

Validação Facial com Histograma de Gradiente Orientado

Lucas dos Santos Silva¹, Yago Pacheco Teixeira¹, D.Sc. Adelson Siqueira Carvalho¹

¹Instituto Federal Fluminense campus Campos-Centro
Rua Dr.Siqueira, 273. Parque Dom Bosco. CEP: 28030-130, Campos dos
Goytacazes/RJ

{[lucas.santos_s](mailto:lucas.santos_s@hotmail.com), [yago_ocdp](mailto:yago_ocdp@hotmail.com)}@hotmail.com, acarval@iff.edu.br

Abstract. *The facial identification has always been a major challenge for computer vision. The main objective of this work was to create an automated system capable of recognizing a face, and through a classifier previously trained and adjusted with a face database to validate the face identifying, the classifier returns if it belongs to the database or not. The feature chosen for the extraction and classification was the Histogram of Orientated Gradient that is usually used for face detection, but it is not used for facial identification. This work sought to verify the applicability of this feature for the identification of registered people in a face database.*

Resumo. *A identificação facial sempre foi um grande desafio para a visão computacional. O objetivo principal deste trabalho foi criar um sistema automatizado capaz de reconhecer uma face, e através de um classificador previamente treinado e ajustado com um banco de fotos, validar a face identificando-a como sendo pertencente ao banco de fotos ou não. A característica escolhida para a extração e classificação foi o Histograma de Gradiente Orientado, que é utilizada para reconhecimento facial, mas pouco utilizada para identificação facial. Este trabalho buscou verificar a aplicabilidade deste tipo de característica para a identificação de pessoas cadastradas em um banco de fotos.*

1. Introdução

Com o desenvolvimento tecnológico cada vez mais avançado, a busca da realização de atividades de maior complexidade é cada vez mais presente, como por exemplo a identificação e validação de rostos humanos, e identificação de padrões para tomadas de decisões. Das diversas funções existentes no corpo humano, dar a uma máquina a capacidade visual de reconhecimento e classificação de objetos e seres é um desafio a ser vencido. Este artigo tem o objetivo de desenvolver um algoritmo capaz de reconhecer rostos provenientes de uma imagem ou vídeo como entrada utilizando a extração de característica chamada Histograma de Gradiente Orientado(HOG).

2. Metodologia

É possível dividir este trabalho em três partes principais: Detecção facial, extração de características e classificação. Cada uma com o seu conteúdo metodológico específico. A figura 1 mostra a relação entre elas.

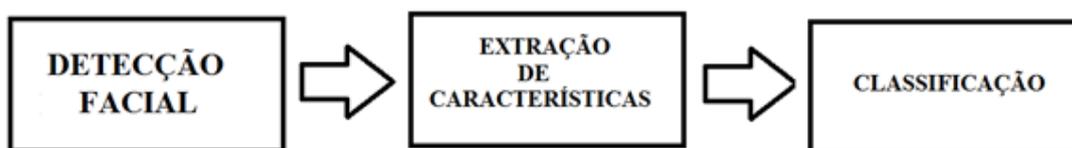


Figura 1. Divisão do trabalho

Fonte: Elaboração própria.

2.1. Detecção Facial

Em 2001, Paul Viola e Michael Jones, criaram um algoritmo de identificação de imagem chamado algoritmo de Viola-Jones. Esse algoritmo é capaz identificar faces usando os padrões que existem no rosto humana, ele também pode identificar qualquer outro objeto, desde que seja treinado (VIOLA; JONES, 2001). Neste trabalho o algoritmo de Viola-Jones já implementado no MATLAB® será utilizado para detecção das faces humanas tanto para cadastro de banco de fotos quanto para a classificação. O algoritmo classifica objetos baseado no valor de características, e faz isso no lugar de usar os pixels diretamente pelo motivo de sistemas baseados em características operam muito mais rápidos que sistemas baseados nos pixels em si (SANTOS, 2011). As características de Haar-like são máscaras retangulares com áreas brancas e pretas com mesma área. A soma dos pixels do lado branco menos os do lado preto representam um a uma variação luminosa que pode ser vista na imagem principal, chamando assim de característica Haar (VIOLA; JONES, 2001). Inicialmente havia quatro tipos de retângulos no algoritmo de Viola-Jones (Figura 2), mas de acordo com as melhorias verificadas por R. Lienhart e J. Maydt esse número aumentou, como (LIENHART; MAYDT, 2002) .

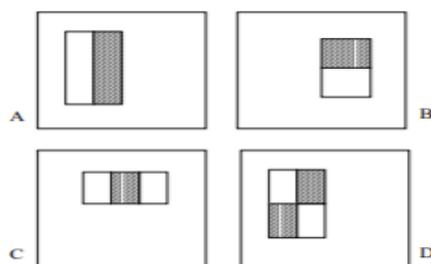


Figura 2. Haar-like

Fonte: VIOLA, JONES, 2001

Uma vez identificadas as características Haar, é possível ver certa simetria entre as mudanças de luminosidade no rosto humano. Por exemplo, as sobrancelhas que são

sempre mais escuras que a região da testa, e o nariz possui uma região mais clara que a dos olhos pois são mais fundos. Usando esse compilado de características o algoritmo de Viola-Jones é capaz de realizar a identificação de faces, após passar por um treinamento, onde vários classificadores são postos em cascata, seguindo uma ordem do mais fraco para o mais forte (LIENHART; MAYDT, 2002). A característica Haar-like extraída atravessa o classificador mais fraco, se não for aprovada pelo classificador como uma característica de uma face ela é descartada, se for aprovada, ela vai para o próximo classificador que possui um critério maior e mais complexo (VIOLA; JONES, 2001). Esse método busca descartar as características irrelevantes rapidamente e encontrar padrões na imagem. O classificador AdaBoost, presente no MATLAB®, é conhecido uma das melhores opções para a escolha das características e treinamento dos classificadores (BRAGA, 2013).



Figura 3. Classificador em Cascata

Fonte: Elaboração própria.

2.2. Extração de Características

O descritor baseado em Histograma de Gradientes Orientados foi proposto por Dalal e Triggs, e é baseado na avaliação de histogramas locais normalizados dos gradientes orientados da imagem (DALAL; TRIGGS, 2005). O vetor de características é formado pela concatenação dos histogramas normalizados dos gradientes orientados presentes em uma janela de detecção.

Para o treinamento do classificador, as imagens devem ser cortadas em um tamanho padrão, quando o HOG de um rosto é extraído, as características da imagem são adquiridas de forma que mesmo sob diferentes condições, como a luminosidade por exemplo, o descritor de imagem mantenha-se o mesmo, ou o mais próximo possível (MCCORMICK, 2013). A extração do HOG de uma imagem começa com a repartição da imagem em pequenas regiões espaciais denominadas como células que podem ser ajustadas de acordo com a necessidade (DALAL; TRIGGS, 2005). Neste trabalho foram usadas inicialmente células com 8 pixels de largura e 8 de comprimento, 64 pixels por célula. A imagem de teste escolhida para exemplificar o processo foi uma pertencente ao banco de fotos produzido especificamente para este trabalho e possui resolução de 80x80 pixels.

A Figura 3 mostra um pixel cujo o valor a esquerda é 56 e a direita 94. Para o cálculo do gradiente (Equação 1), o valor a direita subtrai o valor a esquerda ($94 - 56$), então a taxa de mudança no eixo x neste caso será 38. A conta do gradiente pode ser feita da

direita para a esquerda ou ao contrário; o importante é permanecer com o padrão escolhido por toda a imagem.

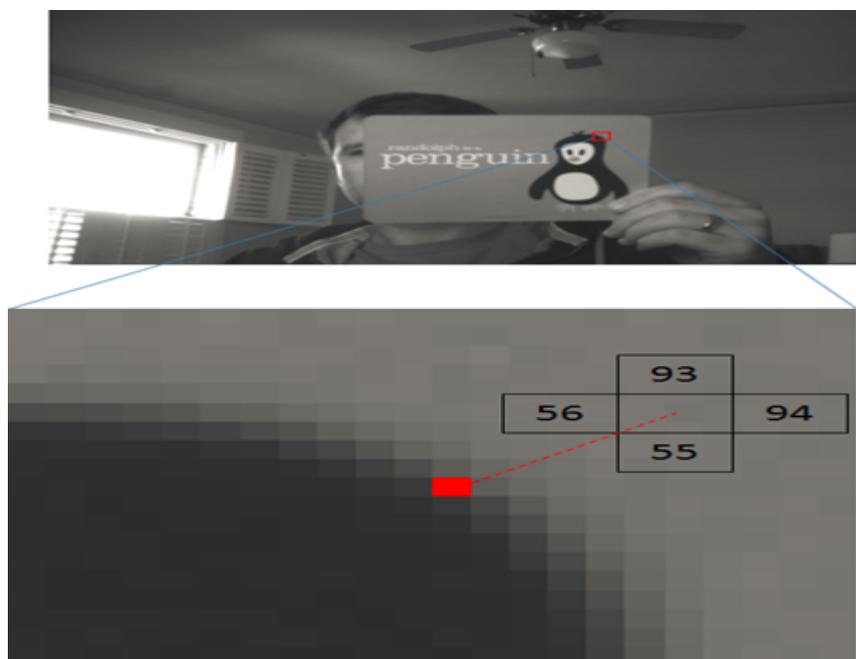


Figura 4. Taxa de mudança

Fonte: MCCORMICK, 2013.

Calculando assim a taxa de variação nos eixos x e y para poder encontrar os valores do gradiente (ângulo e magnitude).

$$\hat{\text{Ângulo}} = \arctan\left(\frac{\text{eixo } y}{\text{eixo } x}\right)$$

$$\hat{\text{Ângulo}} = \arctan\left(\frac{38}{38}\right) = 0,785 \text{ rads} = 45 \text{ degrees.}$$

$$\text{Magnitude} = \sqrt{(\text{eixox})^2 + (\text{eixoy})^2} = 53,74.$$

$$\text{Magnitude} = \sqrt{(38)^2 + (38)^2} = 53,74.$$

Adquirindo esses dois valores um histograma com a relação entre as magnitudes e os ângulos em cada célula é formado. Ainda com base na pesquisa de McCormick, é possível executar este mesmo cálculo sobre esta imagem sob diferentes valores de iluminação, e obter o valor do vetor gradiente igual na maioria das vezes (MCCORMICK, 2013). Os valores dos pixels mudam, mas as taxas de mudanças na direção dos eixos x e y permanece constante. Entretanto, neste projeto é preciso que haja robustez e confiabilidade na resposta do algoritmo, o que torna importante garantir a identificação

do usuário mesmo sob mudanças na iluminação e brilho. Como estratégia para melhorar a robustez do HOG, propõe-se que o valor de magnitude seja utilizado para dividir cada respectivo vetor (MCCORMICK, 2013). Como resultado, o valor do gradiente de cada pixel, torna-se invariante mesmo sobre mudanças de contraste na imagem. No MATLAB® a função de extração de HOG já está implementada como `extractHOGFeatures` no sistema de visão computacional, retornando as características HOG de uma imagem de entrada.

2.3. Classificação

O classificador Support Vector Machine Coarse Gaussian, trata-se de uma das variações de SVM utilizando a técnica de Kernel. Essas variações estão listadas no MATLAB®. A função gaussiana apresenta distinções mais grosseiras entre as classes onde kernel está definida como a raiz quadrada do número de preditores vezes quatro (MATHWORKS, 2015).

3. Escolha do Classificador

Utilizou-se a aplicação Classification Learner do MATLAB® para se testar o comportamento de diversos classificadores quando treinados com as características extraídas do Histograma de Gradiente Orientado do banco de fotos. Foi escolhida a validação por holdout de 20% (retirada de 20% da massa de dados para validação após o treinamento), por se tratar de um grande número de amostras. Foram testados os seguintes classificadores: Complex Tree, Medium Tree, Simple Tree, Linear Discriminant, Quadratic Discriminant, Linear SVM, Quadratic SVM, Cubic SVM, Fine Gaussian SVM, Medium Gaussian SVM, Coarse Gaussian SVM, Fine KNN, Medium KNN, Coarse KNN, Cosine KNN, Cubic KNN, Weighted KNN, Ensemble Boosted Trees, Ensemble Bagged Trees, Ensemble Subspace Discriminant, Ensemble Subspace KNN e Ensemble RUSBoosted Trees.

Muitos classificadores surpreendentemente retornaram uma acurácia considerada alta (acima de 85%), então todos eles foram testados individualmente. Inicialmente executou-se o algoritmo com banco de fotos AT&T para os classificadores que obtiveram maior porcentagem de acurácia anteriormente. Então foram usados rostos de pessoas não cadastradas no banco de fotos utilizado para treinamento dos classificadores para verificar se esses classificadores não apresentariam falsos positivos, ou seja, com a entrada de novas faces não cadastradas apresentar como pertencente a uma classe cadastrada. Para os classificadores treinados com o banco de fotos AT&T, os melhores resultados para esse teste foram o Ensemble KNN, Ensemble Subspace Discriminant, O Ensemble Bagged Trees e o Weighted KNN. Os classificadores baseados em SVM também retornaram um bom resultado, mas para esse teste a variação de resposta para falsos-positivos, e positivos foi muito pequena, dificultando a análise. Decidiu-se então testar os classificadores no banco de fotos produzido pelos autores deste trabalho. Surpreendentemente o que se notou foi a falta de confiabilidade dos classificadores com melhores resultados anteriores quando se tratava de uma entrada não previamente cadastrada, logo se resolveu testar todos os classificadores novamente. Com o teste dos classificadores no novo banco de fotos criado, os melhores resultados para entrada de pessoas não cadastradas foram dos classificadores SVM's, principalmente o Linear, o Medium Gaussian e o Coarse. Estes além de apresentarem ótima resposta a pessoas

previamente cadastradas, conseguiram negar entradas não cadastradas no banco de fotos. Após uma nova análise com os três classificadores melhores destacados até este ponto notou-se que o que apresentava maior confiabilidade com a identificação de pessoas cadastradas e a negação de pessoas não cadastradas foi o Coarse Gaussian SVM.

4. Banco de Fotos

Foi utilizado em banco de fotos padrão usado em trabalhos de reconhecimento de faces que é o AT&T Face Database, pertencente ao AT&T Laboratories Cambridge. Também foram desenvolvidos outro banco de fotos criado pelos próprios autores deste artigo. Com fotos de 15 pessoas diferentes, tendo 200 fotos cada uma em posições e com expressões diferentes em 80x80 de resolução.

5. Testes e Resultados

O detector facial por Viola-Jones foi bastante eficaz, conseguindo reconhecer pessoas em diferentes ambientes e condições, possuindo uma liberdade de 15 a 20 graus de inclinação no rosto das pessoas tanto em fotos quanto em vídeos.

Com o banco de fotos padrão AT&T foi obtido uma acurácia de 96.2% utilizando o Ensemble KNN Subspace. No teste de validação este classificador obteve uma acurácia de 93.75%.A resposta a novas entradas não cadastradas previamente foi de aproximadamente 90%.

Um trabalho correlacionado com o deste artigo é o trabalho de conclusão de curso do Luiz Filipe Zenicola Braga(BRAGA, 2013), que utiliza autovalores para a validação facial. Em um dos seus testes, ele utilizou o mesmo banco de fotos padrão AT&T, e no melhor resultado obteve uma acurácia de 78%, com 5 falsos positivos e 39 fotos negadas pelo critério utilizado.

Outro trabalho utilizando o mesmo banco de fotos padrão é o trabalho de conclusão de curso do Luis Fernando Martins Carlos Junior que visou validar o uso de redes neurais aplicadas no problema do reconhecimento facial , que obteve uma acurácia de 75% em seu melhor teste(CARLOS, 2011).

Para o banco de fotos criado para este trabalho a acurácia na validação foi de 95.83% em média.

O sistema automatizado foi primeiramente testado no mesmo local onde a maioria das fotos do banco de fotos foram cadastradas. Com esse ambiente controlado (câmera, distância e iluminação), a acurácia de identificação foi de 90%. A acurácia de negação de pessoas não cadastradas foi de 84%. Outros testes foram realizados em lugares diferentes do lugar de cadastro, com ambiente controlado, e a acurácia ficou entre 75 a 85%, dependendo da pessoa testada,e conseguiu negar pessoas não cadastradas. Os últimos testes foram realizados em ambientes não controlados (variando bruscamente a iluminação e principalmente a angulação da câmera), e mostraram que o sistema não possui robustez suficiente para essa aplicação, pois retornou pessoas erradas e gerou

falsos-positivos com pessoas não cadastradas no banco. A figura 5 mostra a resposta de saída do sistema automatizado.

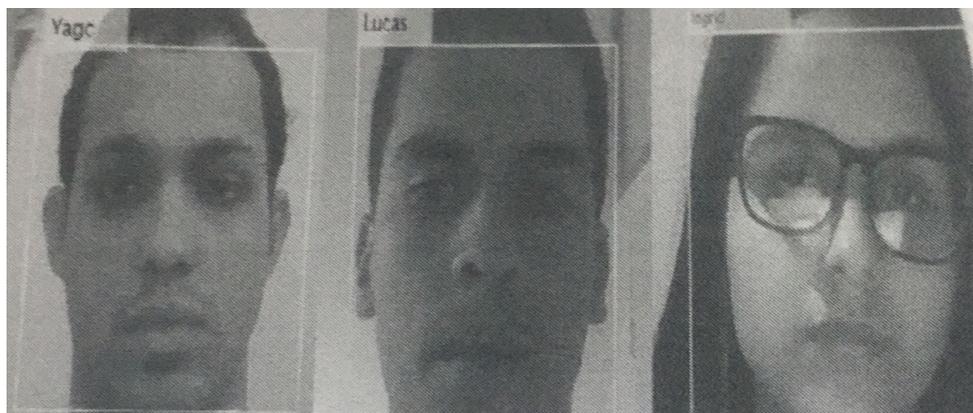


Figura 5. Identificação

Fonte: Elaboração própria.

6. Conclusão

A parte de reconhecimento de um rosto humano foi feita utilizando a função implementada do algoritmo de Viola-Jones no MATLAB e obteve um ótimo resultado encontrando faces em ambientes variados.

Na parte referente a extração de características foi utilizada a função `extractHOGFeatures` já também implementada no MATLAB. As configurações de extração foram mantidas as padrões, sendo o número de orientações igual a 9, e o tamanho das células 8x8 pixels. O vetor concatenado de características neste caso possui um tamanho de 2916 índices para fotos padronizadas utilizadas neste trabalho de 80x80 pixels.

O classificador que obteve melhor resposta tanto para identificar pessoas, quanto para negar pessoas não cadastradas no banco, e uma robustez para variação de ambientes e iluminação foi o SVM Gaussian Coarse que apresentou uma taxa de acurácia de 95.83% em média para o banco padrão criado para este trabalho.

Mesmo com esses bons resultados, o classificador não é robusto o suficiente para variações bruscas no ambiente, como iluminação e principalmente a variação do ângulo da câmera em relação ao ângulo no momento do cadastro. Isso pode ser proveniente de um cadastro sem variações de ambiente e ângulo da câmera.

O sistema automatizado criado é uma união do reconhecimento do rosto humano pelo Viola-Jones com a classificação de uma extração de HOG deste rosto reconhecido através de um classificador treinado com os vetores com índices extraídos do banco de fotos. O sistema possui velocidade de identificação, conseguindo fazer todas as etapas do

processo em uma média de 0.1 segundo em um ultrabook da LENOVO, Intel Core i7, 8GB de memória RAM , e uma placa de vídeo NVIDIA GeForce GT 730M.

Diante destes fatos, pode-se concluir que o Histograma de Gradiente Orientado pode ser utilizado para identificação de pessoas pela face através de uma câmera com boa robustez em relação a mudanças não bruscas de ambiente e iluminação.

Para trabalhos futuros se sugere a criação de um banco de fotos mais padronizado, com variação de ambientes e iluminação, e também variação no ângulo da câmera; Uma interface amigável para melhor utilização dos códigos; Implementação do HOG com outras técnicas utilizadas para reconhecimento e identificação facial como LBP(Local Binary Patterns) e Autovalores; Utilizar técnicas de redução de dimensionalidade do vetor de características como o PCA e o LDA, desta forma poderia-se melhorar as configurações de extração como o número de orientações do gradiente, melhorando assim a classificação.

7. Referências

BRAGA, L. F. Z. (2013). Sistemas de reconhecimento facial. *Escola de Engenharia de São Carlos* p. 84 , 2013.

CAMBRIDGE, AT&T Laboratories (2002). *The Database of Faces*. [S.l.]: Cambridge University Computer Laboratory. Disponível em: <<http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>>.

CARLOS, L. F. M. J. (2011)Reconhecimento Facial Utilizando Redes. *FUNDAÇÃO DE ENSINO “EURÍPIDES SOARES DA ROCHA” CENTRO UNIVERSITÁRIO EURÍPIDES DE MARÍLIA*.

CHAPELLE, O; HAFFNER, P; VAPNIK, V N. (1999). Support vector machines for histogram-based image classification. *Neural Networks, IEEE Transactions on* v. 10, n. 5, p. 1055–1064, 1999.1045-9227.

CORTES, C; VAPNIK, V. (1995) Support-Vector Networks. *Machine Learning* v. 20, n. 3, p. 273–297, 1995.0885-6125.

DALAL, N; TRIGGS, B. (2005). Histograms of oriented gradients for human detection. *Proceedings - 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005* v. I, p. 886–893, 2005.0769523722.

HAKIM, T; COHN, D. *Implementation of HOG for Human Detection*. Disponível em: http://www.geocities.ws/talh_davidc/#cst_extract. Acesso em: 2 jul. 2017.

LIENHART, R.; MAYDT, J. (2002) An extended set of Haar-like features for rapid object detection. *Proceedings. International Conference on Image Processing* v. 1, p. 0–3, 2002.0-7803-7622-6.

MATHWORKS (2015). *Support Vector Machines for Binary Classification* . [S.l: s.n.]. Disponível em: <http://www.mathworks.com/help/stats/support-vector-machines-for-binary-classification.html>. Acesso em: 2 jul. 2017,

MCCORMICK, C. HOG Person Detector Tutorial. *Chris McCormick on WordPress.com*. Disponível em:

<https://chrisjmccormick.wordpress.com/2013/05/09/hog-person-detector-tutorial/>.

Acesso em: 2 jul. 2017.

RAY, S. (2015) Understanding Support Vector Machine algorithm from examples (along with code) *.Analytics Vidhya*. [S.l: s.n.]. Disponível em: <https://www.analyticsvidhya.com/blog/2015/10/understaing-support-vector-machine-example-code/>.

SANTOS, L. (2011) Detecção de faces através do algoritmo de Viola-Jones. *COPPE/UFRJ*, 2011.

VIOLA, P; JONES, M. (2001) Rapid object detection using a boosted cascade of simple features. *Computer Vision and Pattern Recognition (CVPR)* v. 1, p. I-511-I-518, 2001.0-7695-1272-0.