

Sistema de detecção automática de centro de pupila e obtenção de métricas faciais em 2D

Damares S. Cavalcante¹, Luiz R. M. Silva¹, Adonis Born²,
Mário Hozano¹, Tácito T. A. T. Neves¹

¹Núcleo de Ciências Exatas (NCEX) – Universidade Federal de Alagoas (UFAL)
Av. Manoel Severino Barbosa – Bom Sucesso – Arapiraca – AL – Brasil

²Clínica Oftalmo Presente
Rua Dr. Pedro Correia, 433 – Centro – Arapiraca – AL – Brasil

{damaresscavalcante, luizrodrigomarques, adonisborn, hozano}@gmail.com
tacito.neves@arapiraca.ufal.br

Abstract. *The image processing area has been heavily worked on in several areas that require automation of its processes through the extraction of parameters in images. An example is ophthalmology, with the possibility of automating automatic pupil detection and extraction of facial metrics. This work presents an automatic detection system of pupil center in real time images and videos, as well as the extraction of facial metrics in 2D. The metrics were validated in an eye clinic and were compared with the results of the developed algorithm. The results obtained were satisfactory for the set of images evaluated.*

Resumo. *A área de processamento de imagens tem sido fortemente trabalhada em diversas áreas que requerem automatização dos seus processos por meio da extração de parâmetros em imagens. Um exemplo é a oftalmológica, com a possibilidade de automatização da detecção automática de pupila e extração de métricas faciais. Este trabalho apresenta um sistema de detecção automática de centro de pupila em imagens e vídeos de tempo real, assim como a extração de métricas faciais em 2D. As métricas foram validadas em uma clínica oftalmológica e foram comparadas com os resultados do algoritmo desenvolvido. Os resultados obtidos foram satisfatórios para o conjunto de imagens avaliadas.*

1. Introdução

A área de processamento de imagens viveu um grande processo de ascensão nas últimas décadas, viabilizando aplicações em categorias distintas: (1) o aprimoramento de informações pictóricas para interpretação humana; e (2) análise automática por computador de informações extraídas de uma cena [Filho and Neto 1999]. Um grande momento da utilização de técnicas de processamento de imagens veio com o programa americano *Jet Propulsion Laboratory*¹, de sondas espaciais para capturar imagens da lua. Ao longo dos anos, a área foi aprimorada e suas técnicas desenvolvidas em conjunto com outras áreas como a análise de imagens, visão e visualização computacional.

Para processar uma imagem com o objetivo de representá-la e descrevê-la em um ambiente digital, é necessário seguir alguns passos fundamentais, como: (1) aquisição

¹<https://www.jpl.nasa.gov>

da imagem - amostragem, armazenamento e compactação; (2) segmentação - extração dos objetos de fundo; (3) parametrização - determinação de grandezas sobre os objetos; e (4) reconhecimento - classificação dos objetos para o correto tratamento da imagem [de Albuquerque and de Albuquerque 2000].

A utilização de técnicas de processamento de imagens para a extração de características de objetos pode ser aplicada em diversas áreas, principalmente naquelas passíveis de automatização de processos, como por exemplo na área oftalmológica. Diversos processos em uma clínica oftalmológica podem ser automatizados utilizando técnicas avançadas. Entre os exemplos de tarefas que podem ser beneficiadas por esses processos, temos a detecção automática de pupila e a obtenção de métricas faciais, que são de extrema importância para garantir a confecção precisa dos óculos para o paciente.

A extração de informações em uma imagem compreende um processo minucioso que pode envolver técnicas externas ao processamento digital de imagens, tais como aprendizagem de máquina e inteligência artificial. Realizar a aquisição das imagens torna-se também um passo fundamental, uma vez que existem diversos métodos para aquisição, a depender das variáveis do problema a ser considerado. O reconhecimento de faces, por exemplo, tem sido frequentemente trabalhado utilizando-se de redes neurais. A partir de um conjunto de imagens de faces são extraídos pontos de marcação e uma nova base é construída com esses pontos. Essa nova base é submetida à uma rede neural, previamente modelada, para que seja treinada. Quanto mais casos na base, mais precisa se torna a detecção de faces.

Com o objetivo de extrair informações de medidas de pontos de interesse em uma face de imagens e vídeos, foi desenvolvido um sistema de detecção e cálculo de métricas utilizando Python 3 e OpenCV em 2D. O sistema faz uso de um algoritmo capaz de mapear os pontos em uma face por meio do uso de uma rede neural, e através da utilização de técnicas de extração de parâmetros, os pontos são mapeados e outro algoritmo é utilizado para realizar os cálculos para as métricas de distância.

A obtenção das medidas faciais de forma manual não garante a precisão, podendo os valores variar quando diferentes técnicos tentam obter a mesma medida. Assim, a principal contribuição do presente trabalho é a obtenção automática das medidas de distância entre as pupilas.

O restante deste artigo é organizado da seguinte forma: na seção dois são apresentados os trabalhos relacionados, na seção três encontra-se a fundamentação teórica e os métodos utilizados, na seção quatro e cinco são descritas a solução proposta e os resultados obtidos, e por último as conclusões do trabalho atual.

2. Trabalhos relacionados

Tornou-se crescente nos últimos anos a utilização de técnicas de processamento de imagens e visão computacional. Ao considerar as aplicações nas áreas de detecção facial e extração de parâmetros, diversos desafios foram encontrados ao decorrer das pesquisas, principalmente relacionados à precisão, resultando em uma vasta literatura de reconhecimento facial com aplicabilidades distintas a depender do problema apresentado.

[Fujikawa 2016] realiza um estudo sobre o reconhecimento facial por meio de aprendizado não supervisionado utilizando descritores de textura. No desenvolvimento

de sua pesquisa utilizou os chamados Sistemas de Recuperação de Imagens baseado no Conteúdo (CBIR) para modelar um sistema de reconhecimento facial. O sistema funciona apresentando uma imagem base aos algoritmos de modelagem, que retornam uma lista ordenada de imagens de acordo com a similaridade à imagem base.

Já [Mendes 2019] trabalhou o reconhecimento baseado na Análise de Componentes Principais (sendo a sigla em inglês, PCA), no qual cada face é representada como um vetor n -dimensional. A utilização do PCA compreendeu uma série de algoritmos classificados em distância, árvores de decisão, máquina de vetores e redes neurais artificiais (RNA). Sendo esta última a que teve maior efetividade (99%) na base de dados testada, demonstrando a importância de se utilizar técnicas baseadas em aprendizagem de máquina por meio de RNAs.

A literatura atual sobre o reconhecimento facial remonta frequentemente a utilização de PCA e LBP. Os pesquisadores [Amaral and Thomaz 2011] realizaram um estudo comparativo entre LBP e PCA que são classes de algoritmos para reconhecimento facial. O objetivo foi identificar o melhor processo quando se está diante de um problema com restrição amostral. Dessa pesquisa, os resultados apontaram que quando se têm um conjunto de imagens homogêneas o PCA apresenta-se como a melhor solução, enquanto que o LBP se comporta melhor diante de um conjunto de imagens com variações de captura.

3. Fundamentação Teórica

O processo de extração de informações em uma imagem pode utilizar shapes, que são formas conhecidas buscadas dentro de uma imagem. Os shapes podem ser advindos de uma imagem com formas geométricas, uma rede neural treinada com um conjunto de marcações de pontos que normalmente denotam uma forma de um objeto que se deseja conhecer.

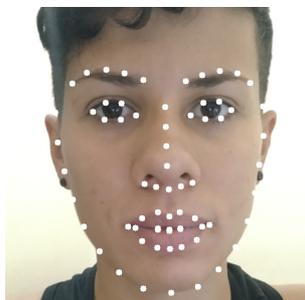
A utilização de *shapes* geométricos corresponde a um dos processos mais simples para identificação de objetos em uma imagem. Basta obter uma imagem com a coleção de objetos (círculo, elipse, quadrado, retângulo, hexágono, triângulo, etc) a serem reconhecidos, aplicar uma função de threshold para obtenção da imagem em *grayscale* e encontrar contornos nas bordas dos objetos.

No processo de identificação de bordas, figuras como eclipse e círculos, têm suas bordas desenhadas como se fossem polígonos, e conseqüentemente apresentam um número de arestas superiores aos objetos que têm suas arestas geometricamente definidas como quadrado, triângulo, retângulo e hexágono. Diante das formas previamente identificadas, o algoritmo pode ser facilmente aplicado em diversas imagens, que tentará identificar através da similaridade entre a quantidade de arestas do *shape*, o objeto na imagem.

Redes neurais são sistemas de comunicação que funcionam semelhante ao cérebro humano, a partir da utilização do conceito de neurônio. Em uma rede neural tenta-se simular o neurônio humano por meio da modelagem matemática, sendo chamado de neurônio artificial. Uma rede neural pode ser utilizada como técnica para o reconhecimento de objetos ou extração de parâmetros em uma imagem. O reconhecimento se dá a partir de um treinamento prévio da rede por meio de padrões conhecidos, esses padrões são agrupados e posteriormente classificados e a rede permanece em constante aprendizado, assim

quanto mais casos na bases mais preciso é o processo de reconhecimento. A Figura 1 exhibe a marcação de 68 pontos faciais a partir de uma rede neural previamente treinada.

Figura 1. Reconhecimento facial utilizando redes neurais



Fonte: Autores (2020).

4. Metodologia

Para o desenvolvimento da solução proposta foi utilizado um processo de desenvolvimento iterativo e incremental, onde a medida em que os módulos eram desenvolvidos ocorriam entregas parciais, permitindo com que a solução fosse constantemente avaliada para verificar as alterações realizadas.

O início do desenvolvimento se deu a partir de uma revisão da literatura em busca das principais técnicas para reconhecimento facial. Nesse sentido, optou-se por utilizar redes neurais a partir de um dataset treinado para o reconhecimento de 68 pontos na região da face. Após a identificação facial foi iniciado o processo de detecção de centro de pupila traçando duas linhas diagonais a partir de quatro pontos na região do olho, e utilizando o conceito de intersecção entre retas foi identificando o ponto principal que marca o centro da pupila, a Figura 2 mostra o processo de detecção.

5. Solução proposta

O sistema calcula a distância entre pupilas (DP), distância da pupila esquerda até o centro do nariz (DNPe) e a distância da pupila direita até o centro do nariz (DNPd). A solução proposta utiliza a linguagem de programação Python, a biblioteca de visão computacional OpenCV e a biblioteca de modelos pré-treinados de *machine learning* dlib.

A identificação facial foi realizada com a utilização da biblioteca dlib, que contém vários modelos prontos de *machine learning*. A biblioteca funciona como um detector de pontos de referências (marcações faciais), a partir de um modelo pré-treinado que estima as posições (x,y) de 68 pontos na região da face. Os pontos identificados foram obtidos através do conjunto de dados iBUG 300-W². A partir da marcação dos pontos, é possível iniciar o processo de detecção do centro da pupila.

Para identificação do centro da pupila, é necessário utilizar apenas quatro pontos dos 6 pontos que são identificados em cada olho pela rede neural. A solução utilizada para detecção foi a utilização do conceito de intersecção entre retas, assim para que o centro da pupila seja devidamente encontrado utilizando o algoritmo, é necessário que os olhos estejam facilmente identificáveis dentro da imagem. O recomendado é que as fotos sejam tiradas olhando 6 metros ao infinito, para que se consiga centralizar a pupila dentro do olho, e obter as coordenadas do seu centro utilizando o ponto de intersecção entre as

²<https://ibug.doc.ic.ac.uk/resources/facial-point-annotations/>

retas criadas na marcação do olho. Assim, a marcação da pupila a partir do método acima, permite com que as métricas desejadas sejam extraídas de forma mais fácil dos demais pontos do rosto.

Figura 2. Identificação de centro de pupila



Fonte: Autores (2020).

5.1. Avaliação

Para testar a validade do algoritmo foi realizada uma medição para a extração da DP, DNPe e DNPd utilizando três formas de medição: (1) através de uma régua milimetrada; (2) através de um aparelho chamado pupilômetro; e (3) através do algoritmo. As três formas de medidas foram utilizadas para comparar a precisão da medição do algoritmo. O procedimento de medição manual e através do pupilômetro foi acompanhado por um oftalmologista, que mediu e validou os valores para cada métrica.

O procedimento para extração de medidas utilizando régua milimetrada contou com o auxílio de armações, para que o centro da pupila pudesse ser marcado na lente. Foram utilizados quatro modelos de armação e 7 voluntários. Cada voluntário permanecia sentado com a armação no rosto e olhava para o infinito, enquanto uma pessoa a aproximadamente uma distância de 6 metros ia em direção ao voluntário e marcava o centro da pupila na armação com um pincel. Posteriormente, era realizada a medição com uma régua milimetrada e as medidas eram armazenadas em uma planilha.

O procedimento para extração de métricas com o pupilômetro foi mais simples, uma vez que o aparelho já imprimia no display a medida da DNPe e DNPd, sendo necessário apenas a soma de ambas para conseguir a DP. O voluntário permanecia sentado, juntamente com a pessoa que iria realizar a medição, o aparelho era encaixado no nariz do voluntário e ele olhava para um ponto escuro dentro do aparelho que centralizava a pupila, com isso, o medidor iria regulando uma linha vertical no centro da pupila do voluntário através de botões no aparelho, este procedimento de regulagem era realizado para os dois olhos. Após a medição, os valores eram armazenados na planilha.

6. Resultados obtidos

Os dados a seguir mostram os resultados da aplicação do algoritmo desenvolvido nesse trabalho para a extração de métricas faciais. A Tabela 1 apresenta os resultados dos cálculos das métricas nas imagens faciais de 7 voluntários. São descritas as medidas para DNP esquerda e DNP direita para cada um dos voluntários. A validação dos resultados foi realizada através de um aparelho oftalmológico chamado pupilômetro, no qual, foi medida a DNP esquerda e DNP direita de cada um dos voluntários.

Como pode ser observado na tabela, os resultados das medições utilizando o pupilômetro e os resultados do algoritmo estão próximos, apresentando uma margem de erro de 5 mm para mais e para menos. A diferença de precisão entre os métodos é decorrente da distância da câmera para os pontos selecionados, considerando que a obtenção dos pontos atuais não inclui a informação de profundidade. Para melhorar a precisão, a informação de distância fornecida pelo sensor da Intel será utilizada.

Tabela 1. Validação das Métricas

VOL	MEDIDORES (mm)				DIFERENÇA (%)	
	PUPILÔMETRO		ALGORITMO			
	DNP_DIR	DNP_ESQ	DNP_DIR	DNP_ESQ	DNP_DIR	DNP_ESQ
1	34.00	31.50	30.00	28.46	11.76	9.65
2	35.00	35.00	34.65	35.42	1.00	1.18
3	35.50	35.50	40.03	39.23	11.31	9.50
4	35.00	34.50	35.39	35.42	1.10	2.59
5	29.50	30.50	29.23	30.00	0.90	1.63
6	29.50	30.50	33.85	33.11	12.85	7.88
7	30.50	31.00	29.29	29.29	3.96	5.51

Legenda: DNP_DIR = Distância Naso-Pupilar Direita;

DNP_ESQ = Distância Naso-Pupilar Esquerda; VOL = Voluntários

7. Conclusões

No presente trabalho foi proposto um sistema de detecção de centro de pupila para extração de métricas faciais em 2D. A partir da detecção da face e centro de pupila é possível extrair parâmetros faciais, que podem ser utilizadas em diversos segmentos: como clínicas oftalmológicas para automatização do processo de obtenção de métricas faciais; óticas para a construção e adaptação das armações nos rostos dos pacientes; e em clínicas que realizam procedimentos estéticos, para obter a melhor aplicação para cada paciente a partir da análise das suas métricas faciais.

Como trabalhos futuros pretende-se melhorar a precisão da detecção do centro da pupila para desenvolver o processo de extração de métricas em um ambiente 3D. Utilizando modelos 3D de rosto e armação, em conjunto com um método de detecção automática da pupila, é possível encontrar a distância da pupila para a base da armação. Essa medida é muito importante no processo de fabricação da lente de cada paciente.

Referências

- Amaral, V. d. and Thomaz, C. E. (2011). Extração e comparação de características locais e globais para o reconhecimento automático de imagens de faces. Dissertação de mestrado, Centro Universitário da FEI, SP, Brasil.
- de Albuquerque, M. P. and de Albuquerque, M. P. (2000). Processamento de imagens: métodos e análises. *Rio de Janeiro, Brasil*, 12.
- Filho, O. M. and Neto, H. V. (1999). *Processamento Digital de Imagens*. Brasport, Rio de Janeiro.
- Fujikawa, C. S. (2016). Reconhecimento Facial utilizando Descritores de Textura e Aprendizado Não Supervisionado. Monografia (Bacharel em Ciências da Computação), UNESP (Universidade Estadual Paulista "Júlio de Mesquita Filho"), Rio Claro, Brazil.
- Mendes, K. B. (2019). Comparativo de Algoritmos Clássicos de Aprendizado de Máquina em um Problema de Reconhecimento de Faces. Monografia (Bacharel em Engenharia Eletrônica e de Telecomunicações), UFU (Universidade Federal de Uberlândia), Patos de Minas, Brazil.