
ESTRUTURA PARA ANOTAÇÃO DE IMAGENS PARA A AVALIAÇÃO DE ALGORITMOS DE DETECÇÃO DE PELE HUMANA

Diogo Rosa Kuiaski, Gustavo Benvenuto Borba, Humberto Remígio Gamba e Hugo Vieira Neto

CPGEI/UTFPR, Curitiba, Brasil

e-mail: diogo.kuiaski@gmail.com

Abstract: We propose new metadata structure for human skin image annotation, whose main features are versatility and accessibility. The metadata are stored in one single file which uses the eXtensible Markup Language – XML. A controlled dataset of 120 skin images was built to populate the database and two well-know skin detection algorithms were used to evaluate the metadata structure. The results demonstrate that the new structure is a useful tool for the assessment of skin detection algorithms on applications such as face detection, recognition of lesions or to simply find people in scenes.

Palavras-chave: anotação de imagens, bancos de imagens, detecção de pele.

Introdução

A detecção de regiões de pele em imagens é atualmente um tópico bastante ativo nas áreas de processamento digital de imagens e visão computacional. Dentre as aplicações existentes, é possível citar a detecção de faces [1, 2], a detecção de conteúdo ofensivo (adulto) na Internet [1, 3, 4] e o uso em sistemas de vigilância para a localização de pessoas [5]. Ainda, importantes aplicações na área médica envolvem a identificação de gestos para o auxílio a pessoas portadoras de deficiência auditiva, através da tradução automática da linguagem de sinais [6] e a segmentação e levantamento de parâmetros de determinadas lesões cutâneas [7].

Uma possível abordagem para a detecção de pixels de pele em imagens é a seleção de um subespaço de um modelo de cores específico, dentro do qual estão contidas as cores típicas de pixels de pele [2, 8]. Neste caso, a identificação das regiões referentes à pele é feita através da simples comparação de cada pixel com os limites do subespaço pré-determinado. Métodos deste tipo são denominados EDSR (*Explicitly Defined Skin Region*) [1]. Também é possível utilizar abordagens estatísticas para a identificação das áreas da imagem contendo pele. Para isso são empregados métodos como classificadores Bayesianos [9] e modelos de misturas Gaussianas [10]. Em geral, estas técnicas requerem uma etapa de treinamento, na qual os pixels de pele (e não-pele) das imagens devem ser previamente conhecidos [1].

As medidas mais utilizadas para a avaliação quantitativa destes algoritmos são as taxas de verdadei-

ros positivos (TPR) e falsos positivos (FPR), além de curvas ROC (*Receiver Operating Characteristic*) [11]. Para estes casos, também são necessárias imagens (máscaras) nas quais cada pixel deve estar previamente categorizado em pele ou não pele.

Assim, pode-se afirmar que, tanto a etapa de desenvolvimento quanto a de avaliação de algoritmos para detecção de pele dependem fortemente da disponibilidade de imagens nas quais os pixels correspondentes à pele sejam conhecidos *a priori*. Estas imagens podem ser organizadas em conjuntos, ou bancos, nos quais para cada imagem original há uma imagem *padrão-ouro* ou *ground-truth* (GT) associada. Neste caso, cada imagem GT, também chamada de máscara, é binária e possui a mesmas dimensões da imagem original. Os pixels brancos (*true*) correspondem aos pixels de pele na imagem original, enquanto os pretos (*false*) aos pixels não-pele. Alguns exemplos de bancos deste tipo são apresentados nas referências [8, 9, 10].

No entanto, bancos de imagens como os citados acima não prevêm a inclusão de outros metadados além das máscaras binárias, não são estruturados e em geral não se encontram disponíveis publicamente.

Neste trabalho, desenvolveu-se um modelo de base de dados estruturada em XML para imagens contendo pele. O modelo proposto prevê a inclusão de metadados relevantes, tais como as condições de iluminação, resolução das imagens, dados do voluntário, região de interesse retangular, dentre outros. Para facilitar a inclusão dos metadados, processo denominado *anotação* das imagens, foi desenvolvido um aplicativo denominado *skinChooser*. A base de dados foi povoada com 120 imagens, divididas em quatro grupos de 30, cada um deles adquirido sob diferentes condições de iluminação.

Finalmente, utilizou-se a base de dados para a avaliação dos algoritmos de detecção de pele propostos em [2, 9]. Os resultados demonstram que a estrutura proposta é adequada para a descrição de imagens contendo pele e as 120 imagens adquiridas, juntamente com as respectivas máscaras binárias, podem ser utilizadas para a avaliação coerente de algoritmos de detecção de pele.

Materiais e Métodos

Estrutura XML – A Figura 1 ilustra o diagrama em árvore XML desenvolvido para a estruturação da base de dados de imagens de pele.

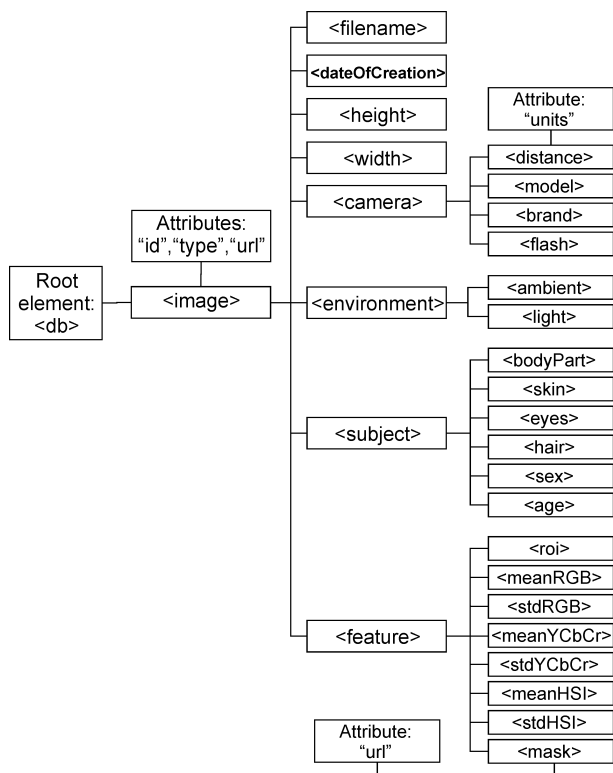


Figura 1: Diagrama em árvore XML desenvolvido.

Cada elemento *image* do arquivo XML suporta os metadados de uma determinada imagem do banco. Dentro do elemento *image*, os elementos *filename*, *dateOfCreation*, *height* e *width* não possuem filhos. Já os outros elementos do mesmo nível hierárquico possuem elementos filhos e são responsáveis pelo encapsulamento das demais informações relevantes de cada imagem.

Dentro de *camera* pode-se anotar em *distance* a distância entre a lente e o objeto, além de outras características da câmera como *model*, *brand* e o estado do *flash* (*on* ou *off*). Estudos realizados por [12] demonstram que, conhecendo-se estes dados, é possível estimar com maior coerência a cromaticidade típica dos pixels de pele.

Em *environment* é possível anotar outras condições de iluminação além da contemplada pelo elemento *flash*, pois sabe-se que as variações de iluminação devem ser consideradas no projeto de algoritmos para detecção de pele [1]. Os parâmetros sugeridos para o elemento *ambient* são *outdoor* ou *indoor*. No caso *outdoor* sugere-se que o elemento *light* receba *sun*, *cloudy* ou *shadow* e, no caso *indoor*, receba *fluorescent* ou *incandescent*.

O elemento *subject* permite a anotação de características da pessoa. Dentre os principais elementos de *subject* destacam-se *bodyPart* e *skin*. Neste último descreve-se o tipo de pele e sugere-se a utilização das seguintes categorias: *white*, *light brown*, *dark brown* ou *asiatic*.

Por fim, no elemento *feature* podem ser anotadas características de cor dos pixels contidos em uma região de interesse (ROI) retangular da imagem. O elemento *roi* indica as coordenadas desta região. Sabe-se que diferentes modelos de cores podem ser utilizados na detecção de pixels de pele [1]. Em *feature* estão contemplados os modelos de cores RGB, HSI e YCbCr [13]. Sugere-se a extração e a anotação, nos elementos *meanX*, do valor médio dos pixels de cada canal do respectivo modelo de cor *X*. Da mesma forma, sugere-se a extração e a anotação, nos elementos *stdX*, do desvio padrão dos pixels de cada canal do respectivo modelo de cor *X*. O elemento *mask* armazena o nome arquivo e o caminho para a imagem GT.

Construção do banco de imagens – Primeiramente, foram adquiridas 120 imagens sob diferentes condições de iluminação, de diferentes partes do corpo e tipos de pele. Há 30 imagens para cada condição de iluminação: *indoor / incandescent*, *indoor / fluorescent*, *outdoor / shadow* e *outdoor / sun*. As partes do corpo escolhidas foram mãos, costas e braços. Ainda, foram geradas manualmente as imagens GT para cada uma das 120 imagens originais. Todas as imagens estão armazenadas no formato JPEG.

Para a anotação das imagens foi desenvolvido uma interface gráfica denominada *skinChooser*. A Figura 2 ilustra a tela principal do aplicativo durante o processo de anotação de uma imagem do banco. Após carregar uma imagem, o usuário aplica os metadados desejados utilizando *combo boxes*, ou digitando-os quando não há esta opção. Os elementos *height* e *width* são obtidos de maneira automática. Para o caso da *roi*, o retângulo é especificado através do *mouse*, clicando e arrastando o ponteiro sobre a imagem. Após este procedimento os elementos *meanX* e *stdX* são extraídos automaticamente e o elemento *roi* recebe as coordenadas de *linha* e *coluna* do canto superior esquerdo e *linha* e *coluna* do canto inferior direito do retângulo. Para efetivamente escrever os metadados no arquivo XML o usuário deve clicar no botão *Salvar*, finalizando assim o processo de anotação de uma imagem.

Detecção de pele – O banco de imagens construído foi utilizado para a avaliação de dois métodos para detecção de pele. O primeiro, proposto em [2] é do tipo EDSR e utiliza o modelo de cores RGB. É implementado aplicando-se uma série de regras, ou condições, sobre os três canais de cor R, G e B para determinar se cada pixel é ou não um pixel de pele. Caso as condições da Equação 1 sejam satisfeitas, o pixel é considerado pele.

$$\begin{aligned}
 &R > 95 \text{ and } G > 40 \text{ and } B > 20 \text{ and} \\
 &\max\{R, G, B\} - \min\{R, G, B\} > 15 \\
 &\text{and } |R - G| > 15 \text{ and } R > G \text{ and} \\
 &R > B
 \end{aligned} \tag{1}$$

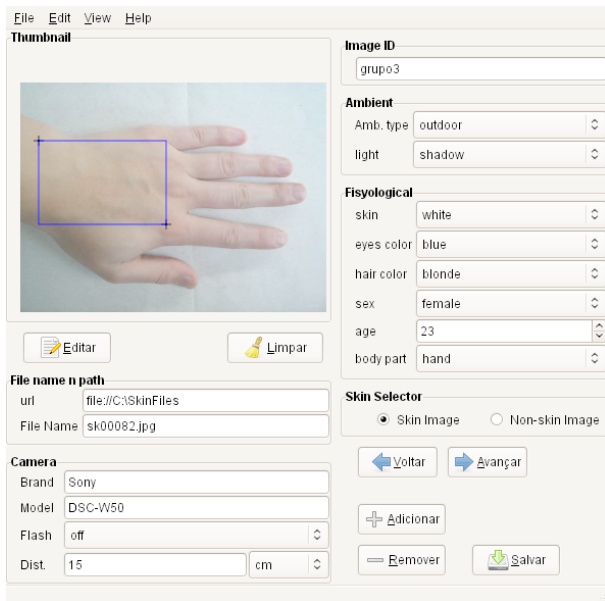


Figura 2: Aplicativo *skinChooser*, desenvolvido para a anotação das imagens contendo pele.

O segundo método avaliado, proposto em [9] é um classificador Bayesiano. Na etapa de treinamento são levantadas as probabilidades condicionais de cor de cada pixel c em relação ao critério pele, $P(c|pele)$, ou não-pele, $P(c|nãoPele)$. Para isto, são utilizadas as imagens GT e os resultados são armazenados em duas *lookup tables*: uma para pixels de pele e outra para pixels de não-pele. Na etapa de classificação, aplica-se a condição expressa na Equação 2, para um determinado valor de limiar Θ . Caso a condição da Equação 2 seja satisfeita, o pixel é considerado pele.

$$\Theta < \frac{P(c|pele)}{P(c|nãoPele)} \quad (2)$$

Resultados

Para validar a estrutura de banco de imagens proposta, preencheu-se o mesmo com as 120 imagens de pele adquiridas, seus respectivos metadados e GT.

A partir desta estrutura, os métodos de detecção de pele descritos anteriormente foram avaliados quantitativamente. Como métrica de desempenho, utiliza-se o histograma cumulativo de cada uma das taxas, TPR e FPR [14].

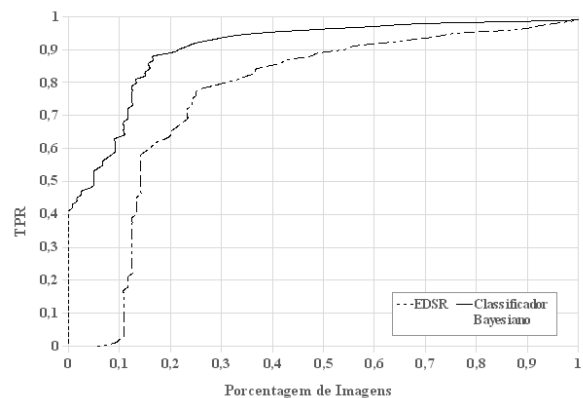
Neste método de avaliação, primeiramente são calculadas a TPR e a FPR para cada imagem, as quais são obtidas da a partir das seguintes relações: $TPR = TP/P$ e $FPR = FP/N$, onde TP é o número de verdadeiros positivos, FP o número de falsos positivos e P e N são os números totais de positivos e negativos, respectivamente, presentes na imagem GT.

Posteriormente, a TPR e a FPR das 120 imagens são normalizadas entre 0 e 1, e quantizadas com resolução de 1/100. Isto possibilita o levantamento de um histograma de 101 posições para cada uma das taxas. Assim,

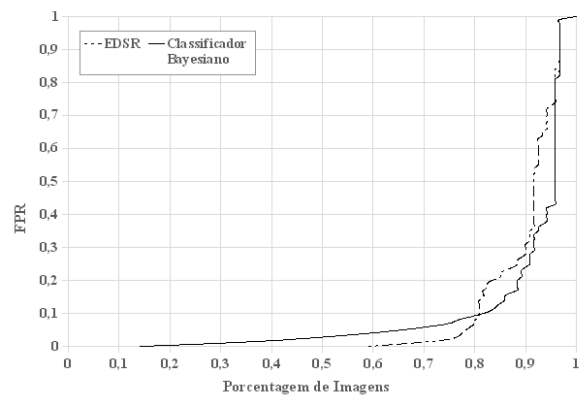
cada posição k do histograma contém a probabilidade do banco apresentar uma imagem com TPR ou FPR igual a k .

Finalmente, cada histograma é acumulado e plotado em função da *porcentagem de imagens*. Dessa forma, um ponto específico da curva (x , $TPR(x)$) indica que 100· x por cento das imagens apresentam TPR menor que $TPR(x)$. Consequentemente, 100·(1 - x) por cento das imagens apresentam TPR maior que $TPR(x)$. A mesma interpretação é válida para a curva (x , $FPR(x)$).

As curvas obtidas para os métodos de detecção de pele implementados são mostradas na Figura 3. As curvas da Figura 3(a) são denominadas *curvas de desempenho TPR* e, as da Figura 3(b), *curvas de desempenho FPR*.



(a)



(b)

Figura 3: Curvas de desempenho para os algoritmos de detecção de pele implementados: (a) TPR, (b) FPR.

Discussão

Com relação aos resultados da avaliação dos métodos para detecção de pele, pode-se constatar que, para as 120 imagens do banco, o classificador Bayesiano apresenta TPR superior à do EDSR. Entretanto, o critério FPR demonstra que os métodos comportam-se de forma parecida. Com isto, é possível afirmar que o classificador Bayesiano apresenta um desempenho superior ao do método EDSR, confirmando os resultados obtidos em outros estudos [1, 8].

Desta forma, demonstra-se que as imagens adquiridas e as máscaras GT geradas, dentro da estrutura de banco de dados proposta, podem ser utilizadas satisfatoriamente para a avaliação de métodos de detecção de pele.

Conclusão

Ao contrário de outros bancos de imagens dedicados ao desenvolvimento e avaliação de métodos de detecção de pele, o banco implementado neste trabalho inclui metadados que complementam as máscaras binárias e é estruturado. Além disso, o banco desenvolvido é implementado em XML, permitindo o uso das facilidades deste tipo de formato. Dentre elas, pode-se citar a fácil visualização, reestruturação (inclusão e remoção de elementos), além da facilidade de publicação na Internet. Há, ainda, linguagens para a busca em arquivos XML, como a Xpath e a XQuery, a partir das quais pode-se desenvolver ferramentas específicas para o acesso aos dados, permitindo pesquisas mais elaboradas.

O aplicativo *imageChooser* desenvolvido permite a anotação das imagens de maneira eficiente e ainda possui recursos que facilitam a visualização e manipulação dos dados. Exemplos deste recursos são a navegação através do banco, união de bancos (arquivos XML) e exclusão de imagens.

Na continuidade do trabalho, a partir da nova estrutura e da ferramenta para anotação propostas será realizada a expansão do banco de imagens disponível publicamente em www.lvc.citec.ct.utfpr.edu.br. Pretende-se ainda realizar a análise e a interpretação das estatísticas de cor extraídas das ROI das imagens, e o levantamento de outras aplicações para a base de dados.

Agradecimentos

O presente trabalho foi realizado com o apoio do UOL (www.uol.com.br), através do Programa UOL Bolsa Pesquisa, processos números 20080129153400 e 20080130200500.

Referências

- [1] Vezhnevets, V.; Sazonov, V. Andreeva, A. (2003), "A survey on pixel-based skin color detection techniques", In: *Proceedings of GRAPHICON 2003*, Ljubljana, Slovenia, 22-24 September, p. 85-92.
- [2] Kovač, J., Peer, P., Solina, F. (2003), "Human skin color clustering for face detection", In: *Proceedings of EUROCON 2003*, vol. 2, Moscow, Russia, 5-10 September, p. 144-148.
- [3] Fleck, M., Forsyth, D., Bregler C. (1996), "Finding naked people", In: *Proceedings of the Fourth European Conference on Computer Vision*, vol. 2, Cambridge, UK, 14-18 April, p. 593-602.
- [4] Ruiz-del-Solar, J., Castañeda, V., Verschae, R., Baeza-Yates, R., Ortiz, F. (2005), "Characterizing objectionable image content (pornography and nude images) of specific web segments: Chile as a case study", In: *Proceedings of the Third Latin American Web Congress*, Buenos Aires, Argentina, 31 October-2 November, p. 269-278.
- [5] Yang, M., Ahuja, N. (1998), "Detecting human faces in color images", In: *Proceedings of the International Workshop on Multi-Media Database Management Systems*, Dayton, USA, 5-7 August, p. 127-130.
- [6] Han, J., Award, G., Sutherland, A., Wu, H. (2006), "Automatic skin segmentation for gesture recognition combining region and support vector machine active learning", In: *Proceedings of the Seventh International Conference on Automatic Face and Gesture Recognition*, Dublin, Ireland, 10-12 April, p. 237-242.
- [7] Celebi, M., Aslandogan, Y., Bergstresser, P. (2005), "Unsupervised border detection of skin lesion images", *Proceedings of the 2005 International Conference on Information Technology: Coding and Computing*, vol. 2, Arlington, USA, 4-6 April, p. 123-128.
- [8] Phung, S., Bouzerdoum, A., Chai, D. (2005), "Skin segmentation using color pixel classification: Analysis and comparison", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, n. 1, p. 148-154.
- [9] Jones, M., Rehg J. (2002), "Statistical color models with application to skin detection", *International Journal of Computer Vision*, vol. 46, n. 1, p. 274-280.
- [10] Caetano, T., Olabarriaga, S., Barone, D. (2003), "Do mixture models in chromaticity space improve skin detection?", *Pattern Recognition*, vol. 36, n. 12, p. 3019-3021.
- [11] Schmugge, S., Jayaram, S., Shin, M., Tsap, L. (2007) "Objective evaluation of approaches of skin detection using ROC analysis", *Computer Vision and Image Understanding*, vol. 108, n. 1-2, p. 41-51 .
- [12] Störring, M., Andersen, H., Granum E. (1999), "Skin colour detection under changing lighting conditions.", *Proceedings of the Seventh Symposium on Intelligent Robotics Systems*, Coimbra, Portugal, 20-23 July, p. 187-195.
- [13] Poynton, C. (1995), "A guided tour of color space", In: *Proceedings of the SMPTE Advanced Television and Electronic Imaging Conference*, San Francisco, USA, February, p. 167-180.
- [14] Ge, F., Wang, S., Liu, T. (2006), "Image-segmentation evaluation from the perspective of salient object extraction", In: *Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Columbia, USA, 17-22 June, p. 1146-1153.