuminant colour abruptly changes (second column).

5 Conclusions

The GW algorithm was conceived to work with static images. In this paper we have presented a dynamic extension of the GW algorithm and a Projective GW Normalisation procedure. These extensions of GW were designed to make it work in real-time with sequences of images with varying environmental conditions. In the experiments conducted, the DGW algorithm, both with Basic and Projective GW Normalisations, performed better than the rg-normalised algorithm, when sudden changes in the illuminant colour take place. As predicted, the Projective GW Normalisation procedure is preferable to the Basic one, because it performs better with illuminant intensity variations and because it works in a two dimensional colour space.

The least favorable case for our algorithm occurs when strong changes in the illuminant geometry take place. In this case, the average geometrical reflectance is not constant and consequently the normalisation procedure is not successful. More research is needed in order to design normalisation procedures invariant to these changes.

In conclusion, a tracker based on a simple cue such as colour, will always have failure situations. In this paper we have analysed some of the weak points of the rg-normalised algorithm. The DGW algorithm is not perfect either, as its performance can be seriously affected by strong and fast changes in the illuminant geometry. Also, in general, colour-based tracking has difficulty with similarly coloured distractions in the background. In spite of these limitations,

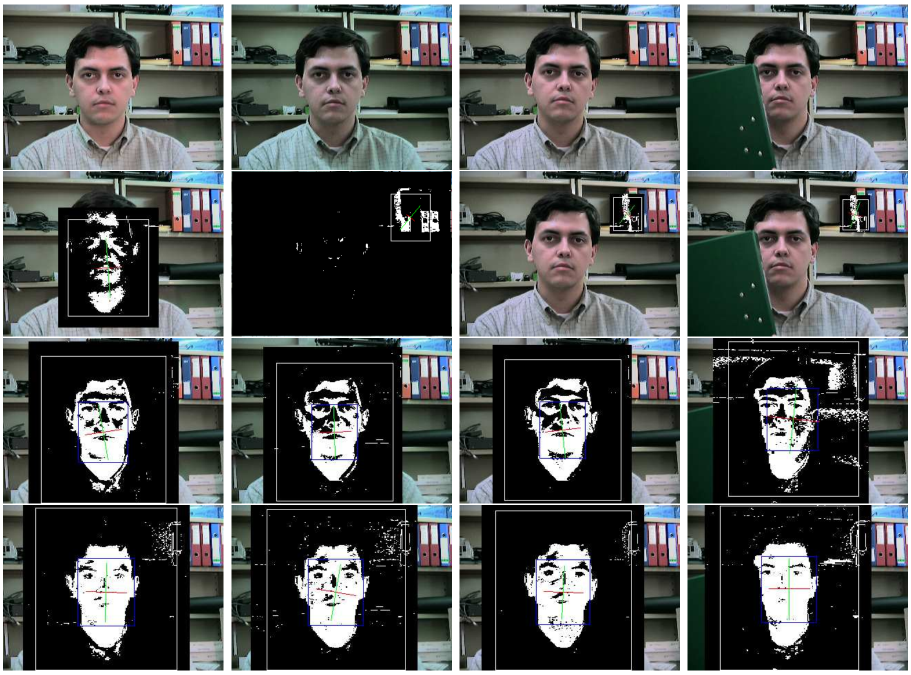

Figure 6: Comparison of DGW and RG-normalisation colour constancy for face tracking.

colour-based trackers are good 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.

Acknowledgements

The authors gratefully acknowledge the Spanish Comisión Interministerial de Ciencia y Tecnología (CICyT) for funding this research under contract number TIC1999-1021. José Miguel Buenaposada was also funded by a FPU grant from the Spanish Ministry of Education.

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] G. Buchsbaum. A spatial processor model for object colour perception. *Journal of the Fanklin Institute*, 310: 1-26, 1980.

[3] G. Bradski. Computer Vision face tracking for use in a perceptual user interface. *Proc. of Workshop on applications of Computer Vision, WACV'98*, 214–219, 1998.

[4] Y. Cheng. Mean shift, mode seeking and clustering. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 17: 790-799, 1995.

[5] 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.

[6] 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.

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

[8] R. Gershon, A.D. Jepson and J.K. Tsotsos. From [R,G,B] to surface reflectance: Computing color constant descriptors in images. *Proc. Int. Joint Conf. on Artificial Intelligence*, 755–758, 1987.

[9] 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.

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

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

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

[13] M. Soriano, B. Martinkauppi, S. Huovinen, M. Laaksonen. Skin detection in video under changing illumination conditions. *Proceedings of the Int. Conference on Automatic Face and Gesture Recognition* (FG'00), 839–842, Grenoble. France. 2000.

[14] 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.

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

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

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

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

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