Reconhecimento de Expressões Faciais em Imagens

Elizabeth Duane Santos da Costa
Universidade Federal de Minas Gerais
elizabeth.duane@dcc.ufmg.br

Resumo

Este trabalho apresenta dois métodos para reconhecimento de expressões faciais: um através do cálculo de Eingenfaces e outro através do método de aprendizado de Máquina de Vetores de Suporte. O trabalho restringe-se na classificação de imagens estáticas. Os resultados dos experimentos mostraram que as abordagens podem ser aplicadas para o reconhecimento de expressões faciais com uma significante taxa de reconhecimento.

1. Introduction

O reconhecimento de expressões facias, estudado inicialmente no campo da psicologia, ganhou campo de estudo na computação principalmente para trazer soluções baseadas em IHC (Interação Humano-Computador) e computação pervasiva, onde pretende-se uma interação com o computador de forma transparente. É uma tarefa desafiadora, uma vez que a classificação depende de um julgamento que é ambíguo até para os seres humanos e a mesma expressão pode variar de pessoa para pessoa [8].

O processo de reconhecimento de expressão facial envolte três etapas que são: detectar e localizar faces na cena, extrair as características faciais e classificar a imagem em uma determinada expressão [4, 11].

As abordagens dos sistemas de reconhecimento de expressões faciais apontam métodos diferentes para reconhecimento em imagens estáticas [8, 11] e vídeos [2, 1, 6]. Conforme apresentado em [4], o processo de extração de características envolvem abordagens holísticas e locais. Na abordagem holística, um modelo pode ser todos os pixels da imagem ou um vetor de características que identifica a imagem como um todo e a abordagem local analisa características específicas da face. Para a classificação pode-se utilizar alguns métodos tais como Redes Neurais Artificiais, Máquinas de Vetores de Suporte, Rede de Bayes, sitemas de regras e distância euclidiana [4, 10].

A classificação das expressões normalmente é baseada no conjunto de emoções básicas que caracterizam de forma única uma expressão facial e são: alegria, desgosto, medo, raiva, surpresa, tristeza, neutra como mostrado na Figura 1 [4].

Figura 1. Exemplo das seis expressões básicas e da face neutra: alegria, desgosto, medo, raiva, surpresa, tristeza, neutra (esquerda para direita)

Neste trabalho pretende-se avaliar dois métodos para o reconhecimento de expressões faciais: utilização de Eigenface e o método de aprendizagem por Máquinas de Vetores de Suporte (SVM - Support Vector Machine). O objetivo do trabalho é verificar se os métodos abordados atendem ao problema de classificar corretamente uma determinada imagem de acordo com as expressões descritas anteriormente. O restante do trabalho segue da seguinte

A artigo está organizado como segue: Na Seção 2 são apresentados os trabalho relacionados e na Seção 3 uma breve introdução de SVM e Eingenface. Na Seção 4 é apresentada a metodologia utilizada neste trabalho e na Seção5 os experimentos e resultados. Finalizando, a Seçao 6 apresenta a conclusão e trabalhos futuros.

2. Trabalhos Relacionados

Com a abordagem de classificação em imagens estáticas, os autores do trabalho [11] salientam que o processo de percepção humana envolve tanto a análise global como local da face e propõe um método combinando as abordagens

de reconhecimento por FACS (Facial Action Coding System), com dois conjuntos de características que completam a caracterização de cada expressão. Utilizando o método estatístico AAM (Active Appearence Model) combinam a forma e a textura dos rostos, gerando um modelo de aparência. São gerados três conjuntos: FACS, EDU (Expression Descriptive Units) e Vetor de Aparência. FACS e EDU são uma média local das características e o Vetor de Aparência apresenta uma informação global da imagem. O modelo apresentado utiliza a informação dos três conjuntos e apresenta com taxa de . Comparando com uma abordagem implementada com Rede Neural a abordagem proposta teve melhores resultados na classificação de expressões

O trabalho de [7] apresenta uma implementação automática para inferir emoção através da classificação de expressão. O sistema recebe como entrada uma sequência de imagens e a classificação é feita em tempo-real. O processo de classificação envolve a detecção e rastreamento da face utilizando fluxo óptico e os parâmetros detectados são introduzidos na Máquina de Vetores de Suporte para análise e classificação. A partir dos resultados da classificação, o sistema apresenta um *e*motion dependendo do estado emocional identificado.

O trabalho apresentado em [8] utiliza o método PCA e Rede Neural para classificar as imagens em um conjunto de sete expressões (seis expressões básicas + face neutra).

Propõe-se um framework utilizando uma rede de função de base radial hierárquica em que as características locais são extraídas utilizando o método PCA. Os autores utilizaram apenas as características dos olhos e boca. Os testes realizados mostraram que com apenas essa duas características e com o modelo proposto pelos, foi possível classificar as imagens de acordo com as expressões definidas.

3. Eigenfaces e SVM

3.1. Eigenfaces

Foi proposto por Turk e Pentland [12] e consiste em calcular uma imagem que melhor represente um dado conjundo de imagens. Uma eigenface contém as características de cada imagem do grupo que representa. Utiliza a técnica de PCA para gerar um conjunto de vetores que são utilizados para a detecção de padrões em imagens. Eigenfaces é considerado o primeiro caso de sucesso para a tecnologia de reconhecimento facial [12].

O PCA é responsável por treinar uma base de dados de imagens, transformando as informações visuais em vetores, que representam os pontos mais marcantes na imagem (auto-vetores). Como o conjunto de imagens proposto são faces, os auto-vetores são chamados de Eigenface. Estes vetores são projetados no subspaço multidimensional. Para classificação de uma nova imagem, calcula-se a similaridade entre a projeção da imagem de teste e da imagem de treino através da distância euclidiana ou as eigenfaces escolhidas podem ser usados como entrada para algum algoritmo de aprendizagem. Os passos básicos para classificação com o método de eigenface são descritos abaixo:

- Dado um conjunto de imagens de treinamento T, subtrair o meio (S = T - média(T)).
- Calcule a matriz de covariância de S
- Calcule os auto-vetores e os auto-valores desta matriz de covariância
- Escolha os componentes principais.
- Projete as imagens de treinamento e teste no subspaço multidimensional

A Figura 2 apresenta um exemplo de quatro eigenfaces calculadas do banco de dados JAFFE.

Figura 2. Eigenfaces

3.2. Máquinas de Vetores de Suporte

Maquinas de Vetores de Suporte é um método de classificação estatística proposto por Vapnik [13] e tem como princípio a minimização do risco estrutural para o treinamento de classificadores. O método tem sido usado para vários problemas de reconhecimento [3, 7]. SVM é um tipo de aprendizado de máquina supervisionado.

Dado um conjunto $(x_1, y_1), (x_2, y_2), ..., (x_l, y_l)$, a idéia básica do SVM é determinar uma função hiperplano que separa os dados em duas classes[3]. É definido uma margem com a menor distância que separa as classes. O SVM procura separar as classes com a maior margem possível como apresentado na Figura 3.

O método utiliza funções de decisão (função kernel) para ajustar os dados em uma das classes determinadas. A função kernel pode ser linear, polinomial, sigmoid entre outras [3].

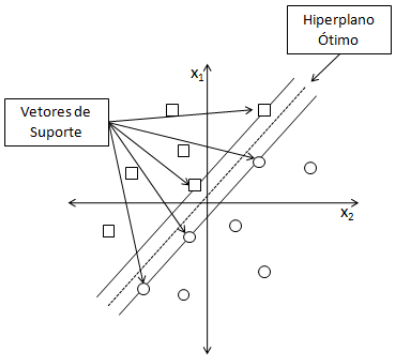

Figura 3. O algoritmo SVM determina o hiperplano que maximiza a distância entre os vetores de suporte [7]

4. Metodologia

A Figura 4 apresenta o processo de classificação utilizando os dois métodos propostos no trabalho atual.

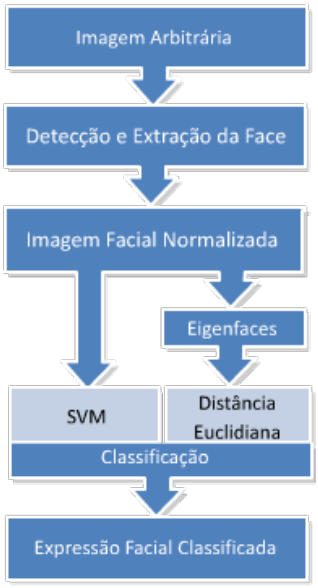

Figura 4. Processo de Classificação

Inicialmente, dado uma imagem arbitrária, é necessário detectar a face para posterior reconhecimento da expressão. O detector de face mais utilizado no processo de reconhecimento de expressões é o proposto por Viola and Jones [14]. A imagem é subdividida em partes que são processadas por classificadores em cascata. Os classificadores são definidos de acordo com o objeto que se deseja localizar e são treinados pelo método AdaBoost. Para localização da face, há vários classificadores disponíveis. Para este trabalho utilizou-se os classificadores disponíveis na biblioteca OpenCV e o método HaarCascade. Um dos resultados da detecção é apresentado na Figura 5.

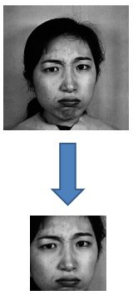

Figura 5. Extração da Face

Os próximos passos do processo foram implementado no MATLAB 7.8.0.

As imagens das faces detectadas foram convoluidas com um filtro gaussiano para atenuação do ruído e a dimensão de todas as imagens foi ajusta. As imagens foram introduzidas como entrada para o método utilizando Eigenface e SVM.

O processo de classificação utilizando Eigenface e SVM compreende os seguintes passos:

- Eigenface
  - Cálculo das Eigenfaces como proposto na Seção 3.1
  - Cálculo da distância euclidiana entre a projeção das imagens de teste e das imagens de treino
  - A menor distância é escolhida
  - Classificação: a imagem de teste estará no mesmo grupo da imagem de treino que obteve menor distância.
- SVM
  - As imagens são vetorizadas formando uma matriz de treinamento onde cada coluna representa uma característica e cada linha uma instância de treinamento
  - Criação de um vetor de resposta que classifica cada linha em uma determinada classe.
  - O treinamento é conduzido de acordo com um função kernel. Para este treinamento foi aplicada a função de kernel linear.

O SVM é um classificador binário e o problema abordado é multiclasse. A forma mais utilizada para esse problema é a decomposição do problema multiclasse em um problema binário [9]. Para se fazer o treinamento através da decomposição utilizou-se o

método “um-contra-todos”. As imagens são divididas em duas classes: a classe treinada e todas outras imagens, logo o treinamente é feito para as K classes do problema. A classificação é feita para cada classe individualmente e após uma votação, a classe com maior pontuação é escolhida.

4.1. Banco de Dados

O conjunto de imagens escolhido para os experimentos foi o JAFFE (Japanese Female Facial Expressions)[5]. O banco de dados contém 213 imagens de 7 expressões faciais (6 expressões básicas + neutra) referentes a 10 modelos japonesas. As imagens são no formato “tiff” em escala de cinza e dimensão 256 x 256. As Figura 1 apresentada na Seção 1 mostra um conjunto de imagens do banco de dados JAFFE.

5. Experimentos e Resultados

- Eigenfaces

  Os experimentos utilizando o método de Eigenfaces constitui no conjunto de 143 imagens para treinamento e 69 imagens para teste. As imagens foram redimensionadas para 100 x 100 pixels. A Tabela 1 apresenta os resultados obtidos utilizando o método.

ER/ERC	**R**	**D**	**M**	**A**	**N**	**T**	**S**
Raiva (R)	**72.7**	9,1	0,0	9,1	0,0	9,1	0,0
Desgoto (D)	18,2	**72,7**	9,1	0,0	0,0	0,0	0,0
Medo (M)	0,0	0,0	**69,2**	0,0	7,7	15,4	7,7
Alegria (A)	0,0	0,0	0,0	**66.7**	16,7	16,7	0,0
Neutro (N)	0,0	0,0	9,1	27,3	**63,6**	0,0	0,0
Tristeza (T)	0,0	11,1	0,0	11,1	11,1	**66,7**	0,0
Surpresa (S)	0,0	0,0	0,0	0,0	0,0	0,0	**100,0**

Unidade dos dados: (%)

Tabela 1. Matriz de confusão representando os valores médios do reconhecimento usando Eigenfaces. ER e ERC representa a expressão de referência e a expressão reconhecida, respectivamente. A taxa média geral de reconhecimento foi 73,1%.

A taxa de reconhecimento com o método de Eigenface ficou um abaixo dos trabalhos de referência (80% -95%). O método obteve um alto índice para classificação da expressão “Surpresa”. Analisando o banco de dados, as imagens da expressão “Surpresa” apresentam uma grande similaridade, por isso o alta taxa de reconhecimento. Porém algumas expressões causam confusão ao classificador tais como “Neutro” e “Alegre” e “Desgosto” e “Raiva”.

- SVM

  Para a classificação com SVM, utilizou-se um conjunto de 139 imagens de treino e 69 imagens de teste. As imagens foram redimensionadas para 50 x 50 pixels. O conjunto de imagens foi subdividido em 3 partes utilizando a validação cruzada K-folds. A Tabela 2 apresenta os resultados obtidos utilizando o método de aprendizagem com SVM. A taxa de reconhecimento foi 82,2%.

ER/ERC	**R**	**D**	**M**	**A**	**N**	**T**	**S**
Raiva (R)	**96,7**	0,0	3,3	0,0	0,0	0,0	0,0
Desgoto (D)	16,7	**76,7**	3,3	0,0	0,0	3,3	0,0
Medo (M)	26,7	6,7	**66,7**	0,0	0,0	0,0	0,0
Alegria (A)	13,3	3,3	0,0	**80,0**	0,0	3,3	0,0
Neutro (N)	13,3	0,0	3,3	0,0	**83,4**	0,0	0,0
Tristeza (T)	13,3	6,7	3,3	3,3	10,0	**63,4**	0,0
Surpresa (S)	3,3	0,0	3,3	3,3	0,0	0,0	**90,0**

Unidade dos dados: (%)

Tabela 2. Matriz de confusão representando os valores médios encontrados através do uso de SVM. ER e ERC representa a expressão de referência e a expressão reconhecida, respectivamente. A taxa média geral de reconhecimento foi 82,2%.

O método de aprendizagem com SVM apresenta uma taxa de reconhecimento satisfatória, próxima dos trabalhos de referência. Para a classificação individual, o método apresenta algumas taxas acima da . As expressões que causaram maior confusão ao classificador foram “Raiva” e “Medo”.

6. Conclusão e Trabalhos Futuros

O trabalho apresentou dois métodos aplicados ao problema de reconhecimento de expressões faciais. Com os métodos apresentados obteve-se uma taxa de reconhecimento aceitável com uma vantagem para o método do SVM. Ambos métodos são fáceis de serem implementados. Para algumas expressões específicas os métodos apresentaram uma taxa de reconhecimento de 100%, porém outras expressões causam dúvidas aos classificadores, tais como “Medo” e “Desgosto”.

O método de classificação deve ser testado em outros bancos de dados, uma vez que os modelos do presente trabalho foram apenas japonesas.

Como trabalhos futuros, pretende-se utilizar outros métodos de aprendizado, principalmente os que permitem o aprendizado on-line, onde o sistema após a fase de treinamente continuar aprendendo a cada iteração.

Pretende-se também a construção de um sistema de reconhecimento automático para sequência de imagens para aplicação em sistemas inteligentes com o objetivo de melhorar a interface Humano-Computador.

Referências

[1] M. J. Black and Y. Yacoob. Recognizing facial expressions in image sequences using local parameterized models of image motion. *International Journal of Computer Vision*, 25(1):23–48, Oct. 1997.

[2] L. S. Davis and Y. Yacoob. Recognizing human facial expressions. In *Image Understanding Workshop*, pages II:827–835, 1994.

[3] M. Dumas. Emotional expression recognition using support vector machines, June 19 2001.

[4] B. Fasel and J. Luettin. Automatic facial expression analysis: A survey, 2002.

[5] M. Kamachi, M. Lyons, and J. Gyoba. The japanese female facial expression (jaffe) database. http://kasrl.org/jaffe.html.

[6] S. Kumano, K. Otsuka, J. Yamato, E. Maeda, and Y. Sato. Pose-invariant facial expression recognition using variable-intensity templates. *International Journal of Computer Vision*, 83(2):xx–yy, June 2009.

[7] B. Lee, J. Chun, and P. Park. Classification of facial expression using SVM for emotion care service system. In *SNPD*, pages 8–12. IEEE Computer Society, 2008.

[8] D.-T. Lin. Facial expression classification using PCA and hierarchical radial basis function network. *J. Inf. Sci. Eng*, 22(5):1033–1046, 2006.

[9] E. M. C. Pimenta. Abordagens para decomposição de problemas multiclasse: os códigos de correção de erros de saída. Master's thesis, Faculdade de Ciência da Universidade do Porto, Dec. 2004.

[10] S. V. P. G. Rosário. Facial emoticons: Reprodução de informação associada a expressões faciais por via do seu reconhecimento. Engenharia informática e de computadores, Universidade Técnica de Lisboa, Dec. 2008.

[11] M. Sorci, J. P. Thiran, J. Cruz, T. Robin, and M. Bierlaire. Modelling human perception of static facial expressions. In *FG*, pages 1–8, 2008.

[12] M. Turk and A. Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1):71–86, 1991.

[13] V. Vapnik. *Statistical Learning Theory*. Wiley, 1998.

[14] P. Viola and M. J. Jones. Robust real time object detection. In *Workshop on Statistical and Computational Theories of Vision*, pages xx–yy, 2001.