

## 2 Reconhecimento Facial

Em termos gerais, o reconhecimento facial é o processo pelo qual se mede o grau de similaridade entre duas imagens faciais com o propósito de identificar a um indivíduo ou de verificar a sua identidade.

Num sistema de reconhecimento a tarefa de identificação visa determinar a identidade de uma pessoa entre registros de um banco de dados. Já a verificação visa determinar se a identidade invocada por um indivíduo é verdadeira.

Em outras palavras, do ponto de vista do reconhecimento facial, a identificação tenta determinar a identidade de uma imagem facial desconhecida (denominada deste ponto em diante no texto pelo termo “prova”), mediante a comparação com outras imagens faciais armazenadas num banco de dados (galeria), fornecendo assim um grau de similaridade a cada comparação. A identidade que se associa à prova é, em geral, aquela que alcança o maior grau de similaridade na comparação.

A identificação supõe a existência da identidade na galeria. A tarefa de verificação, por outro lado, confere se a prova corresponde de fato à identidade alegada, conforme registrada por uma ou mais imagens da galeria.

Em 2002, o NIST<sup>1</sup> por meio de seu programa *The Face Recognition Vendor Tests 2002*<sup>2</sup> (Phillips, et al., 2003), acrescentou uma nova tarefa para os sistemas de reconhecimento facial em vídeo, a qual consiste em fornecer uma lista de identidades (suspeitos) geradas a partir de aplicações de vídeo monitoramento; a lista é formada pelos registros da galeria cujos valores de similaridade superam um determinado limiar.

---

<sup>1</sup> Instituto Nacional de Padrões e Tecnologia, pelo acrônimo em inglês.

<sup>2</sup> Serie de avaliações que medem a capacidade dos sistemas de reconhecimento facial para atender as exigências em aplicações no mundo real a grande escala.

## 2.1. Sistemas de Reconhecimento Facial

De um modo geral, os sistemas de reconhecimento facial automáticos, estão formados por quatro módulos: localização de face, normalização, extração de atributos, e comparação (Li, et al., 2011), conforme é apresentado na Figura 1 e que são descritos nas seções seguintes.

A Figura 1 ilustra o processo pelo qual uma imagem facial é reconhecida de forma automática. Neste processo a imagem ou sequência de vídeo na entrada passa pelo módulo de localização de face, a fim de obter uma imagem facial que logo é transformada pelo módulo de normalização de face, esta imagem transformada (normalizada) permite coletar de forma mais adequada as principais características da face, a qual é conseguida pelo módulo de extração de atributos. Já na correspondência os modelos criados com base nestes atributos são comparados com outros similares previamente cadastrados na galeria.

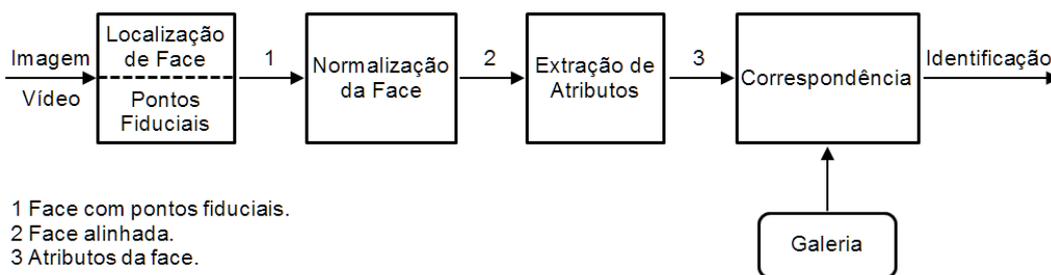


Figura 1. Sistema de Reconhecimento Facial, adaptado de (Li, et al., 2011).

### 2.1.1. Localização de Face

A tarefa de localização de face, usualmente denominada “detecção facial”, procura extrair a região duma imagem na entrada que contém pelo menos uma imagem facial. A partir desta, caso a aplicação o requeira, determinar os pontos referentes aos olhos, nariz, boca, entre outros, denominados “pontos fiduciais” (ver Figura 2).

Quando se trata do reconhecimento facial em vídeo, há ainda mais um passo não apresentado na Figura 1, que consiste em rastrear no vídeo a face detectada a partir dum de seus quadros. Este processo pode se restringir a um conjunto de quadros do vídeo ou ser realizado através do vídeo completo.

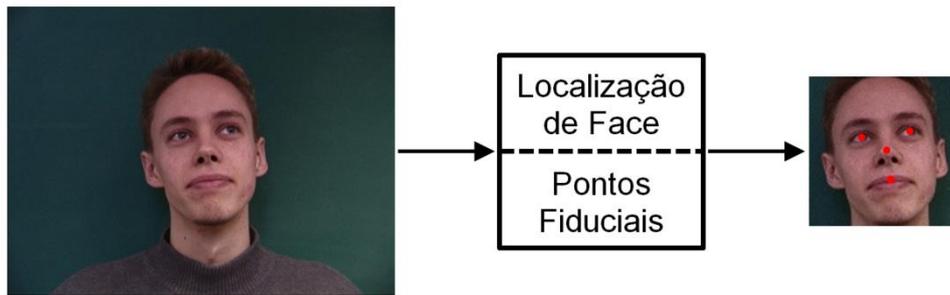


Figura 2. Detecção facial com pontos fiduciais. Imagem adquirida de (Nordstrom, et al., 2004).

### 2.1.2. Normalização

A normalização é o processo pelo qual a imagem facial é transformada geométrica e fotometricamente (ver, Figura 3). Os pontos fiduciais fornecidos pela detecção facial permitem normalizar a imagem facial em rotação e escala no plano da imagem, de modo que a face resultante esteja alinhada ao eixo horizontal do plano com a distância fixa entre os olhos.

As transformações em iluminação, contraste, entre outras propriedades da imagem constituem a chamada normalização fotométrica.

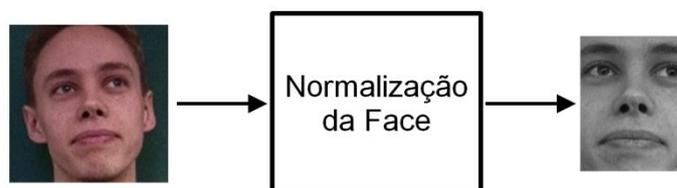


Figura 3. Normalização Geométrica e Fotométrica.

### 2.1.3. Extração de Atributos

O reconhecimento facial é essencialmente um problema de reconhecimento de padrões. Portanto, é necessário o uso de ferramentas que permitam extrair as informações mais relevantes das imagens faciais normalizadas, ou seja, um vetor de atributos, com base em que seja possível diferenciar uma face entre uma variedade de identidades. É desejável que a solução funcione a despeito de variações de iluminação e movimentos da face fora do plano, entre outras.

Vale a pena mencionar que, a acurácia dos sistemas de reconhecimento facial depende fortemente dos atributos que são extraídos das imagens faciais normalizadas, os quais por sua vez, dependem da correta localização da face e da exatidão na determinação dos seus pontos fiduciais.

#### **2.1.4. Correspondência**

Nesta etapa, os atributos extraídos da imagem facial normalizada da prova são comparados com aqueles pertencentes a cada imagem facial normalizada na galeria. Quando o sistema de reconhecimento facial opera em tarefas de verificação, também conhecidas na literatura como “correspondência de 1 a 1”, o classificador fornece duas respostas: “sim” quando a identidade invocada é verdadeira, ou “não” quando é falsa. Quando a tarefa é de identificação, denominada muitas vezes “correspondências de 1 a n”, o classificador fornece uma lista de identidades da galeria ordenada segundo o grau de similaridade com a imagem facial de prova.

Independentemente de que a tarefa do sistema seja verificar ou identificar, um dos grandes desafios nesta etapa é desenvolver métricas de similaridade que permitam comparar os atributos da face com um alto nível de acurácia.

### **2.2. Técnicas de Reconhecimento Facial**

As técnicas de reconhecimento facial podem ser divididas em três categorias, dependendo do método da aquisição das imagens (câmeras fotográficas, de vídeo, sensores infravermelhos, escâneres 3D, etc.), do método que manipulam as imagens, dos métodos aplicados às sequências de vídeo, e em função de requererem ou não dados adicionais sobre o sensor, como informação 3D ou imagens infravermelhas (Jafri, et al., 2009). Geralmente, os sistemas de reconhecimento facial baseados em vídeo processam os dados do sensor a cada quadro, voltando, assim, ao reconhecimento facial em vídeo para um reconhecimento facial baseado em imagens estáticas.

Outra categorização dos métodos de reconhecimento facial, e que é amplamente aceita pela comunidade científica, foi a proposta em (Zhao, et al., 2006), que classifica os métodos do ponto de vista psicológico de como humanos usam atributos locais e globais para o reconhecimento:

- **Métodos de correspondência holística**, são aqueles que procuram identificar faces usando representações globais da imagem facial. Pode-se ainda subdividir esta classe em dois grupos: aqueles que usam abordagens estatísticas (Análises de Componentes Principais, Análises Linear de Discriminante, etc.) e técnicas de aprendizado máquina (Redes Neurais, SVM's).
- **Métodos de correspondência baseadas em atributos**, são os métodos que buscam extrair atributos representativos duma imagem facial não necessariamente normalizada, e.g. pontos dos olhos, nariz, boca, entre outros pontos de referência, com a finalidade de relacioná-los e obter uma representação simplificada da face, a qual será submetida a técnicas estatísticas de reconhecimento de padrões.
- **Métodos Híbridos**, que combinam abordagens dos métodos holísticos e dos baseados em atributos.

Adicionalmente aos métodos já mencionados, a literatura reconhece a mais uma categoria de técnicas de reconhecimento facial, os chamados **métodos de correspondência baseados em componentes**, que visam identificar faces usando sub-regiões (partes ou componentes das imagens faciais) que contêm características locais, por exemplo, os olhos, nariz, boca, etc.

### 2.3. Desafios no Reconhecimento Facial Automático

Um sistema de reconhecimento facial é considerado ideal se, para qualquer imagem de entrada, a identidade fornecida pelo comparador for a correta. Contudo, em condições reais isto não sucede devido a que no mundo existe uma quantidade enorme de fatores que dificultam o reconhecimento.

Em sistemas de reconhecimento faciais bem comportados, as imagens de prova têm características similares a aquelas na galeria, por exemplo: iluminação uniforme, bom contraste, nitidez, expressão neutra e face frontal com respeito da câmera. No entanto, em condições reais as imagens de prova sofrem variações consideráveis nestes fatores, que acontecem de forma não controlada, o que causa um deterioramento no desempenho dos algoritmos de reconhecimento.

Muito embora os sistemas de reconhecimento facial na atualidade ofereçam métodos contra os efeitos de uma iluminação inadequada, contraste muito baixo e imagens borradas, ainda não existe solução totalmente robusta

contra toda a variação que ocorre na imagem facial de um indivíduo em situações reais.

Além das características inerentes aos indivíduos, como envelhecimento, semelhança física entre indivíduos, textura e cor da pele, o uso de acessórios externos, como óculos, chapéus, cachecóis, também gera dificuldades para o reconhecimento, não só automático, mas também para seres humanos.

## **2.4.**

### **Sistemas de Reconhecimento Facial: Aplicação no Vídeo Monitoramento**

O vídeo monitoramento se refere à observação das atividades, comportamento e mudanças que ocorrem numa determinada área ou cena, as que são adquiridas por meio de câmeras de vigilância conectadas a uma central de monitoramento onde pessoal capacitado, como por exemplo, policiais e/ou guardas de segurança, analisam de forma visual as imagens obtidas. No entanto, devido à quantidade de câmeras distribuídas numa área e ao volume de dados que geram, a análise por parte dos funcionários da segurança se torna difícil. A solução para este problema reside no desenvolvimento de sistemas automáticos capazes de analisar as imagens a fim de manter a ordem social e prevenir atividades criminosas.

Assim, o reconhecimento facial automático se perfila como viável para inclusão nos sistemas de vídeo monitoramento. Neste caso, o grande desafio é processar a maior quantidade de dados disponíveis no menor tempo possível e com uma acurácia aceitável para o reconhecimento.

Bonfá, em sua dissertação de mestrado (Bonfá, 2013), propõe uma arquitetura paralela para o reconhecimento facial baseado em vídeo, que pretende identificar os indivíduos numa cena em tempo real. Um protótipo baseado nesta arquitetura, aplicado ao vídeo monitoramento, vem sendo desenvolvida no LVC<sup>3</sup>.

## **2.5.**

### **Contextualização do Trabalho**

Nesta dissertação se estudam diversos algoritmos de reconhecimento facial visando sua comparação a partir de imagens faciais normalizadas

---

<sup>3</sup> Laboratório de Visão Computacional da PUC – Rio.

extraídas de sequências de vídeo, onde as variações em iluminação, contraste, nitidez, expressões faciais, posição da cabeça, textura da pele, entre outras, representam as condições de captura em ambientes reais não controlados.

Pretende-se que a análise dos resultados deste estudo permita entender melhor o processo de reconhecimento facial automático em vídeo, assim como identificar aos algoritmos com melhor desempenho em tarefas de verificação e identificação voltadas ao vídeo monitoramento, e finalmente indicar os algoritmos alternativos ao utilizado em (Bonfá, 2013).