

ANÁLISE DE CARACTERÍSTICAS PARA DETECÇÃO DE PELE EM IMAGENS

Clayton Santos, Kely M. B. Pinto, Eulanda M. dos Santos, Eduardo Souto

Departamento de Ciência da Computação (DCC) - Universidade Federal do Amazonas (UFAM) – Manaus – AM - Brasil
{clayton.maia, kely, emsantos, esouto}@dcc.ufam.edu.br

Abstract – Skin detection is very often used in several applications involved on image and video processing, such as image categorization, face detection, gestures recognition, and nudity. This paper investigates the most frequently features applied in the literature dealing with automatic skin detection on images of nudity. These features are based on information related to color, shape and texture. Our goal in this paper is to conduct an experimental study in order to point out the best features for achieving the highest classification rates. The selection process is performed by measuring the impact of different combinations of features as well as by applying them individually. Focusing on improving skin tone detection and on controlling luminance information (brightness or intensity), we apply YCbCr color space during a segmentation stage. We also use Support Vector Machine (SVM) to classify the samples. Experimental results show that the combination of four features, more precisely, two color-based features and two texture-based features provided the best subset of features for automatic skin classification.

Keywords – images, skin detection, features analysis.

Resumo – Detecção de pele é um método comumente usado em diversas áreas de aplicação de processamento de vídeo e imagem, tais como: caracterização de imagens, detecção de faces, reconhecimento de gestos e de nudez. Este artigo investiga as principais características de cor, forma e textura empregadas em trabalhos que realizam detecção automática de pele em imagens de nudez. O objetivo do trabalho é realizar um estudo experimental a fim de selecionar as melhores características para o processo de classificação das amostras, sendo que a seleção é feita através da análise do impacto individual e de combinações das características. Para melhorar a identificação de tons de pele e controlar a luminosidade, o espaço de cores YCbCr é usado na segmentação das imagens, enquanto *Support Vector Machines* (SVM) é usado para classificação das amostras. Resultados experimentais mostram que a combinação de duas características de cor com duas características de textura forma o melhor subconjunto de características para a classificação automática de pele.

Palavras-chave – imagens, detecção de pele, análise de características

1 Introdução

A massificação de dispositivos pessoais como câmeras digitais, celulares e PDAs (*Personal Digital Assistants*) tem contribuído para o aumento de conteúdo multimídia na Internet. Uma consequência desse fenômeno é a proliferação de imagens consideradas ofensivas, como vídeos e imagens contendo nudez, que podem ser involuntariamente acessadas por crianças ou indevidamente armazenadas em instituições em que o acesso a este tipo de conteúdo não é autorizado [13].

Na literatura, diferentes métodos para filtragem de imagens ofensivas têm sido considerados. Em [7], Forsyth e Fleck apresentam um sistema para verificar se existem pessoas nuas em imagem digitais. Inicialmente, o sistema aplica uma estratégia de detecção de pele. Em seguida, um processamento geométrico é usado para agrupamento de pele em figura humana para a detecção do corpo humano. Wang *et al.* [19] apresentam um sistema de rastreamento de imagens ofensivas. O método proposto usa uma combinação de técnicas como filtros baseados no histograma de cores, textura e forma da imagem. Jones e Reh [11] descrevem um sistema baseado em um modelo estatístico apoiado na cor da pele e em um classificador do tipo rede neural. Wang e Yao [20] apresentam um algoritmo de reconhecimento de nudez a partir da localização do umbigo em imagens digitais. O umbigo é usado como ponto de referência para a construção de retângulos que são usados para computar a proporção de pele.

Um denominador comum a essas e outras abordagens é a utilização de um filtro de pele para detecção de nudez. Arentz e Olstad [1] destacam que o principal elemento para detectar nudez é a identificação de regiões de pele na imagem. Porém, o grande problema em detecção de pele é que diversos objetos do mundo real, tais como, alguns tipos de couro, areia, madeira, peles de animais e objetos de tons avermelhados, possuem cores aproximadas às tonalidades da pele e podem ser mal interpretados por detectores automáticos de pele [12]. Para melhorar a identificação de pele em imagens, a maioria das abordagens usa uma combinação de características extraídas da própria imagem como cor [18], textura [23], formas [17] e localização dos *pixels* de pele na imagem [21]. Grande parte dos trabalhos utiliza elevada quantidade de características e não apresenta uma análise qualitativa das características envolvidas na classificação, além de não oferecer informações importantes sobre condições de luminosidade, aquisição da base de imagens, quantidade e etnia dos indivíduos em estudo [12].

Diante desse contexto, este trabalho apresenta uma análise qualitativa de algumas das principais características envolvidas no processo de classificação de imagens de pele propostas na literatura. O objetivo do trabalho é avaliar a potencialidade e a relevância de cada característica no desempenho da classificação das amostras, com o propósito de reduzir a dimensionalidade do espaço e ao mesmo tempo, aumentar a taxa de classificação. Portanto, trata-se de um processo de seleção

de características do tipo *wrapper* [14], em que serão combinadas características de cor, forma e textura, todas empregadas após filtragem de pele. Vale ressaltar que as características investigadas aqui são as mais comuns nos trabalhos encontrados na literatura que aplicam filtro de pele. Por essa razão, são conhecidas como características locais [12] e há uma quantidade limitada das mesmas. Também é importante destacar que neste artigo, o filtro de pele é aplicado utilizando-se o espaço de cores *YCbCr*, afim de controlar a luminosidade e garantir a eficiência e eficácia na identificação de uma vasta gama de tons de cores de pele.

O restante deste artigo está organizado da seguinte forma. A Seção 2 fornece uma visão geral de alguns trabalhos em detecção de pele em imagens. As características investigadas são descritas na Seção 3, considerando-se a metodologia aplicada neste trabalho. Em seguida, a Seção 4 apresenta os experimentos e discute os resultados obtidos. Por fim, a Seção 5 apresenta as conclusões e trabalhos futuros.

2 Trabalhos Relacionados

Nesta seção são discutidos alguns modelos de detecção de pele em imagens utilizados como etapa inicial em aplicações de detecção de nudez. Vale ressaltar que as características investigadas neste artigo foram levantadas a partir dos trabalhos relacionados nesta seção. Outro fator importante é que os trabalhos aqui relacionados são aqueles que fornecem informações sobre a base de dados, adotam a utilização de alguma técnica de aprendizagem de máquina e utilizam ou não algum espaço de cor para reduzir os efeitos da luminosidade durante a aplicação do filtro de pele.

Zhu *et al.* [23] apresentam um modelo que combina espaço de cores como *YUV* e *YIQ* (*Y* representa a luminância e *UV-IQ* a cromaticidade) para controle da luminosidade e filtragem de pele para extração de informações referentes a cor. Em seguida, informações da textura são utilizadas através da matriz de co-ocorrência e da análise da estrutura geométrica dos seres humanos. Essa combinação de características foi usada pelo classificador SVM. Os autores usaram um conjunto de 400 imagens de nudez e 400 imagens diversas, sendo que foram obtidos 88,9% de sensibilidade e 84,6% de especificidade. A sensibilidade denota a razão entre o número de imagens de pele classificadas corretamente para o número total de imagens de pele, enquanto que a especificidade denota a razão entre o número de imagens que não contém pele classificadas corretamente para o número total de imagens dessa classe. Apesar da elevada taxa de classificação, não foi realizado um processo de análise das características e o modelo utiliza muitas informações baseadas na forma, que demandam alto esforço computacional.

Jiao *et al.* [10] propõem um modelo que utiliza informações de cor e bordas para criar um vetor de características para posterior classificação das imagens através do SVM. O método também usa os espaços de cores *YUV* e *YIQ*. O conjunto de teste é constituído por 1.200 imagens de pessoas e 1.200 imagens diversas, sendo que foram obtidos 89,3% de sensibilidade e 90,6% de especificidade. Porém, esse trabalho, apesar de utilizar informações de formas, analisa apenas imagens contendo nudez em que as pessoas estão em posição frontal, tornando o resultado pouco abrangente. Além disso, o impacto das características não foi investigado.

Duan *et al.* [6] apresentam um novo método de detecção de pele baseado em características da cor. Esse método primeiramente rotula os *pixels* referentes à pele como exemplos de treinamento para posteriormente, converter os valores de *pixels* de RGB para o espaço de cores *YUV* e *YIQ*. Assim, o modelo de distribuição de pele é obtido e o vetor de características criado para classificação com SVM. O conjunto de teste é composto por 312 imagens contendo nudez e 710 imagens que não contém nudez. Esse modelo obteve 80,7% de sensibilidade e 90% de especificidade. Porém, essa abordagem utiliza apenas duas características e a quantidade de amostras é pouco representativa.

Cusano *et al.* [5] descrevem um modelo que utiliza características de borda, textura e o espaço de cores *HSV* (*Hue*, *Saturation*, *Value*), criando um vetor para classificação usando CART (*Classification on Regression Trees*) e SVM. O processo experimental consiste de 1.500 imagens, sendo 750 pornográficas e as 750 restantes sem pornografia. Os resultados alcançaram 90,4% de sensibilidade e 88,4% de especificidade. Apesar desse trabalho utilizar uma grande combinação de características, não foi feita análise qualitativa para selecionar as melhores características.

Lee *et al.* [15] apresentam um algoritmo que utiliza um filtro de pele baseado no controle de luminosidade através do espaço de cor *YCbCr*. Inicialmente, são utilizadas características de cor e formas para detecção de pele, e posteriormente, ocorre um processo de detecção de faces para garantir a eficiência do modelo. Foram consideradas 508 imagens de nudez e 482 imagens que não contém nudez, sendo as imagens foram classificadas utilizando-se uma combinação de classificadores SVM e Redes Neurais (NN). Esse método obteve 86,4% de sensibilidade e 94,8% de especificidade. Entretanto, os autores também não analisaram o impacto das características no classificador.

A principal diferença entre o trabalho descrito neste artigo e as propostas mencionadas nesta seção é a análise qualitativa das características envolvidas no processo de detecção de pele. É de ampla aceitação na área de aprendizagem de máquina que a escolha das características mais relevantes determina o sucesso dos métodos de classificação. Portanto, neste artigo são analisadas 08 características: 02 características de cor (quantidade de *pixels* de pele conectados e proporção da região de pele para a imagem total [5,6,10]); 02 características de forma (área e perímetro [5,10,15,23]); e 04 medidas de textura baseadas na matriz de co-ocorrência (contraste, energia, homogeneidade e correlação [5,23]).

3. Metodologia

Nesta seção, são apresentados os detalhes das características investigadas neste artigo. Essa descrição é feita no contexto de uma metodologia que se baseia em uma arquitetura utilizada em grande parte de abordagens com filtro de pele [1,2,5,6,15,23], como mostrado na Figura 1. A arquitetura é composta por um componente responsável pela normalização das imagens obtidas na Internet; por um mecanismo de filtragem, que é responsável pela segmentação dos *pixels* da imagem; por um extrator de características, que é aplicado a partir do filtro de pele e cujo resultado é usado como parâmetro de entrada para o componente classificador, que por fim, classifica a imagem em pele ou neutra (não pele).

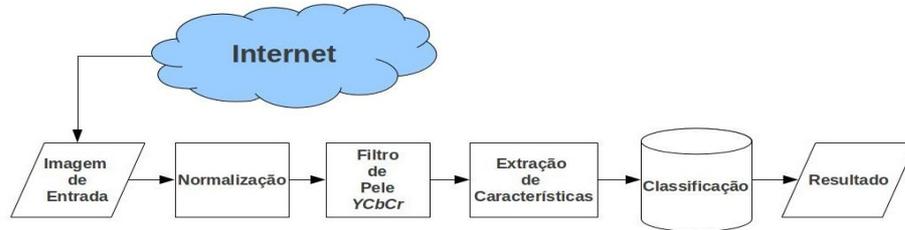


Figura 1 - Visão geral da metodologia utilizada.

3.1 Normalização

Na Internet as imagens estão dispostas em diversos formatos e resoluções, fato que dificulta o trabalho de análise de imagens e o desenvolvimento de um sistema automático de detecção de pele. Logo, é necessário normalizar as imagens, ou seja, padronizá-las em uma mesma dimensão e formato. O processo de normalização das imagens consiste em gerar um novo par de amostras para aplicar a um determinado contexto a ser explorado. Neste trabalho, foram aplicadas técnicas para redução e padronização do formato de entrada das imagens a fim de estabelecer um parâmetro para a geração de vetor de características a ser interpretado pelo classificador. As imagens foram padronizadas em 64 colunas por 42 linhas em formato JPEG.

3.2 Filtro de Pele

A segunda etapa da metodologia corresponde à segmentação das imagens em duas regiões: (1) região de pele; e (2) região de não pele. Segundo Zhu *et al.* [24] a característica mais importante em uma imagem está representada na cor, pois esta contém um alto nível de características para detecção e localização de objetos. Existem diversos espaços de cores, sendo que cada um contém peculiaridades e aplicações específicas. Neste trabalho, o processo de segmentação das imagens foi realizado através da conversão do espaço de cores de RGB para *YCbCr*, onde *Y* é a componente que traz informações de luminosidade da imagem e, *Cb* e *Cr* trazem informações da cromaticidade, obtidas através das equações 1 e 2.

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

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

A escolha do espaço de cor *YCbCr* deve-se à necessidade de tratamento da luminosidade em imagens. Como a pele possui a característica de absorver a luminosidade devido à melanina - substância que envolve e protege o cabelo e a pele humana, e é mais abundante em pessoas de raça negra [11] - existe a necessidade de se utilizar um espaço de cor que isole e trate corretamente a luminosidade em imagem [15]. Outra razão para utilização do espaço de cor *YCbCr* é o fato da pele humana conter uma gama restrita de valores de *pixels* no espaço de cor. Por fim, de acordo com Kelly *et al.* [12], o uso do espaço de cores *YCbCr* proporciona as melhores taxas de detecção de pele, além de reduzir o custo computacional.

Após a conversão das imagens do espaço RGB para o espaço *YCbCr*, as imagens foram pré-processadas através da técnica de processamento digital de imagens conhecida como análise de histogramas, manipulada na ferramenta Matlab [16]. Conforme Gonzalez e Woods [8], o histograma de uma imagem digital representa a função de distribuição de probabilidade (*FDP*) dos níveis de cinza presentes na imagem. Portanto, a *FDP* é construída através da divisão de todos os valores de frequência pelo número total de *pixels* da imagem.



Figura 2 - Exemplos de amostras de pele utilizadas na obtenção dos valores de cromaticidade.

A Figura 2 apresenta algumas amostras de zonas de pele, com dimensões de 30x30 *pixels* cada, extraídas de diversos pontos de faces de pessoas de várias etnias e sob diversas condições de luminosidade. Tais imagens foram convertidas do espaço de cor RGB para *YCbCr*. Em seguida, através da análise de histograma foram obtidos os valores médios de cromaticidade (componentes *Cb* e *Cr*) das imagens. Por fim, após esta análise observou-se que os valores de *pixels* referentes à pele variam em *Cb* de 77 a 127 e, em *Cr* de 133 a 173.

A partir da definição do intervalo de valores dos *pixels* de pele, as imagens foram segmentadas em *pixels* de pele e não pele. A Figura 3 mostra exemplos de imagens e seus correspondentes resultados após a aplicação do filtro de pele.

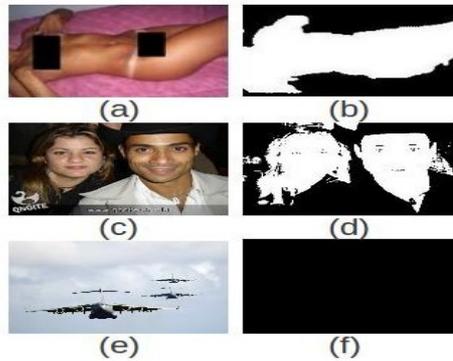


Figura 3 - Imagens que demonstram o resultado do filtro de pele, onde (a) representa uma imagem original de nudez, (b) o resultado do filtro de pele na imagem de nudez, (c) imagem de pessoas com roupa, (d) o resultado do filtro de pele para pessoas com roupa, (e) imagem original que não contém pele e, (f) o resultado do filtro na imagem que não contém pele.

3.3 Extração de Características

Na etapa de extração de características, as informações são extraídas a partir da segmentação da imagem gerada no filtro de pele. Conforme destacado na introdução, características relevantes são fundamentais para qualquer algoritmo de reconhecimento de padrões. As características devem realçar similaridades intra-classe e diferenças inter-classe.

Neste trabalho são analisadas 08 características, sendo 02 características de cor, 02 características de forma e 04 medidas de textura. As características de cor são as seguintes:

- Quantidade de *pixels* de pele conectados. Utilizando-se o espaço de cores *YCbCr*, a cor da pele independentemente de suas variações (branca, negra, amarela, etc), tende a possuir maior conectividade em relação às imagens que não contém pele [5,6,10,22].
- Proporção da área de pele com relação à área total da imagem. Essa métrica fornece informações sobre a proporção de segmentos de pele na imagem (P_s) e pode ser obtida pela seguinte fórmula: $P_s = N_s / N_t$, onde N_s é o número de *pixels* de pele e N_t é o número de *pixels* total da imagem [5,6,10,22].

As características de forma são extraídas a partir de informações da borda da imagem. Dado que $P(u,v)$ representa o valor atribuído ao *pixel* em uma determinada coordenada, as duas características de forma investigadas neste artigo são:

- Perímetro da borda. Em geral, em imagens que não contém pele, os *pixels* classificados como pele estão mais dispersos e possuem um perímetro, relativamente pequeno [5,10,15,23]. Através da seleção de uma região de interesse, esta característica serve para calcular a soma dos *pixels* da região que contém maior concentração de *pixels* de pele. É denotada da fórmula:

$$Per = \sum P(u,v) \quad (3)$$

A estimativa do perímetro é realizada através da aplicação de uma máscara para determinação da região de borda, atribuindo o valor 1 para *pixels* na horizontal e vertical, ou $\sqrt{2}$ para *pixels* na diagonal, percorrendo-se todos os *pixels* com vizinhança 8.

- Área de pele. Fornece informações sobre a área delimitadora da região com maior segmento de pele, onde em geral, as imagens que contém pele possuem uma área consideravelmente grande e, localização dos *pixels* mais ao centro [2,5,10,15,23]. É calculada através da fórmula abaixo,

$$A(L) = \sum_{u=1}^{u=m} \sum_{v=1}^{v=n} P(u,v) \quad (4)$$

onde L é o rótulo do objeto atribuído no processo de classificação, m e n representam o número total de linhas e colunas da matriz respectivamente.

Já as características de textura são extraídas a partir de uma *Gray-Level Co-occurrence Matrix* (GLCM), ou matriz de co-ocorrência. Uma GLCM representa a probabilidade de ocorrência de conjuntos de níveis de cinza em i e j dentro de uma relação espacial definida na imagem, sendo que i e j são os índices da linha e da coluna da matriz vizinhança. Essa relação espacial é definida em termos de uma distância d e um ângulo θ . Então, dado uma GLCM, alguns dados estatísticos podem ser extraídos, tais como: homogeneidade, correlação, contraste e energia. Considerando tais dados estatísticos, este trabalho encontrou a melhor distância para $d = 1$, calculada em 8 direções e $\theta = [0, 45, 90, 135, 180, 225, 270, 315]$. Logo o vetor de características de textura é composto por 32 componentes. As quatro métricas estão listadas abaixo, conforme [5,23]:

- **Contraste.** Retorna uma medida do contraste entre as intensidades de um *pixel* analisado e do *pixel* vizinho. A comparação é realizada em todos os *pixels* da imagem. Para uma imagem constante, o contraste é 0 (zero).

$$Con = \sum_{i,j} P|i-j|^2 p(i,j) \quad (5)$$

- **Correlação.** Retorna uma medida de quão correlacionado está um *pixel* com o seu vizinho. A comparação é realizada em todos os *pixels* da imagem. A correlação é 1 para uma imagem totalmente correlacionada.

$$Corr = \sum_{i,j} \frac{(i-\mu_i)(j-\mu_j)p(i,j)}{\sigma_i\sigma_j} \quad (6)$$

- **Energia.** Retorna a soma dos elementos elevados ao quadrado dentro da matriz de co-ocorrência de tons de cinza. A energia possui valor 1 para uma imagem constante.

$$Enrg = \sum_{i,j} p(i,j)^2 \quad (7)$$

- **Homogeneidade.** Retorna um valor que representa a proximidade da distribuição dos elementos em relação à diagonal da matriz de co-ocorrência dos tons de cinza.

$$Hmg = \sum_{i,j} \frac{p(i,j)}{1+(i-j)} \quad (8)$$

3.4 Classificação

Nesta etapa, a partir da leitura do vetor de características, o classificador tem por finalidade definir se cada imagem pertence à classe pele ou à classe não-pele. Para este fim, um conjunto de imagens de treinamento é usado para a construção de um modelo pelo classificador. Um conjunto de teste, totalmente independente do conjunto de treinamento, é então usado para medir o grau de generalização (desempenho) do classificador treinado.

Neste artigo, o classificador SVM é utilizado por apresentar resultados animadores em problemas de detecção de pele e nudez [12]. Os experimentos e os resultados obtidos são descritos na próxima seção. É importante mencionar que o processo de treinamento/teste do classificador foi realizado através do ambiente Weka [21].

4 Experimentos

Esta seção descreve a composição da base de dados, as métricas utilizadas para obtenção dos resultados e a análise do impacto das características investigadas.

4.1 Base de Dados

Para realizar os experimentos foi utilizada uma base de imagens composta por 2.100 imagens coletadas do repositório de dados da Caltech [4]. Foram selecionadas 550 imagens de amostras da classe pele (pessoas de várias etnias: caucasianos, africanos e asiáticos) e 1.550 imagens da classe neutra (ambientes diversos, carros, motos, aviões, etc). Conforme citado na seção 3.1 todas as imagens são coloridas e foram normalizadas, estando dispostas em formato JPEG.

4.2 Protocolo Experimental

Para análise de desempenho dos classificadores, a base de imagens foi dividida em base de treinamento e base de teste. A base de treinamento foi composta por 800 imagens, sendo 400 imagens de pele e 400 imagens neutras, totalizando 40% do conjunto inicial de imagens. A base de teste foi constituída por 1300 imagens, sendo 150 imagens de pele e 1150 imagens neutras, totalizando 60% do conjunto inicial de imagens.

Foram realizados alguns ajustes de parâmetros para analisar o desempenho dos classificadores durante a fase de treinamento. Para este fim, foi utilizada a estratégia de validação cruzada que atua da seguinte forma: o conjunto de treinamento é dividido em n subgrupos de igual tamanho. A cada iteração, $n-1$ conjuntos são combinados para compor o conjunto de treino, enquanto que o subgrupo restante é usado como base de validação. Esse processo é repetido n vezes. No

final, a média de classificação obtida nas n bases de validação é calculada, sendo que os parâmetros que apresentam os melhores resultados são usados para calcular a capacidade de generalização sobre a base de teste.

O SVM foi usado no processo de classificação. Os resultados da validação cruzada indicaram que o *kernel* RBF superou o *kernel* polinomial. Além do tipo de *kernel*, há necessidade de ajuste do parâmetro de penalização C e do parâmetro γ , definido para o *kernel* RBF. Em nossos experimentos, esses parâmetros foram definidos como $C = 6$ e $\gamma = 0,05$.

4.3 Métricas de Desempenho

Existem na literatura diversas métricas para avaliar o desempenho de um sistema de classificação. Este trabalho adota as métricas propostas por Wang *et al.* [19] que são utilizadas na maioria dos trabalhos descritos na Seção 2. Essas métricas são definidas abaixo:

1. **Especificidade:** definida como a razão entre o número de imagens neutras corretamente classificadas para o número total de imagens neutras.
2. **Sensitividade:** definida como a razão entre o número de imagens de pele classificadas corretamente para o número total de imagens de pele.
3. **Taxa de acertos:** calculada através da razão entre o conjunto total de imagens classificadas corretamente (pele e neutras) para o conjunto total de imagens.
4. **Taxa de erros:** calculada através da razão entre o conjunto total de imagens classificadas incorretamente (pele e neutras) para o conjunto total de imagens.

4.4 Resultados

A análise do impacto das características está dividida em duas partes. Na primeira, as características são avaliadas por grupo, isto é, características de cor, de forma e de textura são aplicadas separadamente. Na segunda parte, as características de todos os três grupos são avaliadas.

A tabela 1 mostra os resultados obtidos por SVM treinado com as duas características de cor: quantidade de *pixels* de pele conectados e proporção da região de pele para a imagem total. Foram obtidos 55,3% de sensitividade, sendo 83 imagens de pele classificadas corretamente e, 90,1% de especificidade, ou seja, 1.043 imagens neutras classificadas corretamente. A taxa de acertos foi de 86,7%, totalizando 1.126 imagens classificadas corretamente e erro de 13,3%, com 174 imagens classificadas incorretamente. Analisando o conjunto de características de cor isoladamente, foi constatado que ambas são totalmente dependentes, pois a partir da quantidade de *pixels* de pele conectados é calculada a porcentagem de pele na imagem. Então, usando apenas uma característica isoladamente, o classificador não aprende e tende a classificar as amostras como pertencentes à classe neutra. Os resultados das avaliações isoladas de cada característica e da combinação (quantidade e porcentagem de pele juntas) são exibidos abaixo.

Tabela 1- Matriz de confusão e análise de características de cor a partir das taxas obtidas pelo classificador SVM

	Quantidade de Pele		Porcentagem de pele		Combinação	
	Pele	Neutras	Pele	Neutras	Pele	Neutras
Pele	37,2%	62,8%	29,4%	70,6%	55,3%	44,7%
Neutras	18,6%	81,4%	16,3	83,7%	9,9%	90,1%

Quanto às características de forma, os resultados indicam que as duas características investigadas neste artigo não são relevantes para treinar SVM para separar corretamente as duas classes do problema. Com essas características, tanto isoladamente quanto combinadas, SVM classificou a maioria das amostras como pertencentes à classe não pele. A Tabela 2 exibe resultados do impacto de cada característica de formas no desempenho do classificador.

Tabela 2- Matriz de confusão e análise de características de formas a partir das taxas obtidas pelo classificador SVM

	Perímetro		Área		Todas	
	Pele	Neutras	Pele	Neutras	Pele	Neutras
Pele	0%	100%	0%	100%	0%	100%
Neutras	0%	100%	0%	100%	0%	100%

Ao analisar o conjunto de características de textura, diversas combinações foram realizadas a fim de encontrar o conjunto mais relevante, e foi verificado que a combinação de homogeneidade e correlação produziu os melhores resultados.

Essa combinação obteve 92% de sensibilidade, sendo 138 imagens de pele classificadas corretamente e 99,7% de especificidade, com 1.147 imagens neutras classificadas corretamente. A taxa de acertos foi de 98,8%, totalizando 1.285 imagens classificadas corretamente e erro de 1,2% com 15 imagens classificadas erroneamente. A tabela 3 apresenta os resultados da análise das características de textura. O uso isolado das características de contraste e energia produziu os mesmos problemas apresentados na análise das características de formas.

Tabela 3 - Matriz de confusão obtida classificador SVM e impacto das características de textura.

	Homogeneidade		Correlação		Contraste		Energia		Todas		Homogeneidade + Correlação	
	Pele	Neutras	Pele	Neutras	Pele	Neutras	Pele	Neutras	Pele	Neutras	Pele	Neutras
Pele	88%	12%	82,7%	17,3	0%	100%	0%	100%	92,6%	7,4%	92%	8%
Neutras	0,8%	99,2%	1,4%	98,6%	0%	100%	0%	100%	0,7%	99,3%	0,3%	99,7%

Por fim, todas as características de cor, forma e textura foram combinadas. A Tabela 4 apresenta os resultados da análise de combinação de todas as 8 características e a melhor combinação resultante. As melhores características resultantes das análises realizadas para cada grupo separadamente, foram combinadas para verificação do desempenho do classificador. Os resultados mostram que a combinação das duas características de cor, mais as duas melhores características de textura, ou seja, homogeneidade e correlação, levou SVM a obter 94,7% de sensibilidade, totalizando 142 imagens de pele classificadas corretamente e 99,6% de especificidade, com 1.145 imagens neutras classificadas corretamente. A taxa de acertos foi de 99%, sendo 1.287 imagens classificadas corretamente e erro de 1%, com 13 imagens classificadas erroneamente.

Tabela 4 - Matriz de confusão obtida pelo classificador SVM e análise de todas as características na base de teste

	Todas as 8 características		Melhor Combinação	
	Pele	Neutras	Pele	Neutras
Pele	(94%) 141	(6%) 9	(94,7%) 142	(5,3%) 8
Neutras	(0,6%) 7	(99,4%) 1143	(0,4%) 5	(99,6%) 1145

Vale ressaltar que, apesar de alguns trabalhos disponibilizarem informações importantes sobre a composição da base de dados, como etnia, quantidade de amostras e condições de luminosidade, um processo de comparação justa e direta é inviável, devido à inexistência de bases de dados de *benchmarking*.

Um ponto importante a ser destacado é que as imagens neutras que foram classificadas como pele apresentavam muitos *pixels* de cor amarela e vermelha. Uma análise detalhada revelou que a maioria dessas imagens correspondia a imagens que contém areia ou alguma formação rochosa. Os resultados também confirmam que o descarte de características irrelevantes melhora o desempenho de classificação das amostras e, conseqüentemente, o custo para processar e extrair as evidências. A combinação das características mais relevantes, identificadas neste trabalho, pode ser adotada por diversos trabalhos futuros.

5 Conclusões e Trabalhos Futuros

Este trabalho apresentou uma análise de 08 tipos de características para a classificação automática de imagens de pele, sendo 02 características de forma, 02 características de cor e 04 características de textura. O impacto de cada característica foi medido em função da taxa de classificação do SVM.

A principal contribuição do artigo é a análise de evidências relevantes, com o objetivo de reduzir a dimensionalidade do espaço de características, diminuindo o esforço computacional, e aumentar o desempenho da classificação das amostras.

Os resultados mostraram que a combinação de duas características de cor, isto é, quantidade de *pixels* de pele conectados e proporção da área de pele com relação à área total da imagem, juntamente com duas características de textura, homogeneidade e correlação, formam o conjunto de características mais relevantes ao problema investigado. As características de forma, porém, apresentaram os piores desempenhos.

Como trabalho futuro, pretende-se avaliar características que informam posição e concentração dos *pixels* de pele através do particionamento da imagem e combiná-las com descritores globais, para complementar o módulo de detecção de nudez.

6 Referências

- [1] Arentz, W. A.; and Olstad, B. "Classifying offensive sites based on image content". Computer Vision and Image Understanding 94. (2004), 295-310.

- [2] Bosson, G.; Cawley, C.; Chan, Y.; and Harvey, R. "Non-Retrieval: blocking pornographic images", ACM CIVR, Vol. 2383, (2002), 60-69.
- [3] Burges, C. J. C. "A tutorial on Support Vector Machines for Pattern Recognition". Data Mining and Knowledge Discovery, Vol. 2, (1998), 121-167.
- [4] Caltech. Computational Vision Laboratory. California Institute of Technology (2005). Disponível em: <http://www.vision.caltech.edu/html-files/archive.html> . Acesso em Abril de 2010.
- [5] Cusano, C.; Brambilla, C.; Schettini, R.; and Ciocca, G. "On the Detection of pornographic digital images", VCIP 2003, 2105-2113.
- [6] Duan, L.; Cui, C.; Gao, W.; and Zhang, H. "Adult Image Detection Method Base-on Skin Color Model And Support Vector Machine". In: ACCV2002, Melbourne, Australia, (2002), 22-25.
- [7] Fleck, M., Forsyth, D.A. and Bregler, C. "Finding naked people", Proc. 4th European Conf. On Computer Vision, vol. 2, (1996), 593-602.
- [8] Gonzalez, R.C.; and Woods, R. E. "Digital Imagem Processing", 3rd Ed., Addison-Wesley Publishing Company, (2008).
- [9] Imagawa, K., Lu, S., Igi, S.: Color-based hands tracking system for sign language recognition. Proc. 3rd. International Conference on Face & Gesture Recognition, Washington, DC, USA, IEEE Computer Society, 462, (1998).
- [10] Jiao, F.; Gao, W.; Duan, L.; and Cui, G. "Detecting Adult Image using Multiple Features", Proc. International Conference Info-tech and Info-net (ICII 2001), 378-383.
- [11] Jones, M. J.; and Rehg, J. M. "Statistical Color Models with Application to Skin Detection". Computer Vision, Vol. 46, (2002), 81-86.
- [12] Kelly, W.; Donnellan, A.; and Molloy, D. "A Review of Skin Detection for Objectonable Images". In Proc. ITT 2007, (2007), 310-319.
- [13] Kherfi, M. L.; Ziou, D.; and Bernardi, A. "Image Retrieval from the world wide web: Issues, techniques, and systems". ACM Computing Survey, Vol. 36, (2004) 35-67.
- [14] Kohavi, R.; and John, G. H. "Wrappers for feature subset selection". Artificial. Intelligence, Vol. 97, (1997), 273-324.
- [15] Lee, J. S.; Kuo, Y. M.; Chung, P.C.; and Chen, E. L. "Naked image detection based on adaptive and extensible skin colour model", *Pattern Recognition*, Vol. 40, (2006), 2261-2270.
- [16] Matlab. "The Language Of Technical Computing" The Mathworks (2009). Disponível em: <http://www.mathworks.com>. Acesso em: Junho de 2010.
- [17] Safar, M.; Shahabi, C.; and Sun, X. "Image retrieval by shape: a comparative study" In: IEEE International Conference on Multimedia and Expo. (2000), 141-144.]
- [18] Yogarajah, P.; Condell, J.; Curran, K.; Cheddad, A.; and McKevitt, P. "A Dinamic Threshold Approach for Skin Segmentation in Color Images". In Proc. IEEE 17nd International Conference on Image Processing (ICIP 2010). Hong Kong, China. (2010), 2225-2228.
- [19] Wang, J.; Li, J.; Wiederhold; and Firschein, G. "System for Screening objectonable Images" Computer Communications. (1998), 20-30.
- [20] Wang, X.; Hu, C.; and Yao, S. "An Adult Image Recognizing Algorithm on Naked Body Detection" on IEEE International Colloquium on Computing Communication, Control and Management. (2009), 197-200.
- [21] Weka. (2010) "Data Mining with Open Source Machine Learnig Software in Java". University of Waikato. Disponível em: <http://www.cs.waikato.ac.nz/~ml/weka>. Acesso em Outubro de 2010.
- [22] Zheng, Z.; Daoudi, M.; and Jedynek, B. "Blocking Adult Images Based on Statistical Skin Detection". Electronic Letters on Computer Vision and Image Analysis, Vol. 4, (2004), 1-14.
- [23] Zhu, H.; Zhou, S.; Wang, J.; and Yin, Z. "An Algorithm of Pornographic Image Detection". In Fourth International Conference on Image and Graphics (2007), 801-804.
- [24] Zhu, Q.; Wu, C-T.; Cheng, K-T.; and Wu, Y-L. "An adaptive skin model and its application to objectionable image filtering", ACM Multimedia, (2004), 56-63.