



Revista Eletrônica
Paulista de Matemática

ISSN 2316-9664
Volume 14, fev. 2019
Edição Ermac

Eduardo Machado Silva

UNESP - Universidade Estadual
Paulista “Júlio de Mesquita
Filho”
eduardo.msilva094@gmail.com

Maurílio Boaventura

UNESP - Universidade Estadual
Paulista “Júlio de Mesquita
Filho”
maurilio.boaventura@unesp.br

**Inês Aparecida Gasparotto
Boaventura**

UNESP - Universidade Estadual
Paulista “Júlio de Mesquita
Filho”
ines.boaventura@unesp.br

Correção de brilho, homogeneização e ajuste de imagens para o reconhecimento facial

Brightness correction, homogenization and adjustment of
images for face recognition

Resumo

O propósito deste trabalho é apresentar técnicas de pré-processamento para correção de brilho e homogeneização de imagens para o problema de reconhecimento facial. As técnicas propostas foram desenvolvidas para lidar com situações em que as imagens apresentem grandes variações de iluminação, poses e ângulos de captura, que são situações muito comum em ambientes não controlados de captura de imagens. Após a aplicação dessas técnicas houve uma melhora no brilho e homogeneização das imagens contidas nos bancos de dados avaliados. Para o processo de reconhecimento facial, primeiramente foram aplicadas as técnicas propostas na fase de pré-processamento e na sequência foi aplicado o método de padrões mapeados localmente em multiescala para o reconhecimento facial, o qual encontra-se descrito em detalhes. Com a utilização das técnicas propostas houve uma melhora significativa no processo de reconhecimento como um todo.

Palavras-chave: Modelagem matemática e aplicações. Computação científica. Reconhecimento de faces. Técnicas de pré-processamento.

Abstract

The purpose of this paper is to present preprocessing techniques for correction of brightness and homogenization of images for facial recognition problem. The proposed techniques presented were developed to deal with situations in which the images have great variations of illumination, poses and capture angles, which often occur in uncontrolled image capture environments. After applying these techniques there was an improvement in the brightness and homogenization of the images contained in the evaluated databases. In the preprocessing phase, the proposed techniques are applied for brightness correction and homogenization of images, then the multiscale localized mapped patterns method for facial recognition is applied, which is described in detail. A significant improvement was reach in the recognition process using the proposed techniques.

Keywords: Mathematical modeling and applications. Face recognition. Scientific computing. Preprocessing techniques.

1 Introdução

Com os avanços de tecnologias e serviços dos dias atuais, atividades e transações humanas em grande escala evoluíram muito e, portanto, a busca por métodos rápidos e confiáveis de identificação de pessoas se torna uma tarefa indispensável. Uma ampla variedade de sistemas requer segurança nos processos de identificação de pessoas, tanto para confirmar a identidade de um indivíduo quanto para tentar identificá-lo. O propósito de tais sistemas é assegurar que o serviço prestado seja realizado somente por pessoas credenciadas. A falta de robustez nos métodos empregados para reconhecimento pode tornar o sistema vulnerável a impostores com propósitos ilícitos ou prejudicar o próprio ato de identificação, o que faz com que o método seja não confiável. Chama-se de **reconhecimento biométrico**, ou simplesmente **biometria**, o reconhecimento automático de pessoas baseado em suas características fisiológicas ou comportamentais (LI; JAIN, 2011).

Mesmo existindo muitas formas de biometria para identificação de uma pessoa, as impressões digitais talvez sejam o exemplo de biometria mais utilizada até o momento. Como já é bem conhecido, pessoas diferentes possuem digitais distintas e elas não se alteram com o passar do tempo, salvo em casos de acidentes ou cirurgias. No entanto, a identificação através de digitais tem como desvantagem a necessidade de contato físico da pessoa com o aparelho leitor. Não existem métodos que captem digitais à distância ou pelo menos nenhum que pareça ser adequado para tal condição (DAUGMAN, 1994). Uma biometria alternativa e eficaz tão usada quanto as digitais é a face.

O reconhecimento facial é uma das tecnologias biométricas mais utilizadas em sistemas automatizados que necessitam garantir a identidade de uma pessoa para acesso autorizado e monitoramento. A grande aceitação do uso da face tem várias vantagens sobre outras tecnologias biométricas: ela é natural, não exige equipamentos sofisticados, a aquisição de dados é baseada em abordagens não invasivas, e pode ser feito a distância, de maneira cooperativa ou não.

Este trabalho tem por objetivo apresentar duas novas técnicas de pré-processamento baseada nos níveis de cinza das imagens, com a finalidade de melhorar a acurácia dos métodos para o reconhecimento facial. Uma delas chamada de correção de brilho e homogeneização de imagens e a outra chamada de ajuste nos níveis de cinza. As metodologias apresentadas se mostraram bastante robustas, houve uma melhora significativa na acurácia do processo de reconhecimento como um todo. Para a tarefa de reconhecimento facial foi utilizado o método de padrões mapeados localmente em multiescala (MSLMP), que foi introduzido por CONTRERAS et al. (2015), com o propósito de identificar impressões digitais fraudulentas, ou seja, impressões digitais obtidas a partir de dedos de silicone, por exemplo. Em SILVA et al. (2017) esse mesmo método foi utilizado para o reconhecimento facial.

Este artigo baseia-se no trabalho apresentado por SILVA et al. (2018), no qual foi apresentada a técnica correção do brilho e homogeneização das imagens. A principal contribuição que difere este trabalho é a proposta de uma técnica para ajustes dos níveis de cinza. Além disso, o texto está completamente reestruturado e há uma descrição mais detalhada do processo de reconhecimento facial utilizado.

Para os testes de avaliação da performance das metodologias propostas, foram utilizadas as seguintes bases de dados: MUCT, ORL, ESSEX96 e YALLE, que são bases complexas de dados e fazem com que os métodos tenham desempenho menor, ao utilizá-las.

Este trabalho está organizado da seguinte forma: na Seção 2 é apresentada a metodologia empregada para o reconhecimento facial; na Seção 3 são mostrados as técnicas para correção do

brilho e homogeneização das imagens; na Seção 4 são apresentados os resultados obtidos; na Seção 5 são expostas algumas conclusões; e finalmente na Seção 6 encontram-se as referências bibliográficas.

2 Metodologia

A ideia deste trabalho foi primeiramente utilizar técnicas de pré-processamento de imagens para correção de brilho e homogeneização das imagens que compõem todo o banco de dados e depois utilizar o MSLMP como extrator de características de imagens de faces para averiguar a sua capacidade de reconhecimento facial. Assim, encontram-se descritos na próximas seções os métodos MSLBP e LMP que deram origem ao MSLMP, que foi o método de reconhecimento facial utilizado neste trabalho e que também encontra-se discutido nesta seção.

2.1 Multi-scale local binary pattern (MSLBP)

O método MSLBP proposto por JIA et al. (2014), consiste na extração do padrão LBP (*Local Binary Pattern*) de cada pixel da imagem em diferentes escalas para um valor fixo de amostragem, após suavização gaussiana. O objetivo do uso de filtros gaussianos de passa-baixa é para que em cada amostra na vizinhança se possa coletar uma quantidade de informação maior que aquela com um único pixel. A Figura 1 exemplifica esse processo, na qual são apresentadas vizinhanças de um pixel quantizadas por oito amostras, como propõe JIA et al. (2014), onde formam-se oito setores que distam um raio R_n do pixel central. Os círculos representam a área para extração do padrão LBP ao redor de cada amostra após filtragem. O raio externo da “área útil” na Figura 1, que é usado para calcular o raio R_n para a extração do LBP, é dado por:

$$r_n = r_{n-1} \cdot \left(\frac{2}{1 - \sin(\pi/P_n)} - 1 \right), n \in \{2, \dots, N\}, \quad (1)$$

na qual N é o número de escalas e P_n a amostragem de cada escala, que no caso do MSLBP foi considerado $P_n = 8 \forall n \in \mathbb{N}$.

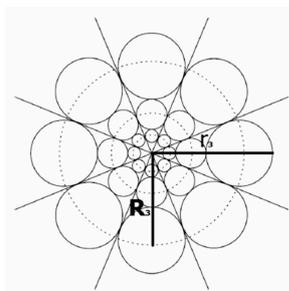


Figura 1: Distribuição ótima dos raios do padrão MSLBP com respeito à redução de redundância da informação capturada pela padrão.

Os raios r_n estabelecem a distância entre o pixel e a borda de uma vizinhança, por isso r_1 é definido como 1,5 que é a menor distância entre o pixel e a borda de uma região 3×3 . Os raios das operações LBP são escolhidos de forma que as “áreas úteis” estejam em contato. Consequentemente, define-se o raio do operador LBP na escala n ($n \geq 2$) como o ponto médio entre r_n e r_{n-1} , assim:

$$R_n = \frac{r_n + r_{n-1}}{2}, \quad n \in \{2, \dots, N\}. \quad (2)$$

Analogamente ao caso dos r_n , os raios dos operadores representam a distância entre o pixel central e os pixels de sua vizinhança, portanto R_1 é definido como 1 por ser a menor distância entre o pixel central e os de sua vizinhança 3×3 . Os raios r_n são usados para calcular o tamanho da janela dos filtros gaussianos, G_n , na escala n como mostra a Equação (3):

$$G_n = 2 \cdot \text{round} \left(\frac{r_n - r_{n-1}}{2} \right) + 1, \quad (3)$$

na qual a função $\text{round} : \mathbb{R} \rightarrow \mathbb{N}^+$ associa a um valor o menor inteiro não negativo maior que ele. Em cada nível, o valor δ_n necessário para aplicar o filtro gaussiano é dado por:

$$\delta_n = \frac{G_n}{\sqrt{-2 \ln(1-p)}}, \quad (4)$$

onde p foi estabelecido como 0,95.

2.2 Local mapped pattern (LMP)

O LMP, que é uma extensão do LBP, assume que a distribuição dos níveis de cinza na vizinhança de um pixel é um padrão local e pode ser representada pela diferença dos níveis de cinza da vizinhança do pixel central (FERRAZ; PEREIRA; GONZAGA, 2014). Dessa forma, cada pixel $g(i, j)$ é mapeado em função de uma região $W \times W$ em torno deste (Figura 2a), através da Equação (5):

$$h(g(i, j)) = \frac{\sum_{k=1}^W \sum_{l=1}^W (f_{g(i, j)}(g(k, l) \cdot P(k, l)))}{\sum_{k=1}^W \sum_{l=1}^W P(k, l)}, \quad (5)$$

na qual $f_{g(i, j)}$ é a função de mapeamento (Figura 2b), responsável pela suavização dos níveis de cinza da vizinhança que mapeará os $W^2 - 1$ vizinhos do pixel central $g(i, j)$ (Figura 2c), P é uma matriz de pesos $W \times W$ pré-definida (Figura 2d) e W é sempre um número ímpar, a fim de que exista um pixel central na região considerada. Este mapeamento define então o micro padrão LMP da região $W \times W$ (Figura 2e).

Através da Equação (5) pode-se deduzir descritores de padrões presentes na literatura. O LBP é obtido ao usar a função de mapeamento como sendo a função degrau de Heaviside (H), dada pela Equação (6):

Note que para a função sigmoide dada por (7), $f_{g(i,j)}(g(i,j)) = 0,5$, independente do valor de $g(i,j)$. Assim, o valor central $P(2,2)$ da matriz de pesos foi considerado igual a 0, evitando esse cálculo e melhorando a performance do método. Dessa forma, o padrão LMP de pixel $g(i,j)$ é dado pela Equação (9):

$$LMP(g(i,j)) = \frac{\sum_{(k,l) \in I - \{(i,j)\}} (f_{g(i,j)}(g(k,l)))}{8}, \quad (9)$$

na qual $I = \{(k,l) : k = i-1, i, i+1 \text{ e } l = j-1, j, j+1\}$.

Como os valores gerados pela função sigmoide estão entre 0 e 1, o padrão *LMP* está contido no intervalo $(0,1)$. Portanto, para obter valores em $[0,255]$ é usada a versão normalizada dos valores obtidos pela Equação (9), assim, o padrão final extraído de cada pixel $g(i,j)$ é dado pela Equação (10):

$$N_{LMP} = \text{round}(255.LMP(g(i,j))). \quad (10)$$

Observe que para que a função de mapeamento seja aplicada nos pixels de fronteira é necessário contruir uma matriz (imagem) aumentada, replicando-se os pontos de fronteira.

2.3 Multi-scale local mapped pattern (MSLMP)

A presença de ruídos na imagem prejudica a extração do padrão LBP das regiões, assim, seguindo os mesmos passos do método MSLBP, proposto por JIA et al. (2014), em CONTRERAS et al. (2015) foi proposta uma extensão multiescala do método LMP descrito na subseção anterior, cujo propósito é atenuar a ação ruidosa presente nas imagens através do uso de aplicações que suavizam as altas frequências presentes nas vizinhanças de um mesmo pixel. Tais vizinhanças seriam representadas por raios que aumentam conforme o nível do processo, caracterizando o método como multiescala. Assim, nesse método a evolução de cada escala se faz de acordo com o aumento do tamanho da dimensão das vizinhanças do pixel $g(i,j)$. Portanto, as dimensões W_n , dadas pela Equação (11), da matriz de pesos P_n aumentam de acordo com a escala.

$$W_n = 2.\text{round}(R_n) + 1, \quad (11)$$

sendo R_n dado pela Equação (2).

Em (BRAVO, 2016) foram realizados testes com matrizes de pesos para classificação de texturas utilizando o MSLMP. Após os experimentos, o melhor resultado de acurácia foi obtido pela matriz P_5 , representada em (12):

$$P_5 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}, \quad (12)$$

na qual estão contidas as distribuições de valores não nulos de peso até a quinta escala do método. Neste trabalho, foi adotada a matriz de pesos P_5 para aplicação do método MSLMP. Considerando o valor $P_5(6,6)$ como centro, o qual representa o pixel da imagem a ser extraída o micro padrão. A região 3×3 denota a P_1 (matriz de pesos para a escala um). A região 5×5 , tomando a região 3×3 como sendo a matriz nula, representa a P_2 . A vizinhança 7×7 , escolhida da mesma forma que a região 5×5 , denota a P_3 e as vizinhanças 9×9 e 11×11 denotam P_4 e P_5 , respectivamente. Além disso, é importante destacar que as matrizes P_n foram construídas de forma a considerar apenas oito amostras de cada região do pixel para o cálculo do MSLMP. Note que, em (13) são mostradas as matrizes P_1, P_2, P_3, P_4 e P_5 sobrepostas.

O padrão MSLMP de um pixel $g(i, j)$ na n -ésima escala é obtido pela Equação (13):

$$MSLMP(g(i, j)) = \frac{\sum_{(k,l) \in I_n - \{(i,j)\}} (f_{g(i,j)}(g(k,l)))}{8}, \quad (13)$$

na qual, $I_n = \{(k, l) : k = i - C, i - 1, i + C - 1, i + C \text{ e } l = j - (C - 1), j + 1, j - 1, j + C - 1\}$, onde $C = \text{round}(R_n)$.

Da mesma forma que o LMP, o padrão MSLMP também está no intervalo $(0, 1)$ e por isso é utilizada a versão normalizada, dada pela Equação(14):

$$N_{MSLMP} = \text{round}(255.MSLMP(g(i, j))). \quad (14)$$

Note que ambos os métodos MSLBP e MSLMP são multiescalas. A diferença básica entre eles, é que o método MSLBP utiliza o LBP como um extrator de características e necessita que as imagens sejam suavizadas por um filtro gaussiano para se obter bons resultados, enquanto que o método MSLMP utiliza o LMP. A função sigmoide, utilizada no método LMP, funciona como um suavizador e isso faz com que não seja necessário utilizar o filtro gaussiano no método MSLMP.

2.4 Extração de características

Os padrões normalizados MSLMP são valores no intervalo $[0,255]$, como visto na Equação (14). Em OJALA et al. (2002) são apresentados alguns valores dentro desse intervalo deno-

minados *uniform patterns*. Esses valores representam uma classe que é responsável, na maioria das vezes, por mais de 90% das características mais representativas da imagem. De maneira geral, o *uniform pattern* é um valor $U \in [0, 255] \cup \mathbb{Z}$, que em sua representação binária $(b_1 b_2 b_3 b_4 b_5 b_6 b_7 b_8)$, com $b_i \in \{0, 1\}$, respeita a desigualdade (15):

$$|b_1 - b_8| + \sum_{i=2}^8 |b_i - b_{i-1}| \leq 2. \quad (15)$$

De acordo com JIA et al. (2014), dadas K amostras da região de um pixel, existem $(K - 1) \cdot K + 2$ *uniform patterns*. Para $K = 8$, que é a quantidade mínima de dígitos necessários para se representar qualquer valor inteiro no intervalo $[0, 255]$ em binário, no método em questão, tem-se 58 valores entre 0 e 255 que satisfazem a relação (15).

Para a extração de características de uma imagem, em cada escala do método, gera-se um histograma dos valores obtidos pela Equação (14) e em seguida toma-se a taxa de ocorrência dos 58 *uniform patterns* nos respectivos histogramas, então somam-se as taxas restantes para formar o vetor $x^n = (x_1^n, x_2^n, \dots, x_N^n) \in [0, 1]^{59}$. O vetor de características é, então, dado pelo vetor $v = (x^1, \dots, x^N)$, cujo número de coordenadas é $N \cdot 59$, em que N é o número de escalas utilizadas.

Esses *uniform patterns* são então utilizados como as características principais de todo o processo. A partir dessas características um processo de classificação deve ser empregado para que seja feito o reconhecimento.

3 Correção de brilho, homogeneização e ajuste de imagens

Nesta seção são apresentadas as duas novas técnicas de pré-processamento baseada nos níveis de cinza das imagens propostas neste trabalho, com a finalidade de melhorar a acurácia de métodos para o reconhecimento facial. Uma delas chamada de correção de brilho e homogeneização de imagens e a outra chamada de ajuste nos níveis de cinza.

Observe que essas técnicas devem ser aplicadas antes do processo de reconhecimento facial descrito na seção anterior.

3.1 Correção de brilho e homogeneização de imagens

A motivação para o desenvolvimento desse método está no estudo de imagens em bases de dados que apresentam grandes variações de iluminação e ângulos de captura, que são situações que ocorrem na prática, quando por exemplo, se deseja reconhecer uma determinada pessoa em meio a tantas outras em ambientes internos ou externos, sem que haja a colaboração da pessoa para isso. Essa é uma aplicação típica da área criminal, na qual se tenta reconhecer criminosos em situações de ataque, com imagens capturadas em ambientes públicos, por exemplo.

A base de Dados MUCT (Milborrow/University of Cape Town) é uma base bastante adequada para avaliar modelos propostos para essa finalidade. Assim, essa foi a principal base de dados utilizada neste estudo. Devido a sua alta variação de iluminação e poses, a média dos níveis de cinza das imagens da base MUCT variam bastante, como pode-se observar na Figura 3, que mostra a quantidade de imagens distribuídas em função das médias dos níveis de cinza.

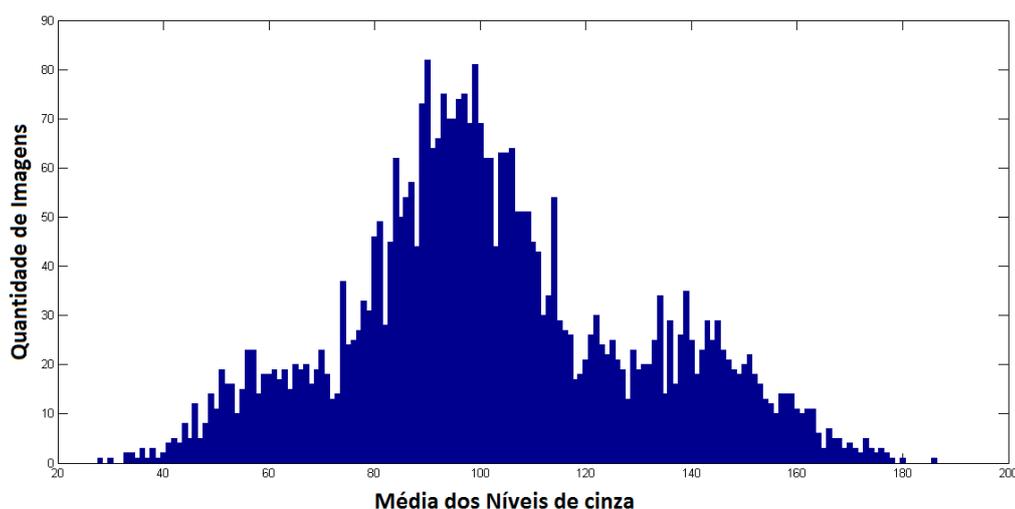


Figura 3: Média dos níveis de cinza da MUCT.

Os valores das médias dos níveis de cinza encontram-se distribuídos no intervalo delimitado por $M_{min} = 28,46$ e $M_{max} = 185,66$. A média dos valores da Figura (3), isto é, a média das médias dos níveis de cinza das imagens é $M_m = 102,44$.

A ideia foi encontrar um valor padrão $M \in [M_{min}, M_{max}]$, de forma que M dependa diretamente das médias das imagens contidas na base de dados e, então, aproximar todas as médias das imagens para esse valor. Após encontrado tal valor, a aproximação é feita através de um parâmetro θ_I , com $\theta_I \in (-1, 1)$, que é unicamente determinado por M e M_I , em que M_I é média dos níveis de cinza da imagem I , a qual se deseja ajustar. Assim, cada imagem possui um determinado parâmetro de ajuste. Considerou-se M , como a média dos valores da Figura 3, ou seja, $M = M_m$, e o valor θ_I foi calculado por meio da Equação (16):

$$\theta_I = \frac{M - M_I}{M_{max} - M_{min}}. \quad (16)$$

Após a obtenção do parâmetro θ_I , cada imagem I , tem seus pixels $I(i, j)$, alterados pelo valor $(1 + \theta_I)I(i, j)$. Dessa forma, a correção é determinada da seguinte maneira:

- se $\theta_I > 0$, significa que a média da imagem I é menor que a média geral, assim, para que haja uma aproximação das médias é necessário somar a cada pixel $I(i, j)$ um valor positivo, ou seja o valor $\theta_I \cdot I(i, j)$;
- se $\theta_I < 0$, a média da imagem I é maior que a média geral e, dessa forma, para que haja uma aproximação das médias é necessário subtrair um valor de cada pixel $I(i, j)$, ou seja o valor $\theta_I \cdot I(i, j)$.

Em termos de visualização das imagens, o reajuste pode ser interpretado da seguinte forma:

- Com relação ao valor de referência, M , se a imagem for “clara”, então a técnica irá “escurecê-la”;
- caso seja “escura”, então o ajuste irá “clareá-la”;

- caso seja uma imagem que não necessite de ajustes, isto é, a média dos seus níveis de cinza estiver próxima a M , a imagem sofrerá alterações mínimas, nestes casos $\theta_I \approx 0$.

Essa técnica pode ser aplicada para outras bases de dados com características similares as da base MUCT. Os efeitos da aplicação dessa técnica podem ser visualizados na Figura 4, mostrada a seguir, na qual pode ser observado que a imagem clara foi levemente escurecida, a imagem escura foi levemente clareada e a imagem neutra se manteve neutra, tornando-as todas mais homogêneas.



Figura 4: Efeito da correção do brilho e homogeneização: (a), (b) e (c) imagens originais e (d), (e), e (f) representam, respectivamente, as imagens reajustadas.

3.2 Ajuste na intensidade dos pixels

Nesta seção é apresentado um método que também pode ser utilizado na fase de pré-processamento, de forma alternativa ao que foi apresentado anteriormente. O objetivo desse método é tratar banco de dados com imagens que possuam pixels predominantemente de alta ou baixa intensidades.

O banco de dados YALE apresenta as características necessárias para a aplicação dessa técnica. Assim, a técnica proposta é discutida utilizando esse banco de dados, porém, vale ressaltar que o mesmo procedimento pode ser aplicado a outros bancos de dados com características similares.

Essa base, em geral, apresenta dificuldades durante o processo de reconhecimento devido a grande variação de iluminação e à alta diversidade de expressões, como foi descrito anteriormente. O fundo branco está presente em quase todas as imagens dessa base.

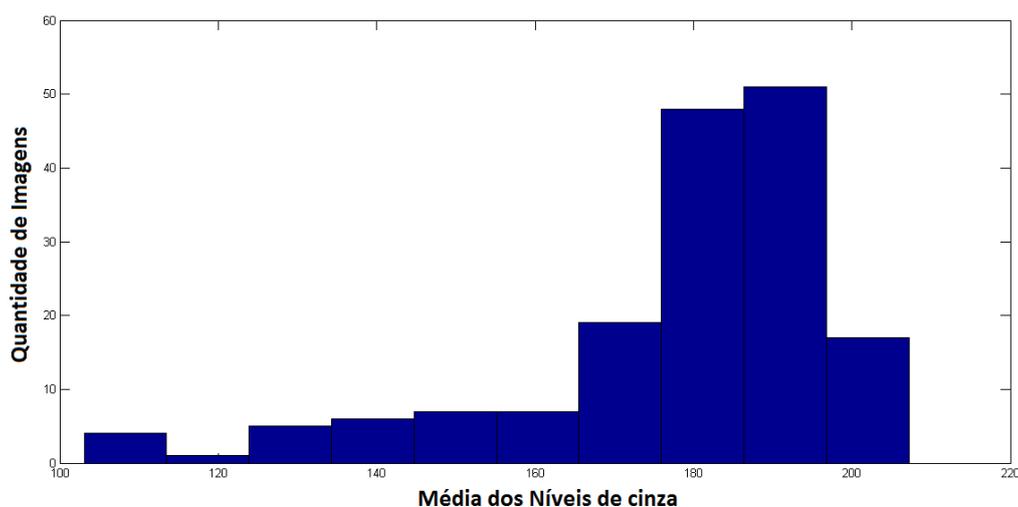


Figura 5: Média dos níveis de cinza da YALE em relação a quantidade de imagens.

A Figura 5 mostra as médias dos níveis de cinza das imagens da YALE. Apesar da base apresentar alta iluminação, as médias das imagens não variam tanto, pois cerca de 80% delas se encontram no intervalo $[177, 197]$, que é um intervalo bastante reduzido, caracterizando as imagens, de um modo geral, como muito claras.

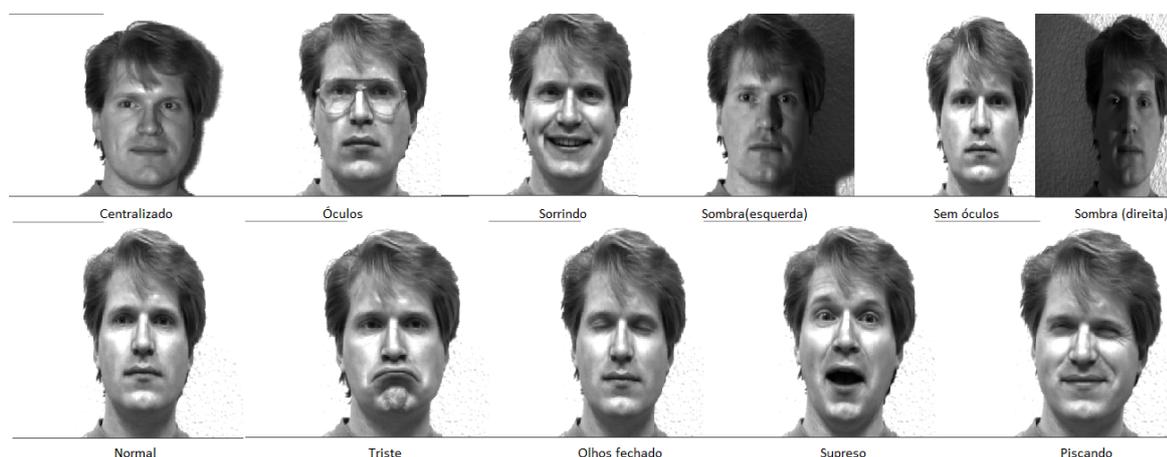


Figura 6: Exemplo das expressões e iluminação de uma classe da YALE.

Observando as imagens da Figura 6, percebe-se que a grande quantidade de pixels brancos no fundo está presente em praticamente todas as imagens da base. Devido a isso, a ideia inicial do ajuste está relacionada com quantidade de pixels de valor 255 (branco). Diferente da técnica anterior, o parâmetro θ_I dependerá unicamente da imagem I . Também foi considerada uma constante $\theta^0 \in (0, 1)$ escolhida empiricamente, que será usada em todas as imagens.

Seja $P_I \in \mathbb{Z}$, a quantidade de pixels da imagem I . Considerando que as imagens de um

determinado banco de dados devem possuir as mesmas dimensões, e caso isso não aconteça, é possível redimensioná-las para uma dimensão padrão, tem-se que P_I é constante para toda imagem I da base de dados. Seja $P_I^b \in \mathbb{Z}$ a quantidade de pixels da imagem I cujo valor é maior ou igual 220, uma vez que não só os pixels de extremo branco (255) contribuem para deixar a imagem clara. Assim, a partir dos valores P_I e P_I^b é possível definir θ_I através da Equação (17):

$$\theta_I = \frac{P_I^b}{P_I}. \quad (17)$$

Claramente percebe-se que, $\theta_I \in (0, 1)$, para todas as imagens da base. Assim, dados θ_I e θ^0 o reajuste, inicial, das intensidades é feito da seguinte maneira:

- Se $\theta_I > \theta^0$ será subtraído o valor $\theta_I \cdot I(i, j)$, de cada pixel $I(i, j)$;
- Se $\theta_I < \theta^0$ será somado o valor $\theta_I \cdot I(i, j)$, de cada pixel $I(i, j)$.

A Figura 7 exemplifica os efeitos da técnica descrita nas imagens da YALE.



Figura 7: Efeito da correção do brilho e homogeneização: (a), (b) e (c) imagens originais e (d), (e), e (f) representam, respectivamente, as imagens reajustadas.

Embora o emprego dessa técnica tenha apresentado bons resultados, ela encontra-se em fase inicial de estudos e alguns ajustes ainda se fazem necessários para torná-la mais robusta. É importante ressaltar que outras técnicas disponíveis na literatura foram testadas, tais como as técnicas baseadas na equalização de imagens, porém estas não surtiram efeito no processo de reconhecimento. Apesar de muito simples, essa técnica surtiu melhores efeitos e resultados.

4 Resultados

Conforme apresentado nas seções anteriores, para cada imagem de dados é extraído o padrão MSLMP, criando o vetor de características e, então, é realizado o processo de classificação utilizando-se essas características. Para analisar a acurácia do método foi empregado o processo de validação cruzada. As próximas subseções apresentam a discussão dos procedimentos de validação cruzada e de classificação, os quais foram utilizados neste trabalho.

4.1 Validação Cruzada

A validação cruzada é um método estatístico utilizado para medir a capacidade de generalização de uma técnica. A ideia geral do método é dividir o conjunto de dados em grupos mutuamente exclusivos (grupos para validação) para a avaliação da performance do método.

Neste trabalho, para cada base de dados utilizada foram separadas 20% das imagens para teste e 80% para o treinamento, em forma de rodízio, para que o processo de validação cruzada fosse executado. Assim, cada banco de dados foi dividido em 5 grupos distintos para teste, com o objetivo de avaliar o *MSLMP* para o reconhecimento de todas as imagens da base. A escolha do grupo de imagens para validação cruzada foi diferente em cada base e por isso, os detalhes dessa escolha encontram-se apresentados juntos aos resultados experimentais de cada banco de dados em suas respectivas subseções. Em geral, todos os grupos foram escolhidos de forma que haja diversidade nas imagens em relação à pose, iluminação e expressões.

4.2 Processo de classificação

O processo de classificação é feito calculando-se a distância Chi-quadrado entre dois vetores de características $v_i, w_j \in \mathbb{R}^n$, através da Equação (18):

$$\chi^2(v_i, w_j) = \frac{1}{2} \sum_{k=1}^n \frac{(v_i^k - w_j^k)^2}{(v_i^k + w_j^k)}, \quad (18)$$

na qual, v_i representa o vetor de características da i -ésima amostra de teste a ser avaliada (teste) e w_j o vetor de característica da j -ésima amostra para comparação (treino). A figura (8) ilustra o processo de classificação.

A acurácia A_i do grupo de validação i , $i \in \{1, 2, 3, 4, 5\}$, é então calculada pela Equação (19):

$$A_i = \frac{\text{Número de acertos}}{\text{Número total de imagens do grupo}}. \quad (19)$$

A taxa de acerto final é calculada pela média aritmética entre todos os grupos considerados.

Para analisar a performance dos métodos propostos neste trabalho, foram utilizadas as base de dados ESSEX 96, ORL, MUCT e YALLE

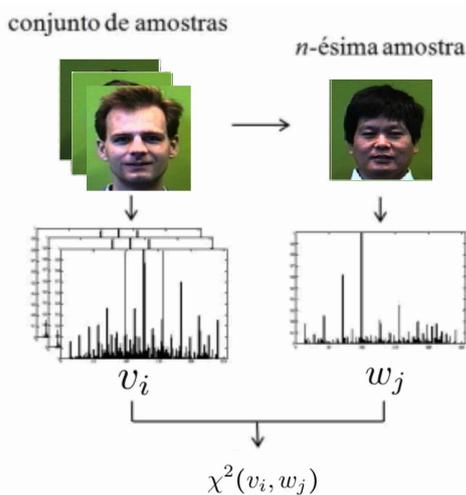


Figura 8: Processo de classificação de faces.

4.3 Base de dados ESSEX 96

Desenvolvida pelo Dr. Libor Spacek, professor da Universidade de Essex esse banco de dados tem como objetivo o encorajamento de comparação entre novos métodos que utilizem faces. As imagens desse banco de dados estão divididas em 4 diretórios, sendo o diretório da ESSEX 96 considerado um dos mais difíceis. Esse repertório contém fotografias de 152 pessoas, sendo cada uma delas fotografada 20 vezes. As imagens tem dimensão 196×196 e foram obtidas em tempos diferentes e enquanto as pessoas davam um passo a frente para introduzir uma certa variação da escala do rosto. Exemplos de imagens que compõem esse banco de dados podem ser visualizados através da Figura 9.



Figura 9: Imagens da base de dados ESSEX 96.

Para esse banco de dados, foi utilizada a técnica de correção de brilho e homogeneização de imagens, dada pela Equação (16). A Figura 10 mostra a média dos níveis de cinza das imagens deste banco de dados. Os valores médios mínimo e máximo são dados por: $M_{min} = 47,06$ e $M_{max} = 122,97$ e, portanto, $M = 85,01$.

As Tabelas 1 e 2 apresentam a acurácia do método proposto para 5 escalas. São mostrados os resultados sem a correção do brilho e com a correção, respectivamente.

Como pode ser observado pelos resultados obtidos, a aplicação da técnica proposta não trouxe melhoras significativas nos resultados, houve sim uma melhora discreta. Entretanto isso já era esperado, pois esse banco de dados não apresenta muitas distorções no padrão das imagens, como pode ser observado no gráfico mostrado na Figura 10, mas mesmo assim, os resultados

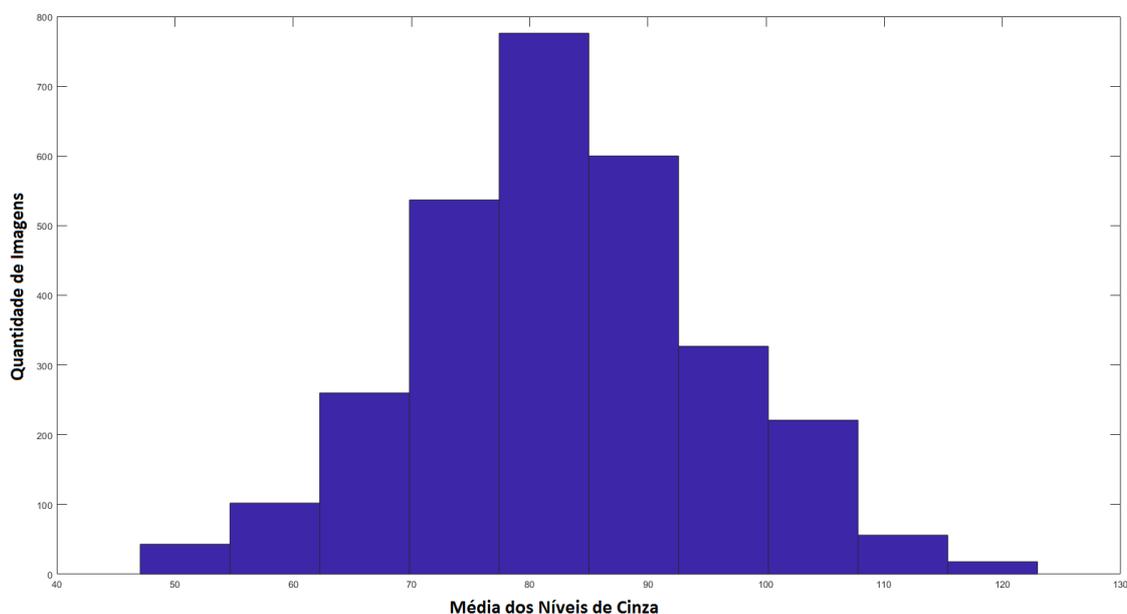


Figura 10: Média dos níveis de cinza da ESSEX 96.

Tabela 1: Taxa de acertos da ESSEX 96 com cinco escalas sem correção do brilho.

Grupo	Acertos
Grupo_a	99,15%
Grupo_b	98,64%
Grupo_c	98,47%
Grupo_d	97,96%
Grupo_e	98,64%
Média	98,57%

Tabela 2: Taxa de acerto da ESSEX 96 para cinco escalas com correção do brilho.

Grupo	Acertos
Grupo_a	98,64%
Grupo_b	98,47%
Grupo_c	98,47%
Grupo_d	98,30%
Grupo_e	99,15%
Média	98,61%

apresentados mostram pequenas melhorias, que de certa forma representam algum avanço.

4.4 Base de dados ORL

O banco de dados ORL foi desenvolvido pelo Laboratório de Computadores da Universidade de Cambridge e contém um conjunto de imagens fotografadas entre 1992 e 1994. Este banco de dados possui dez imagens de 40 pessoas, tendo um total de 400 imagens de dimensão 92×112 . Para alguns indivíduos, as imagens foram tiradas em tempos diferentes, com variação de iluminação, de expressões faciais (olhos/boca fechados, sorrindo ou não) e com objetos cobrindo parte do rosto (óculos). Exemplos de imagens que desse banco de dados podem ser visualizados através da Figura 11.

Para esse banco de dados, também foi utilizada a técnica de correção de brilho e homogeneização de imagens, dada pela Equação (16). A Figura 12 mostra a média dos níveis de cinza das ima-



Figura 11: Imagens do Banco de Dados ORL.

gens. Os valores médios mínimo e máximo são dados por: $M_{min} = 76,86$ e $M_{max} = 147,99$ e, portanto, $M = 112,41$.

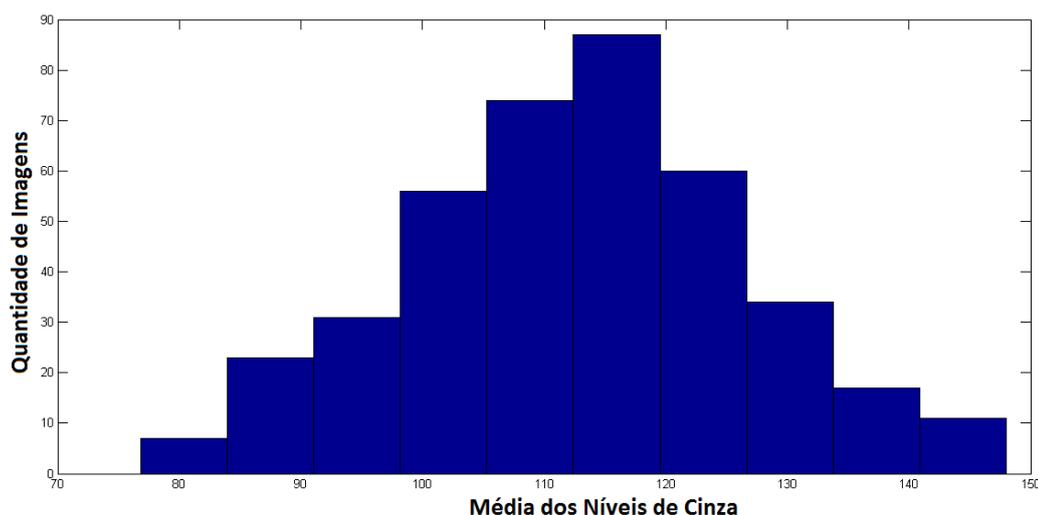


Figura 12: Média dos níveis de cinza da ORL.

As Tabelas 3 e 4 apresentam a acurácia do MSLMP para 3 escalas sem a correção do brilho e com a correção, respectivamente.

Como pode ser observado pelos resultados apresentados, a aplicação da técnica proposta também não trouxe melhoras significativas nos resultados. Entretanto isso também já era esperado, pois esse banco de dados não apresenta muitas distorções no padrão das imagens, como pode ser observado no gráfico mostrado na Figura 12.

4.5 Base de dados MUCT

A base de dados MUCT, contém 3755 imagens, de dimensões 480×640 , de 276 pessoas. Dez tipos de iluminações foram considerados. Comparada a outras bases de dados públicas,

Tabela 3: Taxa de acerto ORL sem correção do brilho.

Grupo	Acertos
Grupo_a	95,00%
Grupo_b	93,75%
Grupo_c	95,00%
Grupo_d	95,00%
Grupo_e	93,75%
Média	94,50%

Tabela 4: Taxa de acerto ORL com correção do brilho.

Grupo	Acertos
Grupo_a	93,75%
Grupo_b	98,75%
Grupo_c	93,75%
Grupo_d	93,75%
Grupo_e	95,00%
Média	95,00%

a MUCT apresenta maior diversidade de iluminação, idade e etnia. É uma base de dados disponível gratuitamente para fins acadêmicos. Nesta base de dados, cada indivíduo foi fotografado utilizando-se cinco câmeras (webcams) e da forma mais natural possível. O objetivo desta base de dados é que seja usada em aplicações que requerem ambientes não controlados, o que dificulta muito o reconhecimento/identificação das pessoas. Devido a essa particularidade, é muito difícil encontrar trabalhos que utilizam essa base de dados para avaliação de performance de métodos destinados ao reconhecimento de faces. A Figura 13 contém exemplos das imagens da MUCT. Cada classe dessa base contém 10 ou 15 imagens, sendo que cada grupo de 5 imagens possui uma específica iluminação. Para esse banco de dados, foi crucial a utilização da técnica de correção de brilho e homogeneização de imagens, dada pela Equação (16), como pode ser observado nos resultados obtidos.



Figura 13: Iluminação e câmeras MUCT.(Extraído de MILBORROW et al. (2010))

As Tabelas 5 e 6 mostram a acurácia do método sem a correção do brilho e com a correção,

respectivamente.

Tabela 5: Taxa de acerto MUCT sem correção do brilho.

Grupo	Acertos
Grupo_a	92,41%
Grupo_b	83,49%
Grupo_c	83,09%
Grupo_d	84,02%
Grupo_e	85,75%
Média	85,75%

Tabela 6: Taxa de acerto MUCT com correção do brilho.

Grupo	Acertos
Grupo_a	94,44%
Grupo_b	92,68%
Grupo_c	84,29%
Grupo_d	87,08%
Grupo_e	89,88%
Média	89,68%

Em BARNOUTI et al. (2016), os autores relatam uma taxa de acerto de 87,50% utilizando essa base de dados, porém foram usadas apenas 7 imagens para treino e 1 para teste não havendo menção quanto à forma da escolha dessas imagens.

Estes resultados mostram que a técnica de ajuste nos níveis de cinza aplicada foi essencial para um aumento na performance do método, com um ganho de aproximadamente 4 pontos percentuais na performance, que é bastante significativo. Isso nos mostra que para bancos de dados que apresentam grandes distorções nas imagens que os compõem, a aplicação de técnicas de correção de brilho e homogeneização são essenciais para se atingir uma boa performance do sistema como um todo.

4.6 Base de dados YALE

O banco de dados YALE, como o próprio nome já diz, foi desenvolvido pela Universidade Yale. Ele contém 11 imagens de 15 pessoas (14 homens e uma mulher) totalizando 165 imagens em 256 escalas de cinza, de dimensões 320×243 .

Esse banco de dados é caracterizado pelas variações de iluminação e de expressão. Não há alteração de ângulo de captura. A Figura 6 exemplifica uma classe desse banco. Para esse banco de dados, foi utilizada a técnica de ajuste de intensidade dos pixels, dada pela Equação (16), para todas as imagens que compõem o banco de dados, o valor escolhido para o parâmetro θ^0 foi $\theta^0 = 0,4$.

As Tabelas 7 e 8 mostram a acurácia para esta base de dados com e sem o ajuste dos níveis de cinza, respectivamente.

Para essa base de dados, em que grandes distorções nas tonalidades das imagens estão presentes, ou seja, contém imagens muito claras ou muito escuras, a técnica de ajuste dos pixels proposta trouxe ganhos significativos na performance do método, aumentando em aproximadamente 3 pontos percentuais a acurácia do método. O mesmo pode ser esperado para bancos de dados que apresentam as mesmas características.

Pode ser observado através das tabelas de resultados apresentadas que há algumas distorções nos percentuais de acertos dos diferentes grupos de imagens utilizados. Os grupos de imagens, como já foi dito anteriormente, foram formados de modo a conter a maior variedade possível tanto para as imagens de treinamento, quanto para as imagens de testes, porém, essas escolhas foram

Tabela 7: Taxa de acerto YALE sem o ajuste nos níveis de cinza.

Grupo	Acertos
Grupo_a	76,67%
Grupo_b	100,00%
Grupo_c	100,00%
Grupo_d	63,33%
Grupo_e	93,33%
Média	86,67%

Tabela 8: Taxa de acerto YALE com o ajuste nos níveis de cinza.

Grupo	Acertos
Grupo_a	80,00%
Grupo_b	100,00%
Grupo_c	96,67%
Grupo_d	73,33%
Grupo_e	96,67%
Média	89,33%

feitas de forma automatizada e não visualmente e, assim, alguns desbalanceamentos podem ter ocorridos e ocasionado tais distorções de percentuais nos diferentes grupos de imagens utilizados.

5 Conclusões

As principais contribuições desse artigo são as propostas da técnica correção do brilho e homogeneização e da técnica para ajuste dos níveis de cinza de imagens. Como pode ser observado, pelos resultados mostrados, o emprego dessas técnicas foi essencial para se obter uma melhora significativa no desempenho do método MSLMP para o reconhecimento facial.

As técnicas propostas apresentaram boa acurácia principalmente nas bases de dados com maior variação de iluminação, poses e ângulos de captura de imagens. Graficamente isto pode ser visto através do desvio da média geral dos níveis de cinza com relação ao ponto em que se encontram a maioria das imagens. Para as bases de dados ESSEX 96 e ORL a melhora não foi tão significativa, pois essas bases, não apresentam grandes variações nas médias dos níveis de cinza em relação a base toda, pode-se notar isso através da quantidade de barras nas Figuras 8 e 9, nas quais menos barras implica dizer que uma quantidade maior de imagens apresentam menos variações do nível de cinza de uma para outra.

Para trabalhos futuros, pretende-se testar outros algoritmos propostos para extrair características, como a entropia wavelet estacionária, biortogonal entropia wavelet, que foram aplicadas no reconhecimento de emoções faciais com grande sucesso. Além disso, pretende-se testar diferentes métodos de classificação, como SVM, Redes Neurais, etc.

6 Referências bibliográficas

BARNOUTI, N. H. et al. Face detection and recognition using Viola-Jones with PCA-LDA and square euclidean distance. **Internacional Journal of Advanced Computer Science and Applications**, v. 7, n. 5, p. 371-377, 2016.

BRAVO, M. J. A. **Análise do descritor de padrões mapeados localmente em multiescala para classificação de textura em imagens digitais**. 2016. 155 f. Dissertação (Mestrado em Matemática)- Universidade Estadual Paulista “Júlio de Mesquita Filho”, São José do Rio Preto, 2016.



CONTRERAS, R. C.; BOAVENTURA, M.; BOAVENTURA, I. A. G. Multi-scale local mapped pattern for image texture analysis. In: WORKSHOP DE VISÃO COMPUTACIONAL, 11., 2015, São Carlos. **Anais...** São Carlos: USP, 2015. p. 346-351.

DAUGMAN, J. G. **Biometrics personal identification system based on iris analysis**. Google Patents. 1994. Disponível em: <<https://www.google.com/patents/US5291560>>. Acesso em: 8 jan. 2018.

FERRAZ, C. T.; PEREIRA JR, O.; GONZAGA, A. Feature description based on center-symmetric local mapped patterns. In: SYMPOSIUM ON APPLIED COMPUTING, 29., 2014, Gyeongju. **Proceedings...** Gyeongju: SAC, 2014. p. 39-44.

JIA, X. et al. Multi-scale local binary pattern with filters for spoof fingerprint detection. **Information Sciences**, v. 268, p. 91-102, 2014.

LI, S. Z.; JAIN, A. K. **Handbook of face recognition**. 2nd ed. New York: Springer-Verlag, 2011.

MAJEED, S. Face recognition using fusion of local binary pattern and Zernike moments. In: IEEE INTERNATIONAL CONFERENCE ON POWER ELECTRONICS, INTELLIGENT CONTROL AND ENERGY SYSTEMS, 1., 2016, Delhi, India. **Proceedings...** Delhi, India: IEEE, 2017. p. 1-5.

MILBORROW, S.; MORKEL, J.; NICOLLS, F. The MUCT Landmarked face database. In: ANNUAL SYMPOSIUM OF THE PATTERN RECOGNITION ASSOCIATION OF SOUTH AFRICA, 21., 2010, Stellenbosch, South Africa, **Proceedings...** Stellenbosch, South Africa: IAPR, 2010, p. 179-184.

OJALA, T.; PIETIKAINEN, M.; MAENPAA, T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. **IEEE Transactions on pattern analysis and machine intelligence**, v. 24, n. 7, p. 971-987, 2002.

SILVA, E. M.; BOAVENTURA, M.; BOAVENTURA, I. A. G. Correção do brilho e homogeneização de imagens aplicadas ao reconhecimento facial. In: ENCONTRO REGIONAL DE MATEMÁTICA APLICADA E COMPUTACIONAL, 5., 2018, Bauru. **Caderno de trabalhos completos e resumos...** Bauru: Unesp, Faculdade de Ciências, 2018. p. 95-101. Disponível em: <<https://drive.google.com/file/d/1-8SZZnioKApdmQG6BaOTltJARzfwAB9P/view>>. Acesso em: 21 dez. 2018.

SILVA, E.; BOAVENTURA, M.; BOAVENTURA, I. A. G. Face recognition based on a multi-scale local mapped pattern method. In: WORKSHOP DE VISÃO COMPUTACIONAL, 13., 2017, Natal. **Anais...** Natal: UFRN, 2017. p. 82-87.