

## 2 Trabalhos Relacionados

Este capítulo está dividido em duas partes, a primeira abarcando estudos sobre as emoções vistas sob um enfoque físico e psicológico. A segunda faz a análise de estudos similares ao trabalho proposto que foram encontrados na literatura.

### 2.1 Aspectos Gerais das Emoções

Darwin(1872) definiu três princípios gerais das expressões, sendo o terceiro o mais relevante para este trabalho: o princípio de ações devido à constituição do sistema nervoso, independente da vontade e até um certo ponto, do hábito. “Esse princípio está associado a estados de espírito no qual quando sensores são fortemente excitados, força nervosa é produzida em excesso e transmitida em certas direções dependentes da conexão das células nervosas e de acordo com o sistema muscular ...” (13, p. 28-29). De tal maneira, diante de uma grande excitação do sistema nervoso, o corpo reage naturalmente através de expressões, de acordo com a própria natureza física do organismo. Em concordância, Jenkis(1998) diz que emoções são repentinas e causam o interrompimento das ações atuais para que possamos lidar com essa nova situação que acarretou tal descarga emocional (20).

Para entender o reconhecimento de emoções, antes é pertinente fazer uma definição coerente sobre o que é uma emoção. De acordo com Adolphs(2002), “[.] neurobiologistas e psicologistas criaram o conceito de emoção como uma mudança concertada, geralmente em fases adaptativas em vários sistemas fisiológicos (incluindo ambos componentes somáticos e neurais) em resposta ao valor de um estímulo”, (1, p. 24).

Para Ekman(1993), sobre a classificação de emoções, “[.] as emoções devem ser consideradas como estados discretos (assim como as que temos nomeadas na linguagem natural e que são definidas como rótulos para expressões faciais)” (14). Russel(1980) também acreditava na categorização das emoções e descreveu um espaço bidimensional de emoções, Figura 2.1, no qual os eixos

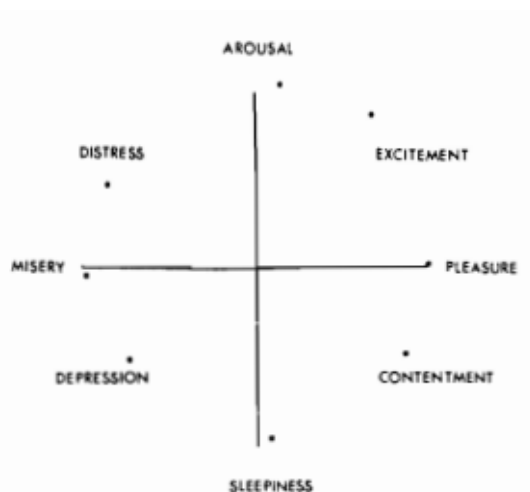


Figura 2.1: Espaço bidimensional de emoções, eixos representam prazer e excitação. Figura de (36).

representam prazer e excitação,(36).

Portanto, a abordagem seguida será em acordo com as pesquisas de Ekman(*op. cit.*) e Russel(*op. cit.*), tomando as emoções básicas como passíveis de serem classificadas discretamente. Existem emoções compostas que não se encaixam em apenas uma categoria, são uma mescla de várias emoções básicas, mas essas não se encontram no escopo do trabalho. As emoções básicas, com uma considerável conformidade entre pesquisadores são: felicidade, tristeza, raiva, medo, surpresa e desgosto.

## 2.2 Reconhecimento Automático de Emoções

Introduzido o ambiente psicológico, parte-se para uma abordagem que contempla o ponto de vista da área da informática, sendo feita, portanto, uma análise de estudos importantes da academia sobre o reconhecimento de emoções de maneira automática. No final desse capítulo, a Tabela 2.1 exhibe os resultados que foram publicados de alguns destes estudos.

O reconhecimento automático de emoções vem sendo assunto de pesquisa desde as três últimas décadas e, apesar de a expressão facial transmitir muita informação sobre um possível estado emocional humano, é importante afirmar que esse não é o único meio de se tentar reconhecer uma emoção. Para citar apenas outra ramificação, estudos (6) (33) (16) vêm utilizando a fala como a fonte das informações relacionadas a um estado emocional.

Além disso, pesquisas com diferentes objetivos de classificação de expressões também foram feitas. Por exemplo, a pesquisa de Nosu *et al.* (2006) (32) tem como enfoque auxiliar no processo de *e-learning* e, por isso, trabalha

com emoções diferentes das seis universais. Naquele estudo, os autores buscaram identificar se um aluno está achando a matéria fácil ou difícil e se está concentrado ou entediado, entre outros comportamentos importantes para o âmbito de aprendizado à distância.

Dentro da análise de emoções através da visualização de expressões faciais, existem também pelo menos dois caminhos. Com o aumento do poder computacional, certas pesquisas têm sido feitas sobre visualizações tridimensionais do rosto de um indivíduo (42). Outras seguem a análise de emoções em imagens bidimensionais, fotos comuns. Existem também estudos que mesclam imagens 2D com sensores 3D (35). O caminho das imagens bidimensionais será mais detalhado, pois é por meio das mesmas técnicas que o presente trabalho foi concluído.

Bettadapura(2012) (2) escreveu uma extensa revisão bibliográfica onde exhibe importantes trabalhos de reconhecimento automático de emoções. Nesta revisão, Bettadapura (*op. cit.*) não se prende a métodos específicos de extração ou de classificação e dá atenção às várias maneiras que estudiosos encontraram para gerar um modelo que tivesse a capacidade de analisar automaticamente as expressões faciais. A presente dissertação tem um enfoque específico na extração de atributos faciais através do uso do ASM e a classificação com uso de SVM e, por isso, estudos que utilizam métodos similares serão mais detalhados a seguir.

Huang *et al.* (1997) (18) usam o *Point Model Distribution* (PMD), que define a forma de um objeto como uma sequência de pontos de referência. Os atributos são criados através da comparação de pontos faciais específicos em várias fotos. O treinamento do modelo foi feito com o uso de *Principal Component Analysis* (PCA) em fotos, a redução dos PDMs para duas dimensões e um algoritmo de agrupamento (*clustering*).

Huang *et al.* (2010) (19) implementaram um algoritmo de ASM modificado, denominado de *Modified Active Shape Model* (MASM), pelos autores. O MASM se difere do algoritmo tradicional por refinar os pontos de referência através da detecção de bordas, com a finalidade de evitar interferência do ambiente. Os autores, então, executam o ASM em uma imagem, extraindo dela diversos pontos faciais que representam o rosto do indivíduo. A partir dos pontos extraídos, 48 triângulos, que têm como vértices índices pré-definidos dos pontos faciais, são montados e os autores utilizam a área desses triângulos como dados de entrada em um algoritmo genético de otimização, de onde os melhores serão selecionados e treinados em uma rede neural com *back propagation* para o reconhecimento dos padrões de cada emoção.

A pesquisa de Wang *et al.* (2010)(41) faz o uso do algoritmo tradicional de ASM. Logo após extrair os pontos faciais para cada imagem, eles são alinhados com a forma média do ASM utilizando três pontos específicos: os cantos internos dos olhos e a ponta do nariz. Então é feito o cálculo do deslocamento de cada ponto para o seu correspondente na forma média do ASM e os resultados são treinados em um SVM multiclasse. Outra questão interessante dessa pesquisa é o foco em tempo real para utilização em robótica. No sistema desenvolvido por estes autores, a execução do ASM durava em torno de 18.39 ms e a predição 0.729 ms, resultando em um algoritmo bastante eficiente.

Gang *et al.* (2009)(15) utilizam o ASM para a extração de pontos faciais e um SVM multiclasse para a classificação. Apesar da similaridade com o trabalho de Wang *et al.* (*op. cit.*), os atributos escolhidos para treinamento são diferentes, pois Gang *et al.* (*op. cit.*) extraem atributos através da comparação entre o indivíduo com emoção e o indivíduo neutro.

Tariq *et al.* (2011)(40) também utilizam SVM para a predição de emoções mas tomam abordagens distintas para os atributos de aprendizado. O vetor de atributos é uma mistura de *Scale Invariant Feature Transform* (SIFT), Gaussianização Hierárquica e deslocamento de pontos, encontrados com o algoritmo de fluxo ótico. Esses atributos são removidos de regiões importantes do rosto: os olhos, a boca e o nariz. Estes autores também apresentam uma tarefa adicional que consiste em fazer a identificação do indivíduo.

Também existem estudos que trabalham com sequência de imagens, como o de Cohen (2010) *et al.*(7), que tem enfoque na etapa de aprendizado de máquina e propõe um *Hidden Markov Model* (HMM) multinível para a predição de emoções. Nos trabalhos tradicionais de HMM para reconhecimento de emoções, é gerado um modelo de HMM para cada emoção e, na hora da predição, aquele com maior probabilidade é escolhido como a emoção correta. No estudo de Cohen *et al.* (*op. cit.*), uma camada HMM acima da convencional é inserida fazendo com que todas as emoções estejam interligadas, criando um HMM de duas camadas. Desse modo, apenas uma única emoção pode estar acontecendo em um determinado momento e todas possuem caminhos e probabilidades de transição entre elas. É importante ressaltar que o estudo de Cohen *et al.* (*op. cit.*) é citado apenas para tornar o capítulo mais abrangente, porque o método de reconhecimento de emoções proposto não se beneficia da temporalidade provida por sequências de imagens para realizar a predição.

Os estudos de Gang *et al.* (15) e Wang *et al.* (41) nos trazem espaço para reflexão. Os dois utilizam o algoritmo ASM para a extração de atributos e

SVM para a classificação, no entanto eles se distinguem no modo com que o vetor de atributos é construído. Enquanto Gang *et al.* (*op. cit.*) utilizam a comparação entre uma imagem e a forma média do ASM, Wang *et al.* (*op. cit.*) fazem a comparação entre a imagem de um indivíduo e a imagem do mesmo indivíduo com expressões neutras. A comparação entre os métodos de extração de pontos faciais propostos por autores diferentes é muito difícil, mesmo quando usam os mesmos princípios. A razão para esta dificuldade é dupla: primeiro, as implementações são diferentes; segundo, os sistemas não são treinados utilizando o mesmo banco de imagens e anotações. Portanto, a seguinte dúvida se revela: qual forma de comparação seria mais apropriada para uma aplicação que necessita de reconhecimento de emoções? Também é importante ressaltar que uma comparação precisa entre os dois métodos de criação de atributos citados acima não existe enquanto houver diferenças na metodologia dos trabalhos. Esta dissertação se propõe a fazer esta comparação utilizando os mesmos dados para os dois métodos, desde o banco de imagens e suas anotações até o algoritmo de extração e a engenharia de atributos faciais e imagens para teste. Desta maneira, cria-se um ambiente homogêneo e favorável para a comparação.

	Extração de pontos faciais	Atributos	Classificador	Banco de Imagens	Emoções	Acurácia
(19)	MASM	Algoritmo genético	Rede Neural	JAFFE	6	70,24%
(19)	Marcação manual	Algoritmo genético	Rede Neural	JAFFE	6	78,24%
(15)	ASM	Engenharia de atributos	SVM	-	4	89,5%
(40)	Mescla de algoritmos	Engenharia de atributos	SVM	GEMEP-FERA	5	79,8%
(18)	<i>Point distribution model</i>	Engenharia de atributos	Clusterização	-	6	84,41%
(7)	-	-	<i>Multilevel HMM</i>	-	6	82,46%

Tabela 2.1: Resultados alcançados por algumas pesquisas no tema de reconhecimento automático de emoções