



José Luiz Buonomo de Pinho

**Medidas de Qualidade de Imagens para
Reconhecimento Facial**

Dissertação de Mestrado

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da PUC-Rio como requisito parcial para obtenção do título de Mestre em Engenharia Elétrica.

Orientador: Prof. Raul Queiroz Feitosa

Rio de Janeiro
Agosto de 2012



José Luiz Buonomo de Pinho

**Medidas de Qualidade de Imagens para
Reconhecimento Facial**

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós-Graduação em Engenharia Elétrica do Departamento de Engenharia Elétrica do Centro Técnico Científico da PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada.

Prof. Raul Queiroz Feitosa
Orientador

Departamento de Engenharia Elétrica – PUC-Rio

Prof. Alvaro de Lima Veiga Filho

Departamento de Engenharia Elétrica – PUC-Rio

Prof. Guilherme Lucio Abelha Mota
UERJ

Prof. José Eugenio Leal
Coordenador Setorial do Centro
Técnico Científico

Rio de Janeiro, 24 de agosto de 2012

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

Luiz Buonomo de Pinho

Graduado em Engenharia Civil ênfase estruturas pela EEVA com Pós-Graduação em Internet, Interface e Multimídia pela UFF com 30 anos de experiência na área de Tecnologia da Informação atuando nas mais diversas áreas como Banco de Dados, Biometria, Suporte, Apoio ao Desenvolvimento, Operação e Produção nos seguintes segmentos de mercado: Bancários Serviços, Terceirização, Governo, Outsourcing, Montagem e Administração de CPD onde desenvolveu e gerenciou trabalhos de grande porte.

Ficha Catalográfica

Pinho, José Luiz Buonomo de

Medidas de qualidade de imagens para reconhecimento facial / José Luiz Buonomo de Pinho ; orientador: Raul Queiroz Feitosa. – 2012.

105 f. ; 30 cm

Dissertação (mestrado) – Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Elétrica, 2012.

Inclui bibliografia

1. Engenharia elétrica – Teses. 2. Reconhecimento facial. 3. Qualidade de imagem. 4. Regressão. 5. Redes neurais. I. Feitosa, Raul Queiroz. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Elétrica. III. Título.

CDD: 621.3

Dedico este trabalho a todos os profissionais e pesquisadores que dedicam parte de seu tempo a temas relacionados à biometria.

Agradecimentos

Agradeço a todos que contribuíram para a realização deste trabalho.

Agradeço ao meu orientador pela atenção e apoio durante o desenvolvimento do projeto assim como pelo meu crescimento intelectual e acadêmico.

Agradeço à minha família pela motivação e apoio às minhas decisões.

Agradeço aos colegas de trabalho que contribuíram com idéias, implementações, críticas e sugestões durante este período intenso de estudos.

Gostaria também de agradecer a M.I. Montreal Informática Ltda pelo apoio financeiro durante o período do curso de mestrado.

Resumo

Pinho, Jose Luiz Buonomo de; Feitosa, Raul Queiroz (Orientador). **Medidas de Qualidade de Imagens para Reconhecimento Facial**. Rio de Janeiro, 2012. 105p. Dissertação de Mestrado - Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

O Reconhecimento Facial é o processo de identificação de uma pessoa a partir da imagem de sua face. Na forma mais usual, o processo de identificação consiste em extrair informações dessa imagem e compará-las com informações relativas a outras imagens armazenadas numa base de dados e por fim indicar na saída a imagem da base mais similar à imagem de entrada. O desempenho desse processo está diretamente ligado à qualidade das imagens, tanto das que estão armazenadas na base de dados, quanto da imagem do indivíduo cuja identidade está sendo determinada. Por isso, convém que a qualidade das imagens faciais seja avaliada antes que estas sejam submetidas ao procedimento de reconhecimento. A maioria dos métodos apresentados até o momento na literatura baseia-se em um conjunto de critérios, cada um voltado a um atributo isolado da imagem. A qualidade da imagem é considerada adequada se aprovada por todos os critérios individualmente. Desconsidera-se, portanto, o efeito cumulativo de diversos fatores que afetam a qualidade das imagens e, por conseguinte, o desempenho do reconhecimento facial. Essa monografia propõe uma metodologia para o projeto de métricas de qualidade de imagens faciais que expressem num único índice o efeito combinado de diversos fatores que afetam o reconhecimento. Tal índice é dado por uma função de um conjunto de atributos extraídos diretamente da imagem. O presente estudo analisa experimentalmente uma função linear e uma rede neural do tipo *back-propagation* como alternativas para a estimativa de qualidade a partir dos atributos. Experimentos conduzidos sobre a base de dados IMM para o algoritmo de reconhecimento baseado em *padrões binários locais* comprovam a o bom desempenho do metodologia.

Palavras-chave

Reconhecimento Facial; Qualidade de Imagem; Regressão; Redes Neurais.

Abstract

Pinho, Jose Luiz Buonomo; Feitosa, Raul Queiroz (Advisor). **Image Quality Metrics for Face Recognition**. Rio de Janeiro, 2012. 105p. MSc Dissertation - Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

Face Recognition is the process of identifying people based on facial images. In its most usual form the identification procedure consists of extracting information from an input face image and comparing them to the records of other face images stored in a face data base, and finally indicating the most similar one to the input image. The performance of this process is directly dependent on the input image quality, as well as on the images in the data base. Thus, it is important that the quality of a face image is tested before it is given to the recognition procedure, either as a input image or as a new record in the face database. Most methods proposed thus far based on a set of criteria, each one devoted to an isolated attribute. The image quality is considered adequate if approved by all criteria individually. Thus, the cumulative effect of different factors affecting the image quality is no regarded. This dissertation proposes a methodology for the design of quality metrics of facial images that Express in a single scalar the combined effect of multiple factors affecting the quality. Such score is given by a function of attributes extracted directly from the image. This study investigates a linear and a non-linear approach for quality assessment. Experiments conducted upon the IMM face database for a *Local Binary Pattern* face recognition algorithm demonstrate the good performance of the proposed methodology.

Keywords

Face Recognition; Image Quality Assessment; Regression; Neural Network.

Sumário

1 Introdução	13
1.1. Organização do Trabalho	18
2 Trabalhos Relacionados	20
2.1. Reconhecimento Facial sob Condições Difíceis de Iluminação	20
2.2. Eliminação de Artefatos em Imagens Frontais de Face	22
2.3. Correção Gama em Imagens de Forma Localizada	26
2.4. Função P.S.F. Baseado na Distribuição de Vetores Gradientes em um Plano Polar	28
2.5. O Método de Landweber	31
2.6. Avaliação de Qualidade de Imagens Faciais	33
2.7. Padronização das Amostras de Face	35
2.8. Impacto de Imagens de Baixa Nitidez em Reconhecimento Facial Baseado em Função de Transferência Modular	40
2.9. Melhores Práticas para Fotografias de Face e Métricas Qualidade de Imagem	44
3 Fundamentos Teóricos	49
3.1. Regressão Linear	49
3.2. Seleção de Atributos	51
3.3. Redes Neurais Artificiais	52
3.4. Reconhecimento Facial	53
4 Método para Medida de Qualidade de Imagens Faciais	59
4.1. Método Proposto	59
4.2. Instanciação do Método	61
4.3. Atributos	62
4.3.1. Arfagem	63
4.3.2. Rolagem	65
4.3.3. Intensidade Média	66
4.3.4. Contraste	66
4.3.5. Gradientes	67
4.3.6. Curtose	69
4.3.7. Não Referência	69
4.3.8. Avaliação dos Efeitos de Iluminação – WMI	71
4.3.9. Ruído 1	72
4.3.10. Ruído 2	74
5 Análise Experimental	76
5.1. Base de Dados	76
5.2. Recursos Computacionais	78
5.3. Experimentos	78
5.3.1. Variação da Qualidade de cada tipo de Imagem	79
5.3.2. Grupos de Atributos Correlacionados	82
5.3.3. Avaliação da Contribuição da cada Atributo Isolado	86
5.3.4. Determinação do Melhor Conjunto de Atributos	90
5.3.5. Estimativa de Qualidade por Rede Neural	93
5.3.6. Comparar a Variação entre os Valores Medidos e Estimados	94
5.3.7. Adequação do Modelo	95

5.3.8. Comentários Sobre Experimentos	97
6 Comentários Finais e Direções Futuras	99
7 Referências Bibliográficas	102

Lista de tabelas

Tabela 2.1 – Correlação média entre MTF-R e médias de nitidez.	44
Tabela 5.1 – Tabela com o módulo dos coeficientes de correlação dos atributos usados.	84
Tabela 5.2 – Grupos de atributos que apresentam coeficientes de correlação muito elevados.	85
Tabela 5.3 – Grupos de atributos que apresentam coeficientes de correlação elevados.	85
Tabela 5.4 – Resultados considerando cada atributo isoladamente.	88
Tabela 5.5 – Melhor conjunto de N atributos.	91

Lista de figuras

Figura 1.1 – Seis tomadas da mesma face.	15
Figura 1.2 – Fluxograma para aplicação da solução proposta.	16
Figura 1.3 – Metodologia proposta.	17
Figura 2.1 – Taxa de reconhecimento facial.	21
Figura 2.2 – Molde de regiões da face.	23
Figura 2.3 – Proposta de melhoria automática em imagens faciais.	24
Figura 2.4 – Janelas sobrepostas de tamanho 128 x 128 pixels e seus valores de Gama associados.	27
Figura 2.5 – Resultados da correção Gama adaptativa.	28
Figura 2.6 – Sistema de imagem em espaço linear invariante.	29
Figura 2.7 – Mapeamento do espectro de amplitude logarítmica da imagem observada ao plano polar.	29
Figura 2.8 – As direções dos vetores gradiente em torno da mínima.	30
Figura 2.9 – Direção dos vetores gradiente.	30
Figura 2.10 – Janelas retangulares e a linha hipotética.	31
Figura 2.11 – Eficiência método ALM versus LM.	33
Figura 2.12 – Mapeamento do angulo da face em função do valor de “pm”.	34
Figura 2.13 – Avaliação dos fatores do algoritmo para avaliação de qualidade de imagens faciais.	35
Figura 2.14 – Fluxograma de avaliação de qualidade de imagens.	36
Figura 2.15 – Simetria facial.	39
Figura 2.16 – Imagens borradas do banco Q-FIRE com escala ótica.	43
Figura 2.17 – Níveis de nitidez.	44
Figura 2.18 – Disposição de luz para estúdio com apenas uma lâmpada	45
Figura 2.19 – Disposição de luz e câmera	45
Figura 2.20 – Posicionamento da cabeça.	45
Figura 2.21 – Uniformidade da iluminação.	46
Figura 2.22 – Efeito do Flash.	46
Figura 3.1 – Representação gráfica de Regressão.	50
Figura 3.2 – Representação de uma Rede Neural Artificial (RNA).	53
Figura 3.3 – Processos de um sistema de reconhecimento facial.	54
Figura 3.4 – Conjunto de vizinhança de pixels para valores de m e R.	55

Figura 3.5 – Cálculo do LBP.	55
Figura 3.6 – Histograma do LBP.	56
Figura 3.7 – Processo de comparação de faces baseado em textura.	57
Figura 4.1 – Abordagem tradicional e abordagem proposta.	60
Figura 4.2 – Região da face.	62
Figura 4.3 – Face no sistema espacial	64
Figura 4.4 – Medida da arfagem.	64
Figura 4.5 – Rolagem.	65
Figura 4.6 – Histograma de uma imagem com baixo e alto contraste.	67
Figura 4.7 – Pesos usados para cálculo de WMI.	72
Figura 4.8 – Imagem de face corrompida com anotações.	73
Figura 4.9 – Imagem recuperada com algoritmo Ruído 1.	73
Figura 4.10 – Evolução algoritmo Ruído 2.	74
Figura 4.11 – Remoção de ruído – algoritmo Ruído 2.	74
Figura 5.1 – Base de dados usada (Rosto N. 2).	77
Figura 5.2 – Fotos normal e embaçada.	77
Figura 5.3 – Gráfico de similaridade 01: Imagem 01 Versus 02,03,04,05,06.	80
Figura 5.4 – Gráfico de similaridade 02: Imagem 01 Versus 08,09,10,11,12.	81
Figura 5.5 – Gráfico de similaridade 03: Imagem 01 Versus 14,15,16,17,18.	82
Figura 5.6 – Escala de cores para o módulo do coeficiente de correlação.	83
Figura 5.7 – Variação da qualidade/similaridade para cada tipo de imagem.	94
Figura 5.8 – Dependência dos valores de similaridade medidos e preditos pelo modelo.	96
Figura 5.9 – Dependência do erro com os valores de similaridade medidos.	97

Introdução

A face apresenta vantagens importantes sobre outras formas de biometria usadas para a identificação e reconhecimento de pessoas. É aceita cultural e socialmente no mundo todo por usar uma informação que já é exposta normalmente e pode ser coletada de forma não intrusiva. O procedimento de coleta é simples e barato se comparado com outros tipos de identificação e de resposta rápida.

Algumas dificuldades técnicas para sua utilização de maneira sistematizada têm dificultado sua disseminação, embora seu potencial de utilização seja imenso tanto para o reconhecimento de imagens estáticas como de imagens dinâmicas obtidas através de câmeras de vídeo de sistemas de vigilância.

No Brasil os sistemas de Identificação com base em tecnologia biométrica por computador implantados em vários Estados e no Departamento de Polícia Federal usam exclusivamente a biometria de impressões digitais e em alguns casos palmares. Estes sistemas contem um grande volume de imagens que ainda não são usadas no processo de identificação.

A possibilidade do uso dessas imagens representa um potencial no aumento da eficácia desses sistemas e o aporte de novas tecnologias investigativas aos usuários dos sistemas de identificação na área policial e de segurança pública.

O aperfeiçoamento dos algoritmos de reconhecimento facial observado nos últimos anos permite antever que, em condições controladas, principalmente de pose e iluminação, poder-se-ão atingir níveis um pouco inferiores aos do reconhecimento por impressão digital, com a vantagem da facilidade de uso e da não intrusividade, podendo funcionar perfeitamente como filtros de primeira seleção. Em um processamento onde existam vários tipos de biometria, o reconhecimento facial pode ser usado como um primeiro item de busca permitindo assim que apenas as amostras biométricas que tem alguma possibilidade de serem encontradas sejam submetidas a outras formas de processamento como, por exemplo, busca por impressões digitais.

Este trabalho tem como finalidade investigar que características doravante

denominadas de “atributos” de uma imagem de face são relevantes no processo de reconhecimento facial.

A qualidade da imagem de uma face afeta diretamente o desempenho do processo de reconhecimento. De uma imagem é possível extrair diversos atributos relacionados à qualidade, que expressam a maior ou menor dificuldade de se reconhecer na imagem o objeto imageado. É importante que a imagem a ser utilizada no reconhecimento facial apresente adequadamente as características faciais discriminantes com o mínimo possível de artefatos e ou ruídos externos.

A literatura contém várias propostas que tentam minimizar os efeitos de fatores físicos e ambientais que agem na formação de uma imagem digital da face, tais como:

- Pose;
- Foco/embaçamento;
- Contraste;
- Iluminação;
- Posicionamento;
- Ruído (Presença de Artefatos);
- Taxas de Compressão (Jpeg e outras técnicas de compressão).

A figura 1.1 mostra seis imagens da mesma face com diferentes tomadas. O grau de similaridade resultante da comparação de uma contra as outras cinco por um sistema de reconhecimento facial produziria diferentes resultados para cada comparação, não só pelo fato de não serem imagens da mesma tomada, mas também por que as imagens estão sujeitas a padrões de iluminação, níveis de contraste e nitidez distintos.

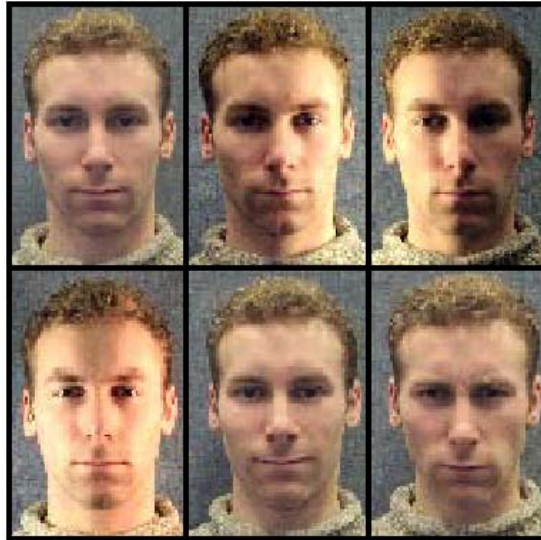


Figura 1.1 – Seis tomadas da mesma face.

No processo de reconhecimento facial, a imagem poderá passar por algum tipo de pré-processamento com objetivo de melhorar e ou restaurar sua qualidade. Após esse pré-processamento, a imagem será submetida a um critério de avaliação onde apenas as imagens que apresentarem um índice mínimo de qualidade serão submetidas ao processo de comparação (*matching* no inglês) já que este processo é computacionalmente muito pesado quando executados sobre grandes bancos de dados. A figura 1.2 apresenta o fluxograma para aplicação da solução proposta que avalia previamente a qualidade da imagem antes de executar o processo de reconhecimento facial propriamente dito.

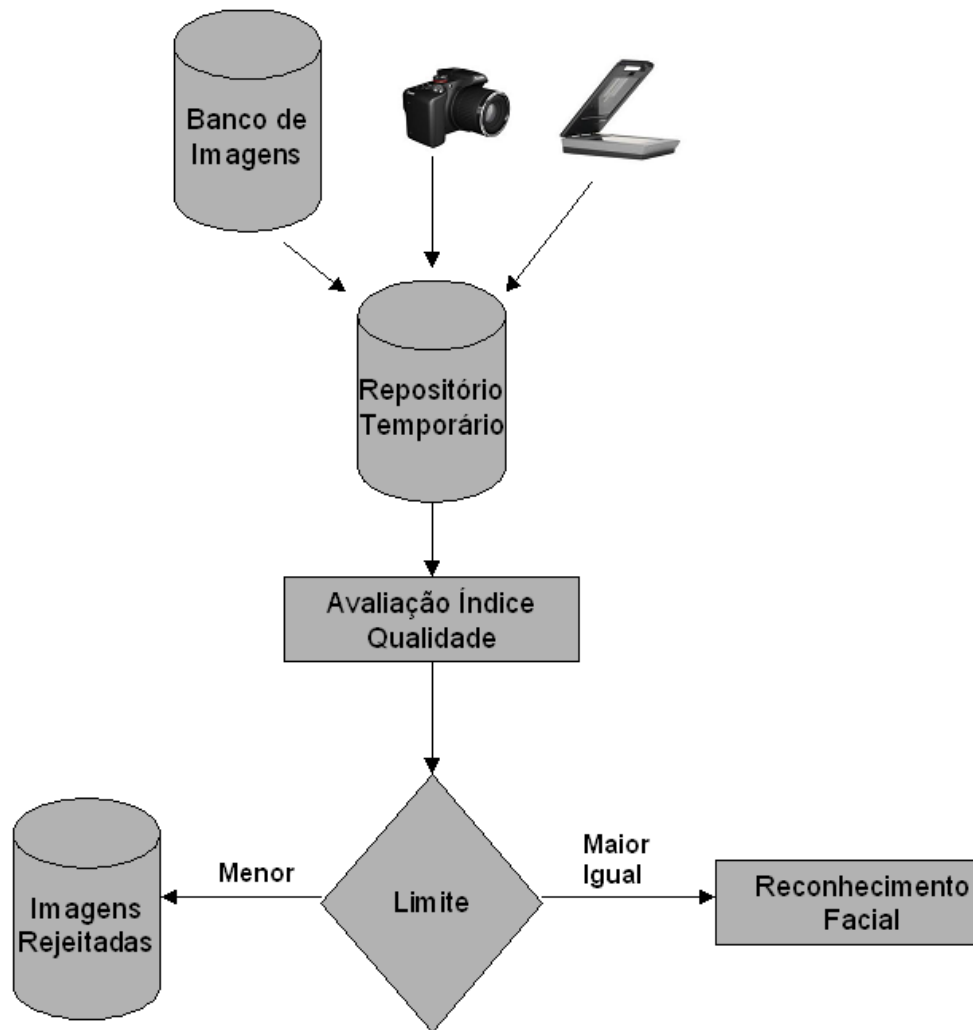


Figura 1.2 – Fluxograma para aplicação da solução proposta.

Como pode ser observado na figura 1.2, as imagens seriam oriundas de bancos de dados, câmeras que podem ser fotográficas ou de vídeos ou até mesmo de scanners usados sobre fichas que contenham uma imagem de face. Essas imagens iriam para um repositório temporário que funcionaria como uma fila de entrada para o processo proposto nessa dissertação para medição da qualidade de cada imagem. Uma vez conhecido o valor da qualidade de cada imagem, elas seriam comparadas a um limiar e as imagens com qualidade superior a esse limiar seriam submetidas ao sistema de reconhecimento facial evitando assim um processamento desnecessário já que imagens de baixa qualidade não gerariam resultados satisfatórios e as demais imagens seriam descartadas gerando assim um

repositório de imagens de baixa qualidade que poderia ser usado em uma análise posterior mais detalhada.

A presente dissertação tem como objetivo geral, propor uma metodologia para o projeto de uma métrica de qualidade de imagens faciais. Mais especificamente, a métrica deve expressar num único índice o efeito combinado de todos os fatores que afetam a qualidade da imagem. Trata-se aqui de uma novidade em relação às abordagens encontradas na literatura, que produzem um índice de qualidade separado para cada um dos fatores (por exemplo, nitidez, contraste, pose) que concorrem para determinar a qualidade da imagem facial. A figura 1.3 ilustra a metodologia proposta.

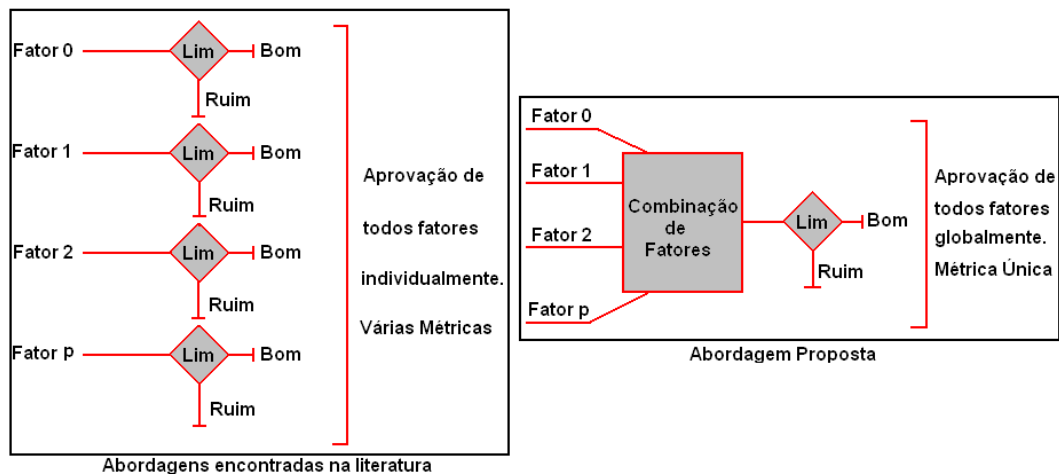


Figura 1.3 – Metodologia proposta.

São ainda objetivos específicos do presente trabalho:

- Validar a metodologia, construindo uma métrica de qualidade orientada ao reconhecimento facial baseado em textura representada especificamente nos chamados *padrões binários locais*,
- Avaliar abordagens alternativas de projeto – Métricas de Qualidade (linear e não linear)
- Identificar o conjunto de atributos mais relevantes para caracterização da qualidade da imagem facial

Cabe mencionar que o trabalho apresentado no capítulo 2 item 2.6 – Avaliação de Qualidade de Imagens Faciais também propõem uma investigação de medidas de qualidade só que bem menos abrangente já que cada atributo é

avaliado individualmente e usando somente métodos lineares. O presente trabalho faz uma análise focando os seguintes pontos.

- Propõe um método que conjuga vários atributos;
- Faz uma seleção do melhor conjunto de atributos;
- Faz uma análise dos resultados através de métodos lineares e não lineares.

Cabe também mencionar que o método de reconhecimento facial em que se baseia este estudo foi desenvolvido pelo autor da presente dissertação em parceria com seu orientador. Um sistema baseado neste método está hoje em fase de implantação na Diretoria de Identificação Civil do Estado do Rio de Janeiro.

1.1

Organização do Trabalho

Essa dissertação está organizada em seis capítulos assim divididos.

Capítulo 1 – Introdução (este).

Este capítulo faz uma introdução da presente dissertação mostrando aplicabilidades para o trabalho aqui desenvolvido e cita alguns pontos que serão explorados nessa dissertação.

Capítulo 2 – Trabalhos Relacionados.

Este capítulo apresenta um sumário dos principais trabalhos existentes na literatura científica que de alguma forma estão relacionados ao mesmo problema aqui estudado.

Capítulo 3 – Fundamentos Teóricos.

Este capítulo detalha os fundamentos teóricos utilizados nos experimentos realizados nessa dissertação.

Capítulo 4 – Métodos de Avaliação de Qualidade.

Este capítulo detalha o método proposto e as informações serão extraídas das imagens de uma face usadas no processo de avaliação da qualidade quando usadas em um sistema de reconhecimento facial.

Capítulo 5 – Análise Experimental.

Este capítulo detalha as análises experimentais realizadas, a base de dados usada nos experimentos, os recursos computacionais utilizados e os resultados obtidos nos experimentos.

Capítulo 6 – Comentários Finais e Direções Futuras.

Este capítulo faz uma síntese e comentários dos experimentos e aponta algumas direções futuras relacionadas ao estudo aqui apresentado.

Ao final da dissertação apresenta-se a Relação das referências biográficas utilizadas para construção dessa dissertação.

2

Trabalhos Relacionados

Muitos pesquisadores têm publicado artigos sobre avaliação e recuperação de qualidade de imagens para o uso em sistemas biométricos. No que diz respeito a imagens faciais, os algoritmos de avaliação e recuperação de qualidade concentram-se predominantemente nas seguintes três características:

- Foco/Nitidez;
- Contraste;
- Artefatos Indesejáveis.

Este capítulo apresenta um sumário dos principais trabalhos existentes na literatura que tratam da avaliação e realce destas características, especialmente em imagens faciais.

2.1

Reconhecimento Facial sob Condições Difíceis de Iluminação

Tan e Triggs (2007)[01] apresentaram uma proposta que consiste de uma sequência de algoritmos de fácil implantação e de grande eficiência que elimina a maioria dos problemas de iluminação não uniforme preservando, entretanto os detalhes necessários ao reconhecimento facial. Essa proposta é constituída de três etapas.

A primeira etapa é uma correção Gama que consiste em uma transformação não linear dos níveis de cinza que substitui o valor $I(i, j)$ por $I^\gamma(i, j)$. Este processo tem o efeito de ampliar a faixa dinâmica local da imagem nas áreas escuras ou sombreadas e de reduzir nas áreas claras e muito iluminadas.

A segunda etapa consiste de uma filtragem por diferença gaussiana já que a correção Gama não remove a influência de todos os gradientes de intensidade, tais como efeitos de sombreamento entre outros.

Essa filtragem é feita utilizando-se dois filtros. Os detalhes finos são criticamente importantes para o processo de reconhecimento facial. O filtro gaussiano interno é tipicamente bem estreito com o valor do desvio padrão menor

ou igual a um pixel, enquanto o externo pode ter o valor do desvio padrão entre dois a quatro pixels ou mais, dependendo da frequência espacial em que a informação de baixa frequência torna-se mais enganosa do que informativa.

Já a terceira etapa consiste em uma equalização de contraste que re-escala globalmente as intensidades da imagem. É importante utilizar um estimador robusto porque o sinal tipicamente ainda contém uma pequena mistura de valores extremos produzidos pelas áreas mais iluminadas da imagem, lixo nas bordas da imagem e regiões escuras tais como as narinas.

Esta técnica de pré-processamento foi examinada em um estudo orientado pelo autor da dissertação e se mostrou bastante satisfatória, conforme mostrado no gráfico da figura 2.1 onde pode ser observado que a taxa de acerto no processo de reconhecimento facial é substancialmente mais alta quando as imagens de faces são pré-processadas usando o algoritmo aqui descrito.

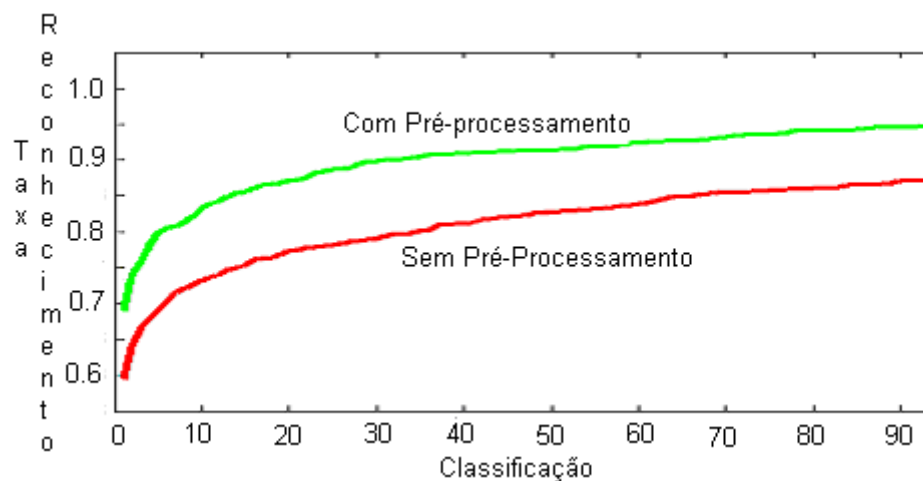


Figura 2.1 – Taxa de reconhecimento facial.

2.2

Eliminação de Artefatos em Imagens Frontais de Face

A implantação de um sistema de reconhecimento facial nem sempre parte do estado da arte em termos de técnicas de aquisição, isto é, muitas vezes faz-se necessário utilizar imagens que foram adquiridas há vários anos e que hoje se encontram coladas em fichas e ou em outros documentos e cuja qualidade já se encontra bastante deteriorada, não só pela passagem do tempo, mas também por artefatos introduzidos manualmente, como por exemplo, anotações, carimbos, etc.

Esta seção apresenta dois algoritmos que visam eliminar artefatos incorporados à imagem por ação humana.

O primeiro algoritmo proposto por André Sobiecki [02] e coautores consiste um arcabouço computacional de segmentação e restauração automática de rasuras baseadas em um modelo estatístico construído a partir de várias imagens faciais frontais. A localização das características faciais (nariz, olhos e boca) é feita com base em uma imagem de referência gerada a partir de uma população amostral que fornece informações *a priori* sobre tons e localização espacial dessas características. As rasuras identificadas na imagem são posteriormente eliminadas por um método que consiste na recuperação de trechos perdidos ou deteriorados de uma imagem. Os experimentos deste estudo mostram que a segmentação baseada no método estatístico apresenta melhores resultados de correlação com a imagem de referência do que outros métodos implementados para comparação, como os métodos baseados em entropia (medida da quantidade de informação na imagem) e no uso do filtro de Sobel. A implementação do método é composta de quatro etapas.

Na primeira etapa um filtro de mediana é aplicado sobre a imagem que se deseja restaurar.

Na segunda etapa onde é feita a identificação de artefatos, duas técnicas de segmentação automáticas são usadas.

A primeira técnica de segmentação consiste em um método estatístico de decisão onde a partir de uma população amostral de imagens frontais com expressão facial neutra e normalizada geometricamente calcula-se a imagem média dessas amostras

A segunda técnica da segunda etapa consiste em um processamento onde seja possível distinguir o plano de fundo da imagem do objeto de interesse. Isso pode ser feito através da aplicação de um limiar segregando áreas com maior ou menor quantidade de informação.

A terceira etapa do processo consiste na aplicação de uma máscara evitando assim que regiões faciais possam ser descaracterizadas devido ao processo de restauração digital adotado. Um molde binário é aplicado para cobrir as regiões dos olhos, do nariz e da boca conforme a figura 2.2. Os pixels ativos (brancos) representam as regiões a serem desconsideradas e os pixels pretos representam as regiões que podem sofrer restauração. Vale lembrar que, para criação do molde as dimensões das imagens foram normalizadas tendo como base a distância inter-ocular (entre os centros dos olhos).



Figura 2.2 – Molde de regiões da face.

Na quarta e ultima etapa a reconstituição de cada pixel é baseada na assimilação das informações contidas nos seus pixels vizinhos. Este método pode ser entendido como uma técnica de restauração de imagens danificadas por meio de interpolação de imagens que busca reconstituir as áreas afetadas por artefatos.

Segundo os autores do trabalho, imagens segmentadas pelo modelo estatístico obtiveram valores maiores de correlação com a imagem de referência, considerando apenas faces frontais e de expressão facial neutra, do que as imagens originais, segmentadas pela entropia e segmentadas pelo método Sobel. No modelo estatístico, cada imagem com artefatos foi comparada com a imagem de referência para um nível de significância estatístico equivalente a 95% de certeza com relação aos defeitos na imagem.

O segundo algoritmo volta ao mesmo tipo de problema e foi proposto por André Sobiecki [03] e coautores, que apresentaram um método para recuperação de imagens de face guardadas em acervos e usadas para reconhecimento facial.

O algoritmo proposto identifica e elimina pequenas anotações, mensagens ou borrões sobre a imagem que impedem uma identificação mais exata da pessoa. Além de identificar e eliminar as partes rasuradas, a imagem é reconstituída a partir das regiões vizinhas, formando uma nova imagem restaurada. Isto significa que, mesmo com as rasuras eliminadas, o resultado preserva as características da imagem, caso a rasura esteja sobre partes do nariz, dos lábios ou sobre qualquer outra região da face humana. Uma das contribuições relevantes deste método é a reconstituição automática das regiões da face humana, sem perder sua forma, pois a reconstituição é aplicada através da cópia e da assimilação das regiões vizinhas.

O algoritmo envolve uma sequência de passos bastante simples e de fácil implementação, conforme mostra a figura 2.3.

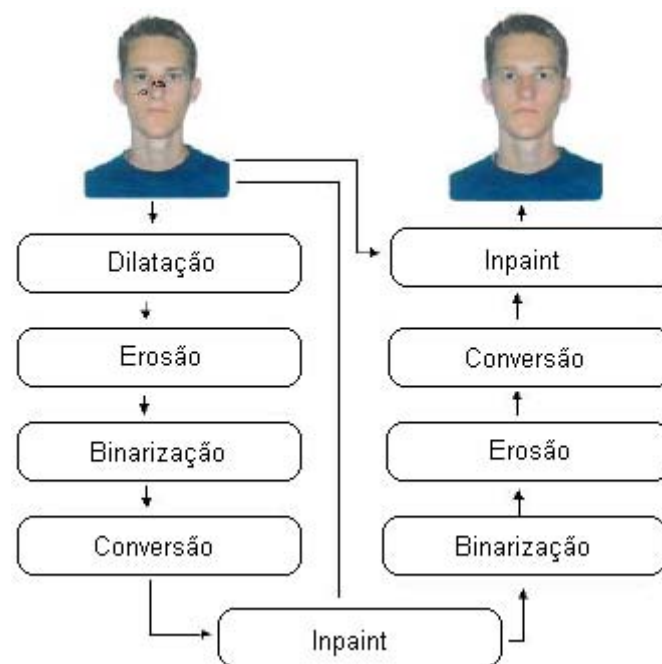


Figura 2.3 – Proposta de melhoria automática em imagens faciais.

Fases do processo de recuperação.

- Dilatação

O processamento da imagem inicia pela dilatação que é a etapa onde as rasuras são preenchidas por pixels de tom claro.

A dilatação torna as regiões escuras menores. Se estas regiões escuras forem pequenas, elas somem por que são preenchidas por regiões claras.

- Erosão

Com a erosão, o cabelo, os olhos, e a camisa irão recuperar seus tamanhos e suas tonalidades perdidos no processo de dilatação.

- Inpaint

Tem como objetivo remover defeitos e reconstituir partes danificadas tornando a imagem mais legível, também é possível remover partes indesejadas e escritas sobre a imagem. Ao remover as partes indesejadas podemos suavizar os contornos. Tendo a área a ser eliminada definida, a técnica “inpainting” cria uma estrutura através de cópias e estimações de regiões vizinhas conseguindo assim reconstruir contornos e tons. Em áreas de falta ou áreas danificadas só podemos esperar uma produção plausível e não uma exata reconstrução. Para que o “inpainting” possa ser razoavelmente bem sucedido é necessário que as regiões de interesse sejam relativamente pequenas.

- Binarização

A binarização é a transformação da imagem de nível de cinza para binária. A imagem binária tem somente duas cores: preto e branco.

- Conversão

É o processo de inversão de cores. A imagem tem suas cores invertidas para se tornar uma máscara. Máscara é a imagem que define as regiões que serão reconstituídas no método “inpainting”.

Os autores do método realizaram vários experimentos obtendo desempenho que julgaram excelentes, chegando a um nível de recuperação de 90 por cento das imagens submetidas ao processo de restauração. O método só se mostrou falho quando as rasuras se encontravam em grandes regiões escuras

como, por exemplo, na região do cabelo o que não é um problema para os sistemas de reconhecimento facial visto que estes sistemas somente levam em conta a região da face.

2.3

Correção Gama em Imagens de Forma Localizada

Visando melhorar o contraste de imagens, Mahmood Farshbaj Doustar e Hamid Hassanpour [04] propuseram um método adaptativo para correção Gama. O método estima diferentes valores de Gama, equação 2-1, para diferentes regiões da imagem e transforma cada pixel da imagem individualmente.

$$S = cr^{\gamma} \quad (2-1)$$

onde $r \in [0,1]$ representa a intensidade dos pixels da imagem, “c” uma constante positiva intrínseca do equipamento e γ uma constante positiva – Gama.

Um dos mais importantes fatores que afetam diretamente o desempenho de muitos sistemas de reconhecimento facial é a direção em que incide a luz sobre a cena.

O método prevê a criação de um banco de dados contendo histogramas de intensidade de pixels. Para isso foram usadas nove imagens padrão onde cada imagem passou por uma correção Gama entre 0.1 e 3.0 com um passo igual a 0.1. Em seguida, vetores de características para cada imagem foram computados a partir de uma janela de tamanho de 128 x 128 pixels que foi deslizando a partir do topo do lado esquerdo até o canto inferior direito com um passo de 16 pixels em cada movimento.

Portanto, um vetor de características e valores de gama entre 0.1 e 3.0 foram gerados para cada janela construindo assim um banco de dados de referência a partir nas imagens usadas para treinamento.

Para determinar o valor de Gama para uma determinada imagem, uma janela deslizante de tamanho 128 por 128 é movida sobre a imagem a partir do canto superior esquerdo ao inferior direito por dezesseis pixels em cada movimento, e então um valor de correção para cada janela é calculado a partir do banco de dados gerado. Depois de calcular o vetor de características para cada

janela como mencionado anteriormente, a técnica do vizinho mais próximo (KNN do inglês K-Nearest Neighbor) compara esses vetores de entrada com vetores conhecidos no banco de dados e atribui um valor de Gama para os pixels da janela em questão para cada pixel em janelas sobrepostas, como mostrado na figura 2.4. O valor de Gama é calculado como a média dos valores de Gama para as janelas que contém o pixel. Isto divide a imagem de entrada em regiões com tamanho de 16×16 fazendo com que todos os pixels nessas janelas tenham o mesmo valor de Gama.

Finalmente, para ter a melhor correção Gama, o valor da correção para cada região é distribuído pelos pixels pelas seguintes regras.

- O valor da correção Gama para pixels de fronteira (entre duas janelas vizinhas de 16×16) é a média de valores de correção Gama para cada janela.
- Em cada janela indo para seu centro, o valor da correção Gama flui linearmente com o valor da correção Gama calculado.

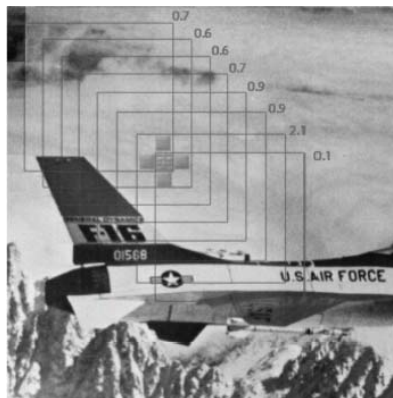


Figura 2.4 – Janelas sobrepostas de tamanho 128×128 pixels e seus valores de Gama associados.

Como pode ser observado na figura 2.5, as imagens corrigidas pelo método proposto têm uma melhor distribuição dos tons de cinza, gerando conseqüentemente um histograma de pixels mais uniforme o que é fundamental para o processo de reconhecimento facial.



Figura 2.5 – Resultados da correção Gama adaptativa.

Segundo os autores, o algoritmo também pode ser ajustado para trabalhar com janelas deslizantes de tamanho variável podendo proporcionar um melhor desempenho.

2.4

Função P.S.F. Baseado na Distribuição de Vetores Gradientes em um Plano Polar

Morihiro SAKANO [05] e coautores elaboraram uma proposta para restauração de imagens com problemas de foco baseada na distribuição de vetores gradientes em um plano polar.

Restauração da imagem é uma das tarefas fundamentais na área de processamento de imagem e vários algoritmos para essa finalidade já foram desenvolvidos. Quando uma imagem é adquirida por um sistema qualquer de captura, a degradação é muitas vezes inevitável. Uma das principais causas desta degradação é a falta de foco devido a uma incompatibilidade da distância focal e distância do objeto.

A grande maioria dos algoritmos de restauração de imagens com problemas de foco partem do princípio que a função P.S.F. (point spreading function) do sistema de aquisição é conhecida tomando como base informações do fabricante do equipamento. Entretanto é muito difícil saber de antemão a função

P.S.F. Os algoritmos não apresentam, portanto na prática, um desempenho satisfatório para o problema de restauração. Conseqüentemente a estimação com exatidão da função P.S.F. se torna de extrema importância na recuperação do foco de imagens desfocadas.

A figura 2.6 apresenta um modelo do processo de aquisição que é representado pelas equações 2-2 e 2-3 no domínio espacial e frequência respectivamente.

$$g(x, y) = f(x, y) * h(x, y) + n(x, y) \quad (2-2)$$

$$G(u, v) = F(u, v)H(u, v) + N(u, v) \quad (2-3)$$

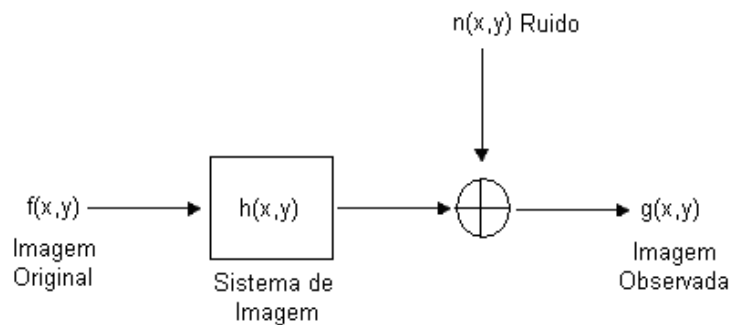


Figura 2.6 – Sistema de imagem em espaço linear invariante.

No algoritmo proposto, o espectro de amplitude $G(u, v)$ é convertido em um periódico linear de “uns” mapeando-os em um plano polar $\mu\phi$ conforme figura 2.7. Para estimar o raio do C.o.C. (círculo de confusão que é o maior círculo em uma imagem dentro do qual se considera que pontos diferentes não são distinguíveis) os mínimos periódicos lineares são encontrados nas direções dos vetores gradientes conforme figura 2.8 e sua implantação consiste de cinco etapas:

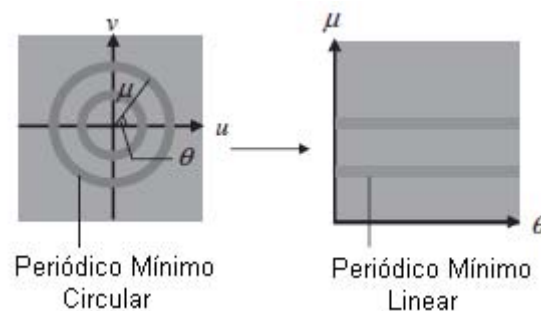


Figura 2.7 – Mapeamento do espectro de amplitude logarítmica da imagem observada ao plano polar.

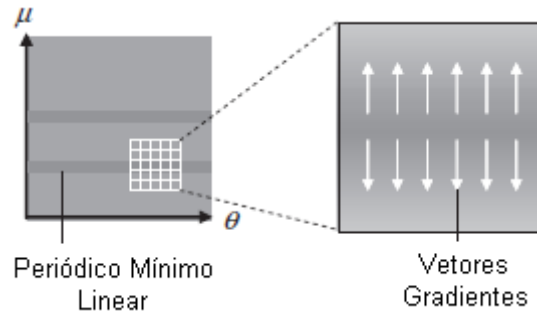


Figura 2.8 – As direções dos vetores gradiente em torno da mínima.

A primeira etapa consiste de um mapeamento em um plano polar. Neste procedimento o periódico circular no espectro de amplitude logarítmica de $G(u,v)$ são convertidos em periódicos lineares sendo mapeados em um plano polar $P(\phi, \mu) = \ln|G(u \cos \phi, \mu \sin \phi)|$.

A segunda etapa consiste na aplicação de um filtro de mediana de tamanho variável aplicado em $P(\phi, \mu)$ para suprimir as flutuações do gradiente causadas por ruídos. O tamanho da janela do filtro deve ser alterado de forma gradual dependendo da frequência e deve ser maior com o aumento da frequência.

A terceira etapa consiste no cálculo das direções dos vetores de gradiente $\phi(\phi, \mu_j)$. Nessa etapa deve ser observado que os vetores gradientes tendem a flutuarem pelo ruído no caso onde o tamanho da janela de cálculo da segunda etapa ser muito pequena.

A figura 2.9 mostra um exemplo das direções dos vetores gradientes onde pode ser observado que em (a) a direção do vetor gradiente é zero e em (b) a direção é 45 graus.

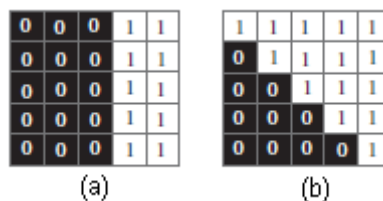


Figura 2.9 – Direção dos vetores gradiente.

A quarta etapa consiste em encontrar o periódico linear mínimo no plano polar usando a ortogonalidade entre os vetores gradientes e a linha hipotética que é paralela ao eixo ϕ figura 2.10.

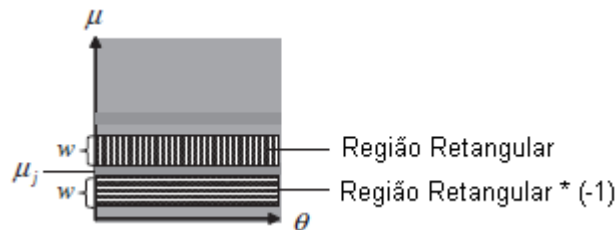


Figura 2.10 – Janelas retangulares e a linha hipotética.

A quinta e última etapa consiste no cálculo do Circulo de Confusão (C.o.C.) que é o maior círculo em uma imagem dentro do qual se considera que pontos diferentes não são distinguíveis.

Os autores concluíram que para estimar a função P.S.F. para imagens fora de foco eles propuseram um algoritmo novo e robusto com especial atenção sobre a direção dos vetores de gradiente sobre o espectro de amplitude mapeado em um plano polar. Usando esse algoritmo, a restauração de imagens fora de foco foi realizada com precisão e firmeza, mesmo em um ambiente altamente desfocado e ruidoso. A eficácia e validade do algoritmo proposto foram verificadas aplicando-a aos problemas de estimação do P.S.F. para imagens fora de foco e com ruídos.

2.5

O Método de Landweber

Diversos métodos de restauração de imagens borradas foram propostos na literatura (p.ex. [06], [07], [08], [09]). Esta seção apresenta um dos métodos mais simples desta classe, método Landweber (LM), e faz menção de uma de suas extensões.

O método de Landweber é um processo iterativo para resolver problemas de foco que utiliza um parâmetro de atualização constante durante um o processo iterativo, o que resulta em uma convergência lenta.

No método ALM ou adaptativo, o parâmetro de atualização é calculado a cada interação, fazendo com que o método convirja mais rapidamente e obtenha

um erro médio quadrático (quadrado da diferença entre a imagem resultante e uma dada imagem referência) menor do que o método LM.

Considere-se, por exemplo, uma imagem \mathbf{I} de tamanho $\mathbf{M} \times \mathbf{N}$ afetada por uma função de suavização gaussiana e também corrompida por um ruído que pode ser representada como $\mathbf{g} = \mathbf{H}\mathbf{f} + \mathbf{w}$, onde \mathbf{f} é o vetor resultante de se empilharem as colunas de \mathbf{I} , \mathbf{H} é a matriz denota o operador que borra a imagem, suposta conhecida e \mathbf{w} representa o ruído introduzido na imagem, este desconhecido.

O método Landweber padrão (LM) tenta achar uma solução como mostrado na equação 2-4.

$$\begin{aligned} \mathbf{f}_{\text{LM}}^0 &= \mathbf{0} \\ \mathbf{f}_{\text{LM}}^{(k+1)} &= \mathbf{f}_{\text{LM}}^{(k)} + \beta \mathbf{H}^T (\mathbf{g} - \mathbf{H}\mathbf{f}_{\text{LM}}^{(k)}) \end{aligned} \quad (2-4)$$

onde β é determinado por $0 < \beta < \frac{2}{\sigma_1^2}$ sendo σ o maior valor singular da matriz \mathbf{H} .

Lei Liang e Yuanchang XU [10], apresentam uma variante adaptativa do método de Landweber (ALM) que converge mais rapidamente do que o método original.

Da equação 2-4 que um valor elevado de β acelera a convergência para a imagem ideal mas aumenta o risco de instabilidade do sistema.

No método ALM as duas primeiras interações são exatamente iguais ao método LM. Os métodos passam a diferir a partir da terceira interação em diante, a partir de onde um novo valor de β é calculado a cada interação com base na razão da magnitude média do gradiente de imagens produzidas em duas últimas interações.

Em ambos os métodos, o critério de parada é satisfeito quando o valor do o erro $\mathbf{e}_{\text{LM}} = \|\mathbf{H}\mathbf{f}_{\text{LM}}^k - \mathbf{g}\|_2^2$ for menor que um determinado limite. O gráfico da figura 2.11 (Número de interações X (Erro Médio Quadrático-M.S.E.) e (Erro Médio Absoluto-M.A.E.)) mostra a vantagem do método ALM sobre o método LM para uma mesma imagem.

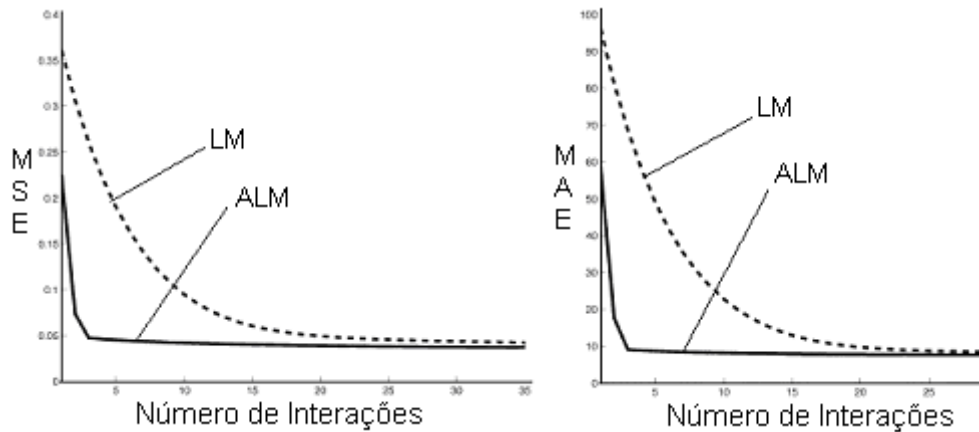


Figura 2.11 – Eficiência método ALM versus LM.

2.6

Avaliação de Qualidade de Imagens Faciais

Mohamed Abdel-Mottaleb e Mohammad H. Mahoor [11] realizaram um trabalho onde foram analisados três fatores que afetam de forma negativa o desempenho de sistemas de reconhecimento facial.

O primeiro desses fatores está ligado ao grau de embaçamento. Os autores propuseram o uso de uma medida de nitidez chamada Curtose, medida essa que também foi usada nessa dissertação e está detalhada no capítulo 4.

O segundo fator usado pelos autores está ligado ao efeito da luminosidade. Neste caso, propõem os autores uma medida que denominaram como WMI (acrônimo do inglês para “weighted sum of the mean intensity”), que também foi usada nessa dissertação e também está detalhada no capítulo 4.

O terceiro fator analisado se refere à avaliação do efeito da pose da cabeça. Os autores usaram uma medida chamada “pm” definida pela equação 2-5

$$pm = \frac{S_L - S_R}{\min(S_L, S_R)} \quad (2-5)$$

onde S_L e S_R são as áreas formadas pelo contorno da face e as linhas que cortam os olhos e centro da boca conforme figura 2.12.

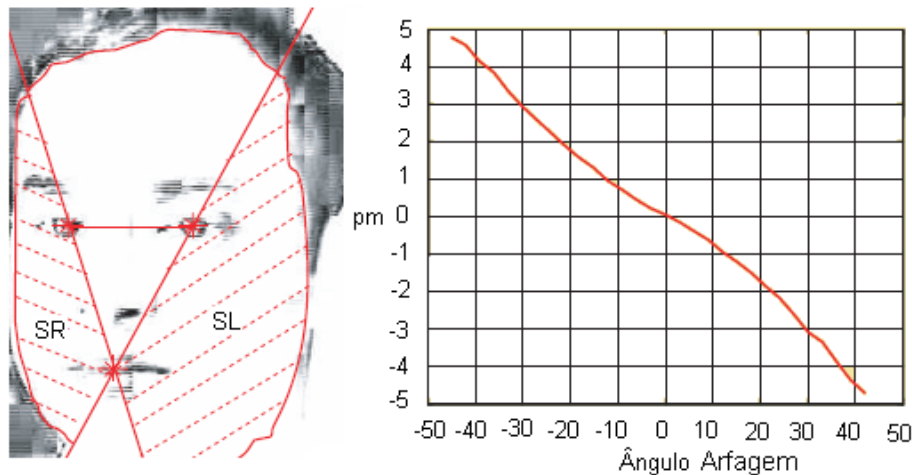


Figura 2.12 – Mapeamento do ângulo da face em função do valor de “pm”.

A partir desta medida pode-se estimar o ângulo conhecido como Arfagem (vide lado direito da figura 2.12) que diz respeito à rotação da face em torno de um eixo vertical que passa pelo centro da cabeça. Trata-se de um atributo crítico no contexto de reconhecimento facial, devido ao fato de que quando maior esse ângulo (independente do sentido) mais a imagem da face será descaracterizada. Essa dissertação faz o uso dessa medida usando uma abordagem um pouco diferente que será detalhada no capítulo 4.

- Os autores analisaram como esses três fatores interferem no desempenho de um sistema de reconhecimento facial. Basicamente concluíram que.
- A medida Curtose (detalhada no capítulo 4) que está ligada ao embaçamento mostra que quanto menor esse valor melhor será o desempenho já que a imagem estará mais nítida.
- A medida WMI (detalhada no capítulo 4) que está ligada a luminosidade também apresenta taxa de reconhecimento maior quando existe melhores condições de iluminação e como era esperado valores do fator “pm” próximos de zero indicam que a face está na posição frontal, isto é, está simétrica e portanto também apresenta melhores taxas de reconhecimento.

A figura 2.13 ilustra os resultados obtidos pelos autores nas três análises realizadas.

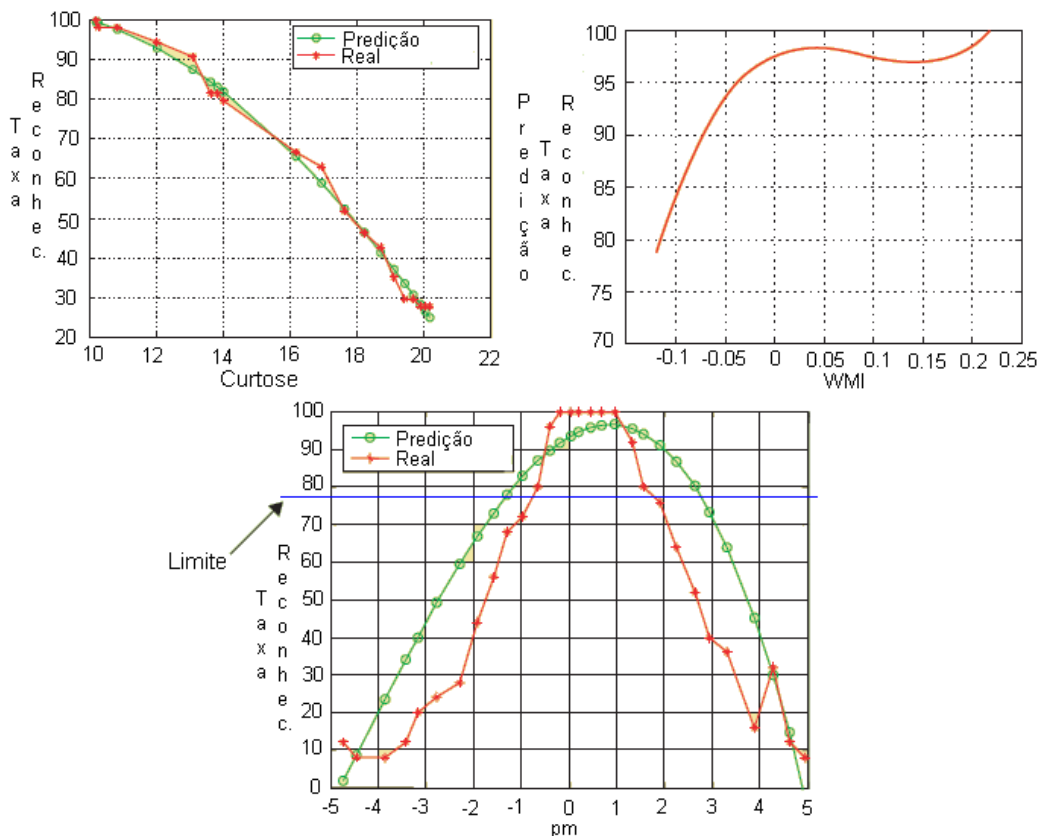


Figura 2.13 – Avaliação dos fatores do algoritmo para avaliação de qualidade de imagens faciais.

2.7

Padronização das Amostras de Face

Para a padronização das amostras de imagens de face, Xiufeng Gao [12] e coautores realizaram um estudo sobre a dependência do desempenho dos sistemas biométricos sobre a qualidade das amostras biométricas. Segundo os autores, a baixa qualidade de uma amostra é a principal causa para incidência de erros em sistemas biométricos e pode ser a principal causa do insucesso em implementações. Neste trabalho, os autores propuseram uma abordagem para padronização da qualidade de imagens facial, e desenvolveram um método que mede assimetrias faciais causadas por iluminação não frontal e poses impróprias.

Conforme também proposto pelo autor dessa dissertação, os autores desse trabalho também afirmam que somente imagens de qualidade aceitável devem ser aceitas em sistemas biométricos conforme o fluxo da figura 2.14.

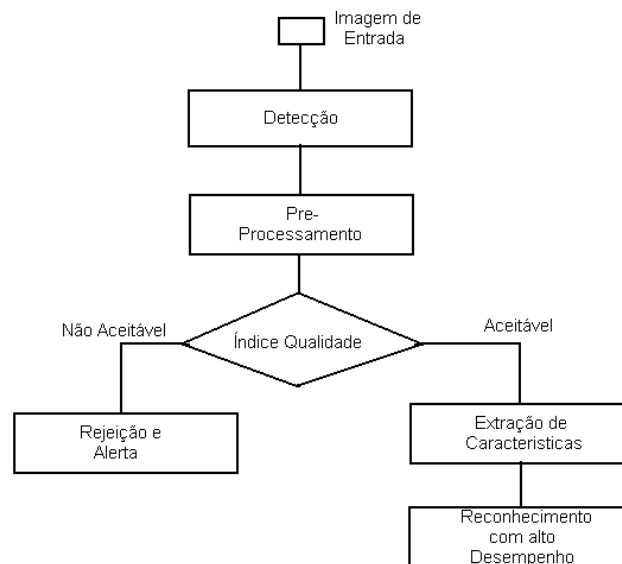


Figura 2.14 – Fluxograma de avaliação de qualidade de imagens.

O trabalho proposto pelos autores é composto de 4 procedimentos detalhados a seguir.

1. Procedimento 01 - Padronização.

- Abordagens de Padronização: As seguintes abordagens de padronização foram adotadas.
 - i. Especificação dos possíveis defeitos do rosto nas amostras biométricas em quatro categorias: ambiente, câmera, usuário e interação usuário versus câmera.
 - ii. Definir os Índices de Qualidade das Faces (IQF) para serem calculados pelo Algoritmo de Avaliação de Qualidade de Face (AAQF) para avaliação dos possíveis defeitos. Essa avaliação analisa uma imagem localmente ao nível de pixel ou como todo e combina os resultados das análises locais em uma região global. Essa avaliação é feita sobre um ou mais defeitos e fornece um indicador de falha.
 - iii. Mapear cada IQF de cada tipo de defeito em um índice de qualidade normalizado gerando um Índice de Qualidade de Face Normalizado (IQFN) que

indicará quão boa ou ruim é a qualidade da amostra e pode ser usada como indicador de possíveis problemas.

- iv. Mapear todos os IQF em um nível de qualidade global normalizada (IQFNG) fornecerá uma avaliação global da qualidade de uma amostra para reconhecimento biométrico.

- Categorização dos Defeitos.

A imagem do rosto obtida a partir de uma câmera estática, câmera de vídeo, foto ou scanner é geralmente imperfeita. Ela pode conter defeitos causados por iluminação deficiente, posicionamento impróprio da face e imperfeição da câmera. Estes fatores podem ser categorizados em quatro aspectos.

- i. Ambiente.

- 1. Variação da simetria da iluminação;
 - 2. Iluminação desigual da face;
 - 3. Iluminação muito fraca ou forte demais;
 - 4. Fundo impróprio.

- ii. Dispositivo de captura;

- 1. Baixo contraste;
 - 2. Distorção geométrica.

- iii. Condições da face;

- 1. Óculos grossos ou escuros;
 - 2. Expressões exageradas;
 - 3. Uso de maquiagem em excesso.

- iv. Posicionamento da câmera;

- 1. Variação da posição frontal;
 - 2. Distância (Imagem muito grande ou pequena);
 - 3. Fora de Foco (falta de nitidez);
 - 4. Oclusão parcial da face;

- Medidas de Qualidade de Face

A qualidade de uma imagem de face pode ser medida em vários níveis. Fato que deve ser observado é que não existe um acordo sobre o que deve ser considerado como um problema de qualidade ou que medidas têm como objetivo final maximizar os resultados dos sistemas reconhecimento facial.

O AAQF usa uma imagem \mathbf{I} como entrada e indica o Índice de Qualidade (IQ) que pode ser um escalar ou um vetor.

O Índice de similaridade entre duas imagens I_1 e I_2 pode ser associado como $IQ(I_1)$ e $IQ(I_2) \Leftrightarrow P(I_1, I_2; q_1, q_2)$.

2. Procedimento 02 - Simetria Facial.

A variação da pose e da iluminação são os dois principais motivos de causa de baixo desempenho em sistemas de reconhecimento facial. Os autores propõem duas análises para avaliação da qualidade da imagem

- Análise da simetria facial.

Essa análise é feita através de distâncias entre histogramas (DH) das regiões esquerda (RE) e direita (RD) da imagem de face conforme a figura 2.15 e tem seu cálculo conforme a equação 2-6 onde quanto maior a distância DH menor a qualidade da imagem.

$$DH_{\ell} = \left| H_{vxb}^{RE} - H_{vxb}^{RD} \right| \quad (2-6)$$

onde ℓ é o índice de um par de sub-janela, DH é alguma forma de distância de histograma, v a dimensão do vetor características e b o número de caixas do histograma.

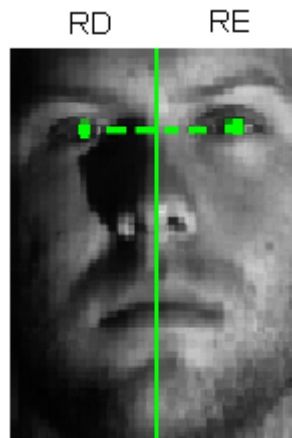


Figura 2.15 – Simetria facial.

- Iluminação e pose assimétrica.

Essa informação é medida usando uma propriedade sensitiva da imagem como, por exemplo, o código LBP (definido no capítulo 3) que é sensível à direção da iluminação e tem sido bastante usado em sistemas de reconhecimento facial. Essa assimetria pode ser medida pela distância entre os histogramas LBP e o valor de AAQF pode ser calculado conforme equação 2-7.

$$AAQF = \sum_{l=1}^{NW} DH \quad (2-7)$$

onde nw é o número de sub-janelas e DH são as distancias das sub-janelas.

Quanto maior for o valor obtido pela equação 2-7 menor será a simetria da face em termos de iluminação.

3. Procedimento 03 - Outras Métricas de Qualidade.

- Distâncias;

De acordo com os autores, a distância da câmera deve estar entre 1.2 e 2.5 metros da face e a distância entre os centros das pupilas deve estar em torno de 120 pixels.

- Intensidade de Iluminação;

Esta medida pode ser definida como a distância entre o histograma de uma imagem com iluminação padrão e o histograma da imagem sendo avaliada.

- Contraste;

Essa análise calcula o índice de qualidade (IQ) para o contraste da imagem.

- Nitidez;

Essa análise se refere a clareza dos detalhes da imagem de face.

4. Procedimento 04 - Normalização dos Índices.

- Pontuações Normalizadas sob o aspecto de Qualidade;

São obtidas para medir a qualidade no que diz respeito ao aspecto de causa de defeitos. As imagens de rostos serão associadas as suas respectivas categorias de qualidade através de pontuações obtidas por um sistema de reconhecimento facial com ajuste subjetivo do ser humano para corrigir possíveis divergências.

- Qualidade Geral Normalizada;

Um índice global normalizado é obtido para medir a qualidade global. A fim de obter uma única saída de todas as métricas de qualidade descritas nas seções anteriores, é necessário combinar a todas as pontuações das métricas de qualidade descritas e produzir uma pontuação de qualidade única.

Os autores realizaram dois tipos experimentos, iluminação assimétrica e posicionamento assimétrico, em ambos os casos o trabalho proposto se mostrou bastante satisfatório ficando os resultados dentro de suas expectativas.

2.8

Impacto de Imagens de Baixa Nitidez em Reconhecimento Facial Baseado em Função de Transferência Modular

Fang Hua [13] e coautores publicaram um artigo cuja a finalidade era de analisar o impacto de imagens com problemas e nitidez em sistemas de

reconhecimento facial. Para isso utilizaram a base de dados Q-FIRE [14] composta de imagens de rostos borrados ao longo de um intervalo de medição contínua gerada pela mudança de foco da câmera.

Para que esse artigo pudesse ser escrito, os autores usaram quatro medidas de qualidade de nitidez da imagem.

A primeira medida usada para examinar a qualidade da imagem é através da Transformada Discreta de Cosseno – DCT. A DCT tem sido amplamente utilizado em sistemas de compressão de imagens, devido à sua resposta uniforme desde baixa frequência até alta frequência.

Um borrão em uma imagem geralmente leva a perda de informações de alta frequência. Diferença entre DCT e IDCT são utilizadas para medir a nitidez em imagens. A DCT em uma imagem de tamanho $M \times N$ é definida como na equação 2-8.

$$C_{DCT}(x, y) = \alpha_x \alpha_y \sum_{i=1}^M \sum_{j=1}^N I_{i,j} \cos \frac{\pi(2i+1)x}{2M} \cos \frac{\pi(2j+1)}{2N} \quad (2-8)$$

onde

$$\alpha_x = \begin{cases} \frac{1}{\sqrt{M}} & , x = 0 \\ \sqrt{\frac{2}{M}} & , 1 \leq x \leq M-1 \end{cases} \quad \alpha_y = \begin{cases} \frac{1}{\sqrt{N}} & , y = 0 \\ \sqrt{\frac{2}{N}} & , 1 \leq y \leq N-1 \end{cases}$$

C_{DCT} são os coeficientes da DCT da imagem. A frequência horizontal aumenta a medida que x aumenta e a frequência vertical aumenta a medida que y aumenta. A maioria da informação da imagem está localizada em valores pequenos de x e y mostrando altas energias no canto superior esquerdo. As frequências espaciais radiais são representativas da nitidez da imagem.

A segunda medida usada pelos autores para medir a nitidez de uma imagem é dada pelo Gradiente Quadrado. A magnitude desse gradiente pode ser computada como a ordem de grandeza das bordas que mede a nitidez da imagem. O Gradiente Quadrado (SG) é definido conforme a equação 2-9.

$$SG(x, y) = \sum_{i=1}^M \sum_{j=1}^N \sqrt{|I_{i,j+k} - I_{i,j}|^2 + |I_{i+k,j} - I_{i,j}|^2} \quad (2-9)$$

onde $I_{i,j}$ representa a intensidade dos pixels de uma imagem de tamanho $M \times N$ e k representa o passo de diferenciação.

A terceira medida usada para avaliar a nitidez de uma imagem através de um método proposto pela empresa Cognitec (definida no capítulo 4) é dado pela equação 2-10 como sendo uma medida de densidade de bordas (ED).

$$D_{\text{Cog}} = \frac{1}{NM} \sum_i \sum_j \text{abs}[I_{i,j} - \bar{I}_{i,j}] \quad (2-10)$$

onde $\bar{I}_{i,j}$ corresponde a imagem de entrada $I_{i,j}$ filtrada por um filtro de mediana 3×3 , M e N as dimensões da imagem.

A quarta medida de nitidez usada é o laplaciano de método gaussiano que combina uma medida laplaciana com um filtro gaussiano e é definido pela equação 2-11.

$$\text{LoG}(x,y) = \frac{1}{\pi\sigma^4} \left[1 - \frac{x^2 + y^2}{2\sigma^2} \right] e^{-\frac{x^2 + y^2}{2\sigma^2}} \quad (2-11)$$

O grau de nitidez é maior se o resultado da convolução é alto. Dado que imagens faciais têm poucas informações de alta frequência, um operador 10×10 e um desvio padrão σ igual a um foram aplicados pelos autores nesse estudo.

O método proposto para a avaliação do impacto de imagens com baixa nitidez – Função de Transferência Modular. A Função de Transferência de Modulação (MTF) é amplamente utilizada como uma maneira padrão para avaliar o desempenho do sistema de imagem, porque pode fornecer uma expressão objetiva e quantitativa de qualidade de imagem, bem como a capacidade de cálculo a partir dos dados oriundos do projeto de lentes.

MTF é definido como a magnitude da função de transferência ótica, que é a transformada de Fourier da Função de espalhamento de um ponto (PSF) usando transformada de Fourier discreta que é uma aproximação numérica da transformada de Fourier, o MTF pode ser definido conforme a equação 2-12.

$$\text{MTF} = \text{DTF}[\text{PSF}] = \sum_{n=0}^{N-1} y_n e^{-ikn \frac{2\pi}{N}} \quad (2-12)$$

onde $K \in [0, N-1]$ e y_n é a posição do n^{th} pixel.

Assim, MTF permite a descrição simplificada das capacidades de resolução espacial no sistema de imagem.

O MTF também é definido como o contraste entre uma dada frequência espacial e as baixas frequências, que geralmente mede a intensidade das linhas pretas e brancas, como mostrado na Equação 2-13 e Figura 2.16. I_{\max} e I_{\min} são a intensidade máxima e mínima a partir da escala ótica.

$$MTF = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} \quad (2-13)$$

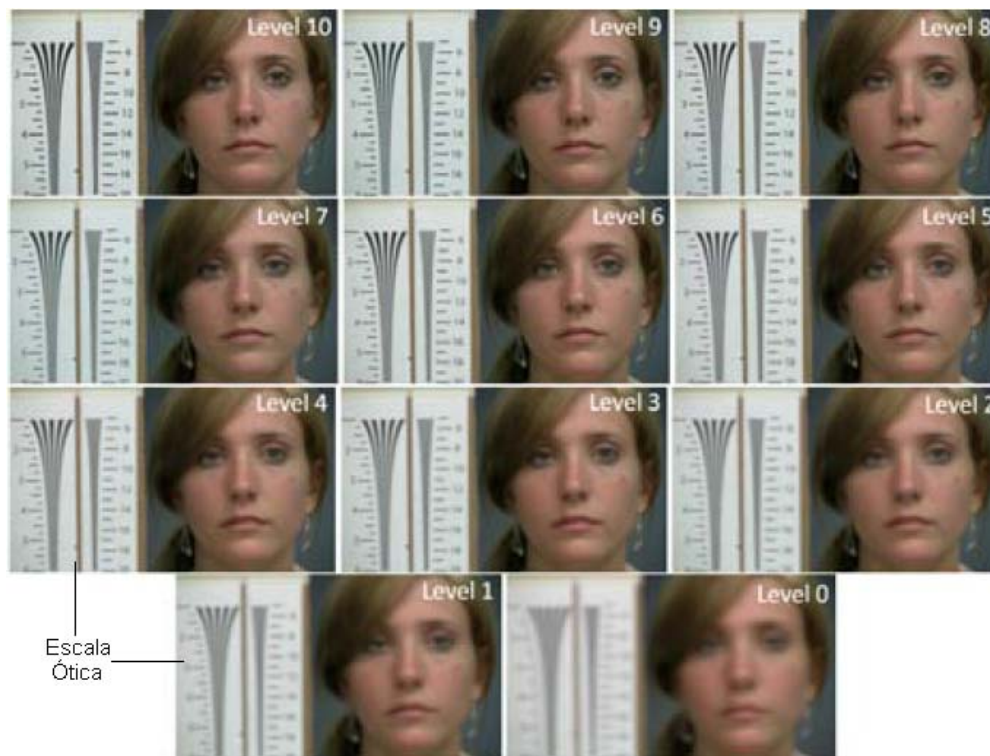


Figura 2.16 – Imagens borradas do banco Q-FIRE com escala ótica.

Os resultados obtidos pelo método são mostrados na figura 2.17 que mostra um exemplo comparando a resposta MTF na escala ótica (MTF-R) com quatro outros métodos mais a métrica de qualidade da SDK FaceIt [15], bem como a medição MTF da imagem da região da face fora de foco. Os pontos mais perto da linha a 45 graus em relação ao eixo X e Y são os métodos mais correlacionados com a MTF-R. O MTF na região do rosto mostra a maior correlação com o MTF-

R. As medidas DCT, SG, LoG mostram maior correlação enquanto a ED e a métrica de qualidade FaceIt são mais baixos como na Tabela 2.1 que mostras as correlações médias.

Tabela 2.1 – Correlação média entre MTF-R e médias de nitidez.

	LoG	DCT	SG	ED	Facelt	MTF (Face)
MTF-R	0.950	0.973	0.963	0.899	0.828	0.985

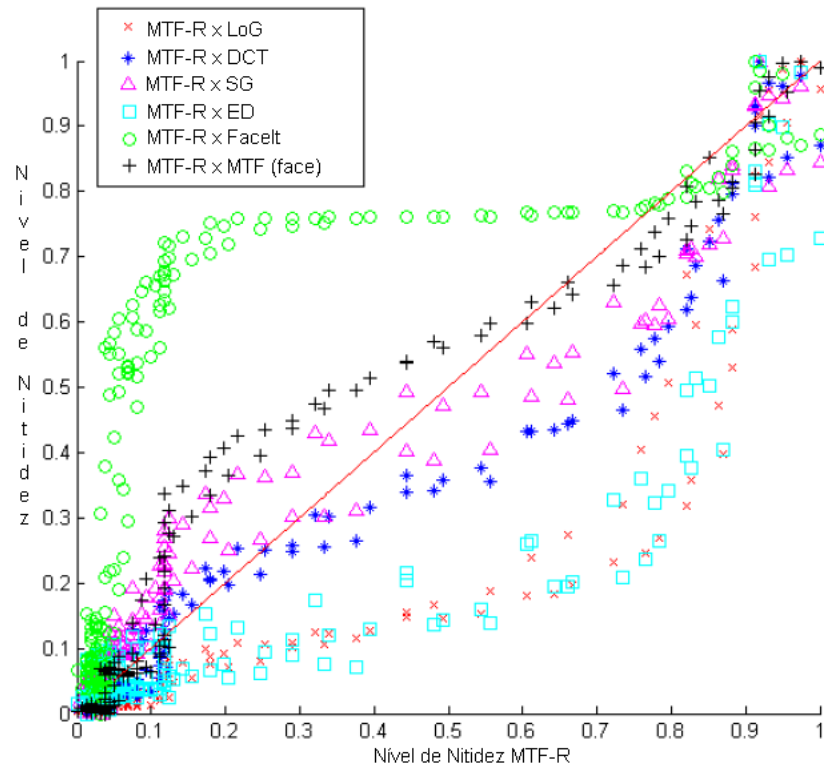


Figura 2.17 – Níveis de nitidez.

2.9

Melhores Práticas para Fotografias de Face e Métricas Qualidade de Imagem

Donald P. D’Amato [16] propôs um procedimento dividido em quatro etapas para assegurar a qualidade de imagens de face sob diversos aspectos.

A motivação para criação desses procedimentos se deve a dois fatos.

- Departamento de Estado dos EUA usando reconhecimento facial automatizado;

- Garantir precisão no reconhecimento facial automático.

Na primeira etapa os procedimentos propostos dizem a respeito ao estúdio fotográfico onde as imagens serão capturadas. As figuras 2.18, 2.19 e 2.20 mostram as posições ideais para câmera, Iluminação e usuário.

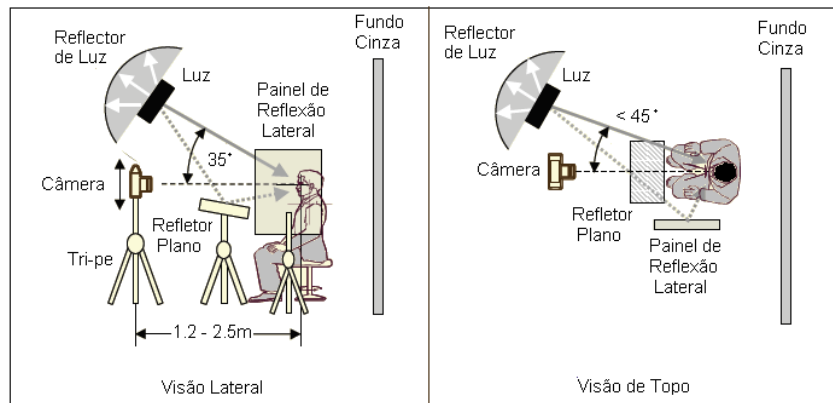


Figura 2.18 – Disposição de luz para estúdio com apenas uma lâmpada

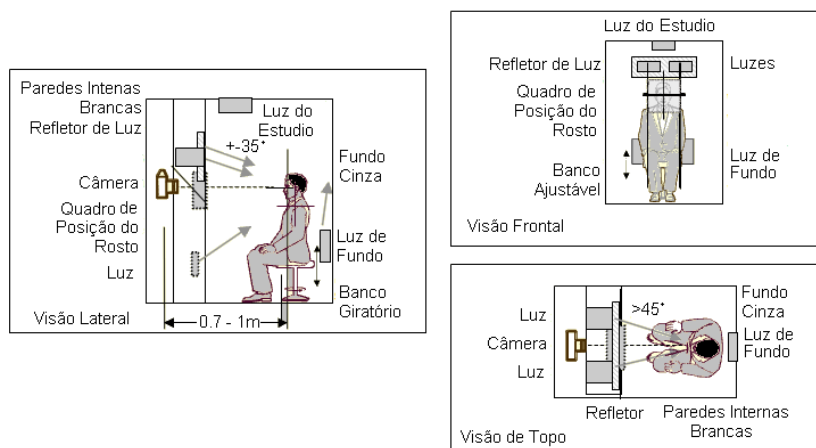


Figura 2.19 – Disposição de luz e câmera

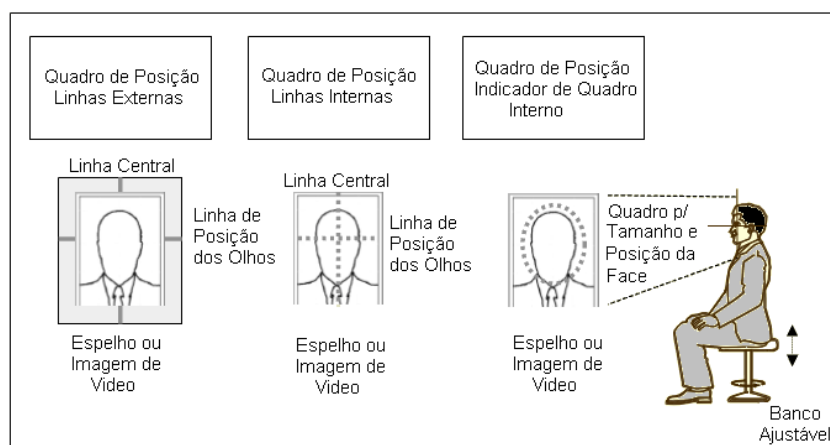


Figura 2.20 – Posicionamento da cabeça.

Na segunda etapa os procedimentos dizem respeito a uniformidade de iluminação.

A figura 2.21 mostra imagens com variações no ambiente de iluminação e detalha um tipo de problema que pode ocorrer quando o ambiente não está de acordo com os procedimentos propostos.

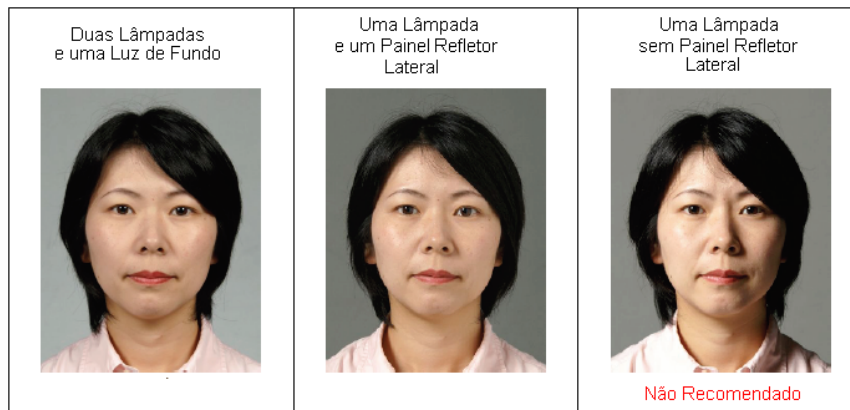


Figura 2.21 – Uniformidade da iluminação.

Na terceira etapa os procedimentos dizem respeito ao painel refletor e ao flash da câmera. A figura 2.22 mostra uma imagem que foi tomada usando apenas o flash câmera como fonte de iluminação onde fica bem evidenciado o problema gerado pela falta de uma iluminação adequada.

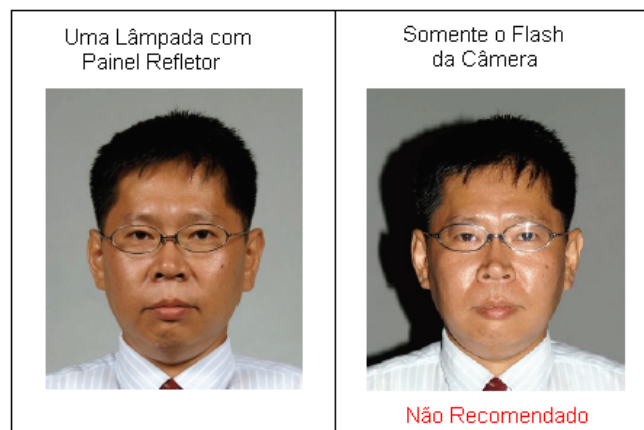


Figura 2.22 – Efeito do Flash.

Na quarta etapa o autor sugere medidas de qualidade de imagens baseadas em normas internacionais tais como:

- Alvos de teste para calibrar câmera (ou scanner) e medir suas características;

- Função de conversão opto-eletrônica (OECF) segundo a norma ISO 14524;
 - Resposta de frequência espacial (SFR) – segundo a norma ISO 12233 ou ISO 16067-1;
 - Medições de ruído segundo a norma ISO 15739;
 - Perfil de cores disponível em www.color.org.
- Especificações da norma ISO 19794-5.

Após o detalhamento das quatro etapas, o autor apresenta um resumo das métricas de qualidade desejadas para imagens de face.

- Paleta de cores;
 - Região facial deve ter pelo menos 7 bits de intensidade de variação (pelo menos 128 valores únicos) no rosto após conversão em escala de cinza.
- Olhos fechados / obstruídos;
 - Medido como um percentual, o valor deve refletir o grau de obstrução dos olhos, devido a óculos, franja, manchas, roupas ou os olhos fechados (neste caso, 100% obstruído).
- Balanceamento de cores;
 - Deve refletir as cores naturais com respeito aos tons de pele. Este valor pode ser afetado por balanço de branco inadequado ou de olhos vermelhos.
- Iluminação uniforme na região da face;
 - Medido como um percentual, o valor deve medir a simetria afetada por sombras ou pontos quentes no rosto.
- Uniformidade de plano de fundo;
 - Medido como um percentual, o valor deve medir a simetria e consistências afetadas por sombras no fundo, cenários texturizados, ou objetos estranhos no plano de fundo.
- Tamanho da cabeça;
 - A razão entre largura da cabeça e a largura da imagem deve estar entre 5:7 e 1:2.

- Centralização;
 - Posição vertical e horizontal da cabeça – Centralizar;
- Distância entre os centros dos olhos;
 - Medida em pixels deve ter pelo menos 90 pixel e tem como número ideal 120 pixels.
- Foco;
 - Nitidez e resolução da região da face. Profundidade de foco deve manter pelo menos 2 mm por pixel de resolução e de preferência de 1 mm por pixel. A imagem não deve ser excessivamente nítida.
- Ângulos;
 - Os ângulos de rotação da face sobre os eixos cartesianos não devem ultrapassar mais ou menos 5 graus.
- Confiança;
 - Confiabilidade de que os olhos foram achados e que se trata realmente de uma imagem de face;
- Brilho e contraste;
 - Resultados insatisfatórios podem ser gerados se as imagens estiverem muito claras ou muito escuras. A textura da pele deve estar visível.

3

Fundamentos Teóricos

Este capítulo descreve sucintamente os fundamentos teóricos usados nesta dissertação. Basicamente apresenta-se sucintamente a técnica de regressão linear, o método utilizado na dissertação para seleção de atributos e se descreve resumidamente o ajuste de funções através de redes neurais artificiais. Na última seção deste capítulo apresenta-se o método de reconhecimento facial utilizado como base neste trabalho.

3.1

Regressão Linear

A análise de regressão é uma técnica estatística para investigar e modelar a relação entre variáveis.

A expressão do modelo linear geral é dada pela equação 3-1.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \varepsilon \quad (3-1)$$

onde

Y é a variável dependente ou explicada, também chamada de resposta;

β_0 é o intercepto ou termo independente;

β_j para $j=0, 1, 2, \dots, p$, é a inclinação de Y em relação à variável X_j , mantendo constantes as variáveis X_k para $k=1, 2, \dots, p$ e $k \neq j$.

ε_i = erro.

Quando se têm várias (q) observações, a equação anterior toma a forma dada pela equação 3-2.

$$Y = X\beta + \varepsilon \quad (3-2)$$

onde

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_q \end{bmatrix} \quad X = \begin{bmatrix} 1 & X_{11} & \dots & X_{1p} \\ 1 & X_{21} & \dots & X_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{q1} & \dots & X_{qp} \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} \quad \epsilon = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_q \end{bmatrix}$$

Y_i é a medida relativa à i -ésima observação e X_{ij} é o valor do j -ésimo atributo na i -ésima observação.

A figura 3.1, ilustra a solução que se procura pela técnica de regressão linear num exemplo em que há 2 atributos. As observações são representadas por pontos azuis, e a solução encontrada é representada por um plano.

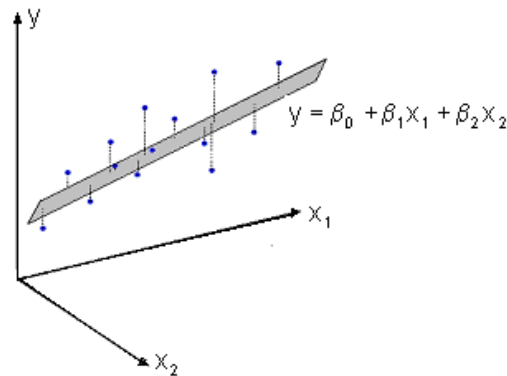


Figura 3.1 – Representação gráfica de Regressão.

A solução do problema de regressão linear consiste em determinar os valores do vetor de coeficientes β , uma vez dados Y e X para $\epsilon = 0$.

Em muitos casos práticos, e em particular no presente trabalho, o número de observações (q) excede o número de incógnitas ($p+1$). Trata-se, em outros termos, de um sistema de equações sobredeterminado, para o qual não existe em geral solução exata. Nestes casos, busca-se a solução que minimiza a soma dos quadrados dos resíduos, dada por $\|Y - X\beta\|^2$, formalmente conforme equação 3-3.

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}} \left[\|Y - X\beta\|^2 \right] = (X^T X)^{-1} X^T Y \quad (3-3)$$

Uma das métricas da qualidade de um modelo produzido por regressão linear é a raiz quadrada do erro médio quadrático (RMSE – do inglês *root mean square error*) que é dado pela equação 3-4.

$$RMSE = \sqrt{\frac{\sum |Y - X\hat{\beta}|^2}{q}} \quad (3-4)$$

onde q é o número de imagens utilizadas para o cálculo de β .

3.2

Seleção de Atributos

Um dos objetivos desta dissertação é selecionar o conjunto de atributos ótimo com base em que se pode estimar a qualidade de imagens faciais. Adotou-se neste trabalho o critério de informação de Akaike [17].

Hirotsugu Akaike [18] desenvolveu o Critério de Informação de Akaike – AIC (do inglês *Akaike information criterion*) que em sua forma geral é dado pela equação 3-5.

$$AIC = -2 * \ln(L) + 2 * p \quad (3-5)$$

onde \ln é o logaritmo natural, L é o valor da verossimilhança e p denota o número de parâmetros do modelo.

Quando se admite que os resíduos sejam distribuídos segundo uma normal, a equação anterior toma a forma da equação 3-6.

$$AIC = -q * \ln(RSS/q) + 2 * p \quad (3-6)$$

onde q é o número de observações no experimento, RSS é a soma dos quadrados dos resíduos.

O critério de cálculo de AIC requer um ajuste quando o número de amostras (q) é pequeno em comparação com o número de parâmetros a estimar (p), mais especificamente quando $q/p < 40$ (Regra de Ouro). Em tais casos o AIC, agora chamado de AIC_c , é corrigido como indica a equação 3-7.

$$AIC_c = -q * \ln(RSS/q) + 2 * p + (2 * p * (p + 1)) / (q - p - 1) \quad (3-7)$$

Vale notar que conforme o número q de observações aumenta em relação ao número de parâmetros p , o termo de ajuste à direita (viés) se torna muito pequeno. Portanto, o AIC_c pode ser utilizado sempre, independentemente da razão q/p . Este foi o procedimento adotado nesta dissertação.

Assim, o valor AIC_c de foi calculado para cada conjunto de atributos. Aquele conjunto ao qual correspondeu o menor valor de AIC_c foi considerado como conjunto ótimo.

3.3

Redes Neurais Artificiais

As Redes Neurais Artificiais – RNAs são sistemas paralelos distribuídos compostos por unidades de processamentos simples (nodos) que computam determinadas funções matemáticas (normalmente não lineares). Tais unidades são dispostas em uma ou mais camadas e interligadas por um grande número de conexões, geralmente, unidirecionais. Na maioria dos modelos estas conexões estão associadas a pesos, os quais armazenam o conhecimento representado no modelo e servem para ponderar a entrada recebida por cada neurônio (nodo) da rede.

Em RNAs, o procedimento usual para a solução de problemas passa inicialmente por uma fase de aprendizagem, onde um conjunto de exemplos é apresentado à rede, que do qual se extraem automaticamente as características do modelo. Essas características são utilizadas posteriormente para gerar respostas para o problema.

A Figura 3.2 representa a estrutura de uma RNA. O sinal de entrada, P , um vetor do espaço \mathbb{R}^r , é aplicado à entrada da rede. A entrada da rede neural consiste de uma camada de nodos que recebe “ r ” valores reais. A soma das entradas multiplicadas por pesos, (não indicados na figura 3.2), à qual é adicionado um valor constante, é aplicada à entrada de cada um dos nodos da camada seguinte, designada por 1ª camada. A saída de cada nodo da 1ª camada é dada por uma função, chamada função de ativação, aplicada ao valor na entrada.

Tipicamente, uma função de ativação é não linear, e, em todos os casos, suave, crescente, e limitada. Exceções são as aplicações em que uma RNA é utilizada para ajuste a um conjunto de observações como ocorre no presente

trabalho. Nestas aplicações a função de ativação usualmente utilizada no neurônio de saída é uma função linear, enquanto nas camadas ocultas utilizam-se funções não lineares. Este foi o esquema adotado nos experimentos descritos no capítulo 5.

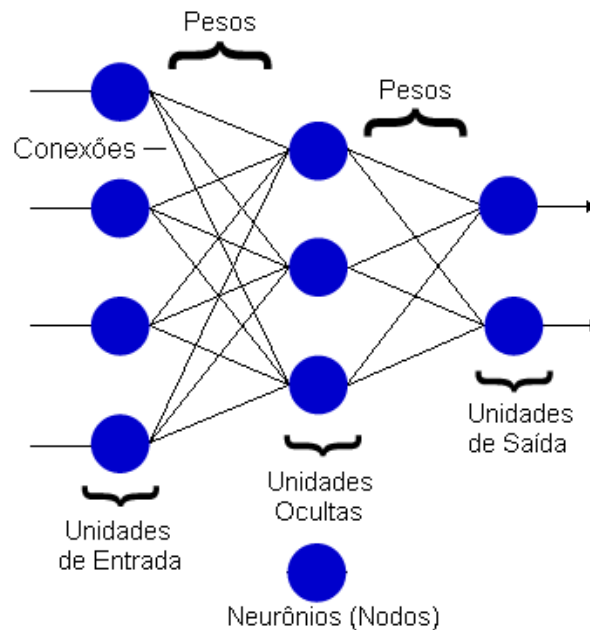


Figura 3.2 – Representação de uma Rede Neural Artificial (RNA).

3.4

Reconhecimento Facial

Os humanos baseiam-se freqüentemente na face para o reconhecimento de indivíduos. Por sua vez, os avanços das últimas décadas na capacidade de computação permitem atualmente um reconhecimento similar, mas de forma automatizada. Os primeiros algoritmos de reconhecimento da face utilizavam modelos geométricos simples, mas o processo de reconhecimento já atingiu um nível de maturidade que lhe permite apresentar-se como uma ciência de representações matemáticas sofisticadas e processos de comparação.

Os grandes avanços e iniciativas registrados nos últimos 10 a 15 anos trouxeram a tecnologia de reconhecimento da face para a ribalta. O reconhecimento de face pode ser utilizado atualmente tanto para fins de identificação, como de verificação.

Um sistema de reconhecimento facial é composto de dois processos conforme ilustrado na figura 3.3.

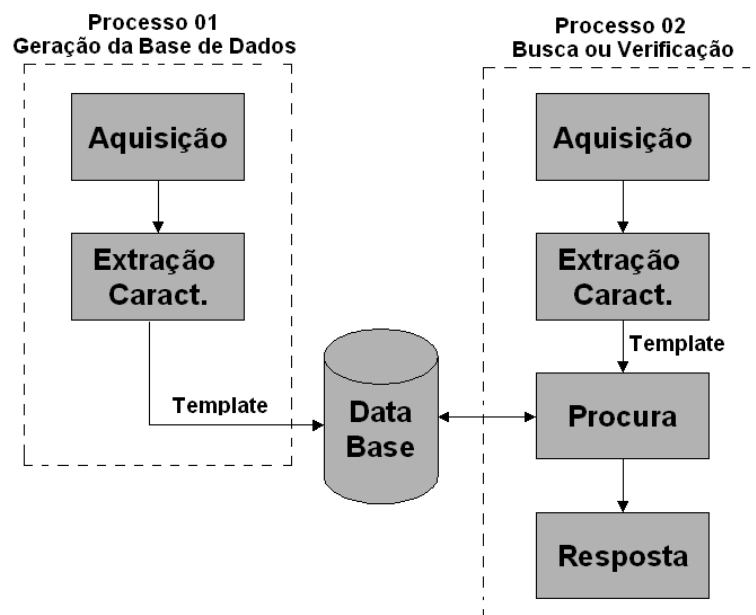


Figura 3.3 – Processos de um sistema de reconhecimento facial.

O processo 01 consiste na geração da base de dados de face – processo de identificação e é composto das seguintes etapas:

- **Aquisição.** Esta etapa consiste da aquisição da imagem de face que pode ser executada através de qualquer método de captura. Nesta fase já devem ser aplicados métodos conforme descritos no capítulo 2 com a finalidade garantir a qualidade das imagens adquiridas.
- **Extração de Características:** Esta etapa consiste da extração das características da face resultado em uma informação organizada conhecida na literatura como *template* que é armazenada em uma base de dados.

O processo 02 consiste na busca ou verificação e é composto das seguintes etapas:

- **Aquisição:** Esta etapa é semelhante a aquisição do processo 01.
- **Extração de Características:** Esta etapa também é semelhante a extração de características do processo 01.
- **Procura:** Esta etapa consiste de comparar o *template* gerado na etapa de extração de características do processo 02 com os *templates* armazenados na base de dados.
- **Resposta:** Esta etapa consiste em relacionar os *templates* existentes na base de dados semelhantes ao *template* gerado na etapa de

extração de características do processo 02 permitindo assim saber se uma pessoa já foi identificada ou não.

Diversos métodos de reconhecimento facial têm sido propostos ao longo de mais de três décadas [19]. Esta seção descreve um método de reconhecimento que se baseia em informações de texturas da imagem.

O processo de extração da textura das imagens faciais pelo sistema de reconhecimento facial usado nessa dissertação foi o LBP (do inglês *Local Binary Pattern*) [20], [21], [22] que é um código binário que representa o arranjo de intensidades dentro de uma vizinhança de cada pixel da imagem. Para o cálculo do LBP de cada pixel de uma imagem, um número de m pixels de uma vizinhança de raio R em torno do pixel é usado, conforme mostra a figura 3.4.

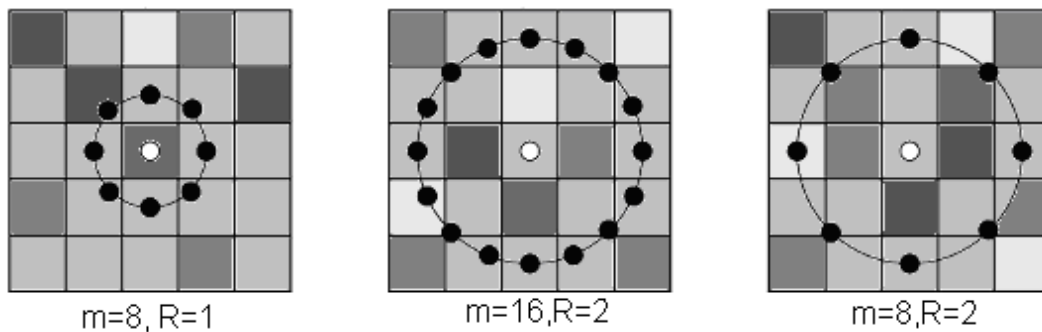


Figura 3.4 – Conjunto de vizinhança de pixels para valores de m e R .

Uma vez decido os valores de m e R , o cálculo do valor do LBP para cada pixel se dá conforme a figura 3.5 que ilustra o processo usando um valor de $m=8$ e $R=1$ usados nessa dissertação.

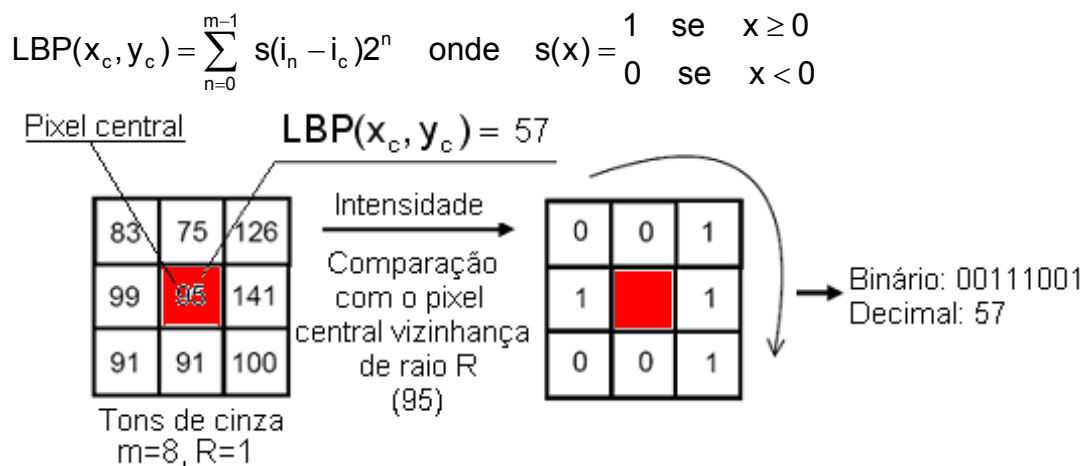


Figura 3.5 – Cálculo do LBP.

Conforme pode ser observado na figura 3.5, o cálculo do valor do LBP para o pixel central da região ($m=8$ e $R=1$) cujo valor é igual a 95 é calculado comparando seu valor com os valores dos pixels da vizinhança e tendo como resultado valores que podem ser zero quando o valor do pixel vizinho é maior e um quando o valor do pixel vizinho é menor ou igual. Após a comparação com todos os pixels da vizinhança temos uma sequência de zeros e uns que representa o código binário do valor do LBP que no exemplo da figura 3.5 temos “00111001” que em decimal é igual a 57 que é o valor do LBP para esse pixel.

Após cálculo de todos os valores LBP para todos os pixels da imagem, a representação da face, chamada *template* pode ser calculada tomando como base os histogramas gerados por esses valores conforme a figura 3.6.

A matriz contendo os códigos LBP de cada pixel é subdividida em blocos não sobrepostos numerados de 1 a B e o histograma desses códigos para cada bloco é computado. Conforme ilustrado na figura 3.6, os histogramas de cada bloco são concatenados um a um (do bloco 1 ao bloco B) formando um conjunto de histogramas conhecido como *template*, que representa a codificação da imagem da face.

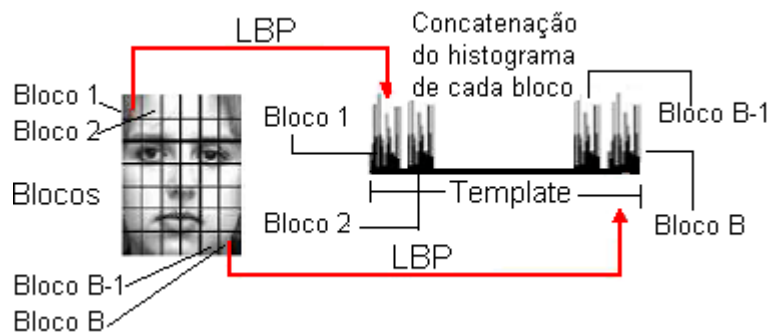


Figura 3.6 – Histograma do LBP.

O processo de comparação desses *templates* envolve o cálculo da dissimilaridade entre em blocos correspondentes, conforme figura 3.7.

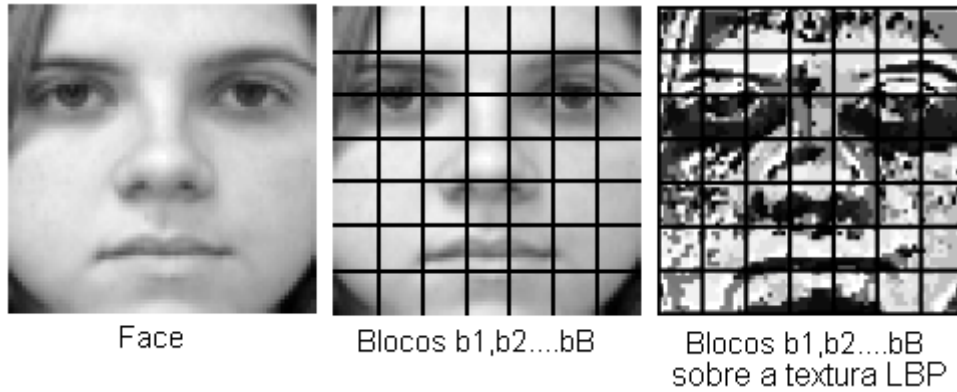


Figura 3.7 – Processo de comparação de faces baseado em textura.

A comparação das imagens de duas faces se dá através do cálculo da dissimilaridade entre os blocos correspondentes das duas imagens e pode ser calculado de várias maneiras diferentes, sendo a distância qui-quadrada (D^2) a mais citada na literatura e a utilizada nesta dissertação. Essa distância é dada pela equação 3-8.

$$D^2(^bH_x, ^bH_y) = \sum_c \frac{[^bH_x(c) - ^bH_y(c)]^2}{^bH_x(c) + ^bH_y(c)} \quad (3-8)$$

onde $^bH_x(c)$ e $^bH_y(c)$ são os histogramas (cth) do bloco para as imagens X, Y e bH_x , bH_y histogramas do bloco b das imagens H_x , H_y .

A distância qui-quadrada final (D_{Final}^2) é dada pelo somatório da distância (D^2) de cada bloco como mostrado na equação 3-9.

$$D_{Final}^2(H_x, H_y) = \sum_b w_b D^2(^bH_x, ^bH_y) \quad (3-9)$$

onde w_b ($b=1, 2, \dots, B$) é um peso associado a cada bloco, que reflete a maior ou menor importância de cada região da face no processo de reconhecimento. O método de ajuste destes pesos utilizado nesta dissertação está descrito detalhadamente em [20], [21].

Uma vez conhecida a distância qui-quadrada entre as imagens das faces que estão sendo comparadas e conhecendo-se o valor da maior distância qui-quadrada que se pode obter, a distância dada pela equação 3-9 é normalizada de

modo a ajustar-se a uma escala de 0 a 100, onde 100 corresponde à maior similaridade possível, o que só ocorre no caso em que as duas imagens comparadas são idênticas.

4

Método para Medida de Qualidade de Imagens Faciais

Este capítulo apresenta o método proposto nesta dissertação para o projeto de métricas de qualidade de imagens faciais. A primeira seção introduz os conceitos em que se fundamenta o método, o qual é descrito em seguida em termos gerais.

As seções subseqüentes descrevem como o método foi aplicado neste trabalho para gerar uma métrica de qualidade específica, cujo desempenho é por fim avaliada no próximo capítulo.

4.1

Método Proposto

De um modo geral, uma métrica da qualidade de uma imagem deve necessariamente levar em conta a aplicação alvo.

Mais especificamente, no reconhecimento facial automático, a qualidade da imagem está relacionada ao método de reconhecimento utilizado. Em outras palavras, a qualidade de uma imagem facial iluminada não uniformemente pode ser “adequada” ao mesmo tempo “inadequada” quando submetida a outro método de reconhecimento sensível ao padrão de iluminação. A métrica de qualidade deve, portanto, ser concebida de modo a capturar as características particulares do algoritmo de reconhecimento em uso.

A métrica de qualidade proposta na presente dissertação é denotada deste ponto do texto em diante pelo símbolo Y , baseia-se nessa idéia. Pressupõe-se para cada indivíduo uma imagem que serve de referência em termos de qualidade para aquele indivíduo e para o método de reconhecimento considerado. O índice de qualidade de outra imagem qualquer do mesmo indivíduo será dado pelo valor da similaridade entre esta imagem e a sua referência, conforme medida pelo algoritmo de reconhecimento utilizado que no caso dessa dissertação é dado pelo valor da distância qui-quadrada conforme equação 3-9.

A medição da qualidade da imagem facial de um dado indivíduo segundo tal métrica não deverá, em princípio, apresentar dificuldade, desde que se

disponha do correspondente sistema de reconhecimento e da imagem referência daquele mesmo indivíduo.

Contudo, na maioria dos casos práticos em que importa determinar a qualidade de uma imagem facial de uma pessoa, não se dispõe nem do sistema de reconhecimento nem da correspondente imagem de referência, mas tão somente da imagem cuja qualidade se deseja medir.

Essa dissertação propõe a criação de uma função f que estima, ou prevê a qualidade da imagem segundo a métrica proposta a partir de atributos medidos direta e exclusivamente na imagem cuja qualidade se deseja avaliar. Prescinde-se, portanto, da imagem de referência e do sistema de reconhecimento.

A figura 4.1 ilustra o processo de obtenção da qualidade de uma imagem de face mostrando as abordagens tradicional e proposta.

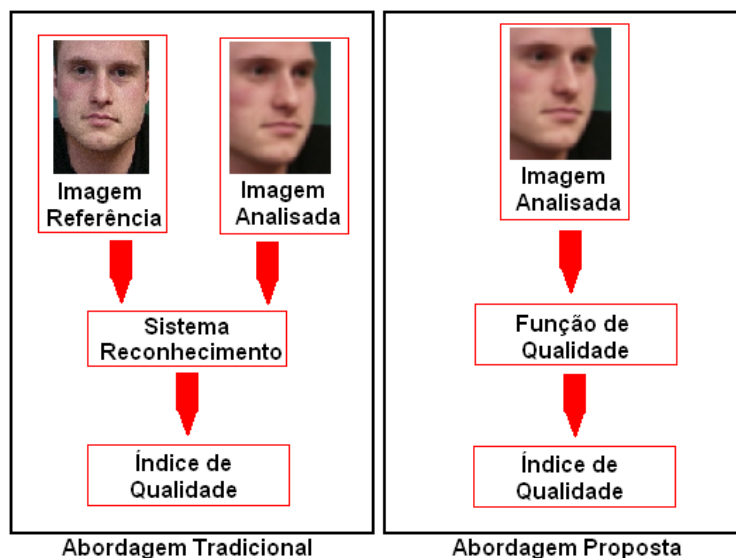


Figura 4.1 – Abordagem tradicional e abordagem proposta.

Formalmente, propõe-se o projeto de uma função do tipo $\hat{Y} = f(a_1, a_2, \dots, a_K)$ onde \hat{Y} é uma estimativa da qualidade (Y), f é a função a ser projetada, e (a_1, a_2, \dots, a_K) são atributos medidos diretamente na imagem sob análise.

Destarte, a realização desta proposta envolve a seleção de atributos (a_1, a_2, \dots, a_K) a partir dos quais se poderá estimar a qualidade da imagem e, ao mesmo tempo, a definição da própria função f . Claramente, ambas as escolhas dependerão do método de reconhecimento utilizado, como se propunha no início desta seção.

O projeto da função f envolve os seguintes passos.

1. O primeiro passo consiste em construir um banco de imagens faciais em que estejam representados os problemas de qualidade que afetam o reconhecimento facial. O banco deve conter para cada indivíduo uma imagem de alta qualidade que será usada como referência, além de outras imagens do mesmo indivíduo com qualidade inferior que serão usadas para teste.
2. No segundo passo, o algoritmo de reconhecimento escolhido mede a similaridade de cada imagem de teste do banco com a correspondente referência. Obtêm-se, assim, para cada imagem de teste as medições de qualidade.
3. No terceiro passo medem-se os atributos selecionados em cada uma das imagens de teste.
4. O quarto passo consiste em ajustar a função escolhida aos atributos e aos valores de similaridade medidos nos passos anteriores.
5. No quinto passo mede-se quão próximos estão os valores previstos pela função encontrada no passo 4 e os correspondentes valores medidos no passo 2.

Esta seqüência de passos deve ser executada para cada conjunto de atributos considerados e para diferentes tipos de funções. A iteração da qual resultar o melhor ajuste no passo 5 elegerá a função f para estimativa da qualidade das imagens faciais.

4.2

Instanciação do Método

O método descrito em termos gerais na seção anterior pode gerar diferentes métricas de qualidade dependendo do algoritmo de reconhecimento, dos atributos e da família de funções escolhidos.

Nesta seção e especialmente no próximo capítulo do presente trabalho projeta-se uma métrica de qualidade de imagens faciais aplicando o método proposto com o propósito de validá-lo.

O algoritmo de reconhecimento utilizado é baseado nos padrões binários locais (LBP) descrito no capítulo 3. Adota-se neste teste uma função linear para predição da qualidade da imagem, que tem a forma conforme a equação 4-1.

$$\hat{Y} = f(a_1, a_2, \dots, a_K) = \beta_0 + \beta_1 a_1 + \dots + \beta_K a_K \quad (4-1)$$

onde β_i ($i=0, \dots, k$) são coeficientes que devem ser ajustados de modo a minimizar o erro quadrático dado por pela equação 4-2.

$$E = \frac{1}{q} \sum_k \varepsilon_k^2 = \frac{1}{q} \sum_k (Y_k - \hat{Y}_k)^2 \quad (4-2)$$

onde $(Y_k - \hat{Y}_k)$ é a diferença entre o valor medido pelo algoritmo de reconhecimento e o valor predito pela função f para a k -ésima imagem de teste, e q é o número total de imagens de teste utilizadas nos experimentos.

Os argumentos da função f serão selecionados de um conjunto de atributos apresentados na próxima seção, utilizando o método de Akaike, descritos sucintamente na seção 3.2.

4.3

Atributos

Para avaliação da qualidade essa dissertação considera dezoito atributos (A01 a A18) extraídos da região da face como ilustra a figura 4.2

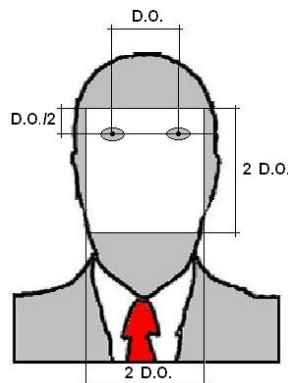


Figura 4.2 – Região da face.

São relacionados a seguir os atributos considerados e os respectivos símbolos que os denotam.

- Arfagem – A01;
- Rolagem – A02;
- Intensidade Média dos Pixels – A03;
- Contraste – A04;
- Gradientes
 - Soma das diferenças absolutas – A05;
 - Soma dos quadrados das diferenças – A06;
 - Soma das máximas diferenças absolutas – A07;
 - Tenengrad – A08;
 - Soma dos Laplacianos – A09;
 - Soma de diferenças na diagonal – A10;
 - Gradiente Cognitec – A11;
- Curtose da Região da Face – A12;
- Não Referência B – A13;
- Não Referência A – A14;
- Não Referência Z – A15;
- Avaliação dos efeitos de Iluminação – A16;
- Ruído 1 – A17;
- Ruído 2 – A18.

os quais estão detalhados neste capítulo nas seções que se seguem.

4.3.1

Arfagem

Para que o processo de reconhecimento facial tenha um bom desempenho, é necessário que a imagem da face esteja posicionada de forma frontal com o mínimo de inclinação em relação aos três eixos do sistema de coordenadas (x, y, z) [23] na figura 4.3.

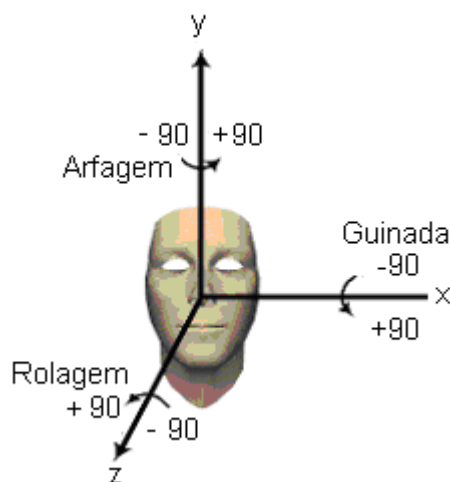


Figura 4.3 – Face no sistema espacial

A arfagem (*yaw* em inglês) [24] é o ângulo de rotação da face em torno do eixo y (Figuras 4.3 e 4.4). Ele é importante no processo de reconhecimento facial já que está associado à simetria da imagem facial – como se sabe, sistemas de reconhecimento facial tem desempenho melhor em imagens frontais.

O cálculo exato da arfagem a partir de uma imagem monoscópica da face é impossível. O presente trabalho propõe o uso de uma medida relacionada com a arfagem, e passível de determinação em uma imagem facial única.

Para o cálculo desta medida de arfagem dois outros ângulos são medidos na imagem, especificamente os ângulos determinados pela reta horizontal que passa pelos centros dos olhos com duas outras retas, cada uma passando pelo centro de um olho e pelo centro do nariz. Tem-se assim os ângulos a_1 e a_2 , como mostra a figura 4.4.

Uma vez calculados os valores dos ângulos a_1 e a_2 , o atributo relacionado à arfagem, denotado pelo símbolo A_{01} é calculado pela equação 4-3.



Figura 4.4 – Medida da arfagem.

$$A01 = \frac{|a1 - a2|}{\text{Min}(a1, a2)} \quad (4-3)$$

Em benefício da simplicidade, o texto que se segue utilizará o termo “arfagem” para designar o atributo A01, muito embora não se trate da arfagem propriamente dita.

4.3.2

Rolagem

Conhecido na literatura como “rolagem” (*roll* em inglês) [24], este ângulo corresponde à inclinação da face em relação ao eixo z (vide figura 4.3).

Durante o processo de análise da imagem, este ângulo é corrigido para que a linha imaginária que passa pelos centros dos olhos fique na horizontal.

Durante o processo de correção deste ângulo (rotação da imagem) haverá alguma interpolação que, salvo melhor juízo, poderá incluir algum artefato na imagem ou suprimir alguma informação relevante, razão pela qual este atributo é levado em conta.

Uma vez determinado o valor da rolagem, indicado na figura pelo símbolo *R* (figura 4.5) o atributo A02 é calculado como mostra a equação 4-4.



Ângulo Eixo z
Rolagem

Figura 4.5 – Rolagem.

$$A02 = 1 - \cos(R) \quad (4-4)$$

Novamente em benefício da simplicidade, o parâmetro A02 será referido daqui para diante no texto como “rolagem”, ainda que não se trate da rolagem propriamente dita, como definida no início desta seção.

4.3.3

Intensidade Média

Este atributo, denotado daqui para diante no texto pelo símbolo A03, é dado pela média das intensidades dos pixels da região da face, conforme a equação 4-5.

$$A03 = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N I_{i,j} \quad (4-5)$$

onde M e N são as dimensões da imagem facial e $I_{i,j}$ é intensidade do pixel na coluna i e linha j da imagem.

4.3.4

Contraste

O contraste [25] está relacionado à distribuição dos tons de cinza de uma imagem entre o preto e o branco. Histogramas de intensidades largos são típicos de cenas com bom contraste, enquanto histogramas estreitos são de imagens com pouco contraste. A figura 4.6 ilustra a relação entre contraste e correspondente histograma de intensidade de uma imagem.

Baixo contraste resulta de problemas de iluminação ou de características do próprio objeto imageado. Fotos obtidas sob neblina ou fumaça terão baixo contraste e fotos obtidas sob sol forte, por outro lado, tendem a apresentar contraste mais alto.

Adota-se neste trabalho o desvio padrão das intensidades da imagem como medida de contraste, conforme mostra a equação 4-6.

$$A04 = \sqrt{\frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (I_{i,j} - A03)^2} \quad (4-6)$$

onde M e N são as dimensões da imagem facial, $I_{i,j}$ é intensidade do pixel na coluna i e linha j e A03 é definida pela equação 4-5.

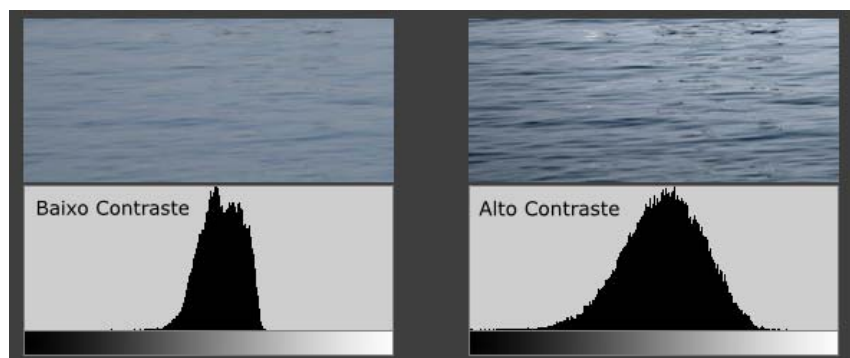


Figura 4.6 – Histograma de uma imagem com baixo e alto contraste.

4.3.5

Gradientes

A nitidez de uma imagem está diretamente ligada às bordas dos objetos na imagem.

A borda de um objeto corresponde às fronteiras entre objetos que aparecem sobrepostos na imagem. Um contorno é uma linha fechada formada pelas bordas de um objeto.

Bordas ocorrem nos picos da magnitude do gradiente, ou seja, nas regiões em que há variações abruptas nos tons de cinza, e, portanto, onde a derivada espacial tem valores extremos.

Admitindo que $g(x, y)$ é uma função contínua de x e y , o gradiente de $f(x, y)$ é definido pela equação 4-7.

$$\nabla f(x, y) = \left[\frac{\partial g}{\partial x} \quad \frac{\partial g}{\partial y} \right]^T \quad (4-7)$$

Para imagens digitais utilizam-se aproximações para o gradiente [26], [27]. Os atributos A05 a A11 são dados pela magnitude do gradiente computado por aproximações distintas, conforme mostrado a seguir através das equações 4-8 a 4-15.

Vale observar que as imagens do banco de dados usado nessa dissertação são normalizadas geometricamente, de modo a terem todas elas o mesmo número de linhas e colunas. Neste estudo consideraram-se as seguintes medidas para o gradiente:

Soma das diferenças absolutas [26]

$$A05 = \sum_i \sum_j |I_{i,j+1} - I_{i,j}| + |I_{i+1,j} - I_{i,j}| \quad (4-8)$$

Soma dos quadrados das diferenças [26]

$$A06 = \sum_i \sum_j \sqrt{(I_{i,j+1} - I_{i,j})^2 + (I_{i+1,j} - I_{i,j})^2} \quad (4-9)$$

Soma das máximas diferenças absolutas [26]

$$A07 = \sum_i \sum_j \max\{|I_{i,j+1} - I_{i,j}|, |I_{i+1,j} - I_{i,j}|\} \quad (4-10)$$

Tenengrad [26]

$$A08 = \sum_i \sum_j [{}^x I_{i,j}^2 + {}^y I_{i,j}^2] \quad \text{para } \sqrt{{}^x I_{i,j}^2 + {}^y I_{i,j}^2} \geq T08 \quad (4-11)$$

onde ${}^x I_{i,j}^2$ e ${}^y I_{i,j}^2$ são as derivadas nas direções x e y da imagem $I_{i,j}$, e T08 é uma constante. Na falta de um valor sugerido pelo autor, nessa dissertação foi usado o valor de 0 para que todas as derivadas fossem somadas.

Soma dos Laplacianos [26]

$$A09 = \sum_i \sum_j |LL_{i,j}| \quad \text{para } |LL_{i,j}| \geq T09 \quad (4-12)$$

onde T09 é uma constante (Na falta de um valor sugerido pelo autor, nessa dissertação foi usado o valor 0) e $LL_{i,j}$ é definido conforme equação 4-13.

$$LL_{i,j} = I_{i,j} * h \quad (4-13)$$

onde “*” é o operador de convolução e h é a molécula computacional de um filtro Laplaciano.

$$h = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

Soma de diferenças na diagonal [26]

$$A10 = \sum_i \sum_j [I_{i-1,j} + I_{i+1,j} - I_{i,j-1} - I_{i,j+1}] \quad (4-14)$$

Uma análise teórica da definição do atributo A10, conforme dada em [26], indica que o atributo não está relacionado ao gradiente. De fato, a avaliação experimental descrita no próximo capítulo demonstrará que este atributo está mais relacionado ao brilho do que à nitidez da imagem.

Gradiente Cognitec [27]

$$A11 = \frac{1}{MN} \sum_i \sum_j \text{abs}[l_{i,j} - \bar{l}_{i,j}] \quad (4-15)$$

onde $\bar{l}_{i,j}$ corresponde à imagem de entrada $l_{i,j}$ filtrada por um filtro mediana 3×3 .

4.3.6

Curtose

Curtose [11], [28] é uma medida que reflete o grau de achatamento da distribuição de intensidades na imagem, e está, portanto, relacionada com o contraste da imagem. É definida pela equação 4-16.

$$A12 = \frac{m_4}{m_2^2} \quad (4-16)$$

onde m_4 é o momento central de quarta ordem e m_2 é o momento central de segunda ordem, isto é,

$$m_4 = \sum_i \sum_j (l_{i,j} - A03)^4$$

$$m_2 = \sum_i \sum_j (l_{i,j} - A03)^2$$

4.3.7

Não Referência

Várias técnicas de avaliação de qualidade de imagem existem hoje na literatura que mostraram uma alta correlação com a avaliação de seres humanos. Uma destas é a técnica chamada Não-Referência [29].

Basicamente, esta técnica procura medir o impacto sobre a qualidade da imagem causado por agentes externos, tais como compressão, embaçamento e ruídos.

Esta técnica, proposta por Wang e coautores [30], calcula três índices de qualidade.

O primeiro índice conhecido como blocagem (*blockness* no inglês) é estimado pela diferença de intensidade média entre os limites de blocos de uma imagem. Para uma imagem de tamanho $M \times N$, esta componente na direção horizontal (B_h) é calculada como mostra a equação 4-17.

$$B_h = \frac{1}{M([N/8]-1)} \sum_{i=1}^M \sum_{j=1}^{[N/8]-1} |d_{h,i,8j}| \quad (4-17)$$

onde $d_{h,i,j}$ é a diferenciação do sinal na direção horizontal como na equação 4-18.

$$d_{h,i,j} = I_{i,j+1} - I_{i,j} \quad \text{para } j \in [1, N-1] \quad (4-18)$$

O segundo índice, conhecido como atividade, captura os efeitos resultantes de compressão e borrões na imagem. Para uma imagem de tamanho $M \times N$, esta componente na direção horizontal (A_h) é calculada como mostra a equação 4-19.

$$A_h = \frac{1}{7} \left\{ \frac{8}{M(N-1)} \sum_{i=1}^M \sum_{j=1}^{N-1} |d_{h,i,j}| - B_h \right\} \quad (4-19)$$

Finalmente o terceiro índice, denominado cruzamento em zero, mede a taxa de cruzamentos em zero de uma imagem. Para uma imagem de tamanho $M \times N$, esta componente na direção horizontal (Z_h) é calculada como mostra a equação 4-20.

$$Z_h = \frac{1}{M(N-2)} \sum_{i=1}^M \sum_{j=1}^{N-2} z_{h,i,j} \quad (4-20)$$

onde $z_h(m,n)$ é definido como em 4-21.

$$z_{h,i,j} = \begin{cases} 1 & \text{se cruzamento por zero em } d_{h,i,j} \\ 0 & \text{caso contrário} \end{cases} \quad (4-21)$$

e $n = [1, N-2]$

As componentes bloqueio, atividade e cruzamento são também computadas para a direção vertical, e denotadas respectivamente como B_v , A_v e Z_v . As estimativas finais destes índices são dadas pelas equações 4-22, 4-23 e 4-24.

Não Referência B

$$A13 = \frac{B_h + B_v}{2} \quad (4-22)$$

Não Referência A

$$A14 = \frac{A_h + A_v}{2} \quad (4-23)$$

Não Referência Z

$$A15 = \frac{Z_h + Z_v}{2} \quad (4-24)$$

4.3.8

Avaliação dos Efeitos de Iluminação – WMI

Para avaliar a qualidade das imagens faciais no que diz respeito às condições de iluminação foi proposto um método denominado “Weighted Sum of the Mean Intensity”, mais conhecido pela sigla WMI [11].

O valor do WMI é calculado da seguinte maneira. Inicialmente os valores de intensidade de uma determinada imagem facial são normalizados de modo a terem média zero e variância unitária. Então, a imagem é dividida em K regiões, de dimensões iguais e não sobrepostas (R_k). Calcula-se em seguida a média das intensidades de cada região. O valor WMI é calculado como uma soma ponderada das médias de cada região, conforme mostra a equação 4-25.

$$A16 = \sum_{k=1}^K w_k \bar{I}^k \quad (4-25)$$

onde $\bar{I}^k = \frac{1}{L^2} \sum_{i \in R_k} \sum_{j \in R_k} I_{i,j}$ e w_k é o peso para a k -ésima região e L^2 é o número de

pixels de cada região. Os pesos decrescem para regiões mais distantes do centro da imagem. Para essa dissertação a figura 4.7 ilustra os pesos usados onde pode

ser observado que as regiões centrais têm pesos maiores variando exponencialmente para as regiões das bordas.

1	2	2	1
2	4	4	2
2	4	4	2
1	2	2	1

Pesos das regiões (16)

Figura 4.7 – Pesos usados para cálculo de WMI.

Quanto maior WMI, maior será a taxa de reconhecimento devido ao fato de este valor estar associado à posição da iluminação na região da face.

4.3.9

Ruído 1

As imagens usadas no processo de reconhecimento facial muitas das vezes são provenientes de banco de dados antigos, coletadas ainda de fichas em papel como nos institutos de identificação existentes em todos os estados do Brasil.

Com o passar do tempo às imagens perdem a qualidade. Este processo pode ser considerado como resultante da agregação de ruído à imagem original à medida que o tempo passa. Imagens faciais (fotos 3x4) arquivadas podem conter pequenos escritos e carimbos (figura 4.8). As rasuras e os carimbos geralmente estão em tons escuros. As regiões de tons escuros, que não são rasuras, correspondem geralmente ao cabelo, à barba e aos olhos. A diferença destas regiões em relação às rasuras e aos carimbos é que estes se apresentam em tamanho menor. É, portanto, em princípio possível diferenciar as rasuras de outras regiões escuras através do tamanho.

Esse atributo de medida de qualidade, que se denominará neste trabalho Ruído 1[03], resulta de se eliminarem pequenas regiões de tom escuro na imagem

através de operações morfológicas. O processo de obtenção do atributo Ruído 1 é composto de três etapas.

A primeira etapa consiste de uma dilatação [31] da imagem. Trata-se aqui de uma operação morfológica em tons de cinza.



Figura 4.8 – Imagem de face corrompida com anotações.

Apesar de prejudicar a nitidez, a dilatação elimina pequenas regiões escuras deixando apenas as médias e as grandes regiões escuras.

A segunda etapa consiste de uma erosão [31], também em tons de cinza.

Com a erosão, o cabelo, os olhos e outras partes escuras recuperam suas dimensões e tonalidades perdidas no processo de dilatação (figura 4.9).

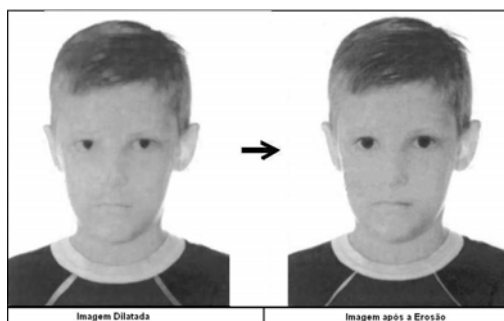


Figura 4.9 – Imagem recuperada com algoritmo Ruído 1.

A terceira e última etapa que resulta no atributo A17 (Ruído 1) é calculada pela subtração da imagem original pela imagem resultante, como definido na equação 4-26.

$$A17 = \sum (I - I') \quad (4-26)$$

onde I é a imagem original e I' é a imagem resultante do processamento morfológico.

4.3.10

Ruído 2

O atributo, aqui denominado Ruído 2, é obtido pela aplicação de um filtro [32] que reduz o ruído em uma imagem. O filtro basicamente compara a intensidade de cada pixel com as de seus oito vizinhos, e se aquela é maior ou menor do que estas, a intensidade do pixel é substituída pela maior ou menor intensidade de seus vizinhos, conforme o caso. Tal filtro remove ruídos que afetam pixels isolados na imagem. A Figura 4.10 mostra dois exemplos, em que se mostra à esquerda uma vizinhança antes e à direita após a aplicação do filtro.

1	9	18		1	9	18
21	44	13	➔	21	32	13
4	32	27		4	32	27
42	29	34		42	29	34
37	9	12	➔	37	12	12
22	17	76		22	17	76

Figura 4.10 – Evolução algoritmo Ruído 2.

A figura 4.11 apresenta outro exemplo da aplicação do mesmo filtro.



Figura 4.11 – Remoção de ruído – algoritmo Ruído 2.

Este índice de ruído tem seu cálculo como o atributo A17 – Ruído 1 subtraindo as duas imagens, e será denotado daqui para diante pelo símbolo A18 conforme equação 4-27.

$$A18 = \sum (I - I'') \quad (4-27)$$

onde I é a imagem original e I'' é a imagem resultante da filtragem

5

Análise Experimental

Este capítulo apresenta a análise experimental realizadas para validação da metodologia descrita no capítulo 4, para o projeto de uma métrica de qualidade de imagens faciais.

5.1

Base de Dados

Para a execução dos experimentos dessa dissertação a base de dados usada foi a base pública IMM [33] que é composta de 240 imagens estáticas de 40 diferentes rostos humanos (6 por pessoa), todos sem óculos sendo 7 pessoas do sexo feminino e 33 do sexo masculino. As imagens foram adquiridas em janeiro de 2001 com uma câmera de vídeo Sony DCR-TRV900E PAL produzindo imagens do tipo JPEG com resolução de 96 por 96 DPI onde cada imagem de cada pessoa contém as seguintes características (uma para cada pessoa).

- Imagem 01 – Face frontal de expressão neutra com luz difusa;
- Imagem 02 – Face frontal de expressão alegre com luz difusa;
- Imagem 03 – Face girada de 30 graus para direita (arfagem) de expressão neutra com luz difusa;
- Imagem 04 – Face girada de 30 graus para esquerda (arfagem) de expressão neutra com luz difusa;
- Imagem 05 – Face frontal de expressão neutra com uma fonte luz adicionada do lado esquerdo do rosto;
- Imagem 06 – Face frontal com uma expressão qualquer (livre) com luz difusa.

Para representar mais variações do que as presentes na base de dados original, para cada imagem existente foram geradas mais duas imagens borradas. Para geração dessas imagens utilizou-se um filtro de convolução cuja molécula computacional tem a forma de um disco de raio 2 e a de um disco de raio 5.

Com isso a base de dados resultante ficou composta de 720 rostos (18 por pessoa) conforme mostra a figura 5.1, que mostra um mesmo rosto da pessoa de número 02 com suas 18 imagens faciais.

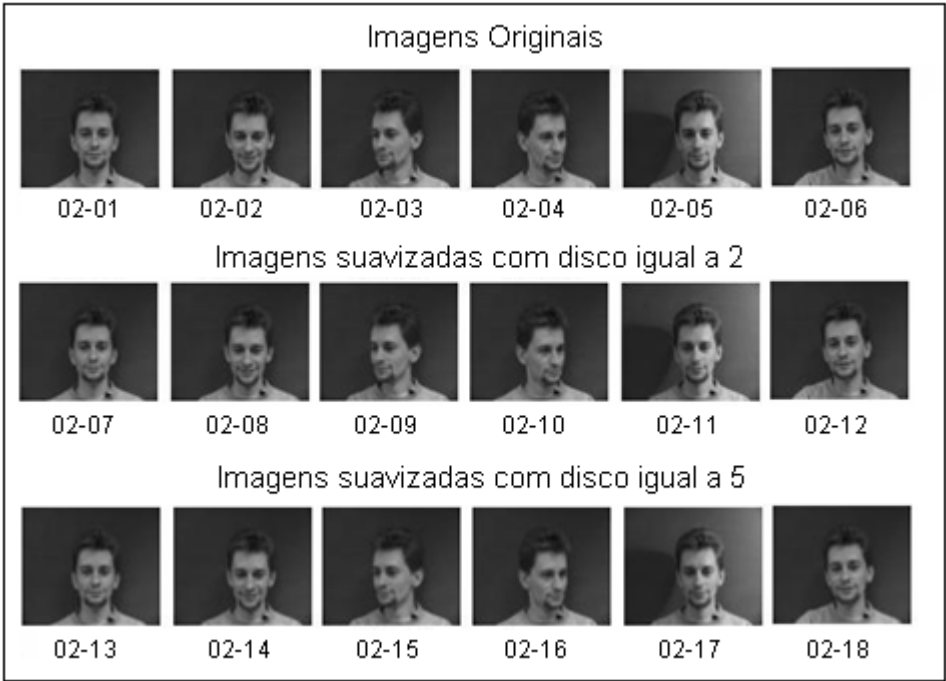


Figura 5.1 – Base de dados usada (Rosto N. 2).

Os dois dígitos mais à direita que aparecem abaixo de cada uma das imagens da figura 5.1 caracterizam o que se chamará daqui para diante no texto de “tipo” de imagem. Assim, as imagens do tipo 01 de cada indivíduo são as imagens de referência, as do tipo 02 são frontais com uma expressão facial não neutra, as imagens do tipo 03 têm arfagem acentuada para a esquerda, e assim por diante.

A título de ilustração, três imagens da pessoa número 02 aparecem ampliadas na figura 5.2.



Figura 5.2 – Fotos normal e embaçada.

5.2

Recursos Computacionais

Para a realização dos experimentos dessa dissertação os seguintes recursos computacionais de hardware e software foram utilizados.

Recursos de Hardware

- Computador pessoal tipo PC com as seguintes características:
 - Processador: I5-2400 3.10GHz;
 - Memória: 4 GB.
- Computador pessoal tipo PC com as seguintes características:
 - Processador: AMD 5200;
 - Memória: 2Gb.

Recursos de Software

- Sistema Operacional Windows Versão XP;
- Sistema Operacional Linux Redhat Versão 2.6.18-194.el5;
- Matlab 7 (Versão Windows e Linux);
- Oracle 10.2 (Versão Linux);
- Visual Basic Versão 6;
- Netbeans 6 (Versão Windows e Linux).

5.3

Experimentos

Foram realizados no total sete experimentos, com o objetivo de validar a metodologia proposta. Os objetivos particulares de cada experimentos, seus resultados e a conclusões deles extraídas são descritos a seguir.

5.3.1

Variação da Qualidade de cada tipo de Imagem

Este experimento teve o objetivo de conhecer como varia a qualidade dentro de cada tipo de imagem do banco de dados. Basicamente, levantaram-se os histogramas da qualidade para cada tipo de imagem.

Vale lembrar que na presente proposta a qualidade de uma determinada imagem é dada pela similaridade medida pelo algoritmo de reconhecimento entre esta imagem e a imagem de referência do mesmo indivíduo. As imagens do tipo 7 e 13 foram produzidas sinteticamente a partir da imagem de referência (tipo 1). Isso pode produzir valores de similaridade atipicamente elevados para as imagens 7 e 13, prejudicando a análise. Para evitar esse risco as imagens do tipo 7 e 13 foram excluídas da análise.

Os resultados são apresentados nas três figuras a seguir (Figuras 5.3, 5.4 e 5.5) que mostram histogramas contendo no eixo horizontal o valor da similaridade encontrada através do sistema de reconhecimento facial que pode variar de 0 a 100 (sendo este o valor máximo), no eixo vertical a frequência normalizada do índice de similaridade.

A figura 5.3 mostra uma imagem de referência (01) e as imagens do mesmo indivíduo dos tipos 02 a 06. As curvas que se vêem no gráfico mais adiante são os histogramas normalizados da similaridade para imagens dos tipos 02 a 06.

Observa-se que a similaridade medida nas imagens do tipo 03 e 04 apresentaram a pior distribuição entre os tipos considerados. Trata-se das imagens em que a arfagem é muito elevada (30 graus). Essas Imagens têm os valores de arfagem iguais em módulo e conseqüentemente os histogramas dessas duas imagens não apresentam diferenças substanciais.

Imagens do tipo 06 apresentam similaridade um pouco superior. Essas imagens estão afetadas por uma arfagem muito menor do que nos tipos 03 e 04, mas ainda assim não desprezível e se nota uma expressão facial intermediária entre “neutra” e “sorriso”.

O próximo tipo de imagens na escala de similaridade com a referência é o 02, em que se tem uma vista frontal (arfagem desprezível) e a expressão facial é claramente “sorriso”.

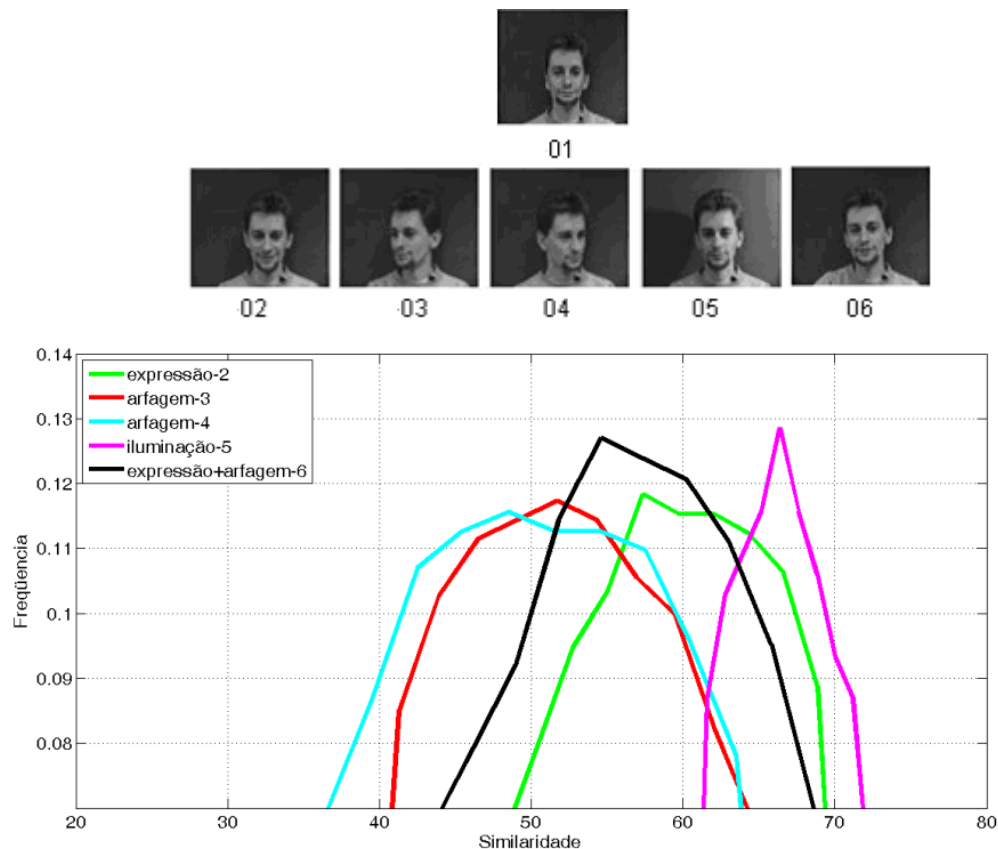


Figura 5.3 – Gráfico de similaridade 01: Imagem 01 Versus 02,03,04,05,06.

O tipo que apresentou melhor similaridade foi o 05, que é caracterizado por uma vista praticamente frontal e expressão neutra. Enquanto que em todos os demais tipos a iluminação é uniforme, nas imagens do tipo 05 o padrão de iluminação é claramente não uniforme.

Os resultados indicam que no banco de dados utilizado, entre todos os fatores presentes a iluminação teve o menor impacto sobre a similaridade da imagem.

Ao mesmo tempo, observou-se que entre os fatores que distinguem cada tipo de imagem no banco de dados, a arfagem é o mais crítico. Embora a maioria dos algoritmos de reconhecimento compense a rolagem, ângulos acentuados de arfagem não podem ser corrigidos por transformações geométricas simples.

É importante salientar que as conclusões acima, estão vinculadas ao banco de dados e ao algoritmo de reconhecimento utilizado e não podem, portanto, ser generalizadas. É fato conhecido que o desempenho de todos os algoritmos de reconhecimento piora para imagens com elevada arfagem e iluminação não uniforme. Já o algoritmo baseado em LBP é especialmente sensível à arfagem de

imagens, e muito robusto contra variações de iluminação, o que explica em parte as observações tecidas sobre aos resultados da figura 5.3.

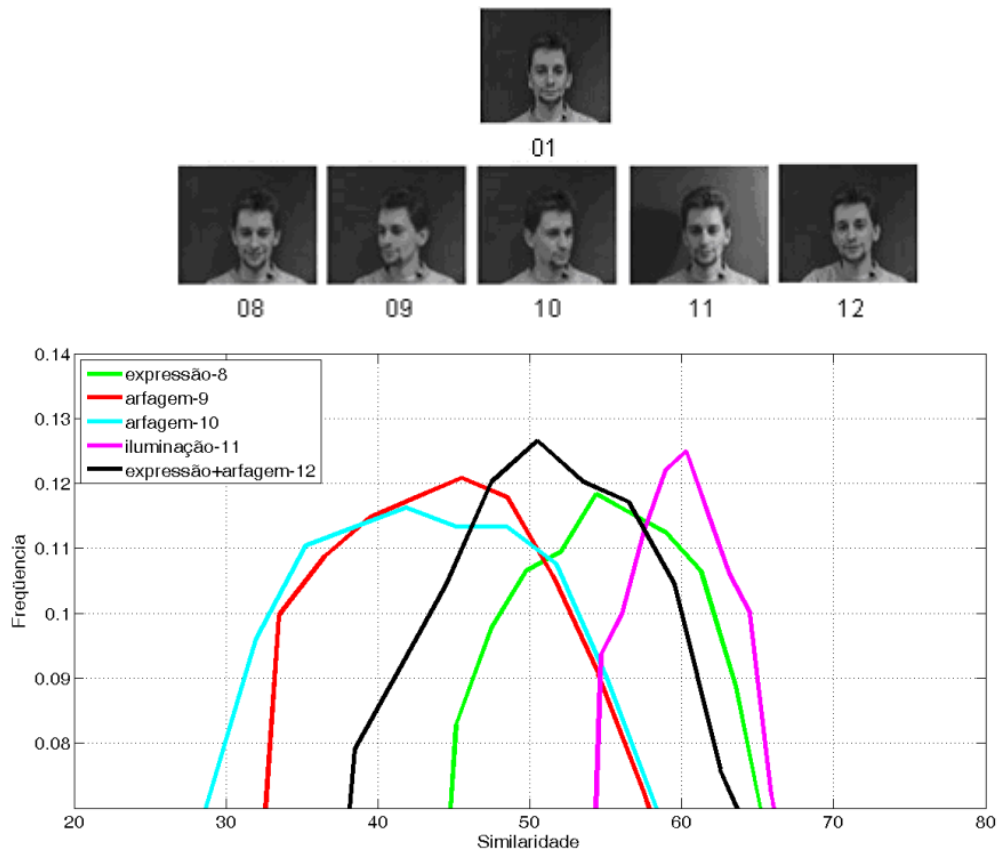


Figura 5.4 – Gráfico de similaridade 02: Imagem 01 Versus 08,09,10,11,12.

As figuras 5.4 e 5.5 apresentam os histogramas relativos às imagens do tipo 08 a 12 e 14 a 18, que são basicamente as imagens do tipo 02 a 06 submetidas a um filtro de suavização que lhes diminui a nitidez e o ruído. Nota-se que as formas dos histogramas não se alteram substancialmente, mas estes se deslocam para a esquerda relativamente à figura 5.3 em virtude da redução da qualidade das imagens decorrente da menor nitidez.

Conclui-se, assim, que os efeitos decorrentes da menor nitidez nas figuras 5.4 e 5.5 se combinaram com os efeitos decorrentes de outros fatores, como arfagem, expressão e expressão facial para degradar a similaridade de cada imagem com a correspondente referência.

Dessa constatação resta clara a limitação das abordagens que analisam atributos de qualidade individualmente e independentemente. Evidencia-se, assim a necessidade de uma métrica de qualidade que combine num único índice todos

os fatores que contribuem para a qualidade de imagens faciais, como se propõe nesta dissertação.

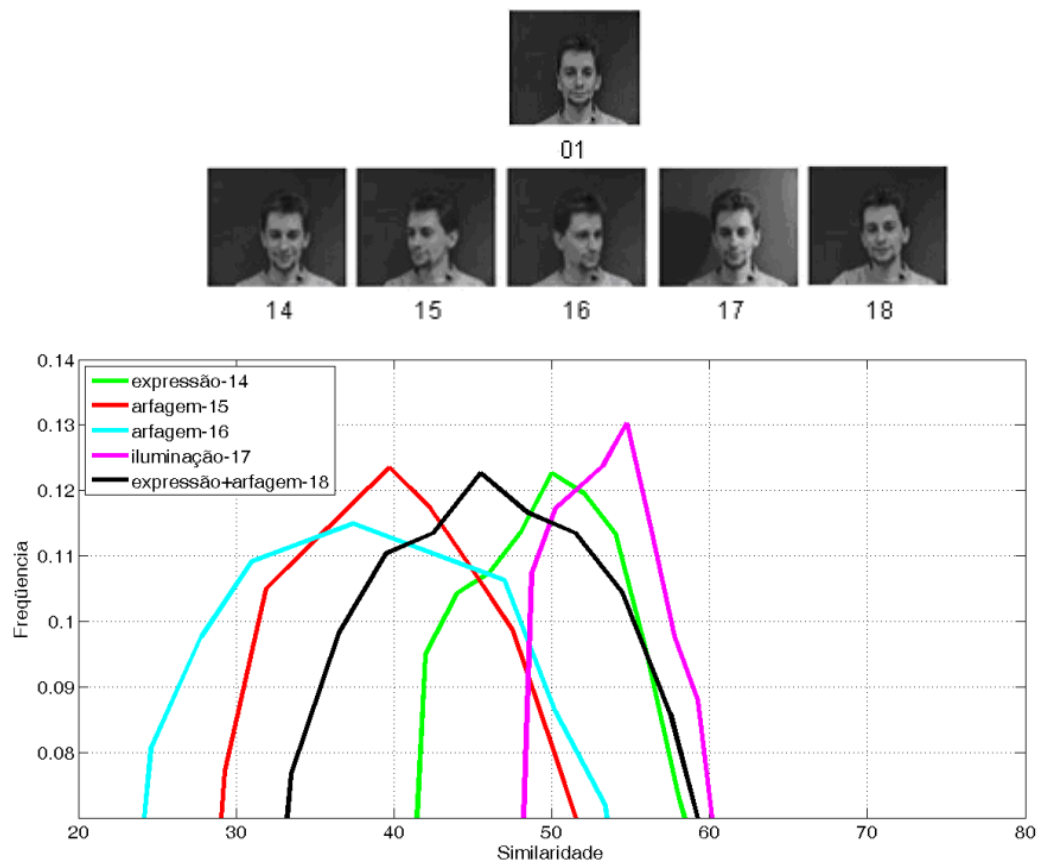


Figura 5.5 – Gráfico de similaridade 03: Imagem 01 Versus 14,15,16,17,18.

5.3.2

Grupos de Atributos Correlacionados

Dentre os dezoito atributos considerados neste estudo há múltiplos, atributos relacionados à pose como o foco e assim por diante.

Este experimento tem como objetivo examinar a correlação entre tais atributos. Pretende-se, desta forma, identificar grupos de atributos a partir dos respectivos coeficientes de correlação.

A tabela 5.1 mostra o valor do coeficiente de correlação medido entre cada par dos dezoito atributos. Evidentemente, os valores ali apresentados podem variar entre -1 e 1.

Para facilitar a interpretação dos resultados, as células da tabela têm cores de fundo relacionadas com o valor absoluto de seus conteúdos, como mostra a figura 5.6.



Figura 5.6 – Escala de cores para o módulo do coeficiente de correlação.

Tabela 5.1 – Tabela com o módulo dos coeficientes de correlação dos atributos usados.

A	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	1,00	0,12	-0,23	-0,01	-0,02	-0,02	-0,03	-0,10	-0,02	-0,24	-0,02	0,03	-0,31	0,31	-0,09	0,55	-0,07	0,00
2	0,12	1,00	0,09	0,00	0,02	0,02	0,02	0,03	0,09	0,08	0,02	0,03	0,07	0,01	0,03	-0,05	0,07	0,01
3	-0,23	0,09	1,00	0,45	0,12	0,12	0,12	0,30	1,00	0,57	0,34	0,35	0,54	0,06	0,00	-0,03	0,85	0,03
4	-0,01	0,00	0,45	1,00	0,25	0,26	0,27	0,54	0,21	0,13	0,12	0,15	0,47	-0,47	0,13	0,07	0,14	0,00
5	-0,02	0,02	0,12	0,25	1,00	1,00	1,00	0,55	1,00	1,00	0,99	0,99	0,93	0,93	0,93	0,93	0,97	0,98
6	-0,02	0,02	0,12	0,26	1,00	1,00	1,00	0,57	0,99	0,99	1,00	1,00	0,60	0,37	0,94	0,06	0,28	0,97
7	-0,03	0,02	0,12	0,27	1,00	1,00	1,00	0,59	0,99	0,13	1,00	0,35	0,54	-0,54	0,93	0,00	0,85	0,96
8	-0,10	0,03	0,30	0,54	0,55	0,57	0,59	1,00	0,52	0,32	0,60	0,69	0,46	-0,46	0,37	-0,06	0,38	0,99
9	-0,02	0,01	0,06	0,21	1,00	0,99	0,99	0,52	1,00	0,06	1,00	0,27	0,51	-0,51	0,94	-0,01	0,99	0,99
10	-0,24	0,09	1,00	0,45	0,12	0,12	0,13	0,32	0,06	1,00	0,08	0,57	0,35	-0,35	0,05	-0,05	0,01	0,03
11	-0,02	0,02	0,08	0,24	0,99	1,00	1,00	0,60	1,00	0,08	1,00	0,32	0,52	-0,52	0,93	0,01	0,86	0,97
12	0,03	0,03	0,57	0,69	0,32	0,34	0,35	0,69	0,27	0,57	0,32	1,00	0,15	-0,15	0,09	0,19	0,03	0,17
13	-0,31	0,07	0,32	0,47	0,53	0,54	0,54	0,46	0,51	0,35	0,52	0,15	1,00	-1,00	0,60	0,27	0,40	0,47
14	0,31	-0,07	-0,32	-0,47	-0,53	-0,54	-0,54	-0,46	-0,51	-0,35	-0,52	-0,15	1,00	1,00	-0,60	0,27	-0,40	-0,47
15	-0,09	0,01	0,04	0,13	0,93	0,93	0,92	0,37	0,94	0,05	0,93	0,09	0,60	-0,60	1,00	-0,08	0,89	0,96
16	0,55	0,03	-0,05	0,07	0,00	0,00	0,00	-0,06	-0,01	-0,05	-0,01	0,19	-0,27	0,27	1,00	1,00	-0,07	0,00
17	-0,07	0,01	-0,03	-0,07	0,87	0,86	0,85	0,28	0,88	-0,03	0,86	0,03	0,40	-0,40	0,89	-0,07	1,00	0,90
18	0,00	0,00	0,03	0,14	0,98	0,97	0,96	0,38	0,99	0,03	0,97	0,17	0,47	-0,47	0,96	0,00	0,90	1,00

Da tabela 5.1 é possível identificar grupos de atributos, como se esclarece a seguir.

Identificam-se três grupos de atributos com coeficiente de correlação muito elevado, isto é, superior em módulo a 0.85, que serão denotados como grupos **G01**, **G02** e **G03**, conforme tabela 5.2.

Tabela 5.2 – Grupos de atributos que apresentam coeficientes de correlação muito elevados.

Grupo	Atributo	Nome do Atributo
G01	A03	Intensidade média dos pixels
	A10	Soma das diferenças na diagonal
G02	A13	Não referência B
	A14	Não Referência A
G03	A05	Soma das diferenças absolutas
	A06	Soma dos quadrados das diferenças
	A07	Soma das máximas diferenças Absolutas
	A09	Soma dos laplacianos
	A11	Gradiente Cogntec
	A15	Não referencia Z
	A17	Ruído 1
	A18	Ruído 2

Identificam-se ainda dois grupos atributos com coeficientes de correlação elevados, ou seja, entre 0.6 e 0.9, que serão denotados como grupos **G04** e **G05**, conforme tabela 5.3.

Tabela 5.3 – Grupos de atributos que apresentam coeficientes de correlação elevados.

Grupo	Atributo	Nome do Atributo
G04	A04	Contraste
	A12	Curtose da região da face
G05	A08	Tenengrad
	A12	Curtose da região da face

Os grupos formados levam a algumas conclusões discutidas a seguir.






- No grupo G01 o atributo A10, que faz parte do grupo de gradientes, tem um coeficiente de correlação muito elevado com o atributo A03. Esta constatação confirma o comentário introduzido no capítulo 4 quanto à definição do atributo A10.
- O grupo G02 formado pelos atributos A13 e A14, ambos pertencentes ao grupo de atributos calculados no item “Não Referência”, apresentam um coeficiente de correlação muito elevado devido ao fato que ambos procuram medir o impacto sobre a qualidade da imagem causado por agentes externos.
- O grupo G03 formado pelos atributos A05, A06, A07, A09, A11, A15, A17 e A18, também com correlação muito alta, é formado basicamente por atributos do grupo gradiente que, como seria de se esperar. A este grupo pertencem também atributos ligados a agentes externos (A15, A17 e A18) que de alguma forma afetam os valores dos gradientes.
- Os grupos G04 e G05 apresentam correlação mais baixa do que os grupos precedentes, mas ainda assim elevada. É formado pelos atributos A04 e A12, (G04) e A08 e A12 (G05). Todos estes atributos tendem a crescer com o contraste das imagens, como se pode inferir de suas definições.

5.3.3

Avaliação da Contribuição da cada Atributo Isolado

Este experimento teve como objetivo determinar a contribuição de cada atributo isolado para a previsão da qualidade das imagens da base de dados utilizada.

A tabela 5.4 detalha os resultados obtidos neste experimento e separa os resultados por grupos baseando-se nos valores de correlação encontrados no experimento anterior. Linhas correspondentes ao mesmo grupo aparecem com sombreamento de mesma cor. Adotou-se a seguinte configuração de cores:

- G01 – cyan 
- G02 – amarelo 
- G03 - laranja 
- G04 e G05 – cinza 
- Demais atributos – branco 

Somente no caso dos grupos G04 e G05, optou-se por utilizar uma mesma cor de fundo, já que estes grupos apresentam correlação significativa. As linhas em fundo branco dizem respeito a atributos não associados a nenhum grupo.

Os resultados da tabela 5.4 pressupõem um modelo linear com um único atributo para estimar a qualidade de uma imagem tomando, como mostra a equação 5-1.

$$Y = \beta_0 + \beta_i X_i \quad (5-1)$$

onde Y é o índice de qualidade, β_0 é o termo independente, β_i é o fator de multiplicação do atributo X_i isoladamente.

Assim, cada linha da tabela fornece o valor de β_0 e de β_i , onde i é índice do atributo considerado – todos os demais elementos ao longo da mesma linha são nulos.

Tabela 5.4 – Resultados considerando cada atributo isoladamente.

Atrib.	AIC	RMSE	R ²	β ₀	β ₁	β ₂	β ₃	β ₄	β ₅	β ₆	β ₇	β ₈	β ₉	β ₁₀	β ₁₁	β ₁₂	β ₁₃	β ₁₄	β ₁₅	β ₁₆	β ₁₇	β ₁₈	
A01	2457	7,71	0,30	52	-5,10																		
A13	2459	7,72	0,30	52	0	0	0	0	0	0	0	0	0	0	0	0	0	5,09	0	0	0	0	0
A14	2459	7,72	0,30	52	0	0	0	0	0	0	0	0	0	0	0	0	0	-5,09	0	0	0	0	0
A16	2466	7,77	0,29	52																-5,02			
A15	2495	7,96	0,26	52	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4,71	0	0	0	0
A07	2509	8,05	0,24	52	0	0	0	0	0	0	4,54	0	0	0	0	0	0	0	0	0	0	0	0
A06	2510	8,06	0,24	52	0	0	0	0	0	4,53	0	0	0	0	0	0	0	0	0	0	0	0	0
A11	2512	8,07	0,24	52	0	0	0	0	0	0	0	0	0	0	4,51	0	0	0	0	0	0	0	0
A05	2513	8,08	0,24	52	0	0	0	0	4,50	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A09	2518	8,11	0,23	52	0	0	0	0	0	0	0	0	4,44	0	0	0	0	0	0	0	0	0	0
A18	2537	8,24	0,20	52	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4,19	0
A17	2567	8,45	0,16	52	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3,75	0	0
A08	2592	8,63	0,13	52	0	0	0	0	0	0	3,32	0	0	0	0	0	0	0	0	0	0	0	0
A04	2631	8,91	0,07	52	0	0	0	2,45	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A10	2663	9,16	0,02	52	0	0	0	0	0	0	0	0	0	1,28	0	0	0	0	0	0	0	0	0
A03	2665	9,17	0,02	52	0	0	1,19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A12	2669	9,20	0,01	52	0	0	0	0	0	0	0	0	0	0	0	0,92	0	0	0	0	0	0	0
A02	2673	9,23	0,00	52	0	-0,58	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Algumas conclusões podem ser extraídas dos resultados da tabela 5.4.

- O atributo que isoladamente melhor explica a qualidade é o atributo A01 – Arfagem. Este atributo, quando elevado, descaracteriza substancialmente a imagem de face. O prejuízo para o reconhecimento é particularmente acentuado para o algoritmo de reconhecimento baseado nos LBPs que requer que as imagens sejam frontais.
- Em seguida temos como fator de importância os atributos A13 e A14, que, como já se mencionou antes, medem o impacto de agentes externos sobre a qualidade das imagens. A tabela 5.1 revela também que esses atributos apresentam substancial correlação com os atributos do grupo G03, relacionados ao gradiente, e, portanto à nitidez da imagem.
- Os atributos do grupo G03 que estão associados à nitidez aparecem em seguida. Este resultado indica que nos experimentos realizados a nitidez, após a arfagem foi o parâmetro mais determinante da qualidade.
- Outros grupos atributos de menor relevância podem ser observados quando usados isoladamente:
 - Atributos A08 e A04 ligados ao contraste;
 - Atributos A10 e A03 ligados ao brilho;
 - Atributo A02 ligado à rolagem.

A relativa pouca importância do contraste e ao brilho decorre pelo menos em parte da elevada robustez do método de reconhecimento utilizado contra variações de iluminação. Os resultados não permitem por si só que se generalize esta conclusão para qualquer método de reconhecimento. Ao contrário, era de se esperar a pouca relevância da rolagem, uma vez que antes da etapa de reconhecimento, usualmente se aplicam procedimentos simples de rotação da imagem colocam a linha que une os dois olhos do indivíduo na horizontal. Este procedimento elimina, desta forma, o impacto da rolagem sobre a qualidade da imagem.

- Em resumo, o modelo que melhor explica a qualidade de uma imagem de face usando um único atributo é dado pela equação 5-2.

$$Y = 52 - 5.10 * A01 \quad (5-2)$$

Vale observar que não se pode afirmar que os atributos com menor RMSE são os mais importantes, pois isso é dependente da base de dados usada. Se, por exemplo, as variações de contraste na base de dados forem mais acentuadas em uma outra base de dados, é possível que os atributos relacionados a contraste sejam mais significativos quando usados isoladamente.

5.3.4

Determinação do Melhor Conjunto de Atributos

Este experimento teve como objetivo determinar o melhor conjunto de atributos e avaliar a qualidade do correspondente modelo de regressão.

A tabela 5.5 detalha os resultados obtidos neste experimento. As colunas da tabela estão coloridas seguindo a mesma convenção de cores indicada na seção precedente. O critério básico para a seleção de atributos foi o critério de informação de Akaike (AIC). A primeira coluna indica o número (N) de elementos do conjunto de atributo. Assim, por exemplo, a linha correspondente a N=10, indica os valores de β dos atributos pertencentes ao conjunto ótimo de 10 atributos, segundo o critério de Akaike. As células ao longo desta linha com conteúdo nulo, indicam os atributos que não fazem parte do conjunto ótimo.

Convém lembrar que os experimentos foram realizados a partir de todas as imagens com exceção das derivadas das imagens de tipo 01, 07 e 13.

Tabela 5.5 – Melhor conjunto de N atributos.

N	AIC	RMSE	R ²	β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	β_{10}	β_{11}	β_{12}	β_{13}	β_{14}	β_{15}	β_{16}	β_{17}	β_{18}
1	2457	7,7	0,30	52	-5,1																	
2	2220	6,3	0,53	52							5									-5,0		
3	2129	5,8	0,60	52	-3,0						4									-3,3		
4	2114	5,8	0,61	52	-2,8			1,1	4											-3,4		
5	2094	5,7	0,63	52	-2,8			2,0			4					-1,6				-3,1		
6	2092	5,6	0,63	52	-2,7			2,0			10		-5,8			-2,2				-3,1		
7	2083	5,6	0,63	52	-2,8			2,5	-4,6		52		-3,5			-1,7				-3,2		
8	2081	5,6	0,64	52	-2,8			2,4	-83	95			-3,1		-6,5	-1,9				-3,2		
9	2079	5,5	0,64	52	-2,8			1,4	-30		34			-5,3		-2,1	2,1			-2,9		
10	2073	5,5	0,64	52	-2,8		47	1,6	-102	108		-2,9		-4,8		-1,8	2,1			-3,0		
11	2074	5,5	0,64	52	-2,8		51	1,4	-68		70	-2,6		-5,2		-1,5	2,6			-3,0		2,9
12	2076	5,5	0,65	52	-2,8		48	1,4	-115	121		-2,2		-4,8	-5,1	-1,6	2,3			-3,0		3,8
13	2077	5,5	0,65	52	-2,8	-0,17	48	1,4	-108	114		-2,4	-5,5	-4,8		-1,6	2,3			-3,0		4,5
14	2079	5,5	0,65	52	-2,8	-0,17	48	1,3	-108	114		-2,4	-5,5	-4,8		-1,6	2,4			-3,0	-0,36	4,8
15	2081	5,5	0,65	52	-2,8	-0,17	48	1,3	-109	115		-2,4	-5,5	-4,9		-1,6	132	129		-3,0	-0,34	4,8
16	2083	5,5	0,65	52	-2,8	-0,16	48	1,3	-100	91	15	-2,4	-5,1	-4,9		-1,5	134	132		-3,0	-0,34	4,9
17	2085	5,5	0,65	52	-2,8	-0,16	49	1,2	-100	91	15	-2,4	-5,0	-4,9		-1,5	136	134	-0,18	-3,0	-0,32	5,0
18	2087	5,5	0,65	52	-2,8	-0,16	48	1,2	-101	92	15	-2,3	-3,9	-4,9	-1,1	-1,5	137	135	-0,17	-3,0	-0,32	4,9

Face aos resultados registrados na tabela 5.5 pode-se chegar às seguintes conclusões:

- Em primeiro lugar, cumpre ressaltar que a superioridade do modelo multivariado sobre o monovariado. Considerando os melhores casos, o RMSE passa de 7,71 no modelo baseado num único atributo para até 5.5 no modelo com múltiplos atributos. Este resultado atesta a validade da proposta desta dissertação de integrar num único índice fatores diversos de qualidade.
- Pelo critério do mínimo valor de AIC o conjunto ótimo possui 10 atributos, conforme indicado na linha correspondente a N=10 com bordas espessas na tabela 5.5.
- Nota-se que o conjunto de valores de R² e RMSE não se degrada significativamente quando se reduz o valor de N até N=3. É interessante notar que deste conjunto constam atributos relacionados à arfagem, iluminação e nitidez, o que, em linhas gerais, é consistente com a importância relativa de cada atributo isolado conforme determinado no experimento 5.3.3.
- É digno de nota, que os atributos relacionados ao contraste (A08 e A04), não aparecem, senão a partir de conjuntos com 9 atributos. Isso se deve provavelmente ao fato de que o algoritmo de reconhecimento baseado em LBP é pouco sensível a transformações afins dos tons de cinza, isto é, $I' = \lambda I + \mu$, onde I e I' representam os tons de cinza da imagem transformada e da imagem original respectivamente, e λ e μ são constantes reais.
- A coluna R² indica que para N=3 o valor de R² está em 60%. Este valor depende muito da base de dados, e não é possível tirar muitas conclusões a partir dele.

Segundo o critério do mínimo AIC entre as alternativas analisadas a função linear que melhor prevê a qualidade de imagem em nossos experimentos é dada pela equação 5-3.

$$Y = 52 - 2.8 * A01 + 47 * A03 + 1.6 * A04 - 102 * A05 + 108 * A06 - 2.9 * A08 - 48 * A10 - 1.8 * A12 + 2.1 * A13 - 3 * A16 \quad (5-3)$$

Observou-se ainda que se pode adotar uma função mais simples contendo apenas três argumentos dada pela equação 5-4.

$$Y = 52 - 3 * A01 + 4 * A07 - 3.3 * A16 \quad (5-4)$$

Encerrando esta seção é relevante a diferença nos valores de RMSE.

5.3.5

Estimativa de Qualidade por Rede Neural

Este experimento teve como objetivo determinar o ganho potencial de se adotar uma formulação não linear como alternativa à função linear estudada nas seções anteriores. Para isso utilizou-se uma rede neural (RN) que teve como entradas os atributos arfagem (A01), máximas diferenças absolutas (A07) e iluminação (A16) (N=3 na tabela 5.5).

A topologia usada para a construção da rede neural foi uma rede do tipo MLP com três camadas tendo como funções ativação.

- Tansig na camada oculta;
- Linear na camada de saída.

Variou-se o número de neurônios na camada oculta entre 3 e 24, de modo a garantir que a razão entre o número de observações para treinamento e o número de parâmetros para estimar os pesos da RNA não excedesse 5.

As observações disponíveis foram divididas aleatoriamente em 60% e 40% respectivamente para treinamento e validação.

Para cada número de neurônios na camada oculta o experimento foi repetido 100 vezes, cada vez com uma seleção aleatória diferente de observações para treinamento e validação. Adotou-se como resultado o correspondente ao melhor valor do RMSE calculado sobre a totalidade das observações.

O valor ótimo ao longo de todos os experimentos apresentou RMSE igual a 5.5 (cinco ponto cinco) para 8 neurônios na camada oculta. Este valor está 0.3 abaixo e, portanto, ligeiramente inferior ao valor obtido com a abordagem linear quando se utilizaram as mesmas entradas (A01, A07 e A16). Resta claro que a maior complexidade da RNA relativamente à abordagem linear não trouxe benefícios correspondentes em termos de acurácia do modelo, pelo que os experimentos favorecem a opção por uma modelagem linear.

5.3.6

Comparar a Variação entre os Valores Medidos e Estimados

Este experimento teve como objetivo determinar a faixa em que varia a qualidade de cada tipo de imagem.

Neste experimento comparam-se os valores estimados pelo método linear (regressão), os valores medidos (calculados pelo sistema de reconhecimento facial) e os valores estimados pelo processo não linear (rede neural).

A figura 5.7 apresenta os resultados. Os círculos vermelhos no meio de cada barra indicam o valor médio.

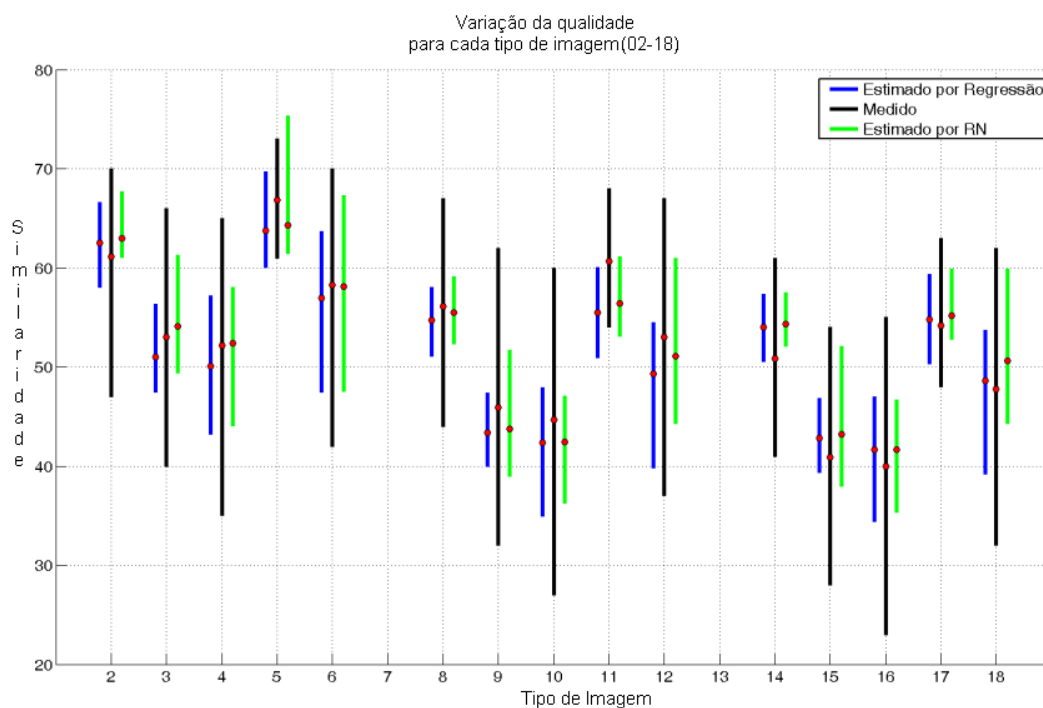


Figura 5.7 – Variação da qualidade/similaridade para cada tipo de imagem.

Em face dos resultados mostrados na figura 5.7, pode-se chegar às seguintes conclusões.

- Os valores estimados por regressão variam menos do que os valores medidos pelo sistema de reconhecimento facial, o que é, por um lado, coerente com a teoria, e por outro lado consistente com o pressuposto de que as imagens de um mesmo tipo têm qualidade semelhante. Parte da variação nas medidas de similaridade deve-se a variações de uma multiplicidade de outras características (p.ex., olhos mais ou menos fechados, penteado, bigode, etc.) tanto na imagem de teste como na sua referência. O método proposto filtra estas variações e captura predominantemente o efeito dos atributos considerados. Neste sentido, a estimativa provida pelo método proposto é potencialmente mais representativa da qualidade do que a medição da similaridade.
- O mesmo comportamento não ocorre sempre para a RNA como, por exemplo, a imagem 05. Do ponto de vista teórico, não existe garantia que a RNA apresentará esta propriedade.
- Os valores estimados por regressão são mais próximos e mais estáveis que os estimados pela RNA o que favorece a escolha do método de regressão em detrimento ao modelo de RNA.

5.3.7

Adequação do Modelo

Os experimentos relatados até aqui foram realizados na amostra (*in sample*). Nesta seção avalia-se o modelo fora da amostra (*out of sample*). Para tanto, aplicou-se neste último experimento a técnica conhecida com validação cruzada. Mais especificamente, adotou-se o seguinte procedimento:

1. Seleccionam-se as imagens referentes ao primeiro indivíduo, dentre os representados na base de dados.
2. Calcula-se o modelo de regressão a partir das imagens dos demais indivíduos, envolvendo os três atributos selecionados na seção 5.3.4.
3. Aplica-se o modelo calculado no passo anterior às imagens do indivíduo selecionado no passo 1, e registram-se os valores do erro.

4. Repetem-se os passos anteriores retirando-se no passo 1 o próximo indivíduo até que todos os indivíduos tenham sido eleitos uma vez.

O valor de RMSE encontrado neste experimento foi 5.9, portanto, ligeiramente superior a 5.8 verificado nos experimentos *in sample* descritos na seção 5.3.4.

O gráfico da figura 5.8 mostra os valores preditos pelo modelo contra os valores medidos. Nota-se que a linha de tendência tracejada passa pela origem com inclinação de 45 graus. A nuvem de pontos sugere que valores estimados e medidos guardam de fato uma relação próxima da linear.

O gráfico da figura 5.9 mostra a dependência do erro com os valores preditos. A forma da nuvem semelhante a uma faixa horizontal indica que a variância do erro é aproximadamente constante e que não há pouca dependência entre o erro e o valor de similaridade predito. A linha de tendência tracejada mostra que o valor médio do erro pouco depende do valor predito.

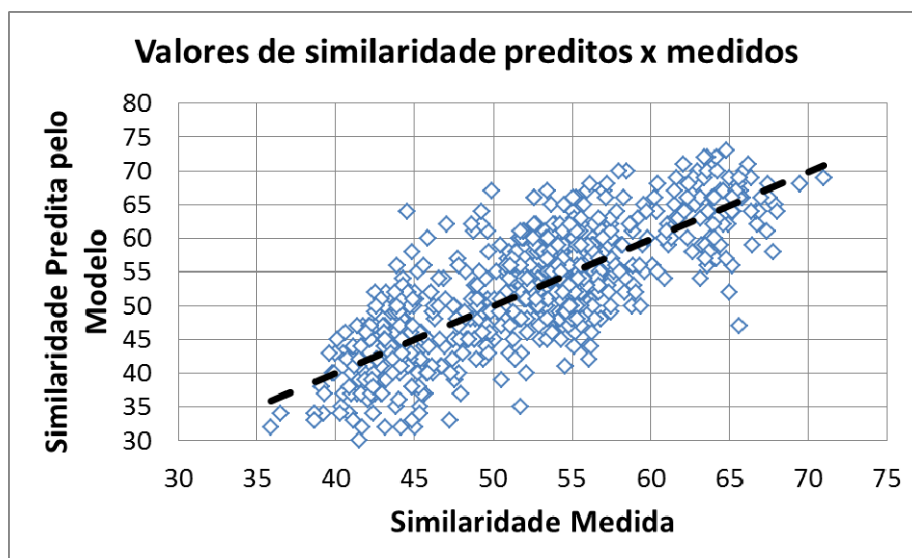


Figura 5.8 – Dependência dos valores de similaridade medidos e preditos pelo modelo.

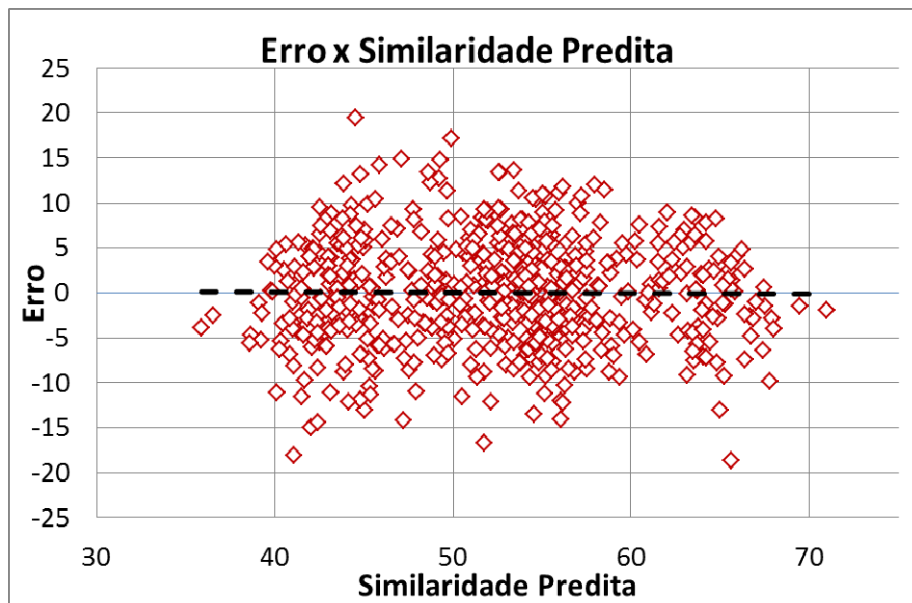


Figura 5.9 – Dependência do erro com os valores de similaridade medidos.

Tais resultados constituem evidência de que o modelo linear proposto é adequado para estimar a qualidade das imagens faciais.

5.3.8

Comentários Sobre Experimentos

Sete experimentos foram realizados com o objetivo de avaliar como diversos atributos diretamente mensuráveis na imagem estão relacionados à qualidade de uma imagem facial e ao mesmo tempo projetar um modelo matemático capaz de estimar a qualidade de uma imagem facial a partir desses atributos.

Em uma primeira análise mediu-se como varia a qualidade das imagens na base de dados utilizada. Os resultados deixaram claro que a qualidade varia substancialmente nas amostras, e que a base de dados se prestava ao estudo em tela.

O passo seguinte foi investigar como os 18 atributos considerados neste trabalho estão correlacionados. A análise de origem a grupos de atributos com elevada correlação. Ficou claro que os grupos encontrados dizem respeito a um fator distinto entre os que concorrem para determinar a qualidade de uma imagem facial.

Uma vez conhecido como os atributos estão correlacionados, investigou-se o projeto de um modelo capaz de estimar a qualidade das imagens como uma função linear de um ou mais atributos extraídos das próprias imagens.

O resultado obtido pelo modelo linear a partir de mais de um atributo mostrou-se mais exato do que o modelo linear usando apenas um atributo. Esta constatação confirma a esperada vantagem do método proposto sobre propostas anteriores que prevêm a aplicação de critérios separados para cada atributos de qualidade.

Ficou também evidenciado que os atributos do modelo selecionado pertencem a grupos distintos, conforme determinados na primeira série de experimentos.

Também foi avaliada a acurácia de um modelo matemático não linear através de redes neurais artificiais, tendo-se verificado que não houve ganho importante relativamente ao modelo linear.

Mediu-se ainda a variação de qualidade dentro de tipos diferentes de imagens. Verificou-se que os valores de qualidade estimados são mais estáveis do que os valores de similaridade medidos, o que é consistente com o propósito a que se prestam métricas de qualidade voltadas para reconhecimento facial.

Finalmente realizaram-se testes fora da amostra para verificar a adequação do modelo. Os resultados obtidos demonstram que o modelo de regressão pode de fato prever adequadamente a qualidade das imagens faciais conforme definida no presente trabalho. Observou-se ainda que nos valores do resíduo independem dos valores estimados pelo modelo. Tais observações constituem evidência de que o método proposto neste trabalho atende ao objetivo de medir a qualidade de imagens faciais visando ao reconhecimento automático.

Comentários Finais e Direções Futuras

Este capítulo faz uma síntese dessa dissertação, tece alguns comentários sobre as conclusões extraídas e aponta direções futuras.

Antes do início deste trabalho, uma pesquisa foi realizada com o objetivo de investigar trabalhos relacionados à restauração de imagens usadas em sistemas de reconhecimento facial. A maioria dos trabalhos encontrados se dispõe a melhorar a qualidade visual, e, portanto subjetiva, das imagens. Além disso, trabalhos que avaliam a qualidade de imagens faciais propõem critérios separados para cada tipo de atributo que contribui para a qualidade da imagem.

Nestes trabalhos, uma imagem tem sua qualidade aprovada quando aprovada individualmente por todos os critérios individualmente. O efeito combinado destes atributos é assim desprezado. A presente dissertação procura superar estas duas limitações e propõe uma medida objetiva de qualidade vinculada ao escore produzido por um sistema de reconhecimento facial. Ao mesmo tempo, a abordagem proposta combina num índice único o efeito combinado de vários fatores que afetam a qualidade da imagem.

Ficou evidenciado que a qualidade de uma imagem pode ser avaliada de forma mais exata quando vários atributos são usados no processo.

Ficou claro também que um grande problema de qualidade em imagens usadas em sistemas de reconhecimento facial está intimamente relacionado à pose da cabeça, isto é, faces giradas para um lado ou para o outro em torno de um eixo vertical (arfagem) prejudicam em muito o desempenho desses sistemas.

Verificou-se ainda que o contraste e a nitidez foram os fatores que mais impactaram na qualidade neste estudo.

De qualquer forma, os experimentos realizados não permitem generalizar tais conclusões. Alterações no banco de dados e, especialmente, do algoritmo de reconhecimento, poderiam afetar os resultados em alguma medida.

Teria sido muito interessante se este trabalho pudesse ter sido desenvolvido usando uma base de dados real e de volume expressivo o que talvez levasse os exemplos e conclusões mais próximas da realidade.

Também se constatou que a abordagem linear e não linear levaram a resultados semelhantes. Esta constatação sugere que a abordagem linear baseada em regressão linear é adequada, devendo ser favorecida em relação a métodos mais sofisticados como redes neurais, por serem mais simples computacionalmente.

O principal produto deste trabalho é uma metodologia para o projeto de métricas de qualidade que podem ser usadas para admitir ou rejeitar imagens a serem incorporadas a bancos de dados. Tais métricas podem orientar o processo de aquisição de imagens faciais para fins de reconhecimento, eliminando amostras de baixa qualidade. Desta forma, poupam-se recursos computacionais que de outra forma processariam dados de baixa qualidade e, se obtém, ao mesmo tempo uma melhoria do desempenho global dos sistemas de reconhecimento facial.

Como direções futuras, vários estudos podem ser iniciados e também completados com o embrião aqui gerado. Dentre esses, alguns podem ser considerados como relevante para o aprimoramento do trabalho já desenvolvido.

- Testar os modelos em outras bases de dados não públicas, contendo uma maior diversidade de amostras.
- Investigar limiares de corte que operassem como pré-filtragem para separação de imagens que serão ou não serão submetidas a um sistema de reconhecimento facial.
- Acrescentar à análise atributos que capturem outros fatores de qualidade. Como exemplo, mencione-se a expressão facial. Muito embora, variações de expressão facial estejam presentes no banco de dados utilizado, nenhum atributo relacionado à expressão facial foi considerado no presente estudo.
- Avaliar o ganho em termos de taxas de reconhecimento resultante de se aplicarem filtros de qualidade para aceitar/rejeitar imagens em um sistema de reconhecimento facial.
- Avaliar o ganho computacional decorrente de se evitar o processamento desnecessário de imagens de baixa qualidade.
- Investigar rotinas de pré-processamento capazes de melhorar os atributos que determinam a qualidade da imagem. Nesta linha pode-

se pensar em realce de nitidez, contraste, redução de ruído, eliminação de anotações, ou transformações geométrica que coloquem a imagem numa posição mais próxima da frontal antes do reconhecimento.

Referências Bibliográficas

[01] Xiaoyang Tan; Bill Triggs. Enhanced Local Texture Feature Sets for Face Recognition under Difficult Lighting Conditions. 3rd International Workshop Analysis and Modelling of Faces and Gestures (AMFG '07) 4778 (2007), pp. 168-182.

[02] André Sobiecki; Luiz A. P. Neves; Gilson A. Giraldi; Gilka J. F. Gattás; Carlos E. Thomaz. Segmentação e Restauração Digital para Eliminação de Artefatos em Imagens Frontais de Face. VII Workshop de Visão Computacional – WVC 2011, pp. 51-56.

[03] André Sobiecki; Carlos Eduardo Thomaz; Luiz Antonio Pereira Neves. Uma Proposta de Melhoria Automática em Imagens Faciais. 2009, pp.1-4;

[04] Mahmood Farshbaj Doustar; Hamid Hassanpour. A Locally-Adaptive Approach for Image Gamma Corretion. 10th International Conference on Information Science, Signal Processing and their Applications (ISSPA 2010), pp. 76-76.

[05] Morihiko SAKANO; Noriaki SUETAKE; and Eiji UCHINO. A Robust Point Spread Function Estimation for Out-of-Focus Blurred and Noisy Images Based on a Distribution of Gradient Vectors on the Polar Plane. Optical Review Vol.14, No. 5 (2007), pp. 297–303.

[06] A. N. Tikhonov and V. Arsenin, Solutions of Ill-Posed Problems. New York: Wiley, 1977.

[07] W. H. Richardson, “Bayesian-based iterative method of image restoration,” J. Opt. Soc. Amer., vol. 62, pp. 55–59, 1972.

[08] L. B. Lucy, “An iterative technique for the rectification of observed distributions,” Astronom. J., vol. 79, 1974.

[09] J. Biemond, R. L. Lagendijk, and R. M. Mersereau, "Iterative methods for image deblurring," *Proc. IEEE*, vol. 78, pp. 856–883, May 1990.

[10] Lei Liang; Yuanchang Xu. Adaptive Landweber Method to Deblur Images. *IEEE Signal Processing Letter*, Vol. 10, NO. 5, May 2003, pp. 129-132.

[11] Mohamed Abdel-Mottaleb; Mohammad H. Mahoor. Algorithms for Assessing the Quality of Facial Images. *IEEE Computational Intelligence Magazine*, May 2007, pp. 10-17.

[12] Xiufeng Gao; Stan Z. Li; Rong Liu²; Peiren Zhang. Standardization of Face Image Sample Quality. 2007, pp. 1-11.

[13] Fang Hua; Peter Johnson; Nadezhda Sazonova; Paulo Lopez-Meyer; Stephanie Schuckers. Impact of Out-of-focus Blur on Face Recognition Performance Based on Modular Transfer Function, pp.1-6.

[14] P. A. Johnson; P. Lopez-Meyer; N. Sazonova, F. Hua; and S. Schuckers. Q-FIRE Database (Quality in Face and Iris Research Ensemble).

[15] Identix Empowering Identification. The Facelt SDK Guidelines for Image-Capturing Devices.

[16] Donald P. D'Amato, Ph.D. Best Practices for Taking Face Photographs and Face Image Quality Metrics. NIST Biometric Quality Workshop. March 8, 2006, pp 1-25.

[17] The Warnell School of Forestry and Natural Resources, University of Georgia, Athens. Workshop on Statistical Sampling and Estimation for Fisheries and Wildlife. Maio de 2006.

[18] Hirotugu Akaike. Estatístico japonês. Na década de 70 formulou o critério de informação de Akaike.

[19] Kresimir Delac; Mislav Grgic. Face Recognition. Published by the I-Tech Education and Publishing, Vienna, Austria 2007 ISBN 3-86611-283-1.

[20] Raul Queiroz Feitosa; Antonio Carlos Censi; Raphael Pithan; José Luiz Buonomo de Pinho. Optimal Region Weighting for Local Binary Pattern Based Face Recognition. In: ISSNIP Biosignals and Biorobotics Conference 2010, 2010, Vitória. Proceedings. Vitória: Universidade Federal do Espírito Santo, 2010. v.1.

[21] Dário Augusto Borges Oliveira; Raphael Pithan; Raul Queiroz Feitosa; Álvaro de Lima Veiga Filho; Antonio Carlos Censi; José Luiz Buonomo de Pinho. A Linear Discriminant Analysis approach for Local Binary Pattern Based Face Recognition. In: IWSSIP 2010 - 17th International Conference on Systems, Signals and Image Processing, 2010, Rio de Janeiro, Proceedings. Niterói: EdUFF, 2010, pp. 474-477

[22] Jo Chang-yeon. Face Detection using LBP features. CS 229 Final Project Report. December 12, 2008. pp. 1-4.

[23] Donald P. D'Amato. Best Practices for Taking Face Photographs and Face Image Quality Metrics, NIST Biometric Quality Workshop Mitretek Systems, Inc., March 8, 2006. pp. 1-26.

[24] Biometric Data Interchange Formats — Part 5: Face Image Data. ISO/IEC JTC 1/SC 37 N 506 (2004),ISO/IEC 2003. pp.1,25,28.

[25] Marques Filho; Ogê; Vieira Neto; Hugo. Processamento Digital de Imagens, Rio de Janeiro: Brasport, 1999. ISBN 8574520098. Cap. 3.

[26] Yi Yao; Bisma Abidi; Mongi Abidi. Quality Assessment and Restoration of Face Images in Long Range/High Zoom Video. pp. 4-5.

[27] Frank Weber; Director Algorithm Development Cognitec Systems GmbH. Some quality measures for face images and their relationship to recognition performance. Março 2006. pp. 1-16.

[28] S. Kotz; N. Johnson. Encyclopedia of statistical sciences. in Wiely, pp. 415–426, 1982.

- [29] Samarth Bharadwaj; Himanshu Bhatt; Mayank Vatsa; Richa Singh; Afzel Noore. Quality Assessment based Denoising to Improve Face Recognition Performance, pp. 16-174.
- [30] Z. Wang; A. Boik; H. Sheikh; E. Simoncelli. Image quality assessment: from error visibility to structural similarity. IEEE Transactions on Image Processing, 13(4):600–612, april 2004. pp.170-173.
- [31] Rafael C. Gonzalez; Richard E. Woods.2002, Digital Image Processing 2 Edição, :ISBN-10: 0201180758, 2002.
- [32] Sun Microsystems. Programming in Java Advanced Imaging, Release 1.0.1, November 1999.
- [33] IMM Database. Disponível em: www.imm.dtu.dk/~aam/aamexplorer. Acesso em fevereiro de 2012.