PERGAMON

Pattern Recognition 36 (2003) 681–690

PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY

www.elsevier.com/locate/patcog

Adaptive skin color modeling using the skin locus for selecting training pixels

Maricor Soriano[a,b,*], Birgitta Martinkauppi[b], Sami Huovinen[b], Mika Laaksonen[c]

[a] *National Institute of Physics, University of the Philippines, Quezon City, 1101 Diliman, Philippines*
[b] *Machine Vision and Intelligent Systems Group, Infotech Oulu, University of Oulu, P.O. Box 4500, FIN-90014 Oulu, Finland*
[c] *NOKIA Mobile Phones, P.O.Box 50, FIN-90571 Oulu, Finland*

Received 15 May 2001; accepted 18 April 2002

Abstract

Techniques for color-based tracking of faces or hands often assume a static skin model yet skin color, as measured by a camera, can change when lighting changes. Therefore, for robust skin pixel detection, an adaptive skin color model must be employed. We demonstrate a chromaticity-based constraint to select training pixels in a scene for updating a dynamic skin color model under changing illumination conditions. The method makes use of the 'skin locus' of a camera, that is, the area in chromaticity space where skin chromaticity under various lighting and camera calibration conditions is observed. Skin color models derived from the technique are compared with that derived by a common spatial constraint and is shown to be more consistent with manually extracted ground truth skin model per frame even as localization errors increase. The technique is applied to color-based face tracking in indoor and outdoor videos and is shown to succeed more often than other color model adaptation techniques. © 2002 Pattern Recognition Society. Published by Elsevier Science Ltd. All rights reserved.

Keywords: Face detection; Tracking; Color; Skin locus; Adaptation; Chromaticity

1. Introduction

Color-based face detection and tracking techniques begin with skin color modeling. Color is usually the first cue sought for locating face candidates in video because skin color is distinct. After locating skin pixel candidates, non-face blobs can be eliminated by using texture, shape or motion cues. Face candidates may then be used to recognize a person or to code the area of interest (the face) with better quality while increasing the compression on the non-essential part (background) as in Ref. [1]. Very often a static skin color model is learned offline or in the first few frames of a sequence. However, skin color, as measured by a camera, may change as illumination condition changes. If the illumination condition is static but non-uniform, movement of the subject can likewise cause the captured skin color to change. Therefore, face detection that uses a static skin color model is certain to fail in unconstrained imaging conditions.

One solution may be to correct image colors first—a body of work in color science related to color constancy [2,3]. Generally, these algorithms rely on a priori conditions such as: (a) that the average color in the scene must be gray (i.e., there are equal amounts of red, green and blue in the scene), (b) that the illumination must be slowly and smoothly varying in time or space, and (c) except for the Retinex algorithm [4,5], that the illumination change must be global. Violations in any of these conditions result in poor color correction.

In real-life cases we are not always in control of the illumination. The lighting variation can be local as, for instance, when a person is indoors and near a window, part of the face may be illuminated by daylight while part may be illuminated by room lighting. Since faces are 3D objects, inter-reflections, shadowing, occlusions and sharp color edges occur on face images. Color constancy algorithms are primarily designed to please the human eye and

* Corresponding author. National Institute of Physics, University of the Philippines, Quezon City, 1101 Diliman, Philippines. Tel.: +63-2-434-4232; fax: +63-2-920-5474.

E-mail address: msoriano@nip.upd.edu.ph (M. Soriano).

PII: S0031-3203(02)00089-4

may not always be appropriate for machine vision applications [6]. Therefore, dynamically adapting the skin color model is a more suitable approach than color correction if the intended face detector or tracker is to be used in unconstrained illumination conditions.

Only a few papers attempt to use dynamically adapting skin color models. In Refs. [7,8], skin color distribution is modeled as a single Gaussian cluster or a Gaussian mixture. Model update is achieved by recursively adapting the mean, covariance and prior probabilities of each Gaussian cluster using color from a subimage within the tracker bounding box. In Ref. [9], Yoo argued that since the face is generally oval in shape pixels from an oval region on the face tracking result may be taken as training pixels. In each case, the selection of training pixels relies on *geometric* constraints. A limitation of such a constraint is that if the training region includes non-skin pixels, the color model can adapt to non-skin color. Tracking failures need to be detected to halt adaptation, for example, by measuring the log-likelihood error, as in Ref. [8].

In this paper we propose to use the knowledge of the range of skin color under the normalized color coordinates (NCC) as a criteria for selecting training pixels when updating the skin model. This paper summarizes the results of previous work wherein we have demonstrated the applicability of the technique [10–12]. Since the technique uses a *chromatic* constraint, its main advantage over geometric constraints is that pixels outside the set of possible skin colors are excluded in the updating process. In NCC-space, the range of skin colors under light sources of varying correlated color temperature follows the curvature of the Planckian locus and we call it the 'skin locus' after Störring [13]. The skin locus can be obtained either by measurements or by simulation if camera sensitivities, skin reflectance and illuminant spectral power distribution are given [13,14]. Once known, it may be used to select skin pixels from a localized face candidate to update the skin color model.

The effectivity of the technique is assessed by comparing it with adaptive skin color modeling which uses geometric constraints, in two aspects: (1) closeness of the computed model with respect to ground truth, and (2) success of a color-based face tracker employing the adapted skin color model.

The rest of the paper is organized as follows: Section 2 discusses the skin locus, ratio histogram and adaptive histogram backprojection. Model error measures, tracking experiments and results are presented in Section 3. The paper ends with conclusions and recommendations in Section 4.

2. Method

2.1. Skin locus

Measurements of skin reflectances, light spectral power distribution and camera channel sensitivities allow the computation of ideal color camera red, green and blue (RGB) output for different skin types [14]. In this study we convert RGB to NCC, r, g, and I where, $I=(R+G+B)$, $r=R/I$, $g=G/I$. Normalizing with intensity, I, reduces brightness dependence in the chromaticity coordinates r and g. Note that $b=B/I$ is no longer unique since $b=1-r-g$. Thus, two values, r and g, are sufficient to represent chromaticity. Hence, the chromaticity coordinates of NCC are called rg-space. One may use either (g,b) or (r,b) as chromaticity components as well. We have chosen to use (r,g) since sensitivity of most cameras in blue are small and may therefore be noisy. Actual measurements have shown that dark, yellowish and pale skin have almost the same chromaticity [15]. Störring [13] has also shown that for a color camera white balanced for one illuminant, the chromaticity of skin follows a curve similar to the Planckian locus when it is imaged under light sources of different correlated color temperatures. Störring termed the set of points in chromaticity space which belong to the range of skin color as the *skin locus*.

NCC is not the only color space that separates pixel brightness and chromaticity but it has a useful property which we will exploit. If two colors are mixed in different brightness proportions, the range of color they can produce together are the colors in rg-space which lie in the line joining the chromaticity points of the two colors. This property is useful in modeling skin color because in real situations, more than one light source may be shining on the face at a time. In Fig. 1a, the left side of the face is near a window and is illuminated by daylight while the right side is mostly illuminated by fluorescent light. The chromaticities of pixels on the face straddling the two light sources (e.g. pixels from the forehead to the chin) are along the line joining the chromaticities of each side of the face, as shown in Fig. 1b. If the person moves her head such that, at times, her face is fully illuminated by one light source, it can be surmised that her skin color distribution will move across the chromaticity region in Fig. 1b. Fig. 1b already shows part of the skin locus of the camera used. Even under only two light sources, other skin colors may appear due to the mixing of these lights on the skin. Therefore, if a skin model was derived from one lighting condition, it may be adjusted to the current skin colors by selecting pixels for training which are known to be instances of skin color under other light sources. Knowing the range of possible skin colors enables us to reject those pixels whose colors do not possibly belong to skin.

2.2. Adaptive histogram backprojection

We propose to use the skin locus as a filter that allows only feasible skin colored pixels in the update of a skin color model for a certain camera. We modify histogram backprojection [16] by adding color histogram adaptation and apply it in color-based face tracking.

An initial estimate of the skin color model, a 2-D histogram $\mathbf{S}(r,g)$, is obtained from cut-out skin regions in the

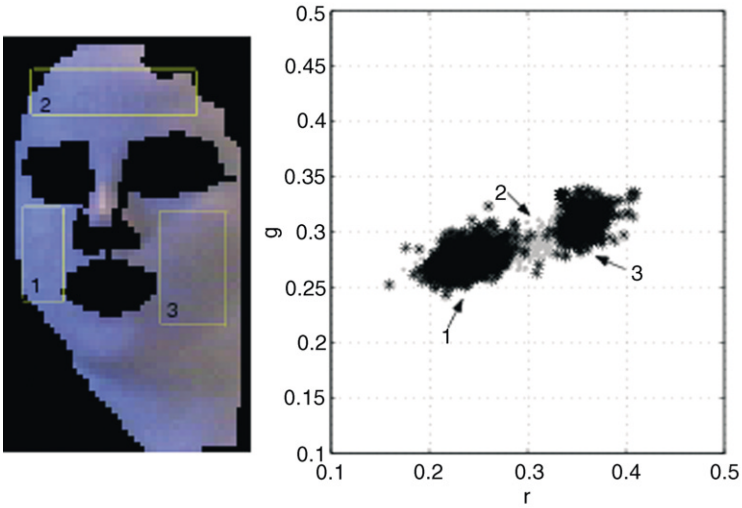

Fig. 1. Chromaticity plots of pixels in different parts of the face: (1) illuminated by daylight, (2) forehead illuminated both by daylight and fluorescent light, (3) cheek illuminated by fluorescent light.

face in the first frame and normalized such that the maximum value is one. The next frame to be segmented is transformed into rg-space and each pixel p_i with chromaticity (r_i, g_i) is assigned the value of the histogram at $\mathbf{S}(r_i, g_i)$ (the chromaticity of a pixel defines its corresponding skin histogram value). A high concentration of values over a predefined threshold in the histogram-backprojected image may then be considered as the likely vicinity of the target.

A variation is to use a ratio histogram $\mathbf{R}(r, g)$, which is $\mathbf{S}(r, g)$ divided by the whole image histogram, $\mathbf{I}(r, g)$, in order to penalize colors which are part of the model but are also present in the background and thereby increase the contrast between skin and background pixels. We use ratio histograms for skin modeling in our experiments. A parametric model for skin color (e.g. Gaussian fitting) tends to smoothen the actual distribution and requires a distance metric (e.g. Mahalanobis) to assign probability of skin color. With ratio histogram and histogram backprojection, no fitting is necessary because the histogram itself is used as the model, and probabilities are assigned by simple table lookup, thus leading to faster labeling.

Further processing may then be performed on the backprojected image. Let a tracking algorithm take as input the backprojected image and deliver as output the bounding box around the face. If illumination conditions cause the measured skin color to change then the current model will only produce few lighted pixels in the face region unless the skin model is updated. We propose to do adaptation by determining pixels in the bounding box that fall under the skin locus and to use these pixels to compute the current ratio histogram. This way the skin pixels are chosen automatically.

For adaptive modeling, the two-dimensional ratio histogram of skin $\mathbf{R}_t(r, g)$ is calculated from each frame. The color of the target (skin) may change rapidly due to varying lighting conditions. Thus, to provide a smooth transition of $\mathbf{R}_t$ between frames, a moving average (MA) is used to calculate a resultant ratio histogram $\tilde{\mathbf{R}}_t$,

$$\tilde{\mathbf{R}}_t = \frac{(1-\alpha)\mathbf{R}_t + \alpha\mathbf{R}_{t-1}}{\max((1-\alpha)\mathbf{R}_t + \alpha\mathbf{R}_{t-1})}, \tag{1}$$

where we set $\alpha = 0.5$. The subscript t denotes the frame index. The denominator in Eq. (1) serves to scale the resulting 2-d histogram such that its maximum element is 1. In addition, using MA reduces noise in backprojected images of low-quality cameras. Due to noise, the color information of pixels may change even if there is no actual change in scene or illumination over the frame. This causes error if color information and ratio histogram from only the previous frame is used to label pixels in the current frame. Note that Eq. (1) is not recursive since the successive frames use $\mathbf{R}_t$ not $\tilde{\mathbf{R}}_t$.

Fig. 2 summarizes the adaptive histogram backprojection algorithm we employed. In the first frame of video, skin regions are manually selected and their colors converted into r–g chromaticities. An initial estimate of the skin color ratio histogram is computed by passing the converted pixels into the skin locus filter to eliminate non-skin colors. The next frame to be processed is transformed into r–g space and histogram backprojected using the current skin color histogram. The outcome is a graylevel image, with pixel values equal to the frequency values of their r–g coordinates in the skin color model. Morphological operations are performed to enhance the face blobs and a bounding box is place around the largest blob. When overlaid on the image, the bounding box frames the face. Adaptation starts on the succeeding frames.

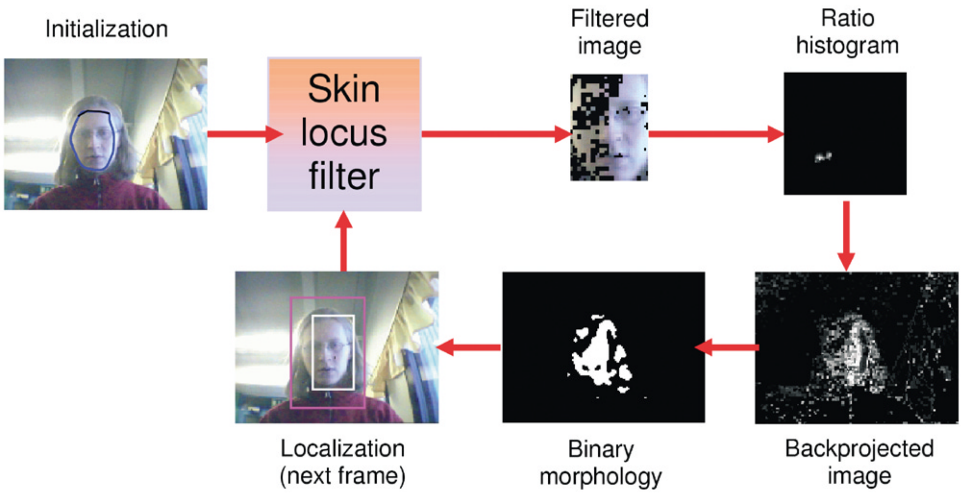

Fig. 2. Flow of adaptive histogram backprojection using the skin locus as chromatic filter.

The pixels from the computed bounding box are screened by the skin locus filter. Using only skin-like pixels from the bounding box, a new skin color histogram is computed and averaged with the previous model. The next frame is backprojected with this new model and the bounding box plus adaptation process begins anew with the succeeding frames.

The cost of using color information from several previous frames with moving average is that the confidence level of ratio information is reduced by a function of number of frames used in MA. This means that even if ratio information of one color is high it is possible that the color does not belong to the target anymore. This may occur sometimes for example when there are too rapid changes in lighting conditions. These characteristics of MA must be balanced and suitable number of previous frames must be selected for MA histogram calculations.

In principle, any color space may be employed for the skin model. The advantage of NCC over other spaces is its speed of computation and the straightforward mixing of chromaticities. Once the skin locus is found, its boundaries may be modeled by functions of, at least, second order [17].

Matas [18] presented a similar idea: possible changes in a color patch are limited to a region in chromaticity space. In our case, we calibrated the camera under four different simulated illuminants and have three illumination changes for each calibration. This allows us a wider operation range where it is possible to adapt to the color change. In Ref. [18], a clustering algorithm is used to find cluster representatives in the chromaticity space for both object and background. Our method is simpler because we only have to define the cluster boundaries and the knowledge of background is not needed. Finally, our work is different in that our approach is a low-level operation, i.e. filtering, while Matas performs a high-level operation, recognition.

3. Experiments

3.1. Creation of skin locus

Faces were imaged in darkroom using four simulated illuminants (2300K horizon daylight, 2856K incandescent A, TL84 fluorescent and D65 daylight). Calibrating the camera for each light source and capturing the image of a face under each light source in turn results in 16 different combinations of current lighting and camera calibration conditions (four calibrations and four illuminants) [14]. Skin color measurements were done for Caucasian and Asian subjects.

When the r and g chromaticities of skin regions from all 16 conditions are plotted in rg-space, skin color occupy a downward opening crescent which is similar to the trend of the Planckian locus. Fig. 3 shows the skin locus of a USB Nogatech 1CCD webcamera used in the experiments. The skin locus is thickest in calibrated conditions (when current illuminant is the same as reference illuminant for white balancing) and thinnest when current color temperature is far from that of the calibrated case. (For comparison of the skin locus of different cameras, see Ref. [17].)

A simple membership function to the skin locus is a pair of quadratic functions defining the upper and lower bound of the cluster. For each r, the maximum and minimum g was used to estimate the upper and lower quadratic functions. Using least-squares estimation, the upper bound quadratic coefficients are found to be $a_u = -1.3767$,

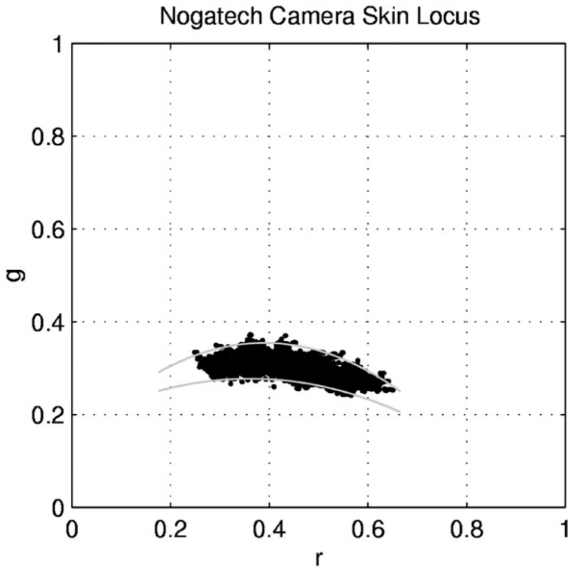

Fig. 3. Skin locus of Nogatech camera in *rg*-space.

$b_u = 1.0743$, $c_u = 0.1452$; the lower bound coefficients are $a_d = -0.776$, $b_d = 0.5601$, $c_d = 0.1766$. For some cameras, the skin locus may partially or wholly include the white point ($r = g = 0.33$). To prevent grayish and whitish pixels from being labeled as skin, a circle with radius 0.02 is drawn around the white point and pixels falling within the circle are excluded from skin membership. Pixels with chromaticity (r, g) are then given skin locus membership value $S(r, g)$ where

$$S = \begin{cases} 1, & (g < g_u) \text{ AND } (g > g_d) \text{ AND } (W_r > 0.0004), \\ 0, & \text{otherwise}, \end{cases} \tag{2}$$

where $g_u = a_u r^2 + b_u r + c_u$, $g_d = a_d r^2 + b_d r + c_d$ and $W = (r - 0.33)^2 + (g - 0.33)^2$.

To capture realistic situations of skin color change, movies of different persons in varying indoor and outdoor lighting conditions were taken (frame rate: 30 fps, frame size: 160×120). Volunteers were asked to walk along selected routes while carrying a laptop upon which a color camera was attached [19]. In the movie NOMOV1, the camera had been white balanced under fluorescent light. A Caucasian volunteer starts near a window and walks to a corridor with fluorescent lighting and then to a window again. In NOMOV2, an Asian volunteer moves in front of a window before turning to the corridor. In NOMOV4, with the camera white balanced for Horizon daylight, a Caucasian volunteer walks from a room out into a corridor. An outdoor movie is shown in NOMOV6 where there was only skylight and direct sunlight. The Caucasian volunteer moves from the Sun to the shade and back to the Sun again.

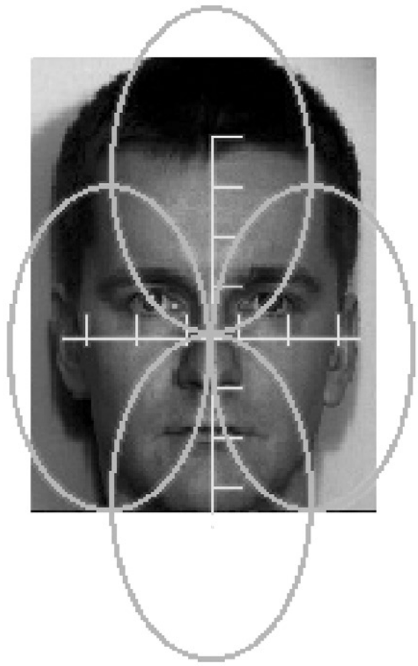

Fig. 4. Displacement of oval search regions to simulate face localization error.

3.2. Agreement of adapted and ground truth skin color model

To compare our chromatic constraint (the skin locus) for selecting training pixels with an existing spatial constraint (elliptical region) we measure how well skin color models derived from each constraint closely resemble the ground truth skin color model.

The ground truth skin color model $\mathbf{R}_{gt}$ was computed from pixels within manually defined face regions in each frame. It should be noted that we included the eyes, mouth, eyebrows and nose in the calculation. The skin color model for our chromatic constraint $\mathbf{R}_l$ was computed from all pixels in the ground truth face bounding box whose chromaticities fall within the skin locus. For the spatial constraint all pixels within an ellipse inside the bounding box where used to compute the skin model $\mathbf{R}_e$.

To simulate errors in face localization we displaced the center of the ground truth bounding box in increments of 10% from the original center in top, bottom, left and right directions, from 0% to 100%. We then computed the skin color models from both constraints using the displaced bounding boxes as new search regions. Fig. 4 illustrates the displacement for the ellipse.

The mean model error D for a movie at certain displacement distance was calculated using following formula:

$$D_{e_B} = \frac{1}{F * 4 * BINS} \sum_{f}^{F} \sum_{dir}^{4} |\mathbf{R}_{gt} - \mathbf{R}_c| \tag{3}$$

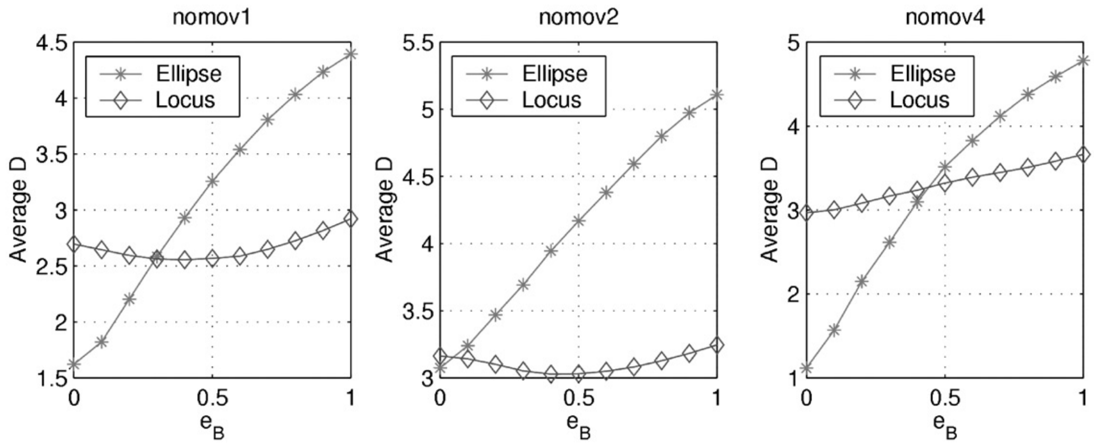

Fig. 5. Histogram error with locus and elliptical constraint for 3 movies, from left to right, NOMOV1, NOMOV2 and NOMOV4.

where $\mathbf{R}_{gt}$ is the ground truth skin ratio histogram and $\mathbf{R}_c$ is the ratio histogram of the constraint used ($\mathbf{R}_l$ or $\mathbf{R}_e$), $|\cdot|$ denotes Euclidean distance, *dir* is direction of displacement (total number is 4 for up, down, left and right), f is frame index, F is the total number of frames, $BINS$ is the total number of bins in rg-space, (we set it to $64 \times 64 = 4096$), and the subscript e_B is the percentage center-to-center displacement from ground truth bounding box.

Fig. 5 shows the average histogram error D for spatial and chromatic constraints for 3 movies. In general, histogram errors from the spatial constraint increases with increasing displacement, which is not surprising. As the percentage displacement of the bounding box increases, less and less of skin falls within the ellipse. Comparing D for both geometric and skin locus constraint, histogram errors from the skin locus constraint are quite stable over the range of displacement as shown by their smaller slope. This implies that a tracking routine that uses the skin locus constraint to update the skin color model will more likely recover from localization mistakes than that which uses only a spatial constraint.

Another observation from Fig. 5 is that for small displacements the spatial constraint appears better than chromaticity constraint in following the ground truth skin model. This is expected because the skin locus will generally exclude pixels from the eyes and lips whereas the ground truth skin color model was computed with these features included. If the eyes, mouth, etc. were removed from the ground truth skin model D would be lesser for chromatic constraint than geometric constraint even during small displacements.

3.3. Tracking success

To test the technique on actual video, a color-based face tracking algorithm was implemented that makes use of adaptive histogram backprojection as described in Section 2.2. Fig. 6 shows the result of the tracking algorithm for every 50th frame of NOMOV1 when the skin model is learned from the first frame and fixed (adaptation turned off). The facial skin color in Fig. 6 first appears bluish and pinkish, then as the volunteer proceeds to the corridor, the color becomes normal skin tone. He passes through dimly lit sections before he approaches another window whereupon his face appears pinkish again. Around the face candidate, the tracking algorithm draws a white box and, for the next frame, limits the search within a slightly larger bounding box (shown in magenta in Fig. 6). The tracker fails in the middle of the sequence, where the skin has normal color, and only recovers towards the end when the skin once again appears pinkish as in the initial condition. In comparison, Fig. 7 shows the result of the tracking when the skin model is adapted. With adaptation using the skin locus, the tracking algorithm recovers well enough to track the face throughout the sequence.

To quantitatively assess the goodness of our localization we introduce an overlap measure A given by

$$A = \frac{|\mathbf{A}_{gt} \cap \mathbf{A}_c|}{\sqrt{|\mathbf{A}_{gt}| \times |\mathbf{A}_c|}}, \quad (4)$$

where $\mathbf{A}_{gt}$ is the ground truth bounding box with $|\mathbf{A}_{gt}|$ as its area in number of pixels, and $\mathbf{A}_c$ is the computed bounding box with $|\mathbf{A}_c|$ its area. The numerator is the area of the overlap of the two bounding boxes. When the area of the ground truth bounding box and the computed bounding box are the same and completely overlap, A equals 1. This quality measure requires that ground truth is available and for movies to be tested ground truth was extracted manually. In addition to A, an error counter is included which increments if no skin cluster is found, and resets to zero when tracking resumes.

Fig. 8 compares the graph of A for the sequence of Fig. 6 when skin color histogram is fixed, no skin locus is employed (top graph), and of Fig. 5 when histogram is adapted with the skin locus constraint (bottom graph). Error count here has been normalized to one. Failure from frames 300 to 770

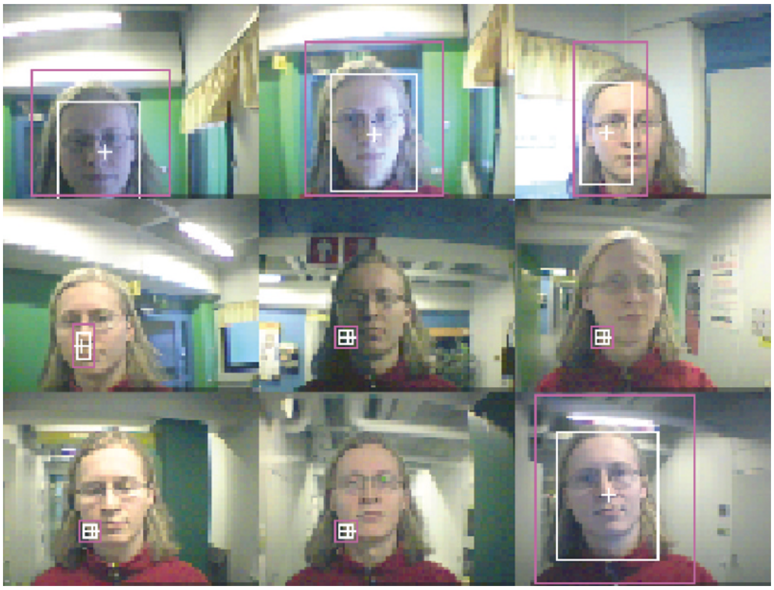

Fig. 6. Tracking results on movie NOMOV1 with fixed skin color model. White bounding box locates the face. Magenta bounding box is search region.

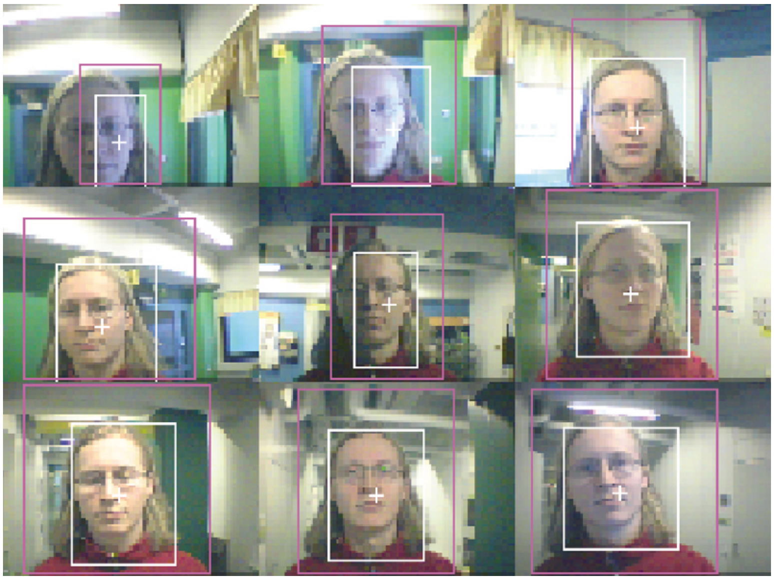

Fig. 7. Tracking results on movie NOMOV1 with adaptive skin color modeling using skin locus constraint. White bounding box locates the face. Magenta bounding box is search region.

is clearly shown when a fixed histogram is used. On the other hand, A is consistently high for the adapted histogram case with only a small range of frames (from 420 to 450) having positive errorcount which was due to the face losing color while passing through a dark region along the corridor.

Finally, we compared tracking performance of a skin color model adapted using the skin locus constraint with that adapted using a geometric (ellipse) constraint. Fig. 9 shows the result of a tracking algorithm for every 50th frame of NOMOV6 where ellipse constraint was used in skin color update. The figure also shows how skin color can change drastically if the subject merely moves from sunlight to the shadows. Without the skin locus, the tracker failed because it adapted to background colors when the person moved to the shadows. Fig. 10 shows the tracking result if the skin locus is used. The skin locus constraint clearly helps

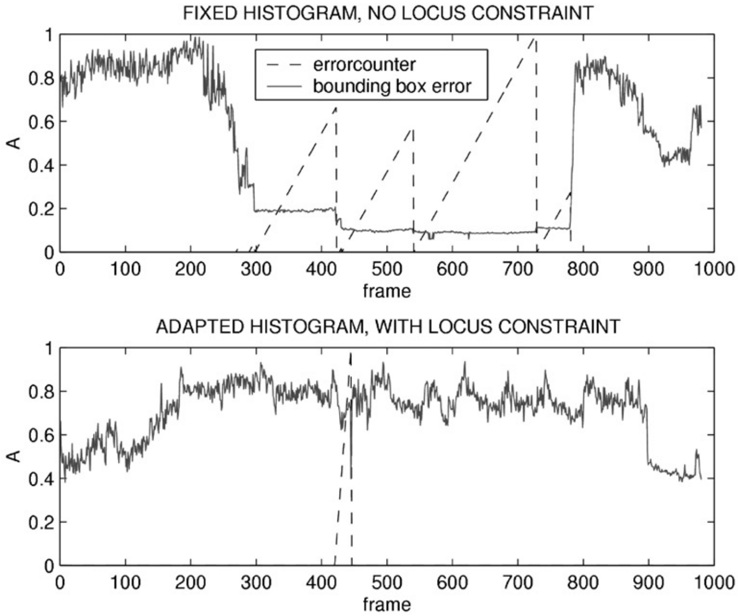

Fig. 8. Tracking success in NOMOV1 measured by overlap A.

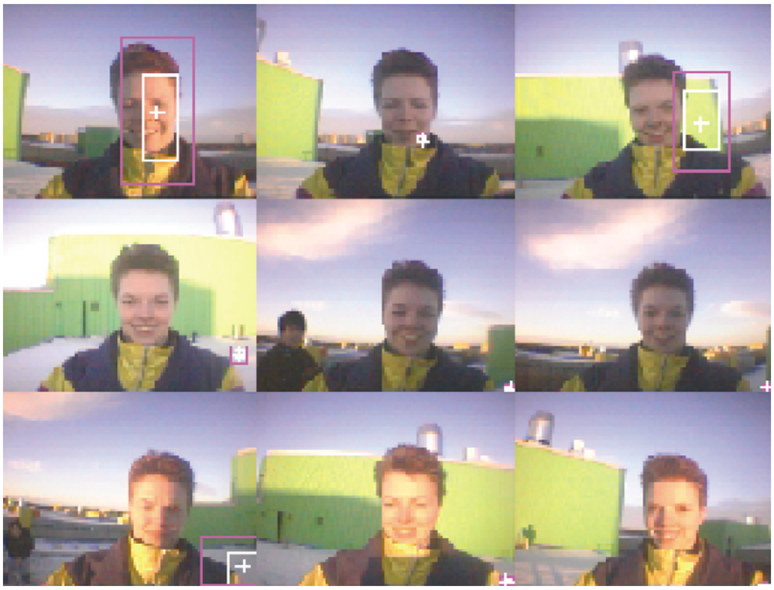

Fig. 9. Tracking result on movie NOMOV6 with elliptical constraint for adaptation. White bounding box locates the face. Magenta bounding box is search region.

the tracking routine cope with the changing illumination condition [19].

4. Conclusions

This paper addresses the issue of how to select pixels for training when a skin color model is to be updated, an important procedure in applications such as color-based face tracking under uncontrolled illumination conditions. The proposed technique makes use of the skin locus, which is the range of skin color in chromaticity space, to choose pixels for training from a tracking bounding box.

Our results prove that the skin locus is an effective filter for selecting skin pixels in changing illumination condi-

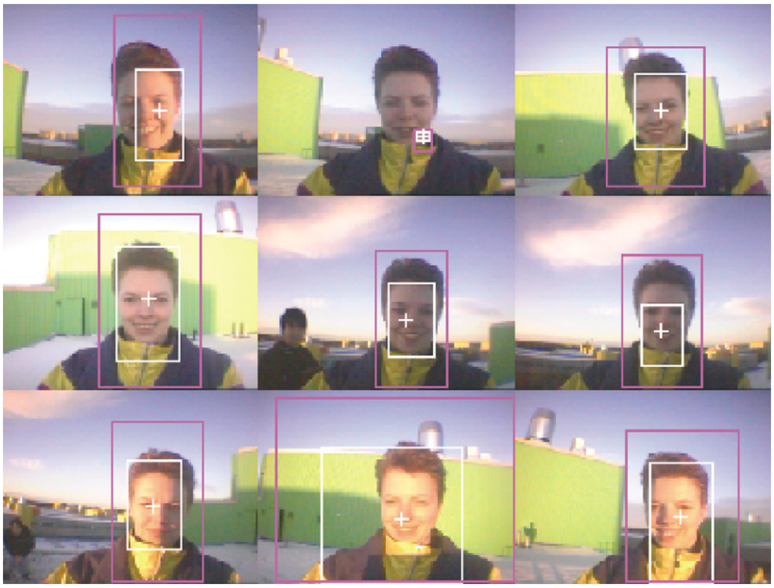

Fig. 10. Tracking result on movie NOMOV6 with skin locus for adaptation. White bounding box locates the face. Magenta bounding box is search region.

tions. We have shown that skin color models generated from the skin locus constraint are robust to localization errors. Color-based tracking experiments demonstrate that adapted skin color models using chromatic constraint succeeds over a fixed model or an adaptive model applying elliptical constraint.

It must be emphasized that the skin locus is camera specific. This technique is suited for a dedicated camera and not for arbitrary sequences from the web. The skin locus may be found in two ways, either by taking images of faces under different illumination conditions, or by calculating the RGB values of the camera given the color signal of skin and illuminant.

As with most face detection techniques, skin color is not enough to locate the face. Other cues such as shape, texture, and motion may still be needed to refine face localization. Once the face is positively detected, our technique may be used to lock the tracker on the face even under changing light.

Acknowledgements

The financial support of Machine Vision Group is gratefully acknowledged.

References

[1] D. Chai, K. Ngan, Face segmentation using skin-color map in videophone applications, IEEE Trans. Circuits Systems Video Technol. 9 (4) (1999) 551–564.

[2] G. Finlayson, B. Funt, K. Barnard, Color constancy with shadows, Perception 23 (1995) 720–725.

[3] D. Forsyth, A novel algorithm for color constancy, Int. J. Comput. Vision 5 (1990) 5–36.

[4] D. Jobson, Z. Rahman, G. Woodell, A multi-scale retinex for bridging the gap between color images and the human observation of scenes, IEEE Trans. Image Process. 6 (7) (1997) 965–976.

[5] E. H. Land, Recent advances in retinex theory, Vision Res. 26 (1986) 7–21.

[6] B. Funt, K. Barnard, L. Martin, Is color constancy good enough? Proceedings of the Fifth European Conference on Computer Vision, 1998, pp. 445–459.

[7] J. Yang, A. Waibel, A real-time face tracker, IEEE International Workshop on Applications of Computer Vision, 1996, pp. 142–147.

[8] S. McKenna, Y. Raja, S. Gong, Tracking color objects using adaptive mixture models, Image Vision Comput. 17 (1998) 225–231.

[9] T.-W. Yoo, I.-S. Oh, A fast algorithm for tracking human faces based on chromatic histograms, Pattern Recognition Lett. 20 (1999) 967–978.

[10] M. Soriano, B. Martinkauppi, S. Huovinen, M. Laaksonen, Using the skin locus to cope with changing illumination conditions in color-based face tracking, Proceedings of the IEEE Nordic Signal Processing Symposium 13–15 June, Kolmården, Sweden, 2000, pp. 383–386.

[11] M. Soriano, B. Martinkauppi, S. Huovinen, M. Laaksonen, Skin detection in video under changing illumination conditions, Proceedings of the 15th International Conference on Pattern Recognition, Vol. 1, Barcelona, Spain, 2000, pp. 839–842.

[12] M. Soriano, B. Martinkauppi, S. Huovinen, M. Laaksonen, Skin color modeling under varying illumination conditions using the skin locus for selecting training pixels, Proceedings

of the Workshop on Real-Time Image Sequence Analysis, Oulu, Finland, 2000, pp. 43–49.
[13] M. Störring, H. Andersen, E. Granum, Skin color detection under changing lighting condition, Proceedings of the Seventh Symposium on Intelligent Robotics Systems, Coimbra, Portugal, 1999, pp. 187–195.
[14] E. Marszalec, B. Martinkauppi, M. Soriano, M. Pietikäinen, Physics-based face database for color research, J. Electron. Imaging 9 (1) (2000) 32–38.
[15] M. Hunke, A. Waibel, Face locating and tracking for human–computer interaction, Proceedings of the 28th Asilomar Conference on Signals, Systems and Computers, Vol. 2, California, USA, 1994, pp. 1277–1281.
[16] M. Swain, D. Ballard, Color indexing, Int. J. Comput. Vision 7 (1) (1991) 11–32.
[17] B. Martinkauppi, M. Soriano, M. Laaksonen, Behaviour of skin color under varying illumination seen by different cameras in different color spaces, in: M.A. Hunt (Ed.), Machine Vision in Industrial Inspection IX, SPIE Vol. 4301, 2001, pp. 102–112.
[18] J. Matas, R. Marik, J. Kittler, Illumination invariant color recognition, in: E. Hancock (Ed.), Proceedings of the British Machine Vision Conference, pp. 469–479. BMVA Press, York, UK, 1994.
[19] These sequences may be viewed at http://www.ee.oulu.fi/research/imag/color.

About the Author—MARICOR SORIANO obtained her Ph.D. in Physics at the National Institute of Physics, University of the Philippines in 1997. Currently she is assistant professor at the University of the Philippines and a part-time senior researcher at the University of Oulu in Finland. Her research interests include physics education, color and texture image analysis, microscopy, and pattern recognition. Since 2000, she is Secretary General of the Samahang Pisika ng Pilipinas (Physical Society of the Philippines) and Program Coordinator of the Instrumentation Physics Lab at NIP.

About the Author—BIRGITTA MARTINKAUPPI received her M.Sc (Eng.) degree with honors from the Department of Electrical Engineering at the University of Oulu in 1997. Since then she has been a researcher in the Machine Vision and Intelligent Systems Group at the University of Oulu. She is a scholar of the Finnish Graduate School in Electronics, Telecommunications and Automation (GETA) and is a member of OSA, SPIE, IS&T and the Pattern Recognition Society of Finland. Her current research interest is illumination robustness for color machine vision.

About the Author—SAMI HUOVINEN is currently taking up M.Sc in Electronic Engineering at the University of Oulu. His research is in color machine vision since 1999.

About the Author—MIKA LAAKSONEN received his M.Sc (Eng.) degree in 1997 from the Department of Computer Science at the Lappeenranta University of Technology, Finland. Currently, he works as Senior Research Engineer at Nokia Mobile Phones and is enrolled in the Ph.D. program at the University of Oulu. His research interest is in computer vision as input modality for human–computer interaction.