

Extração de Atributos para Reconhecimento de Expressões Faciais

Vinícius Assis Saueia da Silva

1

1. Introdução

Interação Homem-Computador é uma das áreas da computação que procuram fazer com que o ser humano tenha uma interação cada vez mais natural com um aparelho computacional, podendo ser através da fala [Katevas et al. 1997, Simpson and Levine 1997], da utilização de sensores, através de dispositivos de captura de imagens (*web-cams*)[Fernandes et al. 2005], etc.

Hoje em dia, a interação com um computador ainda é, de certa maneira, pouco abrangente em relação às necessidades especiais que uma pessoa pode ter. Uma pessoa com dificuldades de movimentar os membros superiores pode sentir dificuldades no manuseio de interfaces que utilizam de maneira exaustiva esses membros, como por exemplo mouses e teclados.

Dessa maneira, a idéia de termos uma ferramenta computacional que entenda comandos gerados através de gestos ou expressões faciais pode ser muito útil para muitas pessoas que sentem dificuldades em interagir com as formas de interfaceamento que são utilizadas hoje em dia. O reconhecimento de expressões faciais também pode ser utilizado para reconhecimento de sinais usados na linguagem de surdos. A possibilidade de se reconhecer expressões faciais também pode ajudar tetraplégicos na construção de uma cadeira de rodas inteligente guiada por expressões faciais, dessa maneira a cadeira de rodas é controlada somente por diferentes expressões do usuário.

Este trabalho tem como objetivo a construção de uma ferramenta capaz de extrair atributos de imagens que contenham faces, para serem utilizados na tarefa de reconhecimento de expressões faciais. Para conseguir realizar o objetivo serão implementados extratores de atributos utilizando diferentes técnicas estudadas e utilizar técnicas já desenvolvidas de seleção de atributos e aprendizagem de máquina para realizar os experimentos necessários ao problema. Os experimentos serão realizados em um conjunto de 500 imagens de faces, contendo diferentes tipos de expressões e diferentes tons de peles, e então será utilizado um parte dessas imagens para ser utilizado no classificador e a outra parte será classificada. Depois de coletar os dados será feita a análise dos resultados e comparados com os resultados de outros trabalhos correlatos.

Por fim, os módulos desenvolvidos neste projeto serão integrados ao projeto SIGUS¹ (Plataforma de Apoio ao Desenvolvimento de Sistemas para Inclusão Digital de Pessoas com Necessidades Especiais) que tem como principal objetivo criar aplicações computacionais capazes de melhorar a qualidade de vida e inclusão digital de pessoas com necessidades especiais.

¹Disponível gratuitamente em <http://www.ec.ucdb.br/pistori/sigus/>

2. Objetivos

2.1. Geral

Desenvolver um módulo para o projeto SIGUS, utilizando códigos-fonte livre, capaz de extrair atributos de imagens da face recebidas através de um dispositivo de captura de imagens, para serem utilizados no reconhecimento de expressões faciais.

2.2. Específicos

1. Implementar extratores de atributos baseados em diferentes métodos.
2. Analisar seletores de atributos aplicáveis ao problema.
3. Analisar técnicas de aprendizagem de máquina.
4. Integração do módulo desenvolvido ao projeto SIGUS.

3. Revisão de Literatura

3.1. Extração de Atributos

Os atributos de uma imagem são as centenas de características apresentadas em uma imagem. Os atributos podem apresentar variações de uma imagem para outra. Essa variação pode modificar a imagem, isto é, com essa variação a apresentação final da imagem é diferente. E através dos atributos de uma imagem então é possível reconhecer vários padrões, como expressões faciais.

Existem várias maneiras de se extrair atributos de uma imagem. Esses podem ser extraídos através de uma análise estatística da imagem calculando vários dados estatísticos da imagem, como centro de massa, excentricidade, distribuição dos pixels, etc [Souza and Pistori 2005, li Tian et al. 2001]. Ou então, pode ser também extraído atributos com base no movimento de algumas regiões da face, como os olhos, boca e sobrancelhas, estas regiões são características da face que são bastante determinantes no reconhecimento de expressões faciais [Essa and Pentland 1995].

3.2. Seleção de Atributos

Uma centena de atributos podem ser extraídos de uma imagem. Algumas vezes um número muito grande de atributos podem ser prejudiciais a alguns algoritmos de aprendizagem automática [Martins et al. 2004]. Podendo gerar conseqüências de erros de classificação e também do aumento do tempo de processamento do algoritmo.

Por isso, muitas vezes é importante fazer uma seleção dos atributos extraídos dando preferência a atributos que possam ser mais importantes no diferenciamento de padrões. Para se escolher esses atributos existem algumas técnicas que fazem uma análise da amostra e assim escolhem os atributos que, para aquela amostra, são mais importantes.

Uma técnica que está sendo utilizada para seleção de atributos no reconhecimento de expressões faciais é a análise discriminante por Campos em [de Campos 2001]. A análise discriminante consistem em uma combinação linear das componentes extraídas procurando encontrar as variáveis mais discriminantes [de Campos 2001].

3.3. Aprendizagem Automática e Reconhecimento de Expressões

Para se reconhecer expressões faciais em uma imagem é possível utilizar fazer uso de várias técnicas de aprendizagem automática ou através de modelos matemáticos que conseguem modelar o problema, como os Modelos de Markov Ocultos. Neste trabalho espera-se utilizar técnicas que utilizam a aprendizagem de máquinas.

Através da aprendizagem automática pode-se criar um classificador com base nos atributos extraídos e, posteriormente, com os atributos selecionados para que possa se reconhecer expressões faciais em imagens. Para isto é necessário utilizar técnicas aprendizagem supervisionada, existem várias técnicas cabíveis ao problema como Redes Neurais Artificiais, Máquina Vetor de Suporte, etc.

Redes Neurais Artificiais é uma técnica clássica na área de aprendizagem de máquina baseada no funcionamento do cérebro humano, cria-se uma rede de neurônios artificiais e, de certa maneira, os ensina a resolver problemas [de Leon F. de Carvalho et al. 2000]. Redes Neurais é uma técnica que tem se mostrado bastante eficiente em muitos problemas em computação porém pode apresentar uma complexidade computacional alta dependendo dos dados na base de treinamento.

A técnica Máquina de Vetor de Suporte é uma técnica de aprendizagem de máquina utilizada para bases de treinamento com dados polinomiais. A idéia principal da técnica consiste em separar as classes em espaços que melhor definem. [Osuna et al. 1997]

Para o reconhecimento de expressões utilizando a técnica de Modelos de Markov Ocultos são calculados as probabilidades de se ocorrer tais expressões e essa probabilidade que é levada em conta na hora se fazer o reconhecimento. A expressão de maior probabilidade de ocorrer que é utilizada como critério de decisão.[Lien 1998].

Através de uma parametrização local pode-se separar a face em regiões rígidas da face e as regiões não-rígidas da face. As regiões rígidas são aquelas que variam pouco conforme uma expressão é executada, tais como posição da cabeça, contorno da face, etc. As regiões não-rígidas são aquelas que tem uma variação mais significativa de acordo com a expressão, como os olhos, a boca e as sobrancelhas. Faz-se essa divisão na imagem para facilitar o reconhecimento. [Black and Yacoob 1997]

4. Metodologia

Para o desenvolvimento do protótipo em um curto prazo de tempo, serão reaproveitados alguns pacotes livres já existentes, como ImageJ, para processamento digital de sinais e o Weka, para aprendizagem de máquina. Os pacotes citados acima são constituídos de programas na linguagem Java, o que facilitará a portabilidade do protótipo para outros sistemas operacionais e até para dispositivos móveis.

Seguem abaixo as etapas metodológicas relacionadas com com cada um dos objetivos específicos.

1. Implementar extratores de atributos baseados em diferentes métodos.
 - (a) Estudo de técnicas de extração de atributos.
 - (b) Criação de programas em Java extratores de atributos baseados nas técnicas estudadas.

- (c) Criação de um banco de imagens, com 500 imagens que contenham face, com diferentes expressões e diferentes tons de peles.
 - (d) Realização testes do extratores de atributos utilizando as imagens do banco de imagens.
 - (e) Análise manualmente os resultados obtidos.
2. Analisar seletores de atributos aplicáveis ao problema.
- (a) Estudo de técnicas específicas de seleção de atributos para reconhecimento de expressões faciais em trabalhos correlatos a este.
 - (b) Utilização de ferramentas, de preferência as que já estão prontas ou sendo desenvolvidas no projeto SIGUS, que façam a seleção dos atributos extraídos pelas técnicas implementadas nesse trabalho.
 - (c) Extrair, com todas as técnicas implementadas, e selecionar atributos de um conjunto de 500 imagens de faces.
 - (d) Geração de bases de treinamento com base nos atributos selecionados para alimentar um classificador de aprendizagem automática.
3. Analisar técnicas de aprendizagem de máquina.
- (a) Estudo de técnicas de aprendizagem automáticas de acordo com técnicas já citadas em trabalhos correlatos.
 - (b) Marcação manualmente do conjunto de 500 imagens utilizados e classificar manualmente as expressões das imagens disponíveis.
 - (c) Escolha de 3 ou mais métodos de extração e seleção de atributos.
 - (d) Escolha de diferentes métodos/ferramentas de aprendizagem de máquina.
 - (e) Separar manualmente 40% das imagens do banco de imagens para base do classificador e 60% das imagens para serem classificadas automaticamente.
 - (f) Treinar o classificador com as imagens separadas.
 - (g) Realização de testes com os diferentes algoritmos de extração e seleção de atributos e com as diferentes técnicas de aprendizagem automática, utilizando as imagens do banco de imagens separadas.
 - (h) Análise dos resultados obtidos.
4. Integração do módulo desenvolvido ao projeto SIGUS.
- (a) Integração do(s) módulo(s) desenvolvidos ao projeto SIGUS.
 - (b) Comparação das técnica escolhidas com outras técnicas através dos resultados obtidos na fase de classificação das expressões com as de outros trabalhos.
 - (c) Análise dos resultados obtidos comparando os resultados obtidos com resultados obtidos em outros trabalhos através da contagem do número de classificações corretas apresentadas por cada uma dos trabalhos.

Referências

- Black, M. J. and Yacoob, Y. (1997). Recognizing facial expressions in image sequences using local parameterized models of image motion. *IJCV*, 25(1):23–48.
- de Campos, T. E. (2001). Técnicas de seleção de características com aplicações em reconhecimento de faces. Master's thesis, Universidade de São Paulo (USP).
- de Leon F. de Carvalho, A. C. P., de Padua Braga, A., and Ludermir, T. B. (2000). *Redes Neurais Artificiais: Teoria e Aplicações*. LTC, 1 edition.
- Essa, I. A. and Pentland, A. (1995). Facial expression recognition using a dynamic model and motion energy. In *ICCV*, pages 360–367.
- Fernandes, B. P., da Silva, V. A. S., and Pistori, H. (2005). Protótipo de um simulador para cadeiras de rodas guiadas por expressões faciais: Estudos preliminares. *III Congresso Catarinense de Software Livre (SOLISC)*.
- Katevas, N. L., Sgouros, N. M., Tzafestas, S. G., Papakonstantinou, G., Beattie, P., Bishop, J. M., Tsanakas, P., and Koutsouris, D. (1997). The autonomous mobile robot scenario: A sensor-aided intelligent navigation system for powered wheelchairs. *IEEE Robotics and Automation*, 4:60–70.
- li Tian, Y., Kanade, T., and Cohn, J. F. (2001). Recognizing action units for facial expression analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(2):97–115.
- Lien, J. (1998). Automatic recognition of facial expressions using hidden markov models and estimation of expression intensity.
- Martins, C. A., Monard, M. C., and Matsubara, E. T. (2004). Uma metodologia para auxiliar na seleção de atributos relevantes usados por algoritmos de aprendizado no processo de classificação de textos. In Solar, M., Fernández-Baca, D., and Cuadros-Vargas, E., editors, *30ma Conferencia Latinoamericana de Informática (CLEI2004)*, pages 21–32. Sociedad Peruana de Computación. ISBN 9972-9876-2-0.
- Osuna, E., Freund, R., and Girosi, F. (1997). Support vector machines: Training and applications. Technical Report AIM-1602.
- Simpson, R. C. and Levine, S. P. (1997). Adaptive shared control of a smart wheelchair operated by voice control. *International Conference on Intelligent Robots and Systems*, 2:622–626.
- Souza, K. P. and Pistori, H. (2005). Implementação de um extrator de características baseado em momentos da imagem. *XVIII Brazilian Symposium on Computer Graphics and Image Processing*.