

Detecção de Faces em Imagens

Marcos Paulo C. Rosa (14/0027131), Rennê Ruan A. Oliveira (14/0030930), Edgar Fabiano S. Filho (14/0019201)

Universidade de Brasília 1/2015 - Departamento de Ciência da Computação

Introdução ao Processamento de Imagens - Professor: Bruno Macchiavello

Abstract—O material desenvolvido neste relatório apresenta o procedimento para criação de um algoritmo que realiza detecção de faces em uma imagem digital disposta. A determinação é feita a partir da correlação entre pele e cabelo. No primeiro caso, isolando a partir da variação de tons de pele e, no segundo, pelo espectro da intensidade. Considerando diversas situações com um contraste minimamente operável entre os objetos (especialmente entre pele e cabelo), a taxa de acerto comprova-se como de, aproximadamente, 80% e a resposta é rápida, abaixo de 15 segundos para imagens de média resolução. Isso indica uma capacidade de reutilização do programa para aplicações simples ou de uma possível melhoria do algoritmo, completando-o com demais métodos de detecção.

I. INTRODUÇÃO

Utilizações de algoritmos de detecção de face se tornam cada vez mais presentes em nosso cotidiano, sendo criados para diversos fins, tal que há diversos meios de construção de tais algoritmos. Para o modelo específico, o principal fator abordado para a criação do algoritmo são as características gerais que compõem uma face, sendo eles: textura, cor de pele e estruturação geral da face.

As operações realizadas em imagens digitais presentes neste relatório foram realizadas utilizando o software de alta performance MATLAB, facilitando assim a manipulação dos *pixels* da imagem disposta, assim como seus níveis de cor. Esse é um software iterativo da MathWorks, voltado para o cálculo numérico, capaz de integrar a análise numérica, o cálculo com matrizes, o processamento de sinais e a construção de gráficos em um único ambiente de desenvolvimento.

Os procedimentos aqui mostrados são de baixa complexidade, aplicação da teoria em [1], tendo um bom tempo de resposta dependendo da imagem digital a ser analisada, tal que, as operações lógicas utilizadas trabalham, originalmente, *pixel a pixel* em seus planos RGB (Padrão de cores vermelho, verde e azul). Ademais, durante as comparações lógicas do programa, transforma os componentes RGB em HSI (espaço de cor representante de matiz, saturação e intensidade) [2].

II. RESUMO

As faces de diferentes imagens possuem características semelhantes, porém as faces que devem ser reconhecidas

estarão sobre um fundo complexo, ou seja, com diversos objetos ou com luminosidade variável. Assim, é necessário aplicar operações lógicas *pixel a pixel* da imagem para verificar a participação de cada um na estruturação de uma face.

Tendo como base o plano de cores, disposto inicialmente como o RGB, pode-se segmentar, da imagem inicial, a pele presente no rosto, uma vez que o espectro de tonalidade da pele humana possui níveis característicos. Após isso, podemos, por meio da quantização, encontrar os limites geométricos (em coordenadas) da pele segmentada previamente. Este mesmo procedimento será aplicado aos cabelos presente nas imagens, mas, nesse caso, diferenciando os objetos do fundo pela variação de brilho. Por conseguinte, com a união das coordenadas obtidas podemos ter certeza da localização da face.

III. METODOLOGIA

A seguir estão as descrições das técnicas utilizadas em cada módulo do programa:

A. Detecção de Pele

O método utilizado para detectar a pele presente em cada imagem se baseou na normalização do plano RGB das imagens, tendo em vista as fórmulas dadas no artigo referencial:

$$r = \frac{R}{R + G + B} \quad (1)$$

$$g = \frac{R}{R + G + B} \quad (2)$$

Com a presença de tais coeficientes, podemos utilizar as fórmulas abaixo para a determinação dos limites do espectro de cor da pele humana, estas que foram determinadas a partir de uma análise estatística da distribuição e concentração dos tons de pele em cada componente RGB vistos em [3].

$$F1(r) = -1.37r^2 + 1.0743r + 0.2 \quad (3)$$

$$F2(r) = -0.776r^2 + 0.5601r + 0.18 \quad (4)$$

A partir dessas duas equações, capacitou-se a definição dos *pixels*, ou seja, determinar as características de pele por meio das condições estabelecidas, aplicado sobre toda a imagem,

pixel a *pixel*. Entretanto, tais condições impedem das cores branca, inseridas nos espectros de cor de pele, não ser captadas. Em vista que os valores de *r* e *g* dessas tonalidades são conhecidos, uma simples comparação é realizada, retirando o módulo da diferença dos *pixels* da imagem por esses valores, ignorando eles caso sejam brancos.

Se o *pixel* analisado não for branco, o fator normalizado *g* dele deverá ser menor que o valor obtido na equação 3, porém deverá também ser maior que o valor obtido na equação 4, na qual esse intervalo, como anteriormente explicitado, representa um modelo de espectro de cores da pele humana. Depois de efetivadas tais operações, é obtida uma imagem em que os *pixels* considerados pele se tornaram brancos e, caso contrário, preto.

Após o tipo de segmentação realizado anteriormente, foi percebido que a imagem possui *pixels* além do desejado e, baseado no artigo tido como referência, foi utilizada a relação entre dois tipos de representação de cor. No caso, para remover tais *pixels* indevidamente extraídos, usou-se a lógica do modelo HSI, tal que o ângulo (Hue/matiz) é dado pela seguinte equação:

$$\theta = \cos^{-1} \left\{ \frac{0.5[(R - G) + (R - B)]}{\sqrt{(R - G)^2 + (R - B)(G - B)}} \right\}$$

$$\begin{cases} H = \theta, & \text{se } B \leq G \\ H = 360^\circ - \theta, & \text{se } B > G \end{cases} \quad (5)$$

Originalmente, os *pixels* considerados pele pelo método com o matiz seriam aqueles que possuem ângulo maior que 240 ou menor que 20. Todavia, empiricamente, foi comprovado que os melhores valores são os menores que 30, pois alguns componentes de pele eram indevidamente retirados dentro dessa pequena margem. Da mesma forma, os valores acima de 240 causavam mais casos de erro do que acerto, em vista que faziam considerar grande parte dos cabelos e objetos ressaltados ao fundo.

As operações envolvendo o aspecto HSI dos *pixels* foram realizadas em um módulo específico chamado de "*mascara_hue*". Sendo assim, a detecção de pele só é dada por completo ao efetuar a interseção de ambos métodos RGB e HSI, realizada na função "*DetectaPele*".

B. Detecção de Cabelo

A segmentação do cabelo se torna semelhante a da pele, porém deve ser incluído um fato essencial para este, a intensidade (brilho) da imagem, dada pela média dos níveis de cor RGB. E usando a interseção entre as condições de luminosidade para cada *pixel* e a união entre algumas faixas de cor, podemos identificar os *pixels* representativos do cabelo, desta vez com limiares mais simples de se calcular, que englobam as cores comuns de cabelo existente variando entre níveis de marrom, sendo as condições as seguintes:

$$\begin{cases} 1, & \text{se } (I < 80 \cap (B - G < 15 \cup B - R < 15)) \\ 0, & \text{caso contrário} \end{cases}$$

Sendo *I* o componente de luminosidade calculado previamente, foi dispensada, na maioria dos casos, a utilização do fator do matiz para detecção de cabelo. No caso, primordialmente utilizava-se a união entre o cabelo determinado anteriormente e os demais valores encontrado pelo matiz entre 20 e 40. Porém, essa faixa fazia com que grande parte dos objetos, vestimentas e, inclusive, tons de pele fossem considerados como parte do cabelo, levando ao erro.

Tendo isso em vista, apenas é considerado o matiz quando, após o processo de filtragem por tamanho explicado mais adiante, é determinado que os objetos indicados como cabelo são menores do que um limiar estabelecido (o dobro da raiz da multiplicação entre a altura e a largura da imagem). Desse modo, o valor não é fixado e fica dependendo de propriedades de cada imagem, aumentando a taxa de acerto.

C. Quantização

O processo de quantização é realizado visando melhorar a classificação das segmentações obtidas, para realmente decidir se o que extraímos de informação da imagem original se trata de cabelo e pele. Foi elaborado um módulo a parte para isso, denominado "*quantização*".

A quantização foi realizada com a utilização de uma máscara 5x5, tal que seria contabilizado quantos *pixels* dentro da máscara possuem valor, em um processo similar a convolução. Ademais, como não foi realizado *padding*, condições foram criadas para os pontos das bordas serem considerados como nulos.

Após tal obtenção, comparamos o valor contabilizado com um limiar, que será 12 tanto para pele quanto para cabelo. Caso a soma dos *pixels* seja maior que o limiar, o valor no centro da máscara será nulificado, caso contrário será 1 (correspondendo a presença do objeto).

Uma vez que a quantização reduz a qualidade da imagem de forma subjetiva, ela serve para realizar a criação dos contornos respectivos do cabelo e da pele, dados pela aplicação da técnica *BoundingBox* por comandos do software MATLAB e, assim, obtendo as coordenadas dos limites dos objetos segmentados e quantizados.

D. Interseção

Sobre o *BoundingBox*, há o módulo "*size_filter*" que o realiza e opera diversos procedimentos para melhor adequação da imagem. Mais especificamente, faz um fechamento de acordo com o tamanho da imagem, para remover possíveis istmos entre parte de um mesmo objeto, marca as regiões distintas, calcula a área destas e remove algumas.

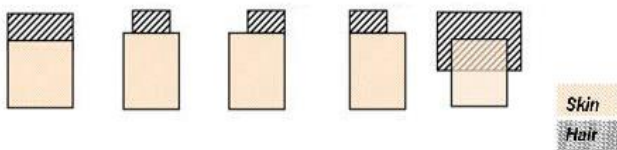
A remoção segue a ideia de filtragem por tamanho, ou seja, se um objeto for menor do que a área estabelecida, ele é desconsiderado. Para casos nos quais apenas uma face é procurada, o limiar é determinado como o valor da área do maior objeto. Enquanto, para casos procurando um valor indeterminado de faces, originalmente seria um valor fixado em 10 (de acordo com o artigo referencial), porém uma nova lógica foi determinada.

A ideia para determinação do limiar é: caso seja pele, será a maior área dividida por três (pois as faces não terão grande variação de tamanho) e, caso seja cabelo, o dobro da raiz da multiplicação entre a altura e a largura da imagem e, se nenhum objeto for maior que esse limiar, será usado o valor da maior área.

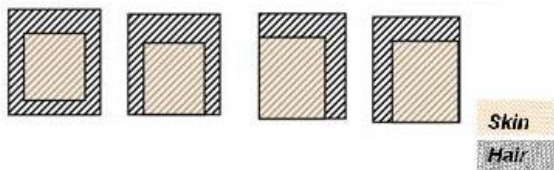
Com os limites em coordenadas da pele e cabelo determinados, deve-se realizar diversas condições de interseção das *BoundingBox* de ambos para abranger os tipos comuns de cabelos existentes. O comportamento das *BoundingBox* é dado a seguir:

[Mínimo x, Mínimo y, Tamanho x, Tamanho y]

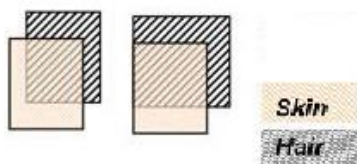
Sendo x coordenadas em relação a horizontal e y coordenadas em relação a vertical, tendo como origem o primeiro pixel da imagem e considerando o limite inferior/mínimo como o mínimo e o limite superior/máximo como a soma do mínimo com o tamanho. O número de condições realizadas é menor que o número de tipos de cabelo, uma vez que ao tratar com os limites dos eixos x e y podemos enquadrar um ou mais exemplos em uma única condição. Para os seguintes tipos de cabelo temos as condições respectivas:



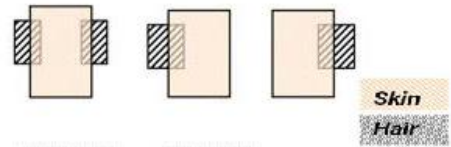
- 1) O mínimo y da pele deverá ser maior ou igual ao mínimo y do cabelo e não poderá ser maior o máximo de y para o cabelo somado a um fator de correção (altura da imagem dividida por cem), este para casos os uma pequena faixa entre o cabelo e a pele não tenha sido encontrada. Além disso, os limites em x do cabelo devem estar contidos no da pele ou o inverso.



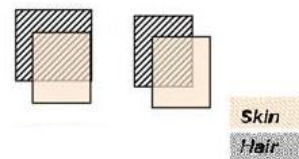
- 2) O limite inferior de x da pele deverá ser maior ou igual ao limite inferior de x do cabelo e o limite máximo em x da pele deverá ser menor ou igual que o do cabelo. As condições anteriores deverão ser realizadas também para os limites em y. No caso, as coordenadas da pele devem estar contidas nas do cabelo.



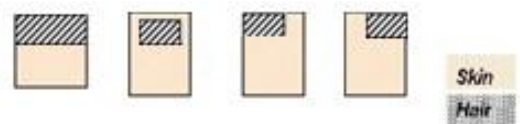
- 3) O limite inferior de x para a pele deverá menor ou igual ao do cabelo, porém o limite inferior em y deverá ser maior ou igual ao mesmo limite para o cabelo. Além disso, o limite superior em x para a pele deverá ser menor ou igual para o mesmo do cabelo e, no eixo y, maior ou igual.



- 4) O limite inferior de y da pele deve ser menor ou igual ao do cabelo e o superior, maior ou igual. Ademais, no eixo x o cabelo deve ter alguma interseção pela direita ou esquerda da pele. Ou seja, limite superior do cabelo maior que o inferior da pele e inferior do cabelo, menor, ou o inferior do cabelo menor que o superior da pele e o superior dele, maior.



- 5) O mínimo x e y da pele necessita ser maior ou igual ao mínimo x e y do cabelo, respectivamente. Assim como os máximos em x e y da pele precisam ser maiores ou iguais aos máximos do cabelo.



- 6) Um sexto caso foi criado, além dos originais do artigo referencial, pois muitas vezes as faces tomavam proporções maiores do que a do cabelo, devido a algum fundo irregular ou a cor do cabelo também ser vista como parte da pele, assim como o caso de ruivos. Para isso, quando os limites do cabelo estão inclusos na parte superior dos da pele, também é considerado como um caso válido.

IV. RESULTADOS

Para a realização dos testes presentes neste relatório foram utilizadas 30 imagens contendo apenas uma face, 12 contendo mais de uma e 7 contendo nenhuma. Como também, apenas foi considerado o modo de procurar diversas faces, para ter certeza de não estar encontrando faces a mais do que as contabilizadas. Para demonstração, utilizamos a seguinte imagem:



Fig. 1. Imagem Original.



Fig. 2. Extração dos *pixels* de pele com base no RGB e no matiz.



Fig. 3. Imagem segmentada após a extração dos *pixels* de pele com base no RGB.

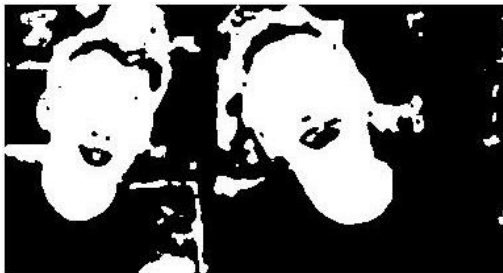


Fig. 4. Quantização dos *pixels* de pele.

A partir das imagens anteriores algumas análises já foram obtidas. Primeiramente, objetos bem contrastados do fundo também foram visto como parte da pele e que o processo de quantização já possibilitou a retirada de alguns pontos pequenos e a homogeneização da pele.



Fig. 5. Processo de segmentação do cabelo, a partir da intensidade e dos componentes RGB.



Fig. 6. Imagem segmentada a partir do cabelo.

Nota-se que há a extração de *pixels* indesejados que, pelas condições, não se encontram no padrão de cores de cabelo. Porém, alguns ainda permanecem, especialmente evidente na camisa do garoto à direita.

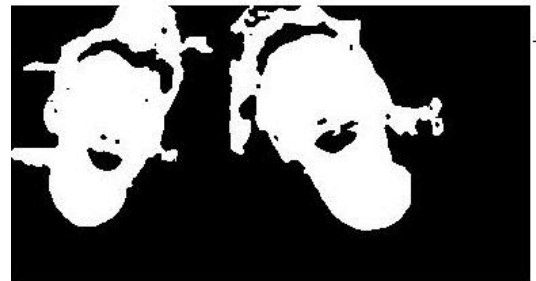


Fig. 7. Aplicação do filtro de tamanho na segmentação pela pele.

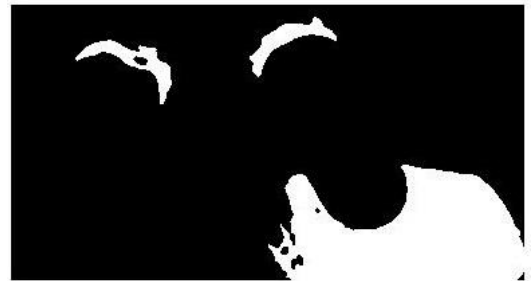


Fig. 8. Aplicação do filtro de tamanho na segmentação pelo cabelo.

Os pedaços soltos da imagem, que não pertenciam nem a pele nem ao cabelo, já foram evidentemente retirados por meio do filtro de tamanho. Durante esse processo, também foram determinados os *BoundingBox* correspondente aos limites de cabelo e pele, obtemos os seguintes limites:

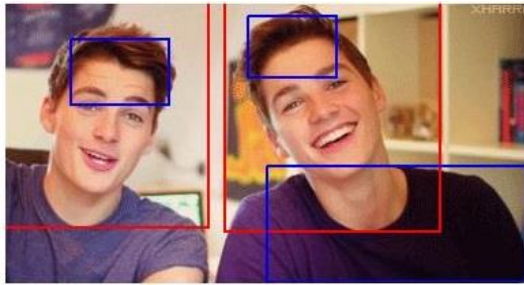


Fig. 9. Em azul, *BoundingBox* do cabelo e, em vermelho, da pele.

Percebe-se que os *BoundingBox* dado em azul, representativo do cabelo, reconhece a camisa presente na foto como cabelo devido a equação utilizada. Porém, para evitar que seja considerada uma face, se põe em prática as condições de interseção já apresentadas neste relatório, tal que, apenas o real cabelo que passe em uma das seis séries de condições irá de fato ser face, resultando no seguinte reconhecimento:

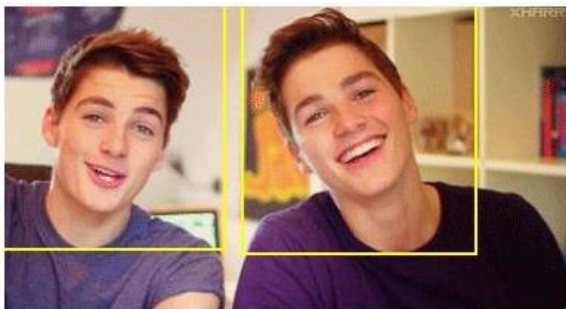


Fig. 10. Resultado da detecção de face.

O caso acima evidenciou o motivo de criação da condição 6, pois as pinturas de fundo também foram vistas como extensão da pele e o cabelo ficou presente na parte superior das coordenadas da pele.

A seguir, há o comparativo para resultados com apenas uma face, mais de duas ou nenhuma:

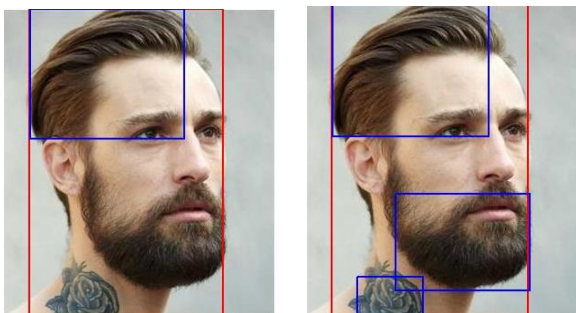


Fig. 11. Em azul, *BoundingBox* do cabelo e, em vermelho, da pele para uma face, à esquerda, e uma ou mais, à direita.



Fig. 12. Resultado da detecção para uma face.

Evidencia-se que, ao procurar apenas o maior objeto (uma face), já são eliminadas a barba e a tatuagem. Ainda assim, isso não se torna um empecilho para a busca de diversas faces, em vista que a posição dessas não levaria a consideração de cabelo, pois não segue um dos tipos de interseção estabelecidos.



Fig. 13. Em azul, *BoundingBox* do cabelo e, em vermelho, da pele para mais de duas face.



Fig. 14. Resultado da detecção para mais de duas faces.

A cerca da detecção para diversas faces, o funcionamento adequado requer uma mínima distância entre os indivíduos para não considerá-los como um único. A parte disso, resultados similares aos com um menor número de pessoas são obtidos.



Fig. 15. Comparativo entre o *BoundingBox* e o resultado da detecção para nenhuma face.

Como comprovado acima, casos em que não há interseção entre as coordenadas da pele e do cabelo são, corretamente indicados como sem face. De fato, na maioria dos casos, esse ponto nem é alcançado, pois (em imagens sem faces) o comum é não encontrar pontos que sejam considerados pele. Isso não impede, entretanto, a ocorrência de alguns erros, como imagens que os objetos tenham tonalidades similares às de pele e cabelo.

Abaixo está uma tabela mostrando os casos testados:

	Corretos	Falsos positivos	Falsos negativos	Total de testes
1 face	34 (89,47%)	1 (2,63%)	3 (7,90%)	38
Mais de 1 face	11 (64,71%)	0 (0,00%)	6 (35,29%)	17
Sem face	7 (63,64%)	4 (36,36%)	0 (0,00%)	11
	52 (78,79%)	5 (7,58%)	9 (13,63%)	66

Evidenciou-se que a maioria dos testes foi um sucesso, entretanto, os seguintes casos podem apresentar problemas:

1. pessoas de pele clara e cabelo ruivo, loiro ou branco e pessoas de pele escura e cabelo preto (o algoritmo, em ambos, não consegue diferenciar a pele do cabelo de forma ideal);
2. indivíduos sem cabelo (vai contra a preposição inicial do artigo de identificar faces pela diferença entre pele e cabelo);
3. imagem com contraste muito baixo (tornando difícil separar o fundo de objetos);
4. fundo com cor similar à da pele ou itens próximos (tornando a pele identificada maior do que simplesmente a face da pessoa);
5. pessoas bastante próximas (torna-se difícil a divisão do reconhecimento de faces distintas);
6. a coloração das roupas utilizadas, as quais podem ser consideradas como cabelo ou pele.

Sem considerar esses casos específicos, as segmentações tendem a separar corretamente a pele do cabelo e as interseções estabelecidas abrangem a maior parte dos casos que podem ser encontrados. Ademais, as mudanças efetivadas pelo grupo em relação à proposta originalmente tiveram, nessa etapa, um retorno, pois diminuiu o caso de falsos positivos e negativos. No caso, originalmente, diversos pontos eram considerados erroneamente, por limiares pouco robustos e

pelas faixas de RGB e HSI consideradas para segmentação serem diferentes das determinadas empiricamente.

V. CONCLUSÃO

Algoritmos de detecção de face têm ganhado cada vez mais aplicações de seu uso em processamento de imagens, desde desbloqueio de *smartphones*, computadores e compras por reconhecimento facial, *smile shutter* (detector de sorrisos em câmeras digitais), até identificação de pessoas em imagens de câmeras de segurança. O método apresentado no relatório é de baixa complexidade, seu objetivo é apenas verificar se tem ou não faces na imagem pela identificação da cor da pele e do cabelo e a verificação da posição de um em relação ao outro.

O código final mostrou-se bastante eficiente para casos em que apenas a face da pessoa estava bem evidente, destacada em relação ao fundo e com poucos objetos escuros próximos a ela. Os problemas de falsos positivos e negativos que apareceram se deram ao fato das limitações do próprio algoritmo, uma vez que o principal parâmetro que é utilizado para detecção de faces é exclusivamente a cor.

No caso, geralmente atrapalham na segmentação imagens com fundo não homogêneo e com cores muito próximas às do espectro de tonalidades de cor de pele humana, com objetos escuros muito próximos à cabeça (simulando uma parte do cabelo) ou em que a pele da pessoa aparece mais do que apenas no rosto (como no pescoço e no peito). Para solucionar o problema, bastava adicionar um pouco mais de complexidade ao algoritmo e colocar os casos para detectar a posição de estruturas na face (olhos, nariz, sobrancelhas, boca e etc.) para verificar se realmente é uma face ou parte dela.

VI. REFERÊNCIAS

- [1] Y. J. Chen and Y. C. Lin, "Simple face-detection algorithm based on minimum facial features," *IECON Proc. (Industrial Electron. Conf.)*, pp. 455–460, 2007.
- [2] M. Thompson, R. C. Gonzalez, and P. Wintz, *Digital Image Processing*, vol. 14, no. 3. 1981.
- [3] M. Soriano and B. Martinkauppi, "Using the skin locus to cope with changing illumination conditions in color-based face tracking," *IEEE Nord. signal ...*, pp. 2–5, 2000.