

Compound Local Binary Pattern para reconhecimento de expressões faciais– São Luís, 2016

André Luiz Dorneles Pires¹, Geovane Menezes Ramos Neto¹

¹Universidade Federal do Maranhão, São Luís, Brasil
andre.lp@hotmail.com, geovannemrn@gmail.com

Resumo

O reconhecimento de expressões faciais é de suma importância para aplicações que necessitam identificar o estado emocional do usuário para obter as informações necessárias a finalidade da aplicação. Este estudo visa utilizar uma técnica baseada em padrões locais binários para caracterização das expressões de tristeza quando comparada a outras seis expressões. A base de imagem utilizada foi a Cohn-Kanade. O melhor resultado obtido foi de 81,90% de acurácia para a proporção treino/teste de 60/40 utilizando como classificador a Máquina de Vetores de Suporte.

Introdução

A expressão facial proporciona uma forma não-verbal de comunicação que facilita o reconhecimento de emoções e intenções humanas. A análise automática de expressões faciais tem atraído muita atenção nos últimos anos devido ao seu potencial de aplicação em muitas áreas, tais como interação humano-computador, animação orientada a dados, e aplicações personalizadas para produtos de consumo.

Neste trabalho, utilizamos o *Compound Local Binary Pattern* (CLBP), que é um método de representação de características baseado no operador LBP original, para o reconhecimento de expressões faciais. Ao contrário do operador LBP original que utiliza P bits para codificar apenas os sinais das diferenças entre o pixel central e os P valores de cinza dos pixels vizinhos, o método proposto emprega $2P$ bits, em que os P bits adicionais são usados para codificar as informações da magnitude das diferenças entre o centro e os valores de níveis de cinza vizinhos em uma vizinhança local utilizando uma limiar. A motivação por trás do esquema de codificação proposto é aumentar a robustez da representação de características por incorporar informações locais adicionais que são descartadas pelo operador LBP original.

Referencial Teórico

O CLBP emprega $2P$ bits, onde os P bits adicionais são utilizados para codificar a magnitude da diferença entre os níveis de cinza já citados. A motivação do CLBP é aumentar a robustez da representação de características incorporando informação local adicional que é descartada pelo LBP tradicional. Além de utilizar os P bits do LBP, o CLBP gera mais P bits.

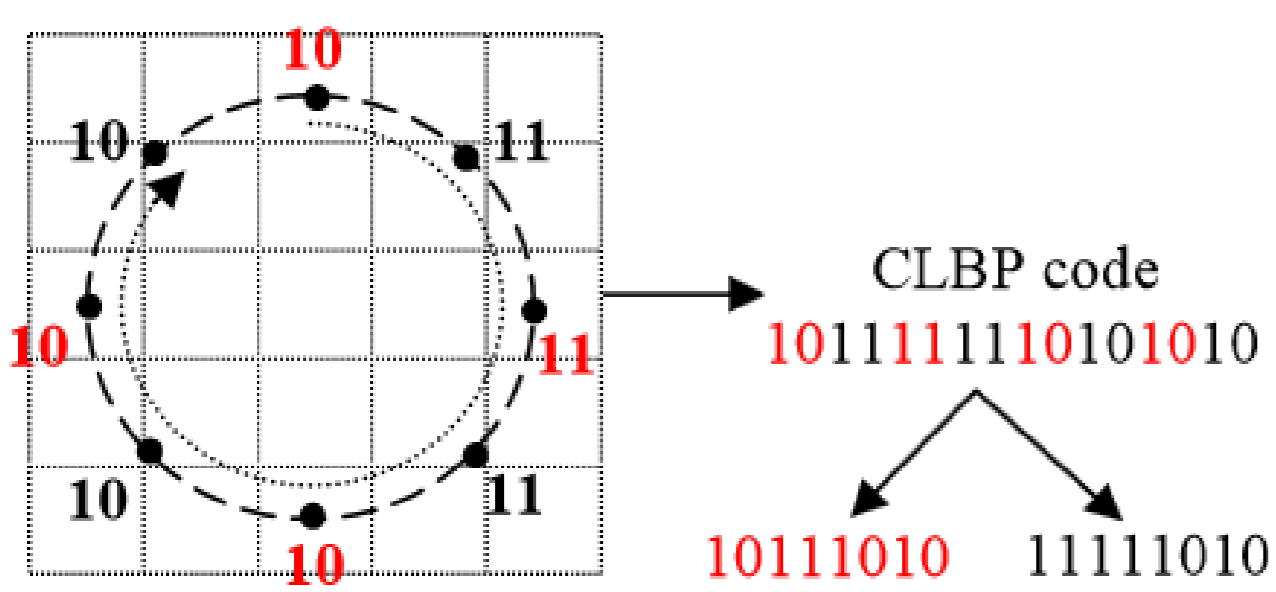


Fig. 1 Geração dos códigos sub-CLBP a partir do código original do CLBP

Em uma vizinhança 3×3 , o CLBP opera com 8 vizinhos ao redor do pixel central, gerando um código de 16 bits. Para reduzir o número de características, esse código é dividido em dois códigos sub-CLBP. Cada sub-CLBP é obtido concatenando os valores da sequência de bits (1, 2, 5, 6, ..., $2P-3$, $2P-2$) e (3, 4, 7, 8, ..., $2P-1$, $2P$), respectivamente do código $2P$ bits original do CLBP, representado na Figura 1. Os dois códigos sub-CLBP são tratados como códigos binários separados sendo estes

calculados separadamente. A partir das imagens geradas, extraímos os histogramas e concatenamos ambos gerando o vetor de características.

Metodologia

Este capítulo descreve os procedimentos utilizados neste trabalho para metodologia da caracterização da expressão de tristeza quando comparada a outras seis expressões faciais (Alegria, Aversão, Medo, Raiva, Neutro, Surpresa) utilizando o Compound Local Binary Patterns. A Figura METODOLOGIA apresenta as etapas do estudo, que são: Aquisição de Imagens, Decomposição da ROI, Extração de Características e Reconhecimento de Padrões.

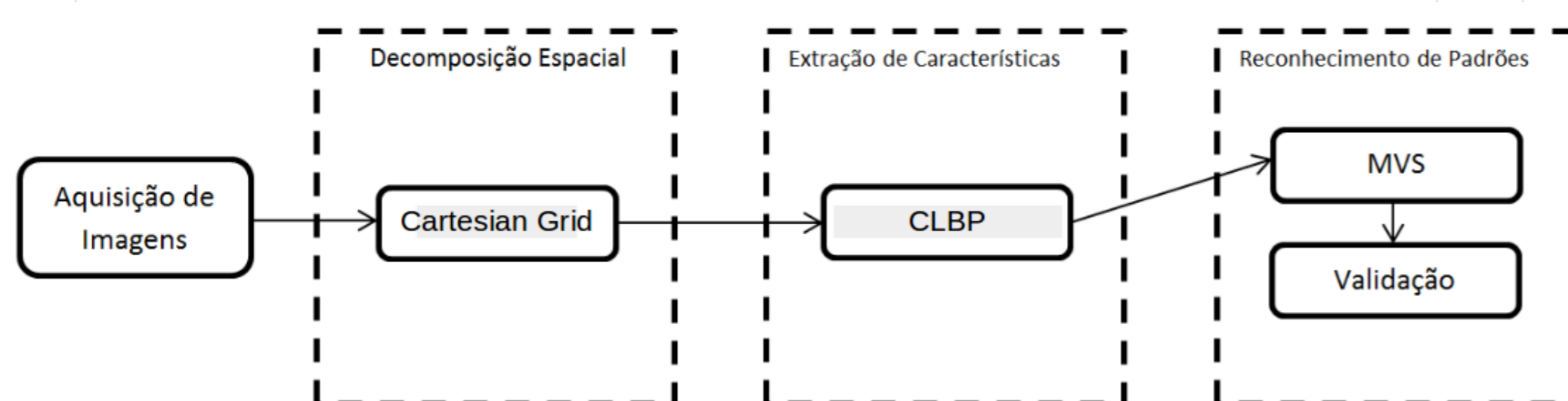


Fig. 2: Representação da metodologia deste estudo

O desempenho do método proposto é avaliado com uma base de imagens bem conhecida, a Base de Dados de Expressão Facial Cohn-Kanade. A base de imagens contém 1224 imagens de faces de um conjunto de seis expressões faciais (Alegria, Aversão, Medo, Raiva, Tristeza e Surpresa), utilizamos além destas um sétimo tipo de expressão denominado Neutro, também contido na base. Neste trabalho utilizamos um subconjunto de 876 Regiões de Interesse, sendo 126 expressões de tristeza e 750 das outras seis expressões. A Figura 3 apresenta exemplos de ROIs utilizadas neste trabalho.



Fig. 3: (a) representa uma expressão facial de tristeza (b) e (c) representam as expressões de aversão e medo respectivamente.

Os experimentos foram realizados de acordo com o fluxo de atividades descrito pela Figura 4. A base de características foi normalizada entre -1 e 1. Esse processo tem como objetivo padronizar a distribuição de valores das variáveis, que podem assumir diferentes domínios. Além disso, busca ajudar o classificador a convergir com maior facilidade na etapa de treinamento. Portanto, utilizou-se o svm-scale presente no pacote libSVM (Este pacote contém a implementação da MVS utilizada neste trabalho) para realizar esta tarefa.

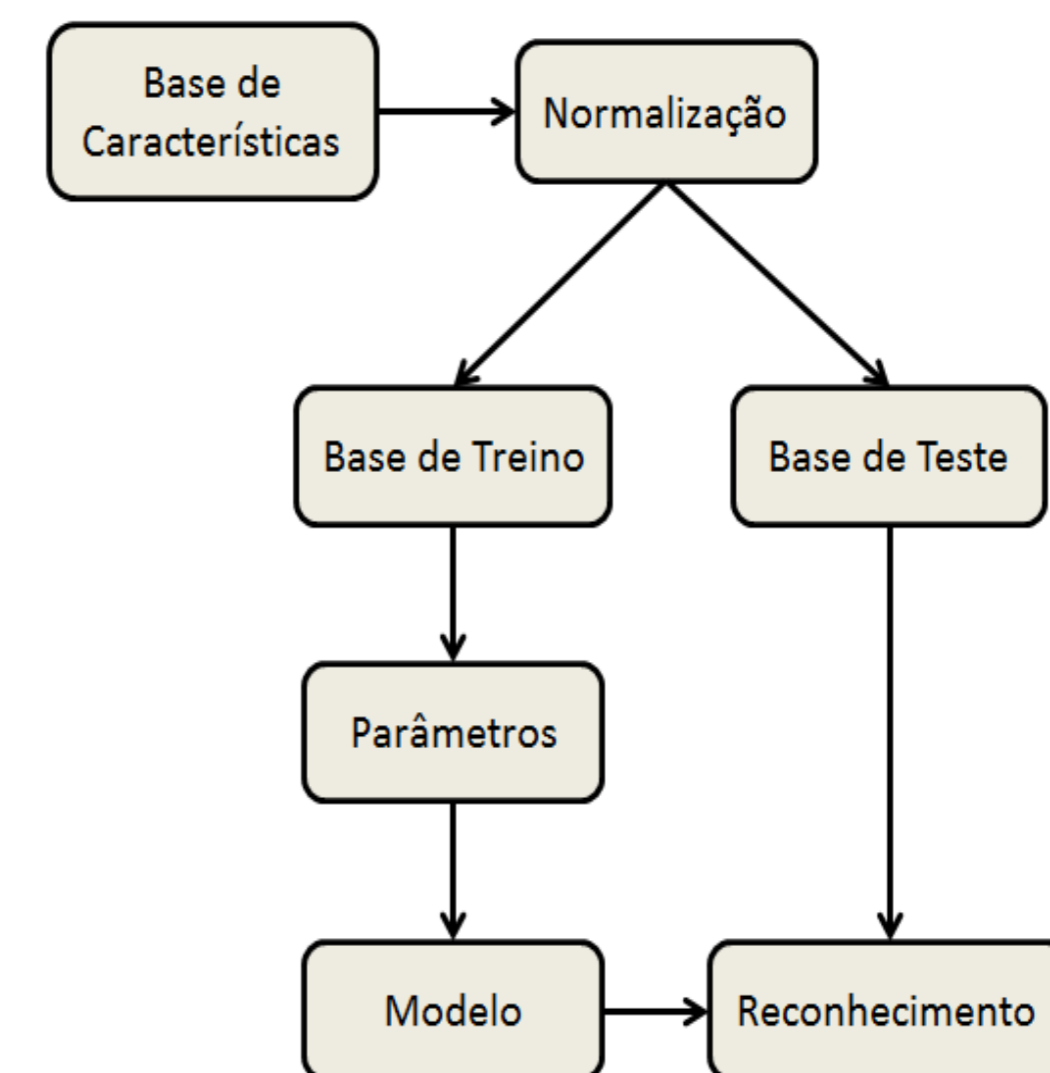


Fig. 9: Fluxo de metodologia para reconhecimento de padrões

Resultados

Para os testes foi utilizada a base Cohn-Kanade [19]. As ROIs foram extraídas usando o Haar Cascade. Os testes foram realizados usando SVM. O SVM foi configurado com núcleo radial e teve o parâmetro de custo ajustado para 8.0 e γ para 2^{-8} . O resultado obtido foi de 81,90% de Acurácia. Os resultados adquiridos se mostram promissores, pois mesmo sem uma escolha otimizada dos parâmetros do CLBP a metodologia obteve bons resultados quando comparada a trabalhos já existentes. Os resultados indicam uma boa discriminação entre expressões de tristeza quando comparada com as outras seis expressões, indicando este trabalho propõe uma boa metodologia e que só tende a aumentar as chances de um reconhecimento bem sucedido.

Conclusões e Próximos Passos

Este trabalho apresentou uma metodologia para descrição dos padrões de expressões faciais usando como extrator de características o CLBP e a técnica de divisão espacial *Cartesian Grid* 3×3 . A utilização da *Cartesian Grid* forneceu certo grau de localização espacial quando utilizada com o CLBP. Se utilizarmos somente o CLBP sem a utilização de um divisor espacial os resultados teriam uma queda de desempenho graças à falta de informações espaciais válidas.

Referências

- [1] da Rocha, S. V. (2014). Diferenciação do Padrão de Malignidade e Benignidade de Massas em Imagens de Mamografias Usando Padrões Locais Binários, Geoestatística e Índices de Diversidade. PhD thesis, Universidade Federal do Maranhão.
- [2] Gonzalez, R. C. and Woods, R. E. (2006). Digital Image Processing. Pearson Prentice Hall, Upper Saddle River, NJ, USA.
- [3] Ahmed, F., Hossain, E., Bari, A., and Shihavuddin, A. (2011). Compound local binary pattern (clbp) for robust facial expression recognition. In Computational Intelligence and Informatics (CINTI), 2011 IEEE 12th International Symposium on, pages 391–395, Obuda University, Budapest, Hungary. IEEE.
- [4] Kanade, T., Cohn, J. F., and Tian, Y. (2000). Comprehensive database for facial expression analysis. In Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on, pages 46–53. IEEE