



**José David Bermúdez Castro**

**Estimativa da Idade a partir de Imagens Faciais**

**Dissertação de Mestrado**

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

Orientador: Prof. Raul Queiroz Feitosa

Rio de Janeiro

Abril de 2015



**José David Bermúdez Castro**

## **Estimativa da Idade a partir de Imagens Faciais**

Dissertação apresentada como requisito parcial para obtenção do título 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

**Profa. Marley Maria Bernardes Rebuzzi Vellasco**

Departamento de Engenharia Elétrica – PUC-RIO

**Prof. Hélio Côrtes Vieira Lopes**

Departamento de Informática – PUC-RIO

**Prof. José Eugenio Leal**

Coordenador(a) Setorial do Centro

Técnico Científico - PUC-Rio

Rio de Janeiro, 30 de abril de 2015

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

**José David Bermúdez Castro**

Nasceu em Barranquilla, Colômbia, no ano de 1987. Em 2009 obteve seu diploma de graduação pela Universidade do Norte de Barranquilla na especialidade de Engenharia Eletrônica. Atualmente está fazendo o mestrado no programa de Engenharia Elétrica da PUC-RIO na área de Processamento de Sinais, Automação e Robótica. Suas principais áreas de interesse são o processamento digital de imagens e visão computacional.

Ficha Catalográfica

Castro, José David Bermúdez

Estimativa da Idade a partir de Imagens Faciais / José David Bermúdez Castro; orientador: Raul Queiroz Feitosa. – 2015.

88 f. : il. (color.) ; 30 cm

Dissertação (mestrado)–Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Elétrica, 2015.

Inclui bibliografia

1. Engenharia elétrica – Teses. 2. Estimativa da Idade. 3. Redução de Dimensionalidade. 4. Modelos Ativos de Aparência. 5. Regressão. I. Feitosa, Raul Queiroz. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Elétrica. III. Título.

CDD: 621.3

A meus avós, Hermenegildo e Carmen, Juan e Josefa.  
A meus pais, Gilberto e Leicy.  
A minha irmã, Maira Alejandra.

## Agradecimentos

Ao meu orientador, o professor Raul Queiroz Feitosa, pela oportunidade, paciência, apoio, orientação e confiança no trabalho desenvolvido.

À PUC-Rio e aos seus professores do Departamento de Engenharia Elétrica.

À CAPES pelos auxílios concedidos.

Aos meus amigos e colegas da PUC-Rio, pela amizade, apoio e alegrias.

À Diana Garcia Martinez, pela amizade e apoio incondicional.

Aos meus amigos e colegas do LVC, pela amizade, apoio e tertúlias na hora do café.

## Resumo

Bermúdez Castro, José David; Queiroz Feitosa, Raul. **Estimativa da Idade a partir de Imagens Faciais**. Rio de Janeiro, 2015. 88p. Dissertação de Mestrado - Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

Esta dissertação tem por objetivo investigar métodos de estimação da idade a partir de imagens faciais. Avalia-se o impacto de distintos fatores sobre a acurácia da estimativa, especificamente, a acurácia da localização de pontos fiduciais, métodos de extração de atributos, de redução de dimensionalidade, e técnicas de regressão. Adicionalmente, foi estudada a influência da raça e do sexo na acurácia da estimação da idade desenvolvido. Consideraram-se cinco métricas de desempenho do sistema, especificamente, o erro médio absoluto (MAE), o erro médio absoluto por década (MAE/D), o erro médio absoluto por idade (MAE/A), o escore acumulado (CS), e os intervalos de confiança (IC). Os experimentos foram realizados empregando dois bancos de dados públicos, cujas imagens estão rotuladas com a idade da face. Os resultados indicaram que o método automático para detecção de pontos fiduciais da face tem uma repercussão moderada sobre a acurácia das estimativas. Entre as variantes analisadas, a que apresentou a melhor acurácia foi o sistema que emprega os AAMs (*Active Appearance Models*) como método de extração de atributos, o PCA (*Principal Components Analysis*) como método para reduzir dimensionalidade, e as SVRs (*Support Vector Regression*) como técnica para fazer regressão.

## Palavras-chave

Idade, Reconhecimento Facial, Regressão Linear, SVR.

## Abstract

Bermúdez Castro, José David; Queiroz Feitosa, Raul (Advisor). **Age Estimation from Facials Images**. Rio de Janeiro, 2015. 88p. MSc. Dissertation - Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

This thesis aims to investigate methods for age estimation from facial images. The impact of distinct factors over the estimate's accuracy is assessed, specifically the accuracy in the location of face fiducial points, feature extraction and dimensionality reduction methods, and regression techniques. Additionally, the dependence on race and gender in the accuracy of age estimation is assessed. Five performance metrics have been considered: the mean absolute error (MAE), the mean absolute error per decade (MAE / D), the mean absolute error for age (MAE / A), the cumulative score (CS) and confidence intervals (CI). The experiments were performed using two public databases, whose images are labeled with the age of the face. The results showed the impact of the automatic method for detection of fiducial points of the face has a moderate impact on the accuracy of the estimates. Among the analyzed variants, the one with the best accuracy was the system that employs the Active Appearance Models (AAMs) as feature extraction method, the Principal Components Analysis (PCA) as dimensionality reduction method, and Support Vector Regression (SVRs) as a technique to do regression.

## Keywords

Ages, Face Recognition, Linear Regression, SVR

# Sumário

1 Introdução	15
1.1. Objetivos	16
1.2. Estrutura da Dissertação	16
2 Trabalhos Relacionados	18
2.1. Modelos Antropomórficos	18
2.2. Modelos Ativos de Aparência	20
2.3. Aprendizagem de Subespaços.	21
2.4. Análise da Aparência Facial.	23
3 Estimativa de Idade e Fundamentos Teóricos	24
3.1. Descrição Geral do Sistema Utilizado	24
3.1.1. Localização da Face.	24
3.1.2. Normalização	29
3.1.3. Extração de Atributos	31
3.1.4. Redução da Dimensionalidade.	36
3.1.5. Regressão	44
4 Análise Experimental	48
4.1. Descrição dos Bancos de Dados	48
4.1.1. FG-NET Aging Database	48
4.1.2. MORPH Database	50
4.2. Métricas	52
4.2.1. Erro Médio Absoluto (MAE)	52
4.2.2. Escore Acumulado (CS)	53
4.2.3. Intervalos de Confiança (CI)	54
4.3. Procedimento Experimental Geral	55
4.3.1. Experimentos	56
4.3.2. Sumário dos Experimentos	78
5 Conclusões	80
5.1. Discussão	80



5.2. Trabalhos Futuros	81
6 Referências Bibliográficas	83

## Lista de figuras

Figura 3.1. Diagrama de Fluxo do Sistema de Estimação da Idade.	24
Figura 3.2. Detecção facial e localização dos pontos fiduciais. A detecção facial é feita através do detector de face de Viola-Jones, e a localização dos pontos fiduciais, por meio de CLM.	25
Figura 3.4 Representação do modelo de formas e o modelo de fragmentos ( <i>patches</i> ).	26
Figura 3.5. Diagrama conceitual do funcionamento dos CLM.	27
Figura 3.6. Normalização Geométrica e Fotométrica. Imagem adquirida em (Ayma 2014)	30
Figura 3.7 Exemplo de pontos fiduciais detectados por meio dos CLM.	31
Figura 3.8 Modelo linear da forma de um AAM.	33
Figura 3.9 Exemplo da instanciação de um AAM.	34
Figura 3.10 Regressão linear unidimensional com a banda $\epsilon$ -insensível.	47
Figura 4.1. Exemplo do Envelhecimento Facial dum Indivíduo no Banco de Dados FG-NET. Obtido de (Lanitis et al. 2002). A idade correspondente se indica debaixo de cada imagem.	49
Figura 4.2. Histograma das Idades no Banco de Dados FG-NET	49
Figura 4.3. Exemplos de Imagens do banco de dados MORPH ; a idade correspondente se indica debaixo de cada imagem.	51
Figura 4.4. Histograma das Idades do Álbum 1 do Banco de Dados MORPH	51
Figura 4.5. Histograma das Idades do Álbum 2 do Banco de Dados MORPH	52
Figura 4.6. Função de Densidade de Probabilidade Normal	55
Figura 4.7. Erro Acumulado (CS) da estimativa da idade no banco de dados FG-NET. Compara-se a acurácia do estimador para pontos fiduciais estimados automaticamente e marcados de forma manual. Também se comparam os três métodos de redução de dimensionalidade.	59
Figura 4.8 Erro Médio Absoluto por Idade (MAE/y) da estimativa da idade no banco de dados FG-NET, usando PCA, LDA e SL, respetivamente.	59
Figura 4.9. Intervalos de Confiança para PCA com nível de confiança 90%.	60
Figura 4.10. Intervalos de Confiança para LDA com nível de confiança 90%	61

Figura 4.11 Resultado dos Intervalos de Confiança para SL com nível de confiança 90%.	61
Figura 4.12 Erro Médio Absoluto por Idade (MAE/y) da estimativa da idade do álbum 2, no banco de dados MORPH. O PCA é usado como método para reduzir dimensionalidade, e o AAM para extrair atributos.	65
Figura 4.14 Erro Acumulado (CS): comparam-se as curvas do erro acumulado dos métodos de redução de dimensionalidade e métodos de extração de atributos de textura.	67
Figura 4.15 Intervalos de confiança para o PCA para a totalidade do bando de dados MORPH. O nível de confiança é de 95% e os atributos foram extraídos unicamente através do AAM.	68
Figura 4.16 Intervalos de confiança para o SL para a totalidade do banco de dados MORPH. O nível de confiança é de 95% e os atributos foram extraídos misturando AAM com LBP.	69
Figura 4.17 Intervalos de confiança para o PCA da raça Africana no bando de dados MORPH. O nível de confiança é 95% e os atributos foram extraídos unicamente através de AAM.	69
Figura 4.18 Resultado dos intervalos de confiança para o PCA da raça Branca - bando de dados MORPH. O nível de confiança é 95% e os atributos foram extraídos unicamente através de AAM.	70
Figura 4.19 Intervalos de confiança para o PCA da raça Hispânica no bando de dados MORPH. O nível de confiança é de 95% e os atributos foram extraídos unicamente através de AAM.	70
Figura 4.20 Erro Acumulado (CS). Comparam-se as curvas do erro acumulado dos métodos de redução de dimensionalidade e gênero das pessoas.	71
Figura 4.21 Resultado dos intervalos de confiança para o PCA no banco de dados FG-NET. O nível de confiança é de 95% e os atributos foram extraídos através de AAM misturado com o LBP.	73
Figura 4.22 Resultado dos intervalos de confiança para o PCA no banco de dados FG-NET. O nível de confiança é de 95% e os atributos foram extraídos através de AAM.	73
Figura 4.23 Erro Médio Absoluto por Idade (MAE/y) da Estimativa da Idade do álbum 2, no banco de dados MORPH. O PCA é usado para redução de dimensionalidade, e se comparam dois métodos para regressão; Regressão Linear e SVRs.	75

Figura 4.24 Intervalos de confiança para PCA + SVR no banco de dados MORPH. O nível de confiança é de 95%.	76
Figura 4.25 Intervalos de confiança ao usar PCA +Regressão linear no banco de dados MORPH. O nível de confiança é de 95%.	76
Figura 4.26 Erro Médio Absoluto por Idade (MAExA) da Estimativa da Idade do banco de dados FG-NET. O PCA é usado para redução de dimensionalidade, e se comparam dois métodos para regressão; Regressão Linear e SVRs.	77
Figura 4.27 Intervalos de confiança utilizando o PCA + SVR no banco de dados FG-NET. O nível de confiança é de 95%.	79
Figura 4.28 Intervalos de confiança utilizando o PCA + regressão linear no banco de dados FG-NET. O nível de confiança é de 95%.	79

## Lista de tabelas

Tabela 4.1. Distribuição percentual das Idades por década nos bancos de dados FG-NET e MORPH. Adquirido em (Geng, Zhou and Smith-Miles 2007)	50
Tabela 4.2 Erro Médio Absoluto (MAE) no Banco de Dados FG-NET; comparam-se os resultados do MAE para cada método de redução de dimensionalidade e para cada estratégia de marcado de pontos de interesse.	57
Tabela 4.3 Erro Médio Absoluto por Década (MAE/D) da estimativa de idade no banco de dados FG-NET.	57
Tabela 4.4. Erro Médio Absoluto (MAE) para estimativa da idade no banco de dados MORPH, Album2. Comparam-se os métodos de extração de atributos de textura usando o PCA para reduzir dimensionalidade com todas as amostras do banco de dados.	63
Tabela 4.5 Erro Médio Absoluto por Década (MAE/D) para estimativa da idade de todas as amostras do banco de dados MORPH, Album2. Comparam-se os métodos de extração de atributos de textura usando o PCA para reduzir dimensionalidade.	63
Tabela 4.6 Erro Médio Absoluto por Década (MAE/D) para Estimativa da Idade da Raça Africana no Banco de Dados MORPH, Album2. Comparam-se os métodos de extração de extração de atributos de textura usando o PCA para reduzir dimensionalidade.	64
Tabela 4.7 Erro Médio Absoluto por Década (MAE/D) para estimativa da idade da raça Branca no banco de dados MORPH, Album2. Comparam-se os métodos de extração de atributos de textura usando o PCA para reduzir dimensionalidade.	64
Tabela 4.8 Erro Médio Absoluto por Década (MAE/D) para estimativa da idade da raça Hispânica no banco de dados MORPH, Album2. Comparam-se os métodos de extração de atributos de textura usando o PCA para reduzir dimensionalidade.	65
Tabela 4.9 Erro Médio Absoluto (MAE) para estimativa da idade no banco de dados FG-NET.	72
Tabela 4.10 Erro Médio Absoluto por Década (MAE/D) para estimativa da idade do banco de dados FG-NET. Utilizou-se o PCA na etapa de redução	

de dimensionalidade.	72
Tabela 4.11. Erro Médio Absoluto (MAE) para estimativa da idade no banco de dados MORPH, Album2. Comparam-se as técnicas de regressão e redução de dimensionalidade.	74
Tabela 4.12 Erro Médio Absoluto por Década (MAE/D) para estimativa da idade do banco de dados MORPH, Album2. Comparam-se as técnicas de regressão e redução de dimensionalidade.	75
Tabela 4.13 Erro Médio Absoluto (MAE) para estimativa da idade no banco de dados FG-NET. Comparam-se as técnicas de regressão e de redução de dimensionalidade.	77

# 1

## Introdução

A estimativa da idade através de imagens faciais é um problema da área de visão computacional que tem atraído especial interesse nos anos recentes devido à grande variedade de aplicações potenciais. Por exemplo, na área de segurança do entretenimento, se pode usar a estimativa de idade para restringir o acesso de menores a sítios físicos ou virtuais exclusivos para adultos. Associada ainda ao reconhecimento do sexo, a estimativa de idade tem grande potencial de aplicação na área de marketing para o levantamento do perfil das pessoas que passam mais tempo olhando uma determinada propaganda, vitrine, ou gôndola de supermercado. Por outro lado, na área da biometria facial, as alterações da aparência da face devido ao envelhecimento podem, em grande parte degradar o desempenho de muitas aplicações de reconhecimento facial. Deste modo, adicionar a informação da idade a estes sistemas pode ajudar a tornar mais robustos e exatos sistemas de reconhecimento facial. Além disso, a estimativa da idade pode ser utilizada para acelerar a busca de pessoas em grandes bancos de faces em aplicações de identificação, restringindo o número de registros a serem comparados.

Não obstante, a grande variedade de fatores externos e genéticos que afetam o processo de envelhecimento facial faz deste problema um desafio para a comunidade científica. Mais corretamente, o envelhecimento facial é um processo incontrollável; ninguém pode antecipar ou retardar o envelhecimento à vontade. Além de ser um processo lento e irreversível, a idade de cada pessoa é determinada tanto por seus genes como por fatores externos, tais como, a saúde, estilo de vida e condições ambientais.

Os primeiros trabalhos realizados na área de estimativa da idade foram os desenvolvidos por (Kwon and da Vitoria Lobo 1994) e (Lobo and Kwon 1998). Esses estudos foram baseados em modelos morfológicos da face. Já a partir dos trabalhos publicados por Lanitis e coautores (Lanitis et al. 2002), a idade da face começou a ser caracterizada através de modelos estatísticos (Cootes et al. 2001), os quais, por sua capacidade para modelar a informação da forma e textura da face, são na atualidade os que melhor descrevem a idade a partir da imagem de um rosto.

Na atualidade, os modelos estatísticos são utilizados como atributos de entrada para um sistema de reconhecimento de padrões. De modo que a pesquisa na área está focada na busca de um subespaço de aprendizado onde seja possível caracterizar o processo de envelhecimento facial (Geng et al. 2013) e (Chen and Hsu 2013).

## 1.1. Objetivos

Neste contexto, a presente dissertação persegue os objetivos abaixo enumerados.

### Objetivo Geral

- Comparar abordagens alternativas para a estimação de idade a partir de imagens faciais.

### Objetivos Específicos

- Determinar a influência da acurácia da localização dos pontos fiduciais das faces sobre a acurácia da estimativa de idade.
- Avaliar a acurácia da estimativa da idade associada a dois métodos de extração de atributos de textura: AAM (*Active Appearance Modes*) e o LBP (*Local Binary Pattern*).
- Comparar métodos de redução de dimensionalidade.
- Comparar técnicas de regressão linear e não linear.
- Avaliar a acurácia do sistema para diferentes raças e gêneros.
- Identificar as dificuldades que mais impactam a acurácia e apontar para desenvolvimentos futuros.

## 1.2. Estrutura da Dissertação

A presente dissertação está estruturada em cinco capítulos organizados como a seguir:

- O capítulo 2 apresenta de forma sucinta diversos trabalhos que representam o estado da arte em estimativa de idade.
- O capítulo 3 apresenta o esquema de estimação construído para esta pesquisa. Descrevem-se em detalhes as etapas envolvidas no processo de



estimação de idade, bem como os fundamentos teóricos em que estão baseados os algoritmos utilizados.

- O capítulo 4 detalha a metodologia experimental empregada e apresenta os resultados obtidos.
- O capítulo 5 resume as conclusões extraídas no desenvolvimento desta dissertação, e aponta para trabalhos futuros.

## 2

### Trabalhos Relacionados

Neste capítulo são apresentados, de forma sucinta, diversos trabalhos que representam o estado da arte em estimativa de idade, e que foram usados como base para o desenvolvimento desta pesquisa. Os trabalhos estão organizados a seguir segundo o esquema proposto em (Fu et al. 2010), que se orienta pelo tipo de atributos extraídos das imagens faciais a partir dos quais se calculam as estimativas.

#### 2.1.

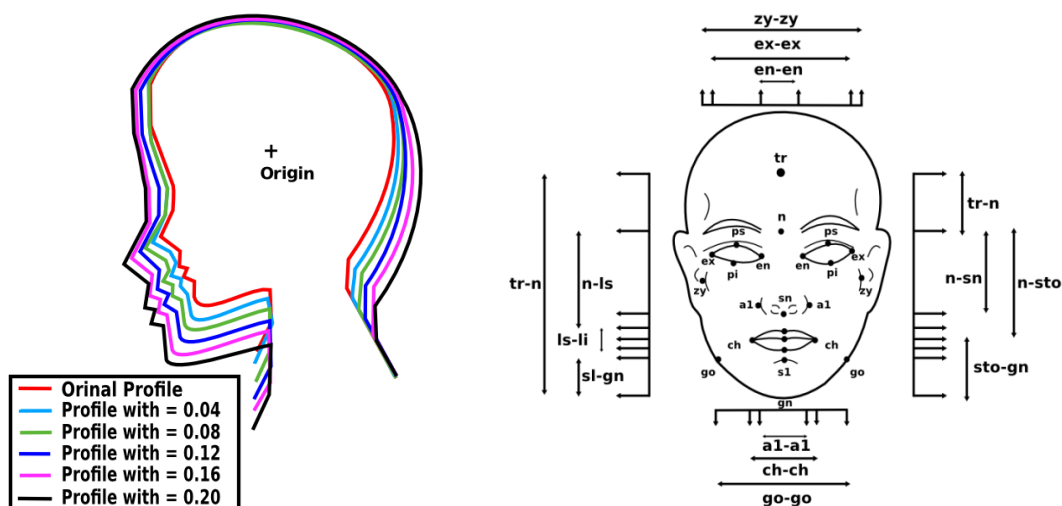
##### Modelos Antropomórficos

Os primeiros estudos publicados na área de estimativa da idade foram os desenvolvidos por (Kwon and da Vitoria Lobo 1994) e (Lobo and Kwon 1998). A ideia principal de ambos os trabalhos está fundamentada nas teorias desenvolvidas em (Alley 1988) para caracterizar o crescimento craniofacial.

Estas teorias utilizam o modelo matemático cardioide para descrever o crescimento da cabeça de uma pessoa desde a infância até a maturidade (Alley 1988). Esta transformação pode ser visualizada como uma série de cardioides.

Duas razões principais fizeram com que esta formulação matemática caísse em desuso para estimar a idade. Primeiro, porque o modelo matemático não é capaz de caracterizar adequadamente a forma da cabeça, especialmente para adultos, e segundo, porque é difícil construir o perfil da cabeça a partir de imagens de faces bidimensionais.

Uma nova abordagem passou a ser utilizada para caracterização do crescimento craniofacial a partir do trabalho publicado em (Farkas 1994), no qual a antropometria da face é definida por 57 pontos fiduciais extraídos da face (ver Figura 2.1). Nesta abordagem são utilizadas as proporções entre distâncias medidas através de pontos de referência da face como alternativa ao modelo matemático proposto em (Alley 1988).



**Figura 2.1** Modelos antropomórficos. a) modelo matemático cardioide do crescimento da cabeça, b) medidas antropométricas da face humana (retirado de (Alley, 1988)).

Em 1998, Lobo e Kwon (Lobo and Kwon 1998) calcularam seis proporções de distâncias em imagens faciais com vista frontal (Ver Figura 2.1), e as usaram para separar os bebês dos adultos. A experiência foi realizada num banco de dados pequeno conformado por 47 imagens faciais. Contudo, os autores não relatam nem desempenho nem a acurácia dos resultados obtidos neste banco de dados, mas mostraram os resultados do uso de cada proporção.

Em (Ramanathan and Chellappa 2006) são utilizadas oito relações de medidas de distância para modelar a progressão da idade em imagens faciais de pessoas jovens, i.e., de 0 até 18 anos. O objetivo é prever a aparência de um indivíduo através da idade, e executar ao mesmo tempo o reconhecimento da face conforme avança a idade. Avaliações do método empregaram Eigenfaces (Turk and Pentland 1991) como algoritmo de reconhecimento facial, num banco de dados com 233 imagens faciais correspondentes a 109 indivíduos, parcialmente coletados do banco de dados FG-NET. Os resultados obtidos apresentam uma melhora na taxa de reconhecimento facial quando é usada a transformação de progressão da idade. Essa transformação determina a aparência de uma pessoa a uma idade específica de acordo as oito relações de distancia e o modelo craniofacial mostrado na Figura 2.1.

As abordagens para a estimativa da idade baseadas no modelo antropomórfico da cabeça (Ver Figura 2.1) mostraram-se razoavelmente adequadas para determinar a idade de pessoas jovens, mas falham para mais

velhos, pois a forma da cabeça humana em adultos permanece quase constante. Em vista disso, em (Lobo and Kwon 1998) foram consideradas as rugas nas faces como indicadores de idade entre adultos. As rugas foram computadas em várias regiões, tais como a testa, próximo aos olhos, e das maçãs da face. A presença de rugas se baseia na detecção de curvas nessas regiões.

## 2.2.

### Modelos Ativos de Aparência

O Modelo Ativo de Aparências (*Active Appearance Model*, AAM) é um modelo estatístico e paramétrico do rosto, proposto em (Cootes, Edwards and Taylor 2001) para a codificação das imagens das faces. O modelo é gerado pela mistura de dois modelos, sendo o primeiro de formas e outro de intensidades, ambos construídos a partir de um conjunto de imagens de treinamento com base na Análise de Componentes Principais (*Principal Component Analysis*). Os parâmetros do modelo são utilizados posteriormente como descritores da imagem facial.

Lanitis e coautores (Lanitis, Taylor and Cootes 2002) estenderam o uso dos AAMs para a estimativa da idade, propondo uma função de envelhecimento que explica a variação da idade, isto é:  $Idade = f(\mathbf{b})$ . Na função de envelhecimento,  $Idade$  é a idade real da pessoa na imagem de face,  $\mathbf{b}$  é um vetor de 50 parâmetros do modelo aprendido com os AAM, e  $f$  é a função de envelhecimento, que define a relação entre a idade dos indivíduos e a descrição paramétrica das imagens de face. Nesse trabalho, os experimentos foram realizados num banco de dados com 500 imagens faciais de 60 indivíduos, sendo que há várias imagens em idades distintas da mesma pessoa no caso de 50 indivíduos. Os autores focaram-se, principalmente, na faixa de idades de 1 a 30 anos, conseguindo um erro médio absoluto (MAE) de 3.94.

Contrariamente às abordagens baseadas em modelos antropométricos, os AAMs apresentam boa acurácia numa ampla faixa de idades, e não apenas para as idades associadas a jovens.

Em função destes resultados, tornou-se consenso na comunidade que pesquisa o tema, que a forma e a textura da face têm ambas um papel importante na captura da evolução de envelhecimento facial. Por isso, os AAMs

são utilizados na maioria dos trabalhos atuais relacionados à estimativa da idade, por exemplo, (Geng et al. 2013, Chen and Hsu 2013, (Lu and Tan 2013) como extrator de atributos para a estimação de idade.

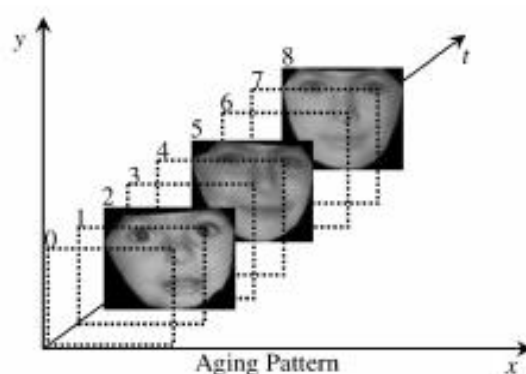
### 2.3. Aprendizagem de Subespaços.

Ao contrário dos métodos em que os atributos extraídos de cada imagem da face são usados diretamente como dados de entrada para algum algoritmo de reconhecimento de padrões, Geng e coautores (Geng, Yin and Zhou 2013; Geng et al. 2007), propuseram a utilização do que chamaram “subespaços de padrões de envelhecimento”. Nesta abordagem, para cada indivíduo do conjunto de treinamento, é determinado um subespaço onde é caracterizado o processo de envelhecimento facial. Depois disto, os dados de entrada ao sistema são mapeados nos subespaços aprendidos, de modo que, o subespaço de menor diferença com a data de entrada, é selecionado como o subespaço adequado para determinar a idade da imagem de face.

Em (Geng, Yin and Zhou 2013; Geng, Zhou and Smith-Miles 2007), o subespaço do padrão de envelhecimento facial é denominado *AGing pattErn Subspace* (AGES). Nesta técnica, um padrão de envelhecimento é definido como uma sequência ordenada cronologicamente de imagens faciais associadas a um indivíduo. Assim, por exemplo, na Figura 2.2, cada imagem de face disponível (2, 5 e 8) é alocada em sua posição correspondente sob o eixo  $t$  que corresponde à idade. Observe-se que as posições para as quais não há imagens disponíveis são deixadas em branco. Supondo que as imagens faciais de uma mesma pessoa em todas as idades estão disponíveis, esta sequência define o que se chama padrão completo de envelhecimento. Quando faltam imagens concernentes a algumas idades, o padrão de envelhecimento é dito incompleto. Neste caso, o método estabelece estimar as imagens faltantes baseando-se nas imagens disponíveis de faces com idades próximas à procurada.

A ideia principal deste método é criar modelos de padrões de envelhecimento facial, de modo que cada modelo representa um subespaço onde é caracterizado o processo de envelhecimento de uma pessoa em particular. Especificamente, a idade de uma face que nunca tenha sido vista pelo sistema é determinada pela projeção dessa imagem em cada subespaço

aprendido durante a etapa de treinamento. Uma vez que o subespaço de maior semelhança à imagem de entrada tenha sido determinado, a posição da face nesse padrão de envelhecimento indicará a idade da face. Em outras palavras, inicialmente os atributos extraídos (AAM) da imagem de entrada são comparados com os respectivos das faces de cada subespaço, seguidamente, é selecionado o subespaço de menor discrepância com face de entrada, finalmente, a mesma comparação é feita no subespaço selecionado, de modo que, a face mais semelhante à imagem de entrada, indicará a idade da imagem de entrada.



**Figura 2.2** Padrão de envelhecimento facial.

Por outro lado, Fu e coautores (Fu et al. 2007) propuseram aprender um padrão de envelhecimento genérico para múltiplas pessoas de diferentes idades, ao invés de aprender um padrão de envelhecimento específico para cada indivíduo. Trata-se de uma abordagem muito mais flexível em comparação com os subespaços de padrões de idades. Em (Fu, Xu and Huang 2007), utiliza-se uma técnica chamada *manifold embedding*, desenvolvida por (Seung and Lee 2000), que visa aprender um padrão de envelhecimento num espaço de baixa dimensão, a partir de muitas imagens faciais de múltiplos indivíduos para cada idade. Admite-se até mesmo que o banco de dados contenha apenas uma única idade para uma pessoa. Essa flexibilidade torna mais fácil a construção de um grande banco de dados para estimação do envelhecimento facial.

Outro método de aprendizado de subespaço foi proposto por Chen e coautores (Chen and Hsu 2013). Neste trabalho, um subespaço de baixa dimensão é criado através da solução de uma sequência de problemas de

otimização com restrições. Este método é explicado com maior detalhe no próximo capítulo.

## 2.4.

### **Análise da Aparência Facial.**

Uma vez que as variações locais na aparência facial durante o envelhecimento são muitas vezes evidentes, em (Ahonen et al. 2006) foram utilizados os Padrões Binários Locais ou *Local Binary Patterns* (LBP) (Ojala et al. 1996), definidos no capítulo 3, para extrair atributos de textura em um sistema automático de estimação de idade. Os autores reportam que alcançaram 80% de exatidão com o banco de dados FERET (Phillips et al. 2000), empregando a classificação do vizinho mais próximo, e de 80% a 90% com AdaBoost (Yang and Ai 2007). Igualmente, em (Yan et al. 2008b) foram utilizados os atributos de Gabor (Liu and Wechsler 2002) no banco de dados FERET, alcançando uma acurácia de 80%.

Do mesmo modo, em (Yan et al. 2008a; Yan, Zhou, Liu, Hasegawa-Johnson and Huang 2008b) foram propostos *Spatially Flexible Patches* (SFP) como descritor de atributos. Este método codifica simultaneamente a informação local da aparência e da posição do bloco analisado. O sistema alcançou um erro médio absoluto (MAE) de 4,95 anos no banco de dados do FG-NET, e de 4,94 e 4,38 anos para mulheres e homens, respectivamente, no banco de dados YGA (Guo et al. 2009).

### 3

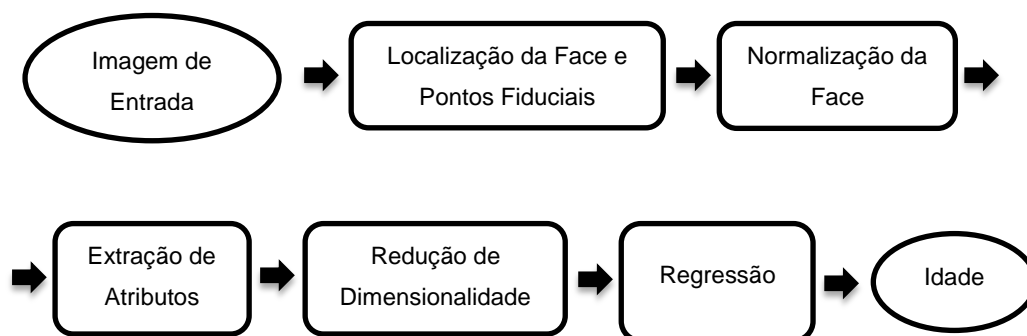
## Estimativa de Idade e Fundamentos Teóricos

Neste capítulo são apresentadas as implementações de métodos de estimação de idade desenvolvidas para este estudo. Descreve-se em detalhe cada uma das etapas do processo de estimação de idade, bem como os fundamentos teóricos em que se baseiam os algoritmos utilizados.

### 3.1.

#### Descrição Geral do Sistema Utilizado

O sistema de estimação da idade desenvolvido neste trabalho é ilustrado na Figura 3.1. O processo inicia com a detecção da face e dos pontos fiduciais na imagem de entrada. Em seguida, a face detectada é normalizada para dimensões padrão. Neste ponto, atributos de textura e de forma são extraídos da face normalizada, e depois, utilizando um método de redução de dimensionalidade, cria-se um conjunto menor de atributos mais relevantes para os propósitos da aplicação. Por último, um modelo de regressão estima a idade da pessoa a partir dos atributos gerados previamente.



**Figura 3.1.** Diagrama de Fluxo do Sistema de Estimação da Idade.

#### 3.1.1.

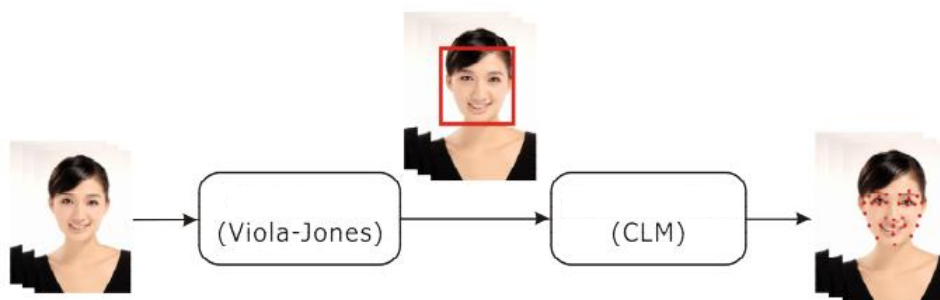
##### Localização da Face.

A localização da face numa imagem é um caso especial da detecção de objetos numa cena (Lewis and Ellis 2003). A tarefa consiste em determinar a

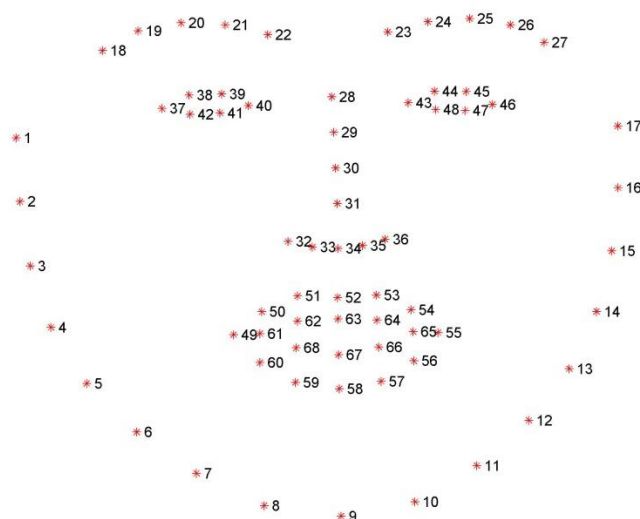


localização e tamanho da face em uma imagem digital, isto é, localizam-se as faces, e ignoram-se os demais objetos presentes na imagem. Com a face detectada, procede-se à localização dos pontos fiduciais, também chamados de pontos de interesse, associados aos olhos, nariz, boca, entre outros (Ayma 2014) (vide Figura 3.2).

Como indicado na Figura 3.2, o detector de Viola-Jones (Viola and Jones 2001) é usado para localizar a face, e os Modelos de Restrições Locais (*Constrained Local Models - CLM*) (Asthana et al. 2013; Cristinacce and Cootes 2006) para determinar os pontos de interesse. Especificamente, os pontos fiduciais a detectar, seguem o esquema de marcação descrito em (Sagonas et al. 2013), e detalhado na Figura 3.3.



**Figura 3.2.** Detecção facial e localização dos pontos fiduciais. A detecção facial é feita através do detector de face de Viola-Jones, e a localização dos pontos fiduciais, por meio de CLM.

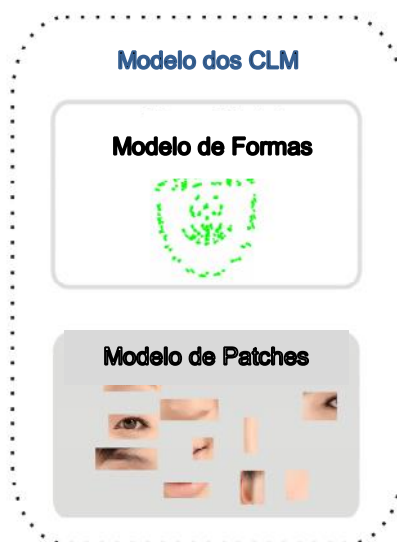


**Figura 3.3.** Esquema da marcação de pontos utilizados. Fonte: (Sagonas, Tzimiropoulos, Zafeiriou and Pantic 2013).

### 3.1.1.1. Modelos de Restrições Locais.

Os Modelos de Restrições Locais ou *Constrained Local Models* (CLM) são modelos que permitem detectar os pontos fiduciais explorando o conhecimento a priori do objeto que está sendo analisado. Por exemplo, para o caso das faces, a partir da detecção do rosto é possível estimar aproximadamente a localização das partes principais da face, como olhos, boca e nariz, guiando-se só pela geometria do rosto. Adicionalmente à informação relacionada à forma, a textura das regiões mencionadas é uma informação útil para a detecção. Nessa linha de raciocínio, os CLM restringem a busca dos pontos fiduciais a regiões onde o item correspondente tem grande probabilidade de estar. Em outras palavras, se utilizam modelos locais para procurar itens individuais, e usar conhecimento prévio quanto à forma da face para restringir a busca.

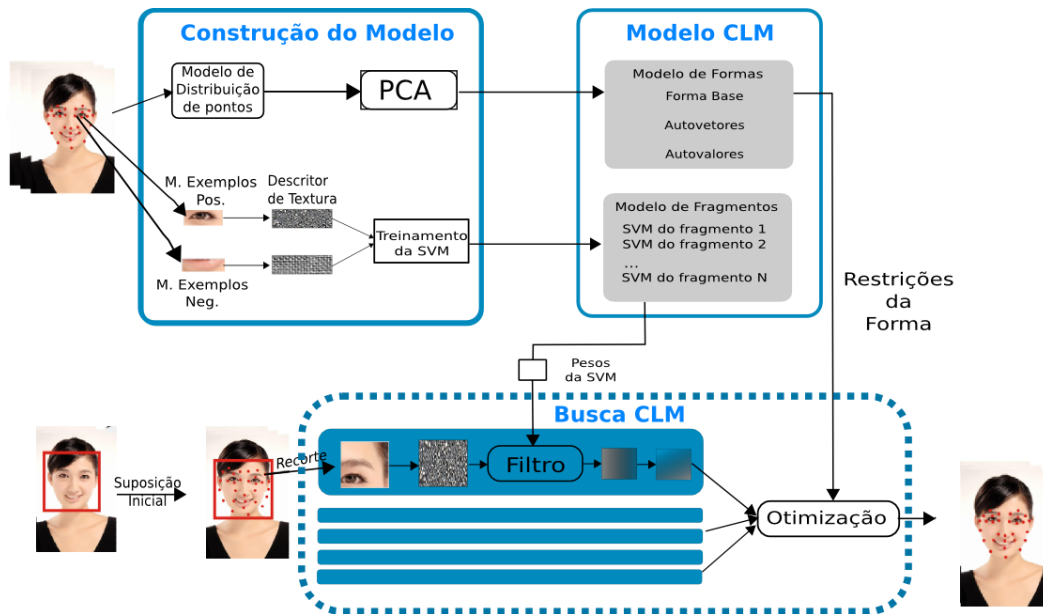
Especificamente, os CLM precisam de duas fontes de informação para operar. A primeira é o modelo individual da aparência de cada um dos itens que se busca localizar, tais como o nariz, olhos, boca, etc. Este modelo também é conhecido como o “modelo local dos fragmentos” (*patches* em inglês). A segunda é a informação relacionada à localização provável desses fragmentos de imagem, denominado modelo de formas. (vide Figura 3.4).



**Figura 3.4** Representação do modelo de formas e o modelo de fragmentos (*patches*).

Estes modelos são construídos seguindo o procedimento que se ilustra de forma sucinta na Figura 3.5. Note-se que os modelos são gerados a partir de um

conjunto de imagens de faces etiquetadas de forma manual. Geralmente, antes de construir os modelos, se acostuma aplicar a Análise de Procustes (Cootes and Taylor 2004) a cada imagem, de modo que, as faces são normalizadas em translação, rotação e escala. A seguir se explica de forma detalhada como são gerados os dois modelos.



**Figura 3.5.** Diagrama conceitual do funcionamento dos CLM.

### a) Modelo de Formas

Matematicamente, se define a forma  $s$  de uma imagem de face como as coordenadas dos  $v$  pontos fiduciais da face.

$$\mathbf{s} = (x_1, y_1, x_2, y_2, \dots, x_v, y_v)^T \quad 3.1$$

Sendo  $m$  o número de imagens de treinamento utilizadas para construir os dois modelos, se define  $\mathbf{S} \in \mathbb{R}^{m \times v}$  como a matriz que concatena os  $m$  vetores de formas  $s$ . Especificamente, a matriz  $\mathbf{S}$  condensa todas as variações de forma das imagens de treinamento. Agora, aplicando a Análise de Componentes Principais ou *Principal Component Analysis* (PCA) nesta matriz, as variações de forma podem ser expressas de uma maneira mais compacta, através de uma forma base  $s_0$  acrescida de uma combinação linear de  $n$  vetores de formas  $s_i$ ,

$$\mathbf{s} = \mathbf{s}_0 + \sum_{i=1}^n p_i \mathbf{s}_i \quad 3.2$$

onde os coeficientes  $p_i$  são os parâmetros do modelo da forma.

Em suma, a forma base  $\mathbf{s}_0$  corresponde à forma média, e os vetores  $\mathbf{s}_i$ , aos  $n$  autovetores associados aos  $n$  maiores autovalores.

### b) Modelo de Fragmentos

Na Figura 3.5 se apresenta o diagrama conceitual de como é construído o modelo de fragmentos. Tal modelo é constituído por um conjunto de  $N$  classificadores do tipo máquinas de vetores de suporte (SVM) (Cortes and Vapnik 1995), ou seja, um classificador SVM por cada ponto, especializado em identificar as características de textura na vizinhança de cada ponto fiducial.

Em particular, como se ilustra na Figura 3.5, um classificador SVM associado a um ponto fiducial é treinado selecionando  $M$  exemplos positivos e  $M$  exemplos negativos. Especificamente, os exemplos negativos fazem referência a outros pontos fiduciais escolhidos geralmente de forma aleatória. Note-se também, que um exemplo é definido como uma janela de pixels centrada no ponto fiducial de interesse, cujas dimensões variam conforme a resolução das imagens de treinamento.

Por outro lado, a saída de um classificador SVM é dada pela Equação 3.3,

$$f(\mathbf{x}, \omega) = \sum_{i=1}^m \omega_i x_i + b \quad 3.3$$

onde o parâmetro  $b$  é o *bias*,  $\mathbf{x}$  é o vetor de atributos de entrada, e  $\omega$  é o plano de separação aprendido pela SVM durante o treinamento. Se os dados de entrada são normalizados com média zero, o *bias*  $b$  pode ser omitido da equação, resultando na seguinte expressão,

$$f(\mathbf{x}, \omega) = \sum_{i=1}^m \omega_i x_i \quad 3.4$$

Observe que a anterior equação corresponde à definição de convolução para o caso de imagens digitais. Neste sentido, os classificadores SVMs podem ser vistos como filtros digitais, com a particularidade de que as imagens resultantes da filtragem terão valores na faixa de -1 até 1. O ponto fiducial em questão estará próximo das posições em que o resultado da aplicação da SVM apresentar for próximo a um.

Em síntese, o modelo de fragmentos está conformado por um conjunto de  $N$  filtros digitais, cuja configuração de pesos é aprendida pelos classificadores SVM.

Uma vez construídos os modelos de forma e de fragmentos, procede-se à detecção dos pontos de interesse sobre a imagem facial, como se detalha a seguir.

Primeiro, a face é detectada utilizando o algoritmo de Viola Jones, por exemplo. Segundo, realiza-se uma suposição inicial quanto à localização dos pontos na imagem de entrada. Geralmente a forma base  $s_0$  é empregada para esse fim. Terceiro, para cada ponto fiducial é selecionada uma janela de pixels de dimensão maior do que a utilizada para construir o modelo de fragmentos. A matriz com os indicadores da localização dos pontos fiduciais provê indícios da localização exata do ponto fiducial procurado. A esta altura, a informação fornecida pelo modelo de fragmentos é comparada com todas as possíveis variações do modelo de formas. A ideia é mover o ponto de interesse para uma coordenada que possivelmente seja mais próxima da posição desejada. Este processo é repetido de forma iterativa até o algoritmo convergir.

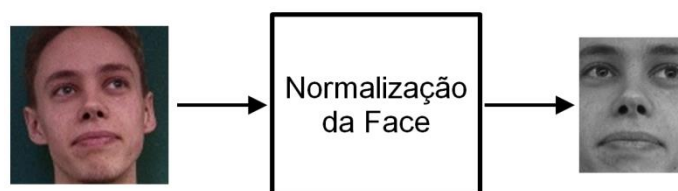
Finalmente, os CLM podem ser vistos como uma variante dos Modelos Ativos de Aparência, os quais se apresentam mais adiante na sessão 3.1.3 que trata da extração de atributos. A implementação de CLM utilizada neste trabalho foi desenvolvida por (Asthana, Zafeiriou, Cheng and Pantic 2013).

### **3.1.2. Normalização**

A normalização é o processo pelo qual a imagem facial é transformada geométrica e fotometricamente de modo a ajustar-se a um dado padrão. Geralmente, em um sistema de reconhecimento facial, a partir dos pontos

fiduciais obtidos na etapa de detecção facial, a imagem da face é normalizada em rotação e escala no plano da imagem, de modo que a face resultante esteja alinhada ao eixo horizontal do plano, e a distância entre os olhos seja fixa (Ayma 2014). A Figura 3.6 mostra um exemplo de normalização frequentemente utilizada em sistemas de reconhecimento facial.

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



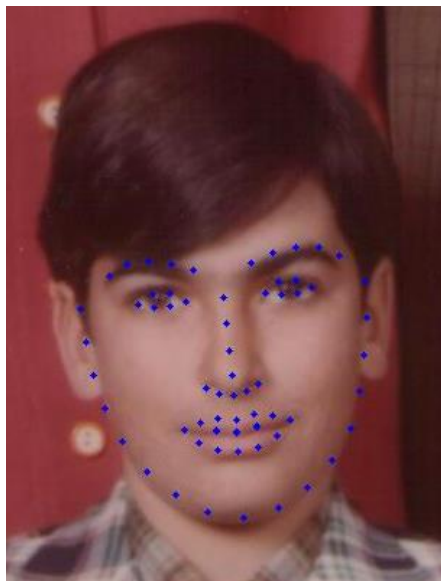
**Figura 3.6.** Normalização Geométrica e Fotométrica. Imagem adquirida em (Ayma 2014)

A abordagem seguida neste trabalho impõe outro padrão à etapa de normalização de face. Mais especificamente, os Modelos Ativos da Aparência (AAM) (Cootes, Edwards and Taylor 2001) utilizados neste trabalho requerem que a imagem normalizada contenha toda a face e ainda certa margem em seu entorno (Ver Figura 3.7).

Para se aplicar a técnica AAM faz-se necessário que a imagem contenha uma janela suficientemente ampla em torno de cada um dos pontos fiduciais no interior da qual o algoritmo que ajusta os pontos fiduciais possa interagir na busca da melhor localização.

Dado que os AAM são invariantes à rotação, translação e iluminação, a normalização consiste em levar as imagens faciais a uma mesma escala, ou seja, as imagens são transformadas de modo a terem um tamanho aproximadamente fixo. Para isto, as imagens são redimensionadas de tal forma que a distância entre os olhos seja igual a um valor determinado. Este procedimento foi realizado neste trabalho porque a implementação de AAM utilizada não é multi-escalar. Existem implementações, não disponíveis publicamente, que são invariantes à escala mediante o uso de pirâmides de imagens <sup>1</sup>.

<sup>1</sup> <http://codecapsule.com/2010/08/12/active-appearance-models-in-c-plus-plus/>



**Figura 3.7** Exemplo de pontos fiduciais detectados por meio dos CLM.

A figura 3.6 ilustra um resultado obtido pelo método. Observa-se que os pontos fiduciais no olho direito foram detectados com um erro perceptível.

### **3.1.3. Extração de Atributos**

A extração de atributos, em termos gerais, é o procedimento que extrai informações relevantes de um objeto a caracterizar, por exemplo, imagens faciais normalizadas, com o intuito de discernir, de algum modo, a classe a que este objeto pertence, no caso do presente trabalho, definidas pelas idades. Como se mencionou anteriormente escolheram-se os Modelos Ativos de Aparência (AAM) para tal fim, porque a diferencia de outras técnicas de extração de atributos, os AAM combinam atributos de forma e textura da imagem facial. Ambos os tipos de atributos têm um papel importante no processo de determinação da idade a partir de imagens faciais (Geng, Yin and Zhou 2013).

Além dos AAM, investigou-se na presente dissertação os Padrões Binários Locais (LBP) (Ojala, Pietikäinen and Harwood 1996) para representação da textura das faces. Com efeito, a informação de textura no AAM é substituída pela obtida por meio do LBP. Desta forma, se gera um modelo alternativo ao AAM que tem como informação de aparência, a imagem gerada pelo LBP. Este estudo é realizado devido a que o LBP é um algoritmo muito poderoso que tem apresentado bom desempenho em muitas aplicações relacionadas á área de reconhecimento facial.

A seguir, se apresentam os fundamentos teóricos dos AAM e LPB.

### 3.1.3.1. Modelos Ativos de Aparência.

Segundo Matthews e Baker (Matthews and Baker 2004), os Modelos Ativos de Aparência são modelos não lineares, generativos e paramétricos de um fenômeno visual. A aplicação mais frequente de AAM tem sido o modelamento da face. Em uma aplicação típica, o primeiro passo consiste em ajustar o AAM à imagem de entrada, isto é, os parâmetros do modelo são determinados de modo a maximizar a correspondência entre uma instância do modelo e a imagem de entrada. Uma vez determinados, os parâmetros podem ser utilizados como atributos de entrada para um algoritmo de reconhecimento facial, por exemplo.

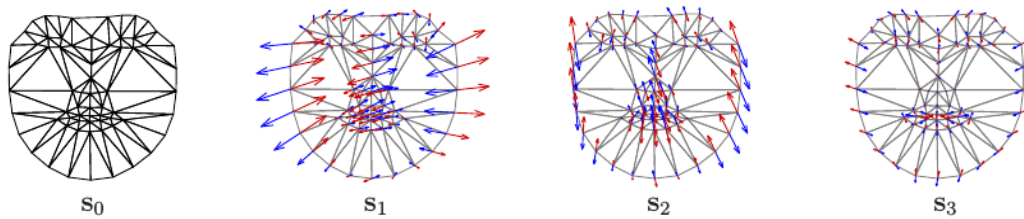
Ao igual que os CLM, os AAM são construídos a partir de um conjunto de imagens etiquetadas de forma manual. Convencionalmente, os AAM consistem de duas componentes: o Modelo de Formas e o Modelo de Aparências ou Textura.

#### c) Modelo de Formas

Matematicamente, o modelo de formas dos AAM segue a mesma definição do modelo de formas dos CLM, explicados de maneira detalhada na seção 3.1.1.1. Em suma, está composto por uma forma base  $s_0$  acrescida de uma combinação linear de  $n$  vetores de formas  $s_i$  (Ver Eq. 3.2), onde os coeficientes  $p_i$  são os parâmetros do modelo da forma.

Na Figura 3.8, se apresenta um exemplo de um modelo de formas de 68 vértices. Nesta figura, a malha base se apresenta à esquerda, e à direita os primeiros três vetores de forma  $s_1$ ,  $s_2$  e  $s_3$ . Observe-se que o modelo de forma refere-se a uma malha triangulada. Finalmente, os parâmetros do modelo de forma do AAM, dados pelo vetor  $\mathbf{p} = (p_1, p_2, \dots, p_n)^T$ , são utilizados como atributos de entrada para um algoritmo de reconhecimento de padrões.





**Figura 3.8** Modelo linear da forma de um AAM.

#### d) Modelo de Aparência.

A aparência de um AAM está definida dentro da malha base  $s_0$ . Admitindo que  $s_0$  também denote o conjunto de pixels  $\mathbf{x} = (x, y)^T$  dentro da malha base  $s_0$ , a aparência de um AAM é uma imagem  $A(\mathbf{x})$  definida sobre os pixels  $\mathbf{x} \in s_0$ . Similarmente ao modelo de forma, a imagem de aparência  $A(\mathbf{x})$  pode ser expressa através duma aparência base  $A_0(\mathbf{x})$  mais uma combinação linear de  $m$  imagens de aparência  $A_i(\mathbf{x})$ :

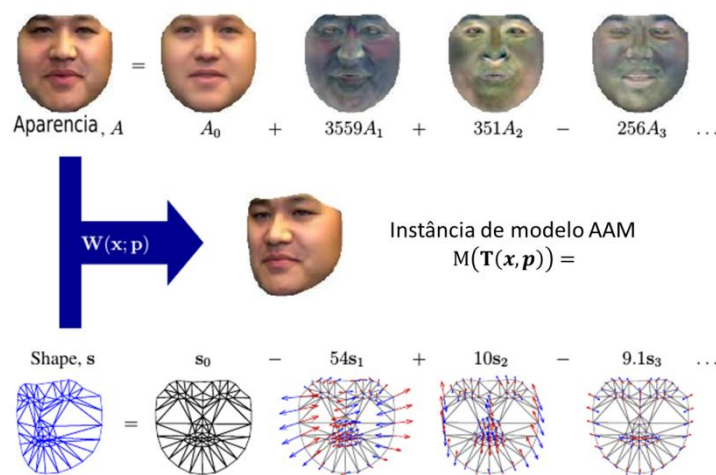
$$A(\mathbf{x}) = A_0(\mathbf{x}) + \sum_{i=1}^m \lambda_i A_i(\mathbf{x}) \quad 3.5$$

onde os coeficientes  $\lambda_i$  são os parâmetros do modelo de aparências. A aparência base  $A_0(\mathbf{x})$  corresponde à imagem média, e as imagens  $A_i(\mathbf{x})$  às  $m$  autoimagens associadas aos  $m$  maiores autovalores.

Uma vez obtidos o modelo de forma e o modelo de aparências, procede-se à instanciação do modelo, como descrito a seguir.

- **Instanciação do Modelo de AAM**

Supondo inicialmente conhecidos os parâmetros  $\mathbf{p} = (p_1, p_2, \dots, p_n)^T$  do modelo de forma do AAM, pode se utilizar a equação 3.2 para gerar a forma  $s$ . Similarmente, se conhecidos os parâmetros do modelo de aparência  $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_n)^T$ , pode se gerar a aparência  $A(\mathbf{x})$  definida no interior da malha base  $s_0$ . Deste modo, a instância do modelo do AAM com parâmetros de forma  $\mathbf{p}$  e parâmetros de aparência  $\boldsymbol{\lambda}$  é criada levando a aparência  $A$  desde a malha base  $s_0$  até o modelo de forma  $s$ . Este procedimento é ilustrado na Figura 3.9 para um conjunto de valores de  $\mathbf{p}$  e  $\boldsymbol{\lambda}$ .



**Figura 3.9** Exemplo da instanciação de um AAM.

Em particular, o par de formas  $s_0$  e  $s$  definem uma deformação afim por trechos desde  $s_0$  até  $s$ . Esta deformação afim consiste no seguinte procedimento,

Cada vértice de cada triângulo em  $s_0$  está associado a um vértice de um triângulo particular em  $s$ , por meio de uma transformação geométrica (translação, rotação, escala). Esta correspondência determina uma deformação afim única que vai de um triângulo em  $s_0$  para outro em  $s$ . Mais precisamente, os vértices do primeiro triângulo mapeiam em vértices do segundo. Deste modo, identificada a correspondência entre cada um dos triângulos, se procede com o conjunto de pixels que compõem as duas imagens, em outras palavras, as aparências das duas imagens são alinhadas. Assim, para cada pixel identificado por seu vetor de coordenadas  $x = (x, y)$  em  $s_0$  deve-se primeiro determinar o par de triângulos associados em  $s_0$  e  $s$ , e aplicar-lhe a função de deformação afim correspondente entre esses dois triângulos. Finalmente, o alinhamento ocorre mediante a aplicação de deformações afim específicas aos pixels que conformam cada triângulo, denotadas por  $T(x, p)$ . Deste modo, a instância final do modelo AAM é calculada deformando a aparência  $A$  de  $s_0$  para  $s$ , utilizando a relação  $T(x, p)$ . Este processo é representado pela seguinte equação:

$$M(T(x, p)) = A(x) \quad 3.6$$

onde  $M$  é uma imagem 2D de tamanho e forma apropriados para comportar a instância do modelo. Esta equação é interpretada da seguinte maneira. Dado um pixel  $x$  em  $s_0$ , seu correspondente em  $s$  é dado por  $T(x, p)$ .

As linhas que se seguem descrevem sucintamente o processo de ajuste de uma imagem ao AAM. A partir desse ajuste, é possível caracterizar essa imagem, baseando-se nos parâmetros calculados ao se ajustar o modelo.

- **Ajuste de uma imagem ao correspondente AAM.**

Esta seção apresenta o procedimento para determinação dos parâmetros de forma  $\mathbf{p}$  e de aparência  $\lambda$  presentes nas equações 3.2 e 3.5, respectivamente.

Os valores ótimos destes parâmetros promovem o melhor ajuste da imagem de entrada  $I(\mathbf{x})$  à correspondente instância  $M(\mathbf{T}(\mathbf{x}, \mathbf{p})) = A(\mathbf{x})$ .

A ideia é, portanto, determinar os valores de  $\mathbf{p}$  e  $\lambda$  que minimizam a discrepância entre  $I(\mathbf{x})$  e  $M(\mathbf{T}(\mathbf{x}, \mathbf{p}))$ .

Para isto, existem dois sistemas de coordenadas onde esta discrepância pode ser minimizada: nas coordenadas da imagem  $I$  ou nas coordenadas do AAM. Tem-se que a melhor escolha em termos da eficiência é o sistema de coordenadas do AAM, ou seja, na base de  $s_0$ , dado que é o sistema de coordenadas comum para cada imagem de entrada que se deseja ajustar ao modelo (Matthews and Baker 2004).

Como já se ressaltou, se  $\mathbf{x}$  é o vetor de coordenadas de um *pixel* em  $s_0$ , o *pixel* que a ele corresponde na imagem de entrada  $I$  está na coordenada  $\mathbf{T}(\mathbf{x}, \mathbf{p})$ . No *pixel*  $\mathbf{x}$  o AAM tem a aparência  $A(\mathbf{x}) = A_0(\mathbf{x}) + \sum_{i=1}^m \lambda_i A_i(\mathbf{x})$  e no *pixel*  $\mathbf{T}(\mathbf{x}, \mathbf{p})$ , a imagem de entrada tem a intensidade  $I(\mathbf{T}(\mathbf{x}, \mathbf{p}))$ . O que se deseja é minimizar a discrepância entre  $A(\mathbf{x})$  e  $I(\mathbf{T}(\mathbf{x}, \mathbf{p}))$  para o conjunto de pixels em  $s_0$ . Usualmente se utiliza como medida de discrepância a soma do quadrado das diferenças entre essas duas quantidades, a saber:

$$\sum_{\mathbf{x} \in s_0} [A_0(\mathbf{x}) + \sum_{i=1}^m \lambda_i A_i(\mathbf{x}) - I(\mathbf{T}(\mathbf{x}, \mathbf{p}))]^2, \quad 3.7$$

onde a soma é realizada sobre todos os pixels  $\mathbf{x}$  na malha  $s_0$ . O objetivo do ajuste de AAM é minimizar a expressão em 3.7 com respeito aos parâmetros de forma  $\mathbf{p}$  e de aparência  $\lambda$ .

Considerável esforço foi investido nos anos recentes visando a uma solução eficiente para o problema de minimização da expressão em 3.7. Em (Matthews and Baker 2004) são resumidas as técnicas desenvolvidas mais importantes. Nesta dissertação foi utilizada a implementação feita por (Vezaro 2010) da técnica *inverse compositional image alignment algorithm* desenvolvida por Matthews e coautores (Matthews and Baker 2004) dado que é uma das soluções mais eficientes que existem na atualidade.

### 3.1.3.2. Padrões Binários Locais (LBP).

O operador LBP introduzido por Ojala e coautores em (Ojala, Pietikäinen and Harwood 1996) é um dos mais amplamente utilizados descritores de textura na atualidade. Embora simples, têm alto poder de discriminação e eficiência computacional, pelo que é extensivamente usado em muitas aplicações.

O operador codifica os pixels duma imagem como uma sequência de bits que indicam se sua intensidade é ou não maior do que a de um conjunto de pixels em sua vizinhança. O resultado é representado por uma matriz de mesma dimensão da imagem de entrada, cujos elementos são os números binários formados pelas sequências de bits assim calculadas.

Mais especificamente, a imagem é codificada pela comparação de cada pixel ( $x_c$ ) com alguns de seus vizinhos ( $x_i$ ). Esta codificação é definida pela equação (3.7), onde,  $LBP(x_c)$  é o número que representa a textura associada à vizinhança do pixel  $x_c$ .

$$LBP(x_c) = \sum_{i=0}^7 S(x_i - x_c) 2^i \quad (3.8)$$

$$S(x_i - x_c) = \begin{cases} 1, & x_i \geq x_c \\ 0, & x_i < x_c \end{cases} \quad (3.9)$$

### 3.1.4. Redução da Dimensionalidade.

A redução da dimensionalidade de um vetor de atributos é um processo necessário e conveniente para muitos problemas de reconhecimento de

padrões. Frequentemente, a dimensão do vetor de atributos é muito elevada, e contém atributos pouco relevantes para o problema ou mesmo redundantes.

Em tais casos a complexidade do modelo e, por consequência, os custos computacionais associados podem ser muito elevados. Na aplicação em tela, espaço de atributos de dimensão elevada implica em longos tempos de processamento para sintonizar os parâmetros do modelo. O impacto dos problemas anteriormente mencionados pode ser diminuído, aplicando alguma técnica de redução de dimensionalidade que permita transformar os dados para um espaço de menor dimensão, e, por conseguinte, reduzir a complexidade do problema.

O presente trabalho investigou três diferentes técnicas de redução de dimensionalidade, a saber, *Análise de Componentes Principais* (PCA) (Jolliffe 2005), *Análise Linear de Discriminante* (LDA) (Fisher 1936) e *Aprendizado de Subespaço através de pares ordenados de atributos* (SL) (Chen and Hsu 2013). Para PCA e LDA, foi imposta como condição que o novo espaço conservasse 99% da informação contida no espaço original. Para o SL, selecionou-se dimensão 100 porque com este valor o método apresentou a melhor performance.

A seção seguinte descreve sucintamente cada um destes métodos.

#### **3.1.4.1. Análise de Componentes Principais (PCA)**

A Análise de Componentes Principais ou *Principal Component Analysis* (PCA) é um procedimento matemático que converte um conjunto de variáveis, potencialmente correlacionadas, em outro conjunto de variáveis linearmente descorrelacionadas, chamadas componentes principais. As novas variáveis, as componentes principais, são combinações lineares das variáveis originais ordenadas segundo sua importância, por exemplo, a primeira componente principal corresponde à combinação linear que captura a máxima variância (Johnson, 1992) dos dados de entrada.

A preservação da variabilidade total contida nos dados de treinamento requer que se mantenham todas as  $K$  variáveis originais. Porém,

frequentemente, a maior parte dessa variabilidade concentra-se em um número menor  $k < K$ , de componentes principais. As  $k$  componentes principais podem então substituir as  $K$  variáveis originais, reduzindo-se desta forma a dimensionalidade do sistema e preservando a maior parte da variabilidade dos dados (Marques, 2005).

Para o cálculo das componentes principais constrói-se inicialmente a matriz  $N \times K$  de treinamento  $\mathbf{X} = [\mathbf{x}_1 \mathbf{x}_2 \dots \mathbf{x}_N]^T$ , onde  $\mathbf{x}_i$  denota daqui para diante no texto o vetor coluna contendo os atributos originais que descrevem uma imagem facial de treinamento. A matriz  $\mathbf{X}$  contém, portanto, em cada uma de suas  $N$  linhas um vetor de treinamento, cada um composto de  $K$  variáveis, correspondentes a cada uma das  $K$  colunas de  $\mathbf{X}$ , onde  $K$  é a dimensão do espaço de atributos original.

O primeiro passo consiste em mover o centróide do conjunto de dados para a origem do sistema de coordenadas. Formalmente, cada vetor  $\mathbf{x}_i$  é substituído por  $\mathbf{x}_i - \bar{\mathbf{x}}$ , onde  $\bar{\mathbf{x}}$  representa a média do conjunto de vetores de treinamento. Este procedimento assegura que a nuvem de dados esteja centralizada na origem dos componentes principais, sem afetar as relações espaciais dos dados nem as variações embutidas nas variáveis. Assim, com certo abuso de notação, adota-se o mesmo símbolo para representar a matriz de dados após a centralização, isto é,  $\mathbf{X} = [\mathbf{x}_1 \mathbf{x}_2 \dots \mathbf{x}_N]^T$ .

No próximo passo se calcula a matriz de covariâncias  $\mathbf{\Sigma} = \mathbf{X}^T \mathbf{X}$ , e desta, se extraem os autovalores  $l_i$  e autovetores  $\mathbf{e}_i$  ( $i = 1, \dots, K$ ). Vale, portanto, a relação

$$\mathbf{\Sigma} = \mathbf{E} \mathbf{D} \mathbf{E}^T \quad 3.10$$

onde  $\mathbf{E} = [\mathbf{e}_1 \mathbf{e}_2 \dots \mathbf{e}_K]$ , é a matriz cujas colunas são os autovetores  $\mathbf{e}_i$  e  $\mathbf{D}$  é a matriz quadrada diagonal de dimensão  $K \times K$  contendo os autovalores  $l_i$  ao longo da diagonal. O vetor  $\mathbf{E}$  é montado segundo a ordem decrescente dos autovalores, ou seja,  $l_i > l_{i+1}$  para  $i = 1, \dots, K$ .

Cada coluna de  $\mathbf{E}$  é uma componente principal. Seus elementos nada mais são do que os coeficientes da combinação linear das variáveis originais que dá origem a uma variável do novo espaço.

Para reduzir a dimensão do espaço de atributos original tomam-se apenas as componentes principais correspondentes aos maiores autovalores. Assim, um vetor de atributos  $x$  definido no espaço de atributos original é representado por um vetor  $y$  definido no novo espaço  $k$ -dimensional, dado por

$$y = \hat{E}^T x_i,$$

onde  $\hat{E} = [e_1 \ e_2 \ \dots \ e_k]$  é a matriz cujas colunas são os  $k$  ( $< K$ ) autovetores correspondentes aos maiores autovalores de  $\Sigma$ .

### 3.1.4.2.

#### Análise Linear de Discriminante

A Análise Linear de Discriminante ou *Linear Discriminant Analysis* (LDA) é um método supervisionado de aprendizado de subespaço. O objetivo é encontrar uma transformação linear definida pela matriz  $W \in \mathbb{R}^{K \times k}$  que mapeie  $x_i$  de um espaço de dimensão  $K$  em um espaço de dimensão  $k < K$ , buscando o melhor compromisso entre elevada dispersão entre classes e baixa dispersão intraclasse. Demonstra-se que a transformação procurada é dada por

$$\hat{W} = \underset{W}{\operatorname{argmax}} \operatorname{tr}((W^T S_w W)^{-1} (W^T S_b W)), \quad 3.11$$

onde  $S_b$  e  $S_w$  são as matrizes de covariância entre classes e intraclasse, respectivamente, definidas como:

$$S_b = \sum_{c=1}^C n_c (\mu_c - \mu)(\mu_c - \mu)^T, \quad S_w = \sum_{c=1}^C \sum_{i \in C_c} (x_i - \mu_c)(x_i - \mu_c)^T \quad 3.12$$

onde  $C_c$  representa a  $c$ -ésima de um total de  $C$  classes do problema,  $\mu_c$  e  $n_c$  são o vetor de médias e o número de padrões de treinamento da  $c$ -ésima classe no espaço de dados de entrada.

Finalmente, seguindo o mesmo procedimento utilizado no PCA, a redução de dimensionalidade se obtém aplicando-se transformação.

$$y = \hat{W}^T x.$$

### 3.1.4.3.

#### Aprendizado de Subespaço através de pares ordenados de atributos

Uma das dificuldades importantes que retardam o desenvolvimento de sistemas mais exatos de estimativa de idade é a escassez de bases de imagens faciais rotuladas, ou seja, em que a idade da pessoa no momento da aquisição da imagem facial é informada.

Nesta seção apresenta-se um método de redução de dimensionalidade proposto em (Chen and Hsu 2013) que visa tirar maior proveito das bases de imagens disponíveis levando em conta a relação mais velho/mais jovem entre pares de imagens faciais com idades distintas.

Na técnica de redução de dimensionalidade em questão utilizam-se para treinamento, além de imagens faciais individualmente, pares ordenados de imagens faciais de um mesmo indivíduo, tal que a primeira imagem é mais recente do que a segunda imagem do par, ou seja, na primeira imagem o indivíduo é mais velho do que na segunda. Tais pares de imagens dão origem a restrições consideradas no cálculo da matriz de transformação. O método que calcula a matriz de transformação busca atender aos seguintes objetivos:

- i) Cada atributo deve satisfazer um subconjunto das restrições.
- ii) Não se admitem atributos redundantes (correlacionados).
- iii) O conjunto de restrições deve ser explorado ao máximo.

Estas condições podem ser representadas matematicamente pelas considerações introduzidas a seguir.

Seja  $\mathbf{X} = [x_1 x_2 \dots x_N]^T$  a matriz já definida anteriormente que contém ao longo de cada linha os vetores de atributos extraídos de um conjunto de  $N$  imagens de treinamento, e  $\mathbf{a} = [a_1, \dots, a_N]$ , as correspondentes idades. Portanto, cada imagem é representada por meio de um vetor coluna de atributos  $\mathbf{x} = \{x_i | i = 1, \dots, K\}$  de dimensão  $K$ .

Seja  $\Omega$  o conjunto de restrições composto de pares de imagens de treinamento ordenadas por idade



$$\Omega = \{(s, t) | a_s > a_t, 1 \leq s, t \leq N\}. \quad 3.13$$

O método cria os vetores  $\mathbf{v}_i$  que compõem a matriz de transformação  $\hat{\mathbf{V}} = [\mathbf{v}_1, \dots, \mathbf{v}_k]$  sequencialmente. Cada novo vetor dá origem a uma variável, cujo valor deve ser coerente com a ordem das idades em cada par de imagens que compõe certo subconjunto de restrições. Em termos formais

$$\mathbf{v}_d^T \mathbf{x}_s - \mathbf{v}_d^T \mathbf{x}_t > 0, \forall (s, t) \in \Omega_d, \quad 3.14$$

onde  $\Omega_d$  representa um subconjunto de  $\Omega$ . Claramente, a relação acima atende o objetivo i). Já a condição ii) é satisfeita ao se impor ortogonalidade entre os vetores  $\mathbf{v}_i$ , ou seja:

$$\mathbf{v}_{d1}^T \mathbf{v}_{d2} = 0, \text{ para } 1 \leq d1, d2 \leq k \text{ e } d1 \neq d2 \quad 3.15$$

Finalmente, a condição iii) visa maximizar o uso dos dados de treinamento disponíveis e orienta a subdivisão de  $\Omega$  em subconjuntos disjuntos, isto é,

$$\Omega_{d1} \cap \Omega_{d2} \neq \emptyset, \quad 1 \leq d1, d2 \leq k \text{ e } d1 \neq d2 \quad 3.16$$

Uma vez representados matematicamente os objetivos que norteiam a criação do subespaço de aprendizado, se apresenta a seguir o algoritmo que resolve esse problema.

A ideia principal subjacente ao algoritmo é minimizar a distância Euclidiana entre imagens com mesma idade sujeito às restrições quanto à ordem de idades definida no conjunto de restrições. Deste modo, a partir de um conjunto de dados rotulados  $\{(x_1, a_1), \dots, (x_N, a_N)\}$  e dos subconjuntos de restrições  $\Omega_1, \dots, \Omega_k$ , pode se formular o problema de estimar o vetor de transformação  $\hat{\mathbf{V}} = [\mathbf{v}_1, \dots, \mathbf{v}_k]$  como um problema de otimização com restrições.

O problema é dividido em subproblemas, onde os vetores de projeção  $\mathbf{v}_1, \dots, \mathbf{v}_k$  são determinados de forma sequencial. Seguindo esta linha, a primeira direção de projeção  $\mathbf{v}_1$  é dada por:

$$\mathbf{v}_1 = \underset{\mathbf{v}}{\operatorname{argmin}} \sum_a \sum_{\substack{p, q \in \\ a_p = a_q = a}} \|\mathbf{v}^T \mathbf{x}_p - \mathbf{v}^T \mathbf{x}_q\|_2^2 \quad 3.17$$

sujeito a  $\mathbf{v}^T \mathbf{x}_s > \mathbf{v}^T \mathbf{x}_t, \forall (s, t) \in \Omega_1.$

Em notação matricial,

$$\mathbf{v}_1 = \underset{\mathbf{v}}{\operatorname{argmin}} \mathbf{v}_1^T \mathbf{X} \mathbf{L} \mathbf{X}^T \mathbf{v}_1 \quad 3.18$$

$$\text{sujeito a} \quad \mathbf{C}_1 \mathbf{X}^T \mathbf{v}_1 > 0,$$

onde  $\mathbf{L}$  é a matriz Laplaciana (vide algoritmo 1), o operador  $>$  indica que todos os elementos do vetor à esquerda são maiores do que o escalar à direita (no caso, zero) e  $\mathbf{C}_1$  é uma matriz de dimensões  $n_1 \times N$ , em que  $n_1$  indica o número de pares ordenados que compõem o subconjunto de restrições  $\Omega_1$ . Mais especificamente, dado o subconjunto  $\Omega_1 = \{(s_1, t_1), \dots, (s_{n_1}, t_{n_1})\}$ , assim se define a matriz de restrições  $\mathbf{C}_1$ :

$$\mathbf{C}_1 = \begin{cases} 1, & \text{para o } (c, s_c) - \text{ésimo elemento} \\ -1, & \text{para o } (c, t_c) - \text{ésimo elemento} \\ 0, & \text{para os demais casos} \end{cases} \quad 3.19$$

Para obter os demais vetores de projeções  $\mathbf{v}_d$  ( $2 \leq d \leq k$ ) adota-se o mesmo procedimento utilizado no cálculo de  $\mathbf{v}_1$ , acrescentando-se a restrição de ortogonalidade com todos os vetores calculados até o passo corrente, ou seja:

$$\mathbf{v}_d = \underset{\mathbf{v}}{\operatorname{argmin}} \mathbf{v}^T \mathbf{X} \mathbf{L} \mathbf{X}^T \mathbf{v} \quad 3.20$$

$$\text{sujeito a} \quad \mathbf{C}_1 \mathbf{X}^T \mathbf{v}_d > 0, \text{ e} \\ \mathbf{V}_{d-1}^T \mathbf{v}_d = \mathbf{0}_{d-1},$$

onde  $\mathbf{V}_{d-1}^T = [\mathbf{v}_1, \dots, \mathbf{v}_{d-1}]$  e  $\mathbf{0}_{k-1} \in \mathbb{R}^{k-1}$  denota o vetor de zeros de dimensão  $d - 1$ .

Note-se que as equações 3.18 e 3.20 descrevem um problema típico de programação quadrática.

A seguir se descreve o algoritmo 1, que gera a matriz de projeção  $\hat{\mathbf{V}}$ , e o algoritmo 2, onde se determina o conjunto de restrições.

Como nas duas técnicas anteriores, a redução de dimensionalidade se obtém aplicando-se transformação

$$\mathbf{y} = \hat{\mathbf{V}}^T \mathbf{x}.$$

**Algoritmo 1:****Entrada:**

- banco de dados rotulado  $\mathbf{X} = [x_1 x_2 \dots x_N]^T$ ,
- os correspondentes rótulos das idades  $\mathbf{a} = [a_1, \dots, a_N]$ ,
- a dimensão  $k$ ,
- conjunto contendo todos os possíveis pares de imagens do mesmo indivíduo  $\Omega$ .

**Passos:**

- 1) Define-se a matriz Laplaciana  $\mathbf{L}$  como  $\mathbf{L} = \mathbf{D} - \mathbf{S}$ , onde
  - o  $(i, j)$ -ésimo elemento de  $\mathbf{S}$  é igual a 1, se  $a_i = a_j$  e zero, caso contrário,
  - $\mathbf{D}$  é a matriz diagonal, tal que seu  $(i, i)$ -ésimo elemento é igual à soma dos elementos ao longo de toda a  $i$ -ésima linha de  $\mathbf{S}$ .
- 2) Retira-se de  $\Omega$  o subconjunto  $\Omega_1$ , utilizando o algoritmo 2; em seguida se atualiza das restrições remanescente, ou seja,  $\Omega = \Omega - \Omega_1$ .
- 3) Computa-se a primeira projeção  $\mathbf{v}_1$  através da equação (3.18) e cria-se  $\mathbf{V} = [\mathbf{v}_1]$ .
- 4) Inicializa-se  $d = 2$ .
- 5) Extrai-se de  $\Omega$  o subconjunto  $\Omega_d$ , e se atualiza  $\Omega = \Omega - \Omega_d$ , novamente utilizando o algoritmo 2.
- 6) Computa-se a  $d$ -ésimo vetor de projeção  $\mathbf{v}_d$  através da equação (3.20).
- 7) Se  $d < k$  e  $\Omega \neq \emptyset$ , atualiza-se  $\mathbf{V} = [\mathbf{V} \ \mathbf{v}_d]$ ,  $d = d + 1$ , e volta ao passo 5.
- 8) Retorna a matriz de projeção  $\mathbf{V}$ .

**Saída:** A matriz de projeção  $\hat{\mathbf{V}} = \mathbf{V}$ .

**Algoritmo 1: cálculo da matriz de projeção  $\hat{\mathbf{V}}$** **Algoritmo 2:****Entrada:**

- conjunto de restrições contendo os pares ordenados de idades

$$\Omega = \{(s_1, t_1), \dots, (s_r, t_r)\}.$$

**Passos:**

- 1) Definem-se os conjuntos  $\Theta = \emptyset$ ,  $E = \{1, 2, \dots, N\}$  e  $Y = \{1, 2, \dots, N\}$  representando estes últimos pela ordem o conjunto das imagens que não ocorrem em nenhum dos pares de  $\Theta$  respectivamente como a mais velha e a mais jovem do par.

- 2) Ordenem-se os  $r$  pares em  $\Omega$ , segundo a ordem crescente das diferenças de idade entre as imagens do par. Deste modo, se obtém

$$\tilde{\Omega} = \{(\tilde{s}_1, \tilde{t}_1), \dots, (\tilde{s}_r, \tilde{t}_r)\}.$$

- 3) Monta-se o conjunto de restrição

**FOR**  $i = 1$  **UNTIL**  $r$

**IF**  $\tilde{s}_i \in E$  **AND**  $\tilde{t}_i \in Y$

$\Theta = \Theta \cup \{(\tilde{s}_i, \tilde{t}_i)\}, E = E - \{\tilde{s}_i\}, Y = Y - \{\tilde{t}_i\}$

**END**

- 4) Retorna o conjunto de restrições  $\Theta$ .

**Saída:** O conjunto  $\Theta$

---

**Algoritmo 2:** Determina o conjunto de pares ordenados.

### 3.1.5. Regressão

O problema da estimativa da idade pode ser abordado basicamente de duas formas: como um problema de classificação ou como um problema de regressão (Chen and Hsu 2013). Neste trabalho optou-se por abordar a estimativa da idade como um problema de regressão, porquanto o envelhecimento é um processo contínuo que se ajusta mais a ideia de uma curva de regressão.

Desta forma, o objetivo nesta seção consiste em determinar uma função  $f(\mathbf{y})$ , que permita estimar a idade  $z$  a partir do vetor  $\mathbf{y} = [y_1 \dots y_k]$  resultante da redução de dimensionalidade aplicada ao vetor  $\mathbf{x}$  de atributos extraídos das faces. A seguir se apresentam os fundamentos teóricos das duas técnicas de regressão investigadas nesta dissertação.

#### 3.1.5.1. Regressão Linear

A análise de regressão é uma ferramenta estatística que permite modelar a relação entre uma variável dependente e uma série de variáveis explicativas ou variáveis independentes. Se  $z$  denota a variável dependente, e  $y_1 \dots y_k$  são as variáveis independentes, cria-se um modelo que prevê o valor de  $z$  com base nos valores das variáveis independentes  $y_1 \dots y_k$ .

Assim, o valor  $z_i$  da variável dependente na observação  $i$  é determinado através da equação linear:

$$z_i = \beta_0 + \beta_1 y_{1i} + \beta_2 y_{2i} + \dots + \beta_k y_{ki} + \varepsilon_i$$

onde os  $\beta_i$ 's são coeficientes e  $\varepsilon_i$  representam o erro, modelado como uma variável aleatória com distribuição de probabilidade normal de média nula.  $\beta_0$  é o valor esperado de  $z$  quando todos valores das  $y_j$  são nulos, e  $\beta_i$  é o coeficiente associado à variável  $y_j$ . Os coeficientes  $\beta_i$ 's e o desvio padrão associado a  $\varepsilon_i$  constituem os parâmetros do modelo, ou, em certo sentido, o próprio modelo.

O valor previsto pelo modelo para  $z_i$  a partir dos valores correspondentes de  $\mathbf{y}$  é, portanto:

$$\hat{z}_i = \beta_0 + \beta_1 y_{1i} + \beta_2 y_{2i} + \dots + \beta_k y_{ki}$$

Os parâmetros do modelo são estimados a partir de um conjunto de treinamento, contendo para cada valor  $\{\mathbf{y}_i\}$  do vetor de variáveis dependentes o correspondente valor observado  $z_i$  da variável dependente. O conjunto de valores dos parâmetros é usualmente estimado de modo a minimizar a soma dos quadrados das diferenças (*SSD*) entre os valores preditos e observados no conjunto de treinamento, formalmente.

$$SSD = \sum_i (\hat{z}_i - z_i)^2$$

### 3.1.5.2. Máquinas de Suporte Vetorial para Regressão.

As Máquinas de Suporte Vetorial ou *Support Vector Machines* (SVM) representam uma técnica de aprendizado supervisionado introduzido por Vladimir Vapnik e coautores em (Vapnik, 1995). Utilizadas inicialmente para problemas de classificação, foram adaptadas posteriormente para tratar problemas de regressão, conseguindo excelentes resultados em muitas aplicações.

As máquinas de suporte vetorial para regressão (SVR) funcionam como se a entrada  $\mathbf{y}$  fosse primeiro mapeada por uma função não linear em um espaço de atributos de dimensão elevada, mesmo infinita, e depois, um modelo linear é gerado nesse espaço de atributos. Matematicamente, o modelo linear no espaço de atributos  $f(\mathbf{y}, \omega)$  é dado por:

$$f(\mathbf{y}, \omega) = \sum_{j=1}^m \omega_j g_j(\mathbf{y}) + b \quad 3.21$$

onde  $g_j(\mathbf{y}), j = 1, \dots, m$  denota um conjunto de funções não lineares, e  $b$  é o termo constante. Sem prejuízo para a generalidade, se admite que os dados tenham média zero (isto pode ser realizado pré-processando os dados), o que permite desconsiderar o termo  $b$ .

Atribuem-se aos parâmetros do modelo os valores que minimizam uma função de perda  $L(z, f(\mathbf{y}, \omega))$  estimados sobre um dado conjunto de treinamento. No caso das SVR, a função de perda mais largamente utilizada foi a proposta por Vapnik, denominada função de perda  $\varepsilon - \text{insensível}$ , assim definida

$$(z, f(\mathbf{y}, \omega)) = \begin{cases} 0, & \text{se } |z - f(\mathbf{y}, \omega)| \leq \varepsilon \\ |z - f(\mathbf{y}, \omega)| - \varepsilon, & \text{caso contrário} \end{cases} \quad 3.22$$

Especificamente, as SVR aplicam a regressão linear no espaço de atributos de alta dimensionalidade, a partir da função de perda  $\varepsilon - \text{insensível}$ , buscando ao mesmo tempo um modelo de baixa complexidade medida pelo quadrado do módulo do vetor de coeficientes  $\|\omega\|^2$ .

As chamadas variáveis de folga (não negativas)  $\xi_i, \xi_i^* i = 1, \dots, n$ , permitem medir o desvio das amostras de treinamento para além da zona  $\varepsilon - \text{insensível}$ . (ver Figura 3.10). Deste modo, a regressão com as SVR é formulada minimizando a seguinte expressão:

$$\min\left(\frac{1}{2}\|\omega\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*)\right) \quad 3.23$$

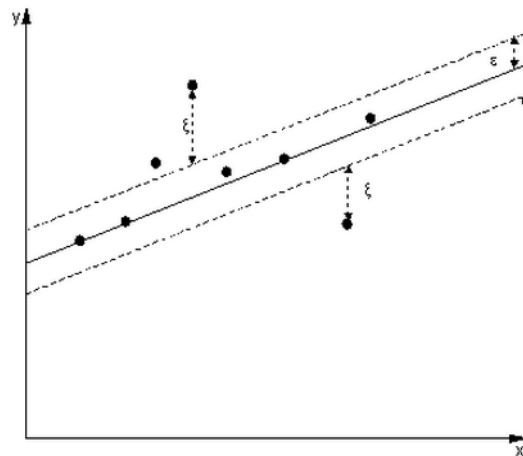
$$\text{sujeito a} \quad \begin{cases} z_i - f(\mathbf{y}_i, \omega) \leq \varepsilon + \xi_i^* \\ f(\mathbf{y}_i, \omega) - y_i \leq \varepsilon + \xi_i \\ \xi_i, \xi_i^* \geq 0, i = 1, \dots, n \end{cases}$$

Este problema de optimização tem uma formulação dual, cuja solução é dada por:

$$f(\mathbf{y}) = \sum_{i=1}^{\eta_{SV}} (\alpha_i - \alpha_i^*) \mathbf{K}(\mathbf{y}_i, \mathbf{y}), \quad \text{s.t. } 0 \leq \alpha_i^* \leq C, 0 \leq \alpha_i \leq C \quad 3.24$$

onde  $\eta_{SV}$  denota o número de vetores de suporte e  $\mathbf{K}$  uma função Kernel adotada:

$$K(\mathbf{y}, \mathbf{y}_i) = \sum_{j=1}^m g_j(\mathbf{y}) g_j(\mathbf{y}_i) \quad 3.25$$



**Figura 3.10** Regressão linear unidimensional com a banda  $\varepsilon$  – *insensível*.

Existe uma farta literatura sobre SVR onde os detalhes conceituais de operacionais são descritos (Hsu et al. 2003) (Smola and Schölkopf 2004).

## 4 Análise Experimental

Neste capítulo, descreve-se o procedimento experimental realizado nesta dissertação. Primeiro, se detalham as características dos bancos de dados usados no desenvolvimento desta pesquisa; segundo, se explicam as métricas empregadas para avaliar o desempenho do sistema de estimação de idade; terceiro, se apresenta a metodologia experimental utilizada para estimar a idade a partir de imagens estáticas; e finalmente, se expõem, analisam e interpretam os resultados obtidos.

### 4.1. Descrição dos Bancos de Dados

Os bancos de dados usados nesta pesquisa foram o FG-NET *Aging Database*<sup>2</sup> e o MORPH<sup>3</sup> (Ricanek and Tesafaye 2006). Ambos os bancos são disponíveis publicamente. Suas características gerais são apresentadas a seguir.

#### 4.1.1. FG-NET Aging Database

O banco de dados FG-NET está composto por 1.002 imagens faciais pertencentes a 82 indivíduos com idades compreendidas entre 0 a 69 anos, sendo que de cada indivíduo se provêm de 6 a 18 imagens faciais associadas às diversas idades. Na figura 4.1 se apresentam as imagens faciais, cronologicamente ordenadas, de um indivíduo pertencente ao banco de dados.

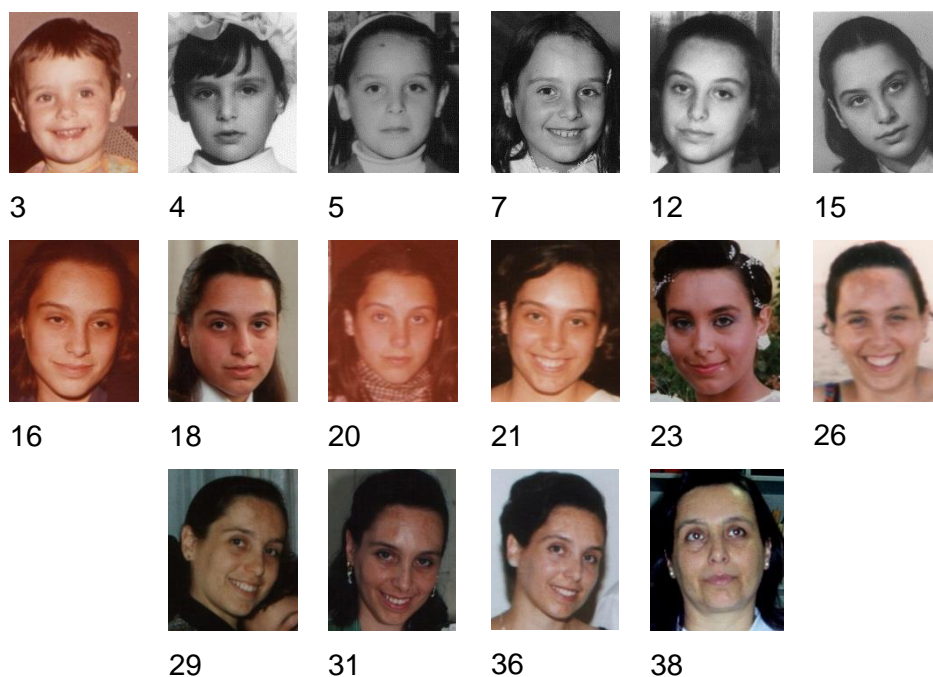
Por outro lado, na Figura 4.2 e Tabela 4.1, descreve-se como estão distribuídas as idades no banco de dados. Na Figura 4.2 se mostra o número de ocorrências por idade, e na Tabela 4.1, a porcentagem de ocorrências por década. Observa-se que o FG-NET não é um conjunto de dados uniformemente distribuído, pois 70% das idades estão concentradas na faixa de 0 a 20 anos.

---

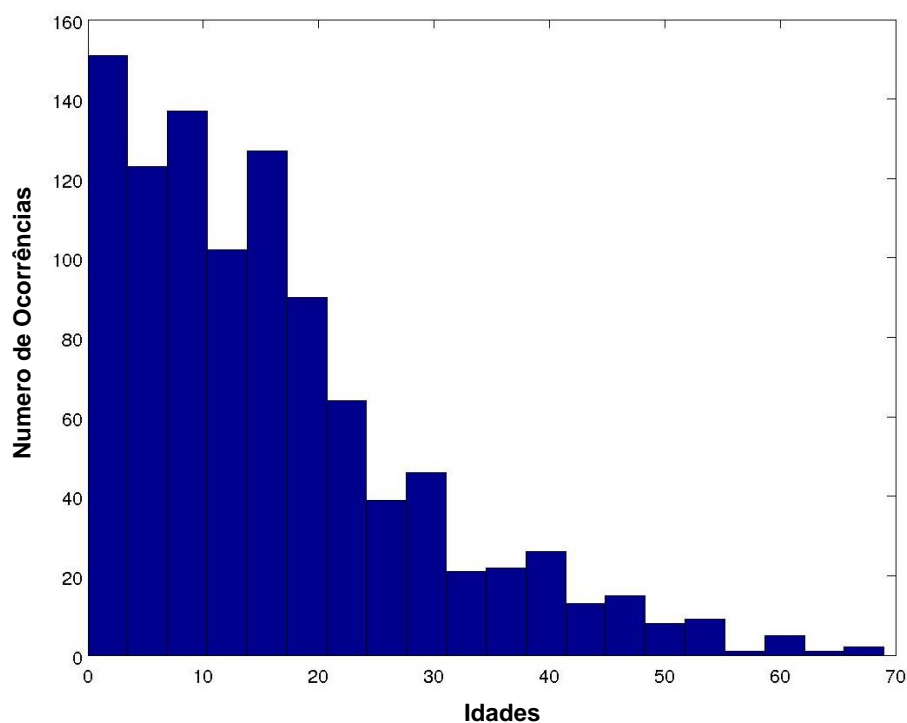
<sup>2</sup> Disponível em: <http://www.fgnet.rsunit.com>.

<sup>3</sup> Disponível em: <https://ebill.uncw.edu>





**Figura 4.1.** Exemplo do Envelhecimento Facial dum Indivíduo no Banco de Dados FG-NET. Obtido de (Lanitis et al. 2002). A idade correspondente se indica debaixo de cada imagem.



**Figura 4.2.** Histograma das Idades no Banco de Dados FG-NET

É importante ressaltar que, além da idade, as imagens faciais apresentam outros tipos de variações, como mudanças significativas na pose, iluminação e expressões faciais.

FAIXA DAS IDADES	FG-NET(%)	MORPH(%)
0 - 9	37,03	0,00
10 – 19	33,83	8,94
20 – 29	14,37	26,94
30 – 39	7,88	32,16
40 – 49	4,59	24,58
50 – 59	1,50	7,37
60 – 69	0,80	0,82
70 – 79	0,00	0,09

**Tabela 4.1.** Distribuição percentual das Idades por década nos bancos de dados FG-NET e MORPH. Adquirido em (Geng, Zhou and Smith-Miles 2007)

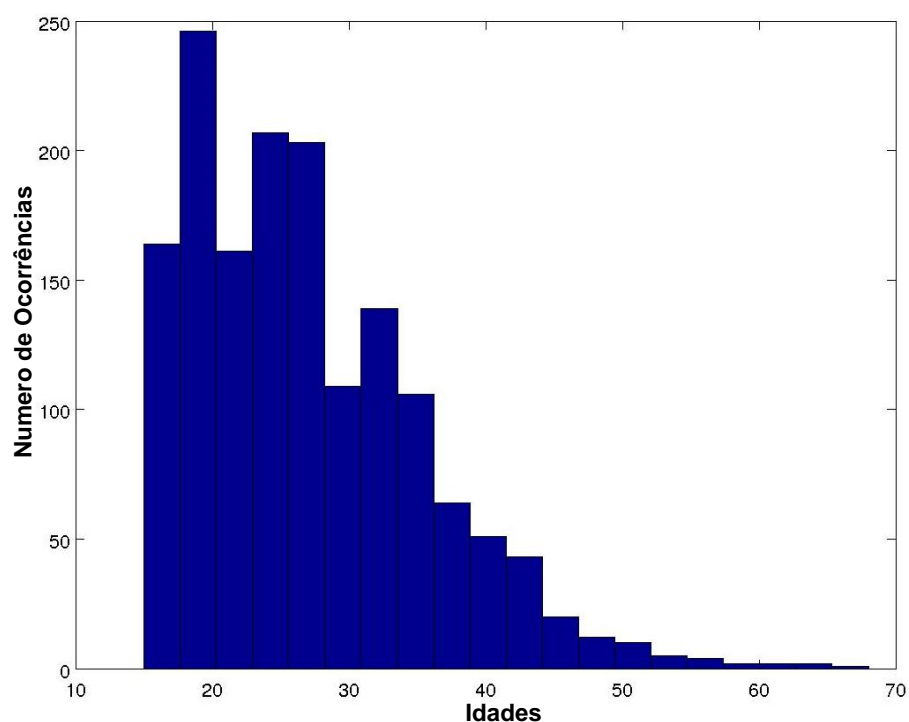
#### 4.1.2. MORPH Database

O banco de dados MORPH está dividido em dois álbuns. O *álbum 1* contém 1.690 imagens faciais de 628 indivíduos com idades compreendidas entre 15 e 68 anos, e o *álbum 2* tem 55.608 imagens faciais de mais de 13.000 indivíduos entre 16 e 77 anos. Adicionalmente, as imagens do segundo álbum estão organizadas em seis raças: Africanos, Brancos, Hispânicos, Indianos, Asiáticos e Outros. Destaque-se que 77% das imagens correspondem a pessoas de raça Africana, 19% a pessoas de raça Branca e o restante 4%, a pessoas das raças Hispânicas, Asiática, Indiana e Outras (Geng, Yin and Zhou 2013).

Diferentemente do FG-NET, o número médio de imagens por pessoa no conjunto de dados MORPH é de apenas quatro. Além disso, são comuns pessoas com uma única imagem no banco. Na Figura 4.3 exibem-se indivíduos que compõem este banco de dados, onde se pode observar o anteriormente exposto.



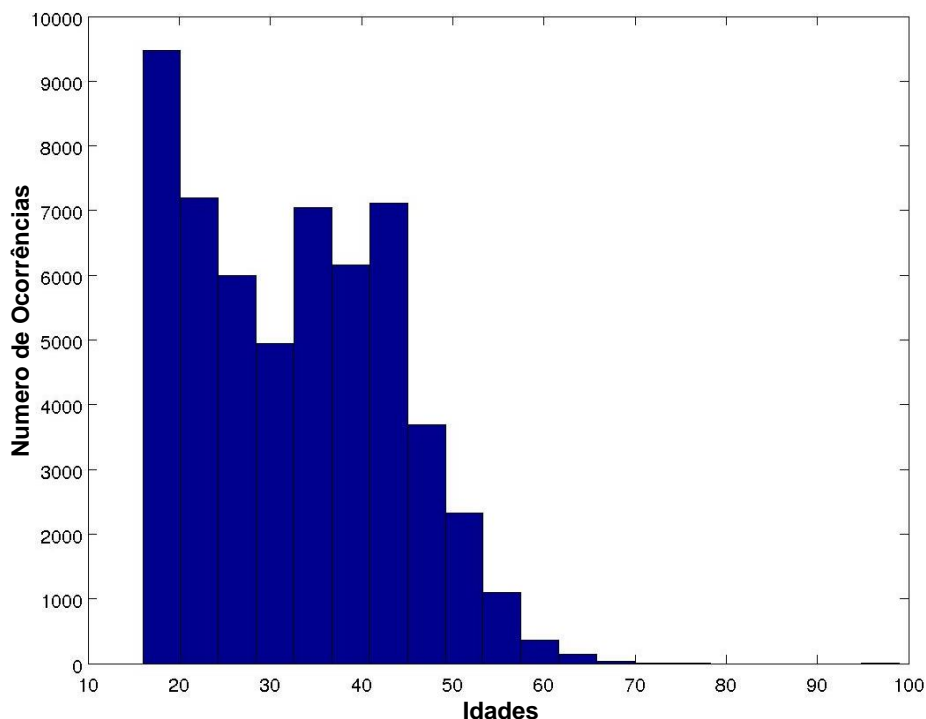
**Figura 4.3.** Exemplos de Imagens do banco de dados MORPH ; a idade correspondente se indica debaixo de cada imagem.



**Figura 4.4.** Histograma das Idades do Álbum 1 do Banco de Dados MORPH

Por outro lado, a Figura 4.4 e a Figura 4.5 mostram os histogramas que descrevem a distribuição das idades em ambos os álbuns. Na Tabela 4.1, se apresenta a mesma informação, mas em porcentagem de ocorrências por

década. Observa-se que, semelhantemente ao FG-NET, o MORPH é um banco de dados em que as imagens não estão distribuídas uniformemente entre as idades; 70% dos dados estão na faixa de 20 a 40 anos de idade.



**Figura 4.5.** Histograma das Idades do Álbum 2 do Banco de Dados MORPH

## 4.2. Métricas

As métricas comumente usadas na literatura para avaliar o desempenho dos algoritmos que estimam a idade a partir de imagens faciais, e empregadas de igual forma neste trabalho, são o erro médio absoluto (MAE), o erro médio absoluto por década (MAE/D), o erro médio absoluto por idade (MAE/y), e o Erro acumulado (CS) (Geng et al. 2006), (Lanitis et al. 2004). Adicionalmente, nesta dissertação apresentam-se os intervalos de confiança (IC) para cada idade estimada por cada algoritmo analisado.

A seguir se descreve cada uma destas métricas.

### 4.2.1. Erro Médio Absoluto (MAE)

O erro médio absoluto (MAE) é definido como:

$$MAE = \frac{1}{N} \sum_{i=1}^N |a_i - \hat{a}_i| \quad 4.1$$

onde,  $\hat{a}_i$  representa a idade estimada,  $a_i$  a idade real da  $i$ -ésima amostra, e  $N$  o número total de amostras.

Observa-se que o  $MAE$  mede o desempenho global do estimador de idade. Por conseguinte, através do  $MAE$  não é possível inferir como os erros se distribuem entre as idades. Para completar a informação expressa pelo  $MAE$ , é usual calcular o erro médio absoluto por década ( $MAE/D$ ) e o erro médio absoluto por idade ( $MAE/y$ ), assim definido:

$$MAE/D(d) = \frac{1}{N_d} \sum_{i=1}^N f(d) * |a_i - \hat{a}_i|, 10(d-1) \leq a_i < 10d \quad (4.2).$$

onde,  $\hat{a}_i$  representa a idade estimada,  $a_i$  a idade real da  $i$ -ésima amostra,  $N$  o número total de amostras,  $N_d$  o número de amostras na década  $d$ , e  $f$  é uma função assim definida:  $f(d) = 1$ , se  $10(d-1) \leq a_i < 10d$  e  $f(d) = 0$ , para os outros casos, sendo que  $d = 1, 2, \dots$

$$MAE/A(a) = \frac{1}{N_a} \sum_{i=1}^N g(a) * |a_i - \hat{a}_i| \quad 0 \leq a < N \quad (4.3).$$

onde,  $\hat{a}_i$  representa a idade estimada,  $a_i$  a idade real da  $i$ -ésima amostra,  $N$ , o número total de amostras,  $N_a$  o número de amostras da idade  $a$ , e  $g$  é uma função assim definida:  $g(a) = 1$ , se  $a = a_i$  e  $g(a) = 0$ , para os outros casos.

#### 4.2.2. Escore Acumulado (CS)

O escore acumulado (CS) é definido como,

$$CS(e) = \frac{1}{N} \sum_{i=1}^N h(|a_i - \hat{a}_i| - e) \quad (4.4)$$

onde,  $h$  é uma função assim definida:  $h(t) = 1$ , se  $t \leq 0$  e  $h(e) = 0$ , para os outros casos, e  $e$  denota o nível da tolerância do erro.

O escore acumulado representa a acurácia do estimador de idade a diferentes níveis de tolerância de erro (Chen and Hsu 2013). Nos experimentos realizados nesta pesquisa, escolheram-se níveis de erros entre zero e dez, já que um erro estimado maior do que dez é usualmente considerado inaceitável.

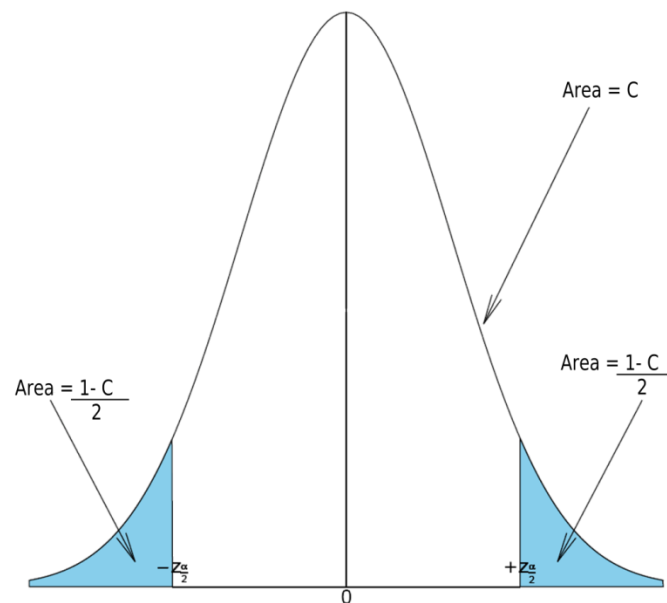
#### 4.2.3. Intervalos de Confiança (CI)

Um intervalo de confiança é uma faixa estimada de valores que provavelmente inclui o valor real de um parâmetro desconhecido relativo a uma população (Easton and McColl 1997). Esta faixa é estimada a partir de um conjunto de dados amostrais.

A notação comum para o parâmetro em questão é  $\theta$ . Sua média amostral é denotada pelo símbolo  $\bar{\theta}$ .

O *nível de confiança*  $C$  de um intervalo de confiança determina a probabilidade de que o valor real do parâmetro  $\theta$  esteja contido no intervalo de confiança. Normalmente, escolhem-se níveis de confiança  $C$  iguais a 0,90, 0,95, e 0,99. O nível de confiança define a área sob a curva da função de densidade de probabilidade do parâmetro  $\theta$  (*Probability Density Function - pdf*) compreendida pelo intervalo de confiança. Por exemplo, um nível de confiança igual a 95%, determina o intervalo de valores de  $\theta$  que compreende 95% da área abaixo da *pdf*, ou seja, a probabilidade de se observar valores de medida dentro desta área é igual a 95%.

Se o número de amostras for elevado, pelo teorema do limite central  $\theta$  tem uma distribuição que se aproxima a uma Gaussiana (Montgomery 2013). Assim, dado que a distribuição normal é uma curva simétrica em relação à média  $\bar{\theta}$ , a metade da área desses 95% está à esquerda da média, e a outra metade, à direita da média, como é apresentado na Figura 4.6, onde as áreas, representadas pela cor azul, nas caudas da normal, representam a probabilidade de  $\theta$  assumir valores fora do intervalo esperado à direita ou à esquerda, dada por  $(1-C)/2$ . Por exemplo, para um intervalo de confiança de 95%, a área em cada cauda é igual a  $0,05/2 = 0,025$ .



**Figura 4.6.** Função de Densidade de Probabilidade Normal

Imagem adquirida em (Christou and Dinov 2011); indicam-se as áreas de interesse para a seleção dos níveis de confiança= $C$ .

Neste trabalho calculam-se os intervalos de confiança para cada idade estimada pelos algoritmos analisados.

### 4.3. Procedimento Experimental Geral

Esta seção detalha o procedimento experimental. O alvo é comparar o desempenho do sistema encarregado de estimar a idade utilizando diversas técnicas de redução de dimensionalidade, extração de atributos, e regressão, respectivamente. De igual forma, se estuda a influência da raça e do gênero das pessoas na acurácia dos algoritmos.

Nos experimentos realizados, se segue o mesmo protocolo experimental utilizado na maioria dos trabalhos relacionados à estimativa da idade (Chen and Hsu 2013; Gehrig et al. 2011; Geng, Yin and Zhou 2013; Geng, Zhou and Smith-Miles 2007; Geng, Zhou, Zhang, Li and Dai 2006; Lanitis, Draganova and Christodoulou 2004).

Para o banco de dados FG-NET, se adota o protocolo *deixar uma pessoa de fora* (Leave One Person Out - LOPO), e para o MORPH, a validação cruzada. No LOPO, as imagens de uma pessoa são selecionadas como conjunto de teste,

e todas as imagens restantes do banco, como o conjunto de treinamento. Este procedimento é repetido 82 vezes, o que equivale ao total de pessoas no banco, de modo que cada indivíduo é usado uma vez para teste. Finalmente, se calculam os resultados finais, em termos das métricas selecionadas como baseados em todas as estimativas realizadas.

No caso do banco MORPH, se utiliza a validação cruzada de 10 partições nos experimentos, porque este banco contém comparativamente um número elevado de padrões, mais de 13.000 pessoas, e no melhor dos casos, 4 imagens por pessoa. Uma estratégia LOPO implicaria em mais de 13.000 rodadas, cada uma envolvendo um conjunto de treinamento de mais de 55.000 imagens e apenas 4 imagens para teste, o que implicaria num tempo de processamento muito alto.

#### **4.3.1. Experimentos**

A seguir se descrevem os experimentos realizados e os resultados coletados em cada um deles. Para cada experimento apresentam-se os objetivos estabelecidos, e detalhes do procedimento experimental.

Os experimentos foram realizados a partir de um protótipo criado em MATLAB. A implementação para MATLAB de AAM utilizada foi obtida de (Musti et al. 2014), a do CLM de (Asthana, Zafeiriou, Cheng and Pantic 2013) e a do SVR de (Chang and Lin 2011). A versão de MATLAB utilizada foi a R2013a (8.10.604) de 64-bits, e executando-se sob Linux 3.13, num computador Intel(R) Core(TM) i7-3930k CPU @ 12x3,20Ghz e com 32 GB de memória RAM.

##### **4.3.1.1. Experimento 1.**

**Objetivos:** Este experimento teve os seguintes objetivos:

- a) Determinar a influência da acurácia da localização dos pontos fiduciais sobre a acurácia do sistema de estimação de idade.
- b) Comparar os métodos de redução de dimensionalidade



**Protocolo Experimental:** Este experimento baseou-se no banco de dados FG-NET porque fornece de antemão os pontos de referência detectados manualmente, considerados como exatos nestes experimentos.

As experiências realizadas seguem a estratégia experimental LOPO, e adicionalmente, as técnicas de redução de dimensionalidade, descritas no capítulo anterior, são empregadas tanto para pontos fiduciais detectados manualmente, como automaticamente.

Por último, dado que o alvo de este experimento é determinar de forma isolada a influencia dos pontos fidúcias na acurácia do sistema de estimação de idade, somente um algoritmo de regressão foi empregado nas experimentações: o SVR. Este método é configurado com os parâmetros adequados mediante validação cruzada de 5 partições, em conjunto com o algoritmo de busca Grid (Wenwen et. 2014).

**Resultados:** Na Tabela 4.2 e Tabela 4.3 se resumem, nesta ordem, os resultados do Erro Médio Absoluto (MAE) e o Erro Médio Absoluto por Década (MAE/D) de todos os testes realizados neste experimento.

Pontos F. Marcados			Pontos F. Estimados		
PCA	LDA	SL	PCA	LDA	SL
5,61	5,56	5,60	6,15	6,29	6,80

**Tabela 4.2** Erro Médio Absoluto (MAE) no Banco de Dados FG-NET; comparam-se os resultados do MAE para cada método de redução de dimensionalidade e para cada estratégia de marcado de pontos de interesse.

Década	Pontos F. Marcados			Pontos F. Estimados		
	PCA	LDA	SL	PCA	LDA	SL
00 – 09	3,70	4,00	3,58	4,23	4,45	4,87
10 – 19	3,80	4,00	3,71	3,75	3,92	3,25
20 – 29	4,75	4,37	4,84	4,61	5,00	6,36
30 – 39	9,39	9,21	10,72	11,30	11,16	14,98
40 - 49	15,21	14,90	17,30	19,50	18,70	22,77
50 – 59	26,93	24,93	27,26	29,53	29,33	32,06
60 - 69	38,75	35,70	34,00	39,28	35,85	36,42

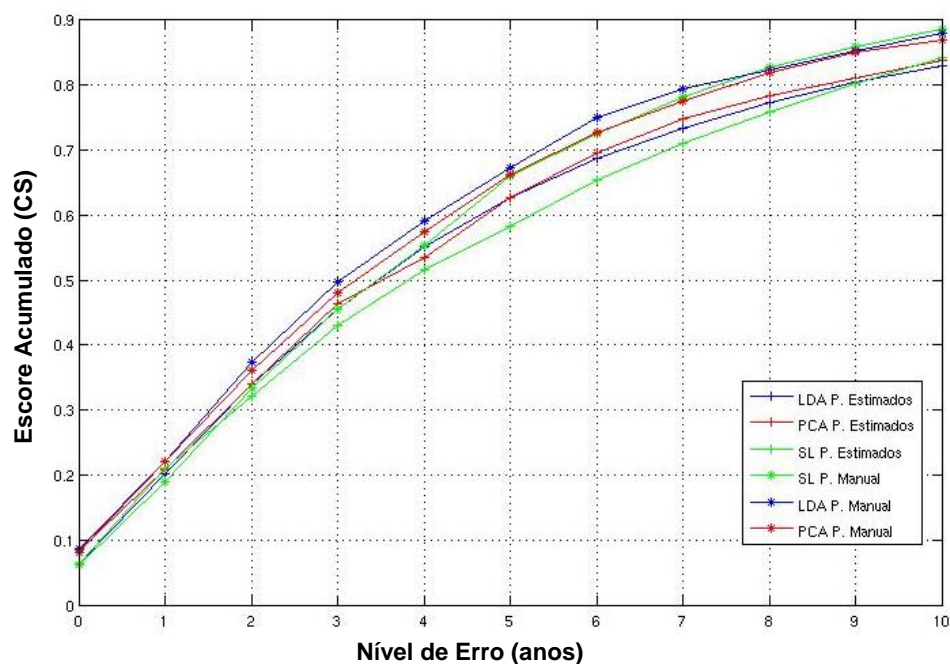
**Tabela 4.3** Erro Médio Absoluto por Década (MAE/D) da estimativa de idade no banco de dados FG-NET.

Observa-se que em todos os casos, o erro médio aumenta em aproximadamente 10% quando se utilizam pontos fiduciais determinados de forma automática.

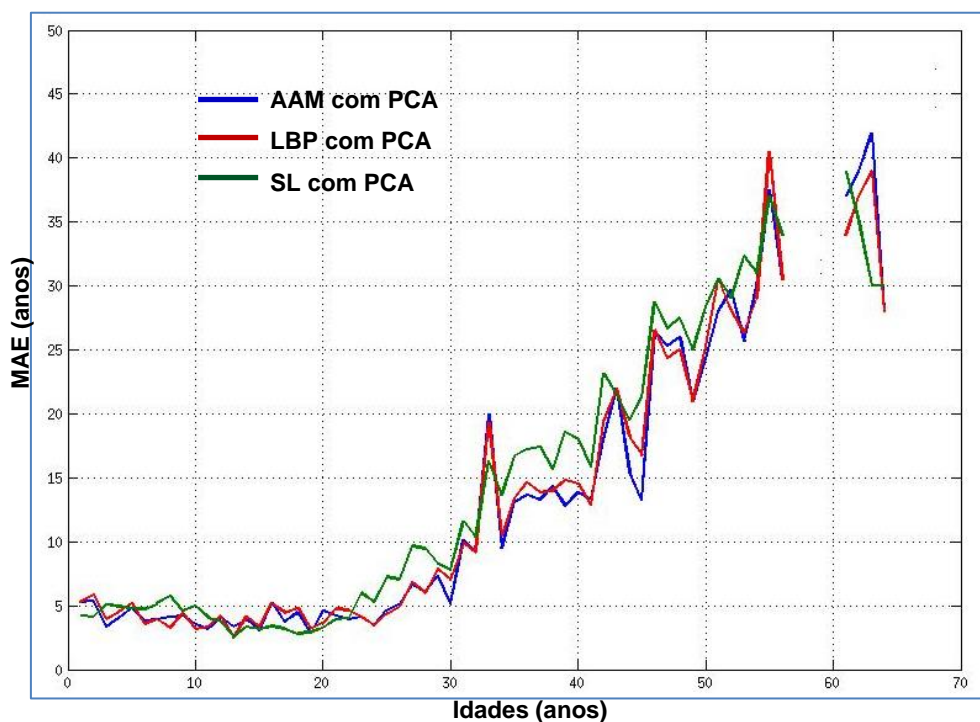
Na Figura 4.7 vêm-se as curvas de acurácia dos métodos empregados. Especificamente, cada curva indica a proporção de padrões de teste cujo erro na estimativa foi inferior a certo nível do erro. Idealmente, a curva seria uma reta horizontal correspondente com a acurácia do sistema igual a 1. Desta forma, pode-se observar que o melhor desempenho é alcançado pelo LDA junto com pontos fiduciais marcados manualmente, e no outro extremo, o pior desempenho foi registrado para o *Subspace Learning* com pontos estimados automaticamente. Nota-se que em todos os casos, as curvas relacionadas às dos pontos fiduciais marcados manualmente, superam em desempenho às dos pontos fiduciais estimados. Este resultado indica que a acurácia do sistema de estimação de idade efetivamente é influenciada substancialmente pela localização dos pontos fiduciais.

Na sequência, analisam-se os resultados relacionados ao segundo objetivo proposto para este experimento, apresentando-se unicamente resultados obtidos a partir de pontos fiduciais estimados automaticamente, uma vez que são mais representativos de condições operacionais de sistemas de estimativa de idade.

A **Erro! Fonte de referência não encontrada.** mostra o Erro Médio absoluto por idade (MAE/Y) para os três métodos de redução de dimensionalidade analisados. Note-se que nos três casos considerados, o erro médio aumenta de forma considerável com a idade. Este comportamento está relacionado à forma como estão distribuídas as idades no banco de dados FG-NET. De fato, a maioria das imagens do banco de dados estão entre 0 e 30 anos, e o número de imagens por idade diminui conforme aumenta a idade (ver Figura 4.2). Em consequência, os algoritmos perdem exatidão para as idades elevadas, uma vez que se têm poucas amostras relacionadas a estas idades em cada conjunto de treinamento.

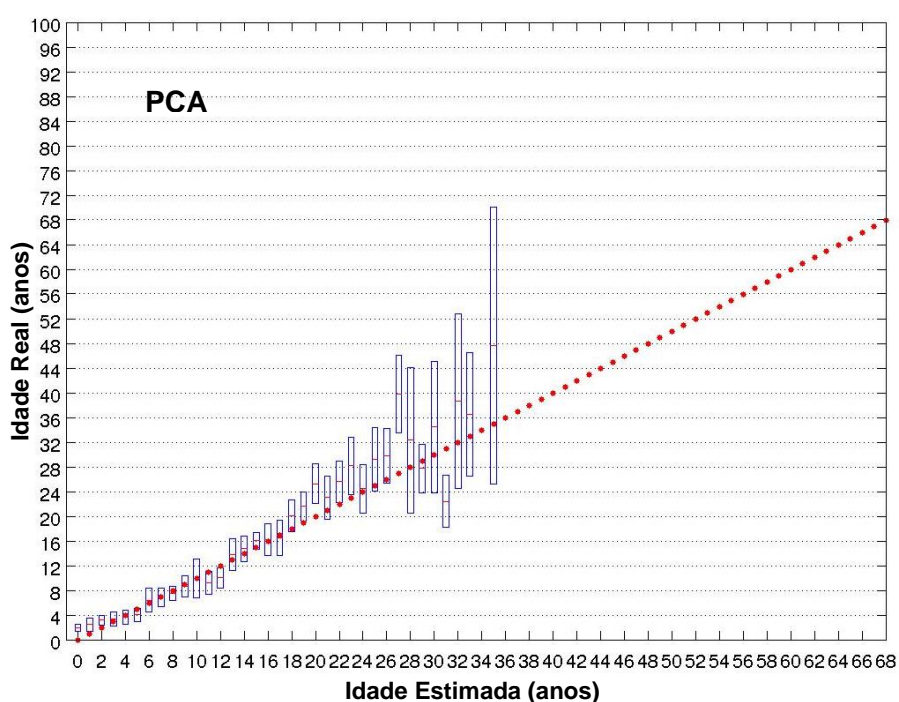


**Figura 4.7.** Erro Acumulado (CS) da estimativa da idade no banco de dados FG-NET. Compara-se a acurácia do estimador para pontos fiduciais estimados automaticamente e marcados de forma manual. Também se comparam os três métodos de redução de dimensionalidade.



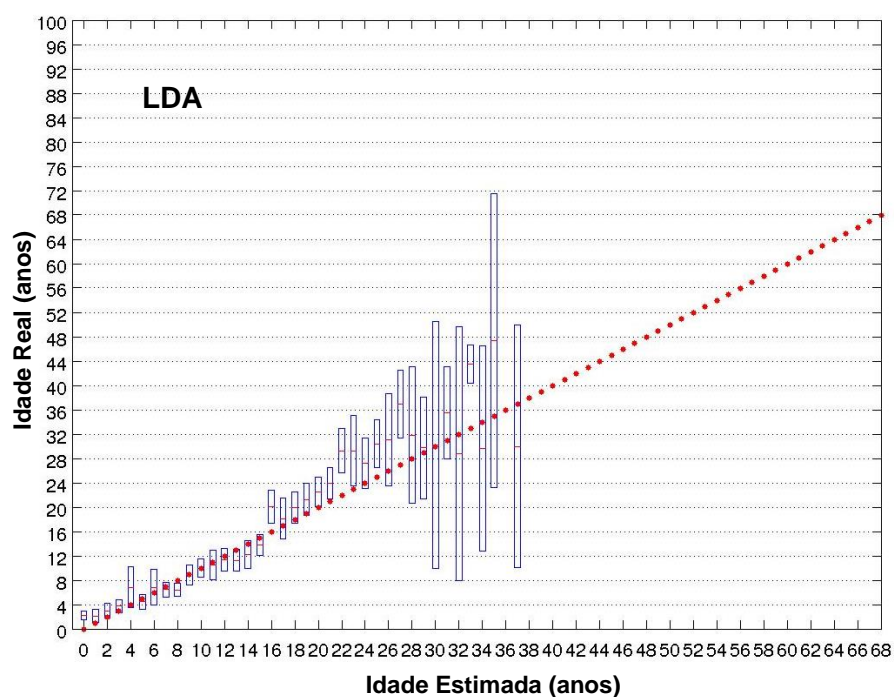
**Figura 4.8** Erro Médio Absoluto por Idade (MAE/y) da estimativa da idade no banco de dados FG-NET, usando PCA, LDA e SL, respectivamente.

As Figura 4.9 a Figura 4.11, mostram para cada idade estimada (eixo horizontal) o intervalo de confiança (eixo vertical). Em todas as figuras, se apresentam os resultados para um nível de confiança de 90%. Note-se que a linha pontilhada vermelha nos gráficos de intervalos de confiança, representa o que seria o desempenho ideal do algoritmo, ou seja, os intervalos de confiança deveriam idealmente reduzir-se cada um a um ponto sobre a linha vermelha. Observe que o algoritmo não é capaz de estimar idades depois dos 38 anos aproximadamente. Este comportamento está relacionado ao fato de ter poucas imagens a partir dessa faixa de idades, de modo que, o estimador não aprende esse conjunto de idades.

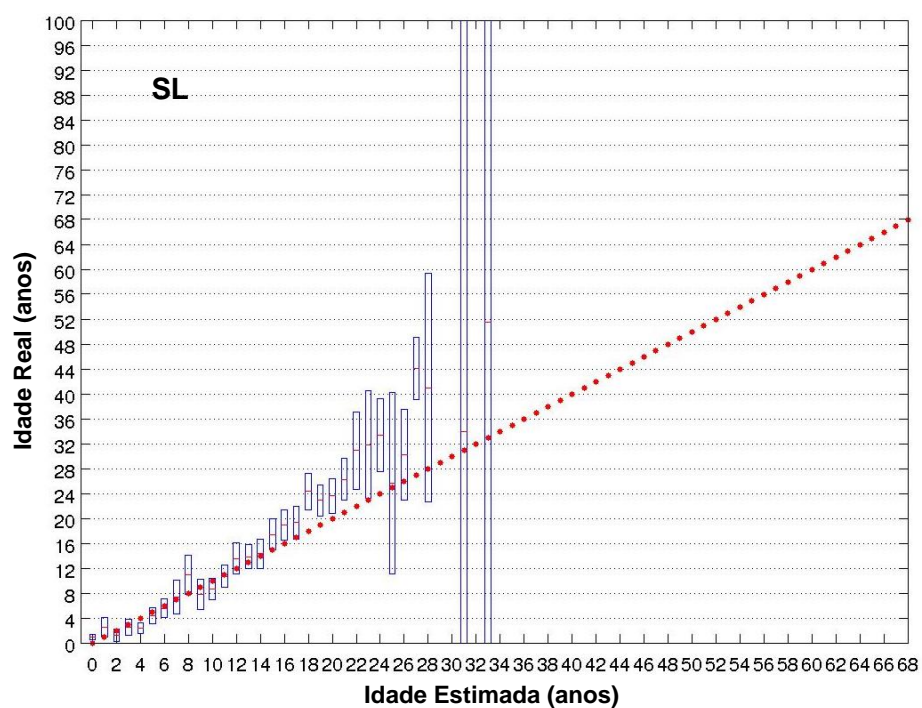


**Figura 4.9.** Intervalos de Confiança para PCA com nível de confiança 90%.

Finalmente, analisando-se unicamente os resultados do MAE/y da **Erro! fonte de referência não encontrada.**, não é fácil discernir quais dos métodos de redução de dimensionalidade avaliados apresenta o melhor desempenho. No entanto, se se comparam os intervalos de confiança nas Figura 4.9, Figura 4.10 e Figura 4.11, se pode observar que o PCA apresenta o melhor desempenho; os intervalos de confiança são mais estreitos e mais centrados nas idades verdadeiras.



**Figura 4.10.** Intervalos de Confiança para LDA com nível de confiança 90%



**Figura 4.11** Resultado dos Intervalos de Confiança para SL com nível de confiança 90%.

#### 4.3.1.2. Experimento 2.

**Objetivos:** Este experimento tem os seguintes objetivos:

- a) Avaliar a acurácia da estimativa associada a dois métodos de extração de textura: o método padrão do AAM e o LBP.
- b) Avaliar a acurácia do método para diferentes raças e gêneros.

**Protocolo Experimental:** Neste experimento foram usados ambos os bancos de dados, e os pontos fiduciais utilizados foram os estimados de forma automática. Também em decorrência das conclusões extraídas do experimento anterior utilizou-se nestes experimentos o PCA como método de redução de dimensionalidade. SVR foi utilizado como método de regressão. Além disso, os experimentos foram realizados separando as imagens da base de dados por raça e gênero com o intuito de avaliar a influência destes fatores nos resultados finais.

**Resultados:** Os resultados são organizados a seguir em dois grupos em função do banco de dados utilizado nos experimentos.

- **Banco de dados MORPH**

Nas Tabela 4.4 e Tabela 4.5 se apresentam os valores do MAE e MAE/D, respectivamente, obtidos a partir do banco de dados MORPH, álbum 2. Mais adiante, nas Tabela 4.6, Tabela 4.7 e Tabela 4.8, se apresentam, nesta ordem, o MAE/D da raça Africana, Branca e Hispana. Os resultados do álbum 1, foram similares aos do álbum 2, por isso são apresentados os resultados do álbum 2.

Constata-se das tabelas do MAE e MAE/D, que o método de codificação de textura do AAM, apresenta melhor desempenho do que a técnica que codifica textura utilizando o LBP. Na maioria dos casos, o erro médio do modelo original do AAM é cerca de 10% menor do que do LBP.

<b>Método para Extrair Atributos</b>	<b>Raça</b>	<b>PCA</b>
AAM	Todas	4,57
AAM + LBP	Todas	5,31
AAM	Africanos	4,70
AAM + LBP	Africanos	5,60
AAM	Branços	4,37
AAM + LBP	Branços	5,70
AAM	Hispânicos	3,93
AAM + LBP	Hispânicos	4,87

**Tabela 4.4.** Erro Médio Absoluto (MAE) para estimativa da idade no banco de dados MORPH, Album2. Comparam-se os métodos de extração de atributos de textura usando o PCA para reduzir dimensionalidade com todas as amostras do banco de dados.

<b>Década</b>	<b>AAM</b>	<b>AAM + LBP</b>
10 – 19	5,48	6,92
20 – 29	4,14	4,98
30 – 39	3,78	4,08
40 – 49	4,95	5,51
50 – 59	7,42	5,50
60 – 69	12,94	14,79
70 – 79	22,07	29,53

**Tabela 4.5** Erro Médio Absoluto por Década (MAE/D) para estimativa da idade de todas as amostras do banco de dados MORPH, Album2. Comparam-se os métodos de extração de atributos de textura usando o PCA para reduzir dimensionalidade.

É interessante notar que o erro para a raça hispânica é o menor entre todas as raças. Isto está relacionado ao fato de que há no banco de dados um número reduzido de imagens de hispânicos compreendidas no intervalo de 60 a 70 anos, e nenhuma acima de 70 anos, onde os erros tendem a ser maiores.

Década	AAM	AAM + LBP
10 – 19	5,60	7,45
20 – 29	4,13	4,94
30 – 39	3,73	4,11
40 – 49	5,11	5,93
50 – 59	7,72	9,32
60 – 69	12,55	14,03
70 – 79	20,42	24,14

**Tabela 4.6** Erro Médio Absoluto por Década (MAE/D) para Estimativa da Idade da Raça Africana no Banco de Dados MORPH, Album2. Comparam-se os métodos de extração de atributos de textura usando o PCA para reduzir dimensionalidade.

Década	AAM	AAM + LBP
10 – 19	5,60	8,44
20 – 29	4,09	5,83
30 – 39	3,56	3,77
40 - 49	4,17	5,37
50 – 59	6,45	8,99
60 - 69	12,75	16,60
70 - 79	24,12	33,25

**Tabela 4.7** Erro Médio Absoluto por Década (MAE/D) para estimativa da idade da raça Branca no banco de dados MORPH, Album2. Comparam-se os métodos de extração de atributos de textura usando o PCA para reduzir dimensionalidade.

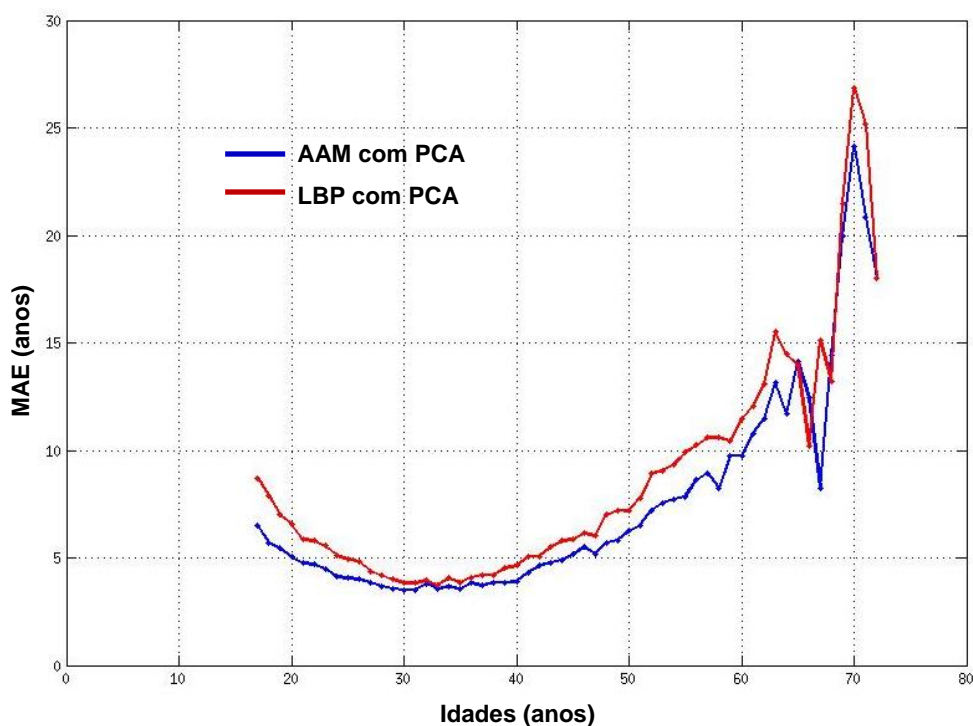
Similarmente aos resultados obtidos no banco FG-NET, os resultados do MAE/D e MAE/y obtidos com o banco de dados MORPH seguem a tendência de apresentar um erro baixo nas idades para as quais há um maior número de imagens de treinamento. Da mesma forma, o erro tende a aumentar quando o número de imagens de treinamento por idade diminui.

Na **Erro! Fonte de referência não encontrada.**, pode-se comparar o AE/y resultante do uso de AAM e AAM + LBP, respectivamente. Observa-se que o erro por idade é maior quando o AAM + LBP é usado como método para extrair informação de textura.



Década	AAM	AAM + LBP
10 – 19	4,13	5,46
20 – 29	2,94	2,88
30 – 39	4,04	5,23
40 - 49	7,11	11,17
50 – 59	11,06	17,83
60 - 69	24,75	26,00

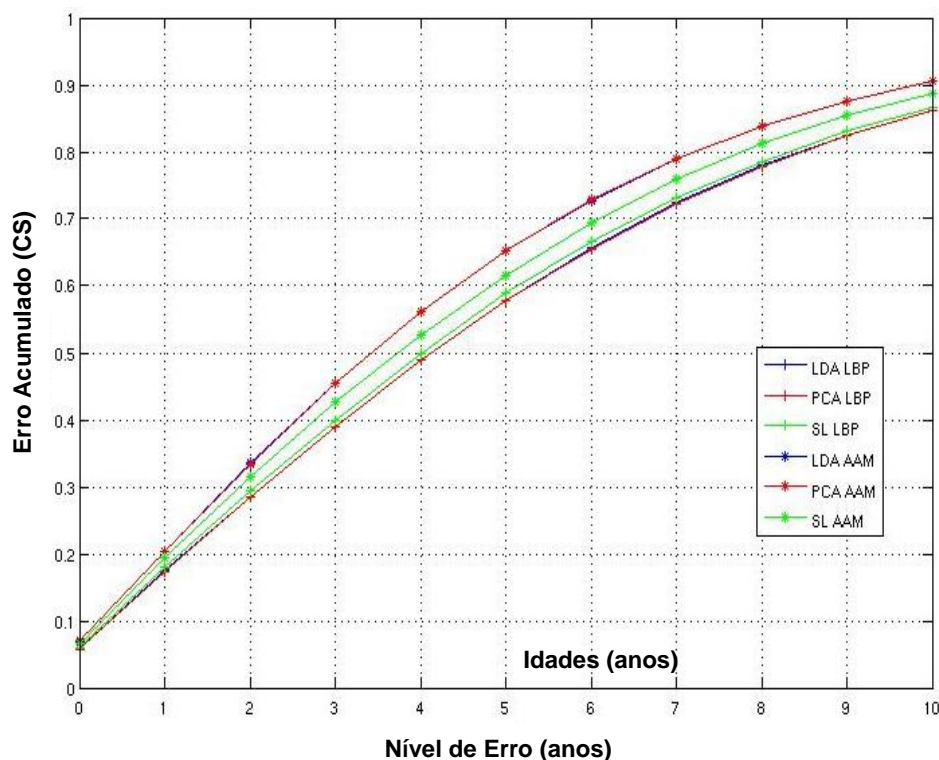
**Tabela 4.8** Erro Médio Absoluto por Década (MAE/D) para estimativa da idade da raça Hispânica no banco de dados MORPH, Album2. Comparam-se os métodos de extração de atributos de textura usando o PCA para reduzir dimensionalidade.



**Figura 4.12** Erro Médio Absoluto por Idade (MAE/y) da estimativa da idade do álbum 2, no banco de dados MORPH. O PCA é usado como método para reduzir dimensionalidade, e o AAM para extrair atributos.

Do mesmo modo, a **Figura 4.13** evidencia que o desempenho do estimador de idade é melhor quando é usada a forma original do AAM para codificação de textura. A figura mostra o escore acumulado contra o nível de erro. Nota-se que em todo momento as curvas que correspondem ao AAM,

combinado com o LBP, estão sempre abaixo das curvas correspondentes relativas ao AAM original.

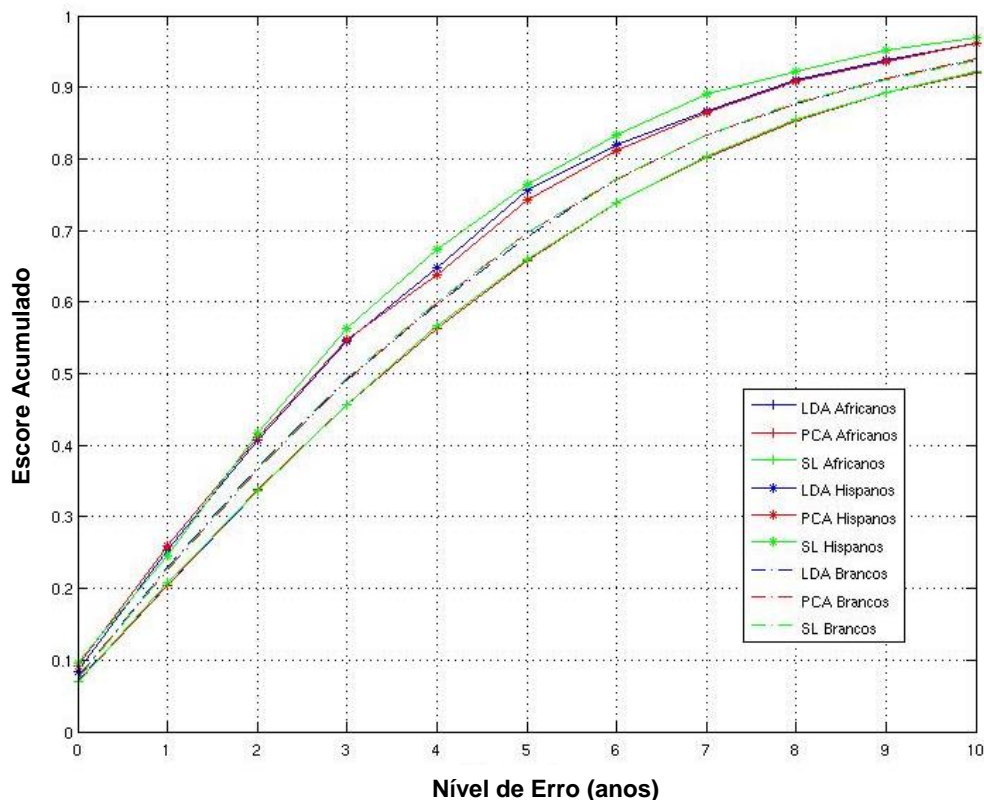


**Figura 4.13** Erro Acumulado (CS): comparam-se as curvas do erro acumulado dos métodos de redução de dimensionalidade e métodos de extração de atributos de textura.

O baixo desempenho por parte do método que mistura AAM com o LBP possivelmente esteja relacionado ao fato de que todos os métodos de redução de dimensionalidade são lineares, enquanto o LBP é notoriamente não linear. De qualquer forma, uma afirmação definitiva quanto à causa do pior desempenho do LBP no caso, requererá estudos específicos que excedem o escopo desta dissertação.

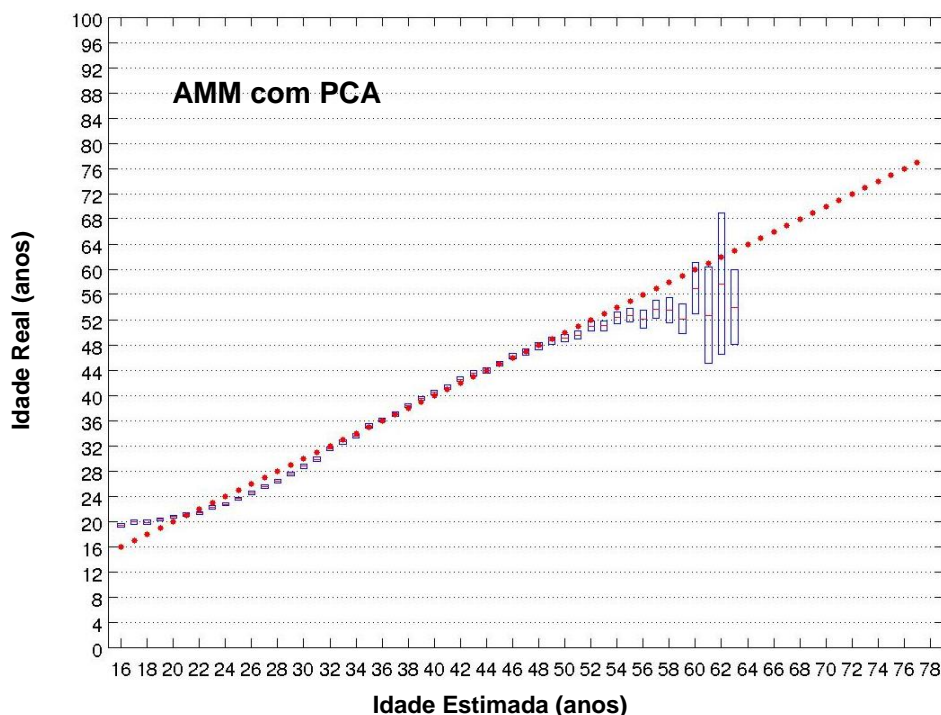
Por outro lado, na Figura 4.14 se apresenta o desempenho do estimador de idade para as três raças com maior ocorrência no banco de dados MORPH: Africanos, Brancos e Hispânicos. Para estes experimentos foi usado unicamente o AAM como método para extrair atributos de forma e textura. Muito claramente, a acurácia do estimador de idade aumenta consideravelmente para raça Hispânica, e diminui, nesta ordem, para as raças Branca e Africana. Por exemplo, para um nível de erro de 5 anos, a acurácia para os Hispânicos é de

0.77 usando SL, para os Brancos é de 0.7 usando PCA, e de 0.66 para os Africanos usando qualquer das técnicas de redução de dimensionalidade.



**Figura 4.14** Erro Acumulado (CS): comparam-se as curvas do erro acumulado dos métodos de redução de dimensionalidade e métodos de extração de atributos de textura.

Os resultados em termos de intervalos de confiança para o banco de dados MORPH, se expõem nos gráficos da Figura 4.15 e Figura 4.16. Na Figura 4.15 se mostram os resultados obtidos através do sistema de estimação de idade que utiliza AAM como técnica para extrair atributos, e PCA para reduzir dimensionalidade. Na Figura 4.16, ao contrário, se apresentam os resultados para o caso de AAM combinado com o LBP e SL para reduzir dimensionalidade. Em consonância com os resultados da Figura 4.14, ambas as configurações, conseguem a melhor acurácia na faixa de 16 até 32 anos de idade (eixo horizontal).

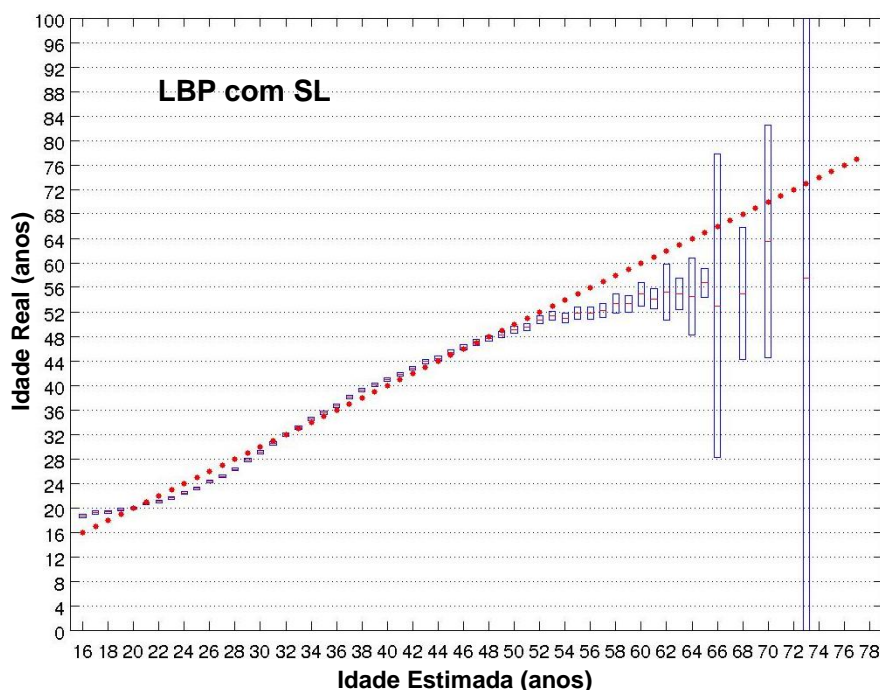


**Figura 4.15** Intervalos de confiança para o PCA para a totalidade do bando de dados MORPH. O nível de confiança é de 95% e os atributos foram extraídos unicamente através do AAM.

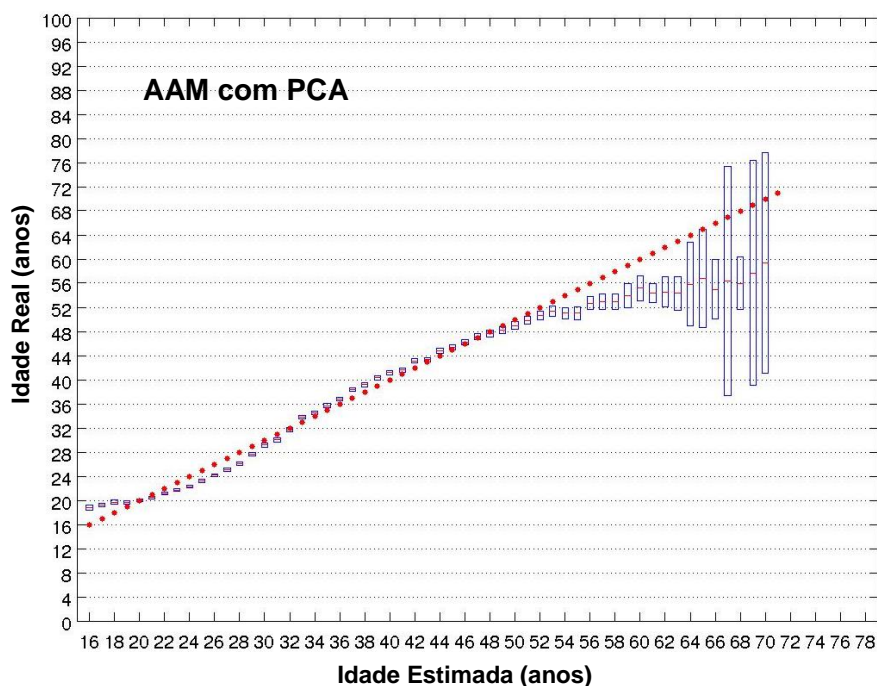
Observa-se que o valor médio dos intervalos de confiança às vezes está ora acima ora abaixo da idade estimada (ponto vermelho). Esta observação sugere que a estimativa de idade seja corrigida por um valor distinto dependendo da própria estimativa. Esta possibilidade não foi explorada neste trabalho, e consta dos trabalhos que se anteveem como continuação desta pesquisa.

Adicionalmente, na Figura 4.17, Figura 4.18 e Figura 4.19, se mostram os gráficos dos intervalos de confiança para as raças Africana, Branca e Hispanos, respectivamente. Observa-se que as larguras dos intervalos de confiança variam de uma raça para outra. Tais resultados são consistentes com os resultados que se apresentam na Figura 4.14, onde se observa que as curvas de acurácia contra o nível de erro são diferentes para cada raça. Estes resultados sugerem a concepção de um método que, com conhecimento a priori da raça, aplique uma função de regressão e uma função que ajuste os valores das idades estimadas de forma especializada para cada raça. É importante ressaltar que nestes experimentos, treinou-se e testou-se separadamente para cada raça separadamente porque 70% das imagens neste banco de dados correspondem

à raça Africana. Deste modo se se treinasse com o banco de dados completo, os resultados por raça estariam influenciados por este desequilíbrio.

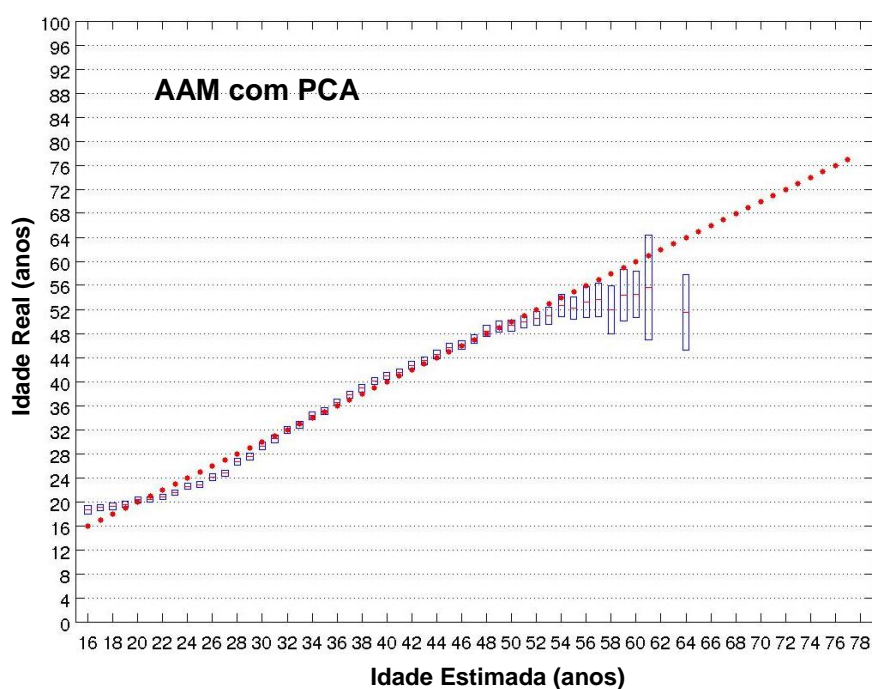


**Figura 4.16** Intervalos de confiança para o SL para a totalidade do banco de dados MORPH. O nível de confiança é de 95% e os atributos foram extraídos misturando AAM com LBP.

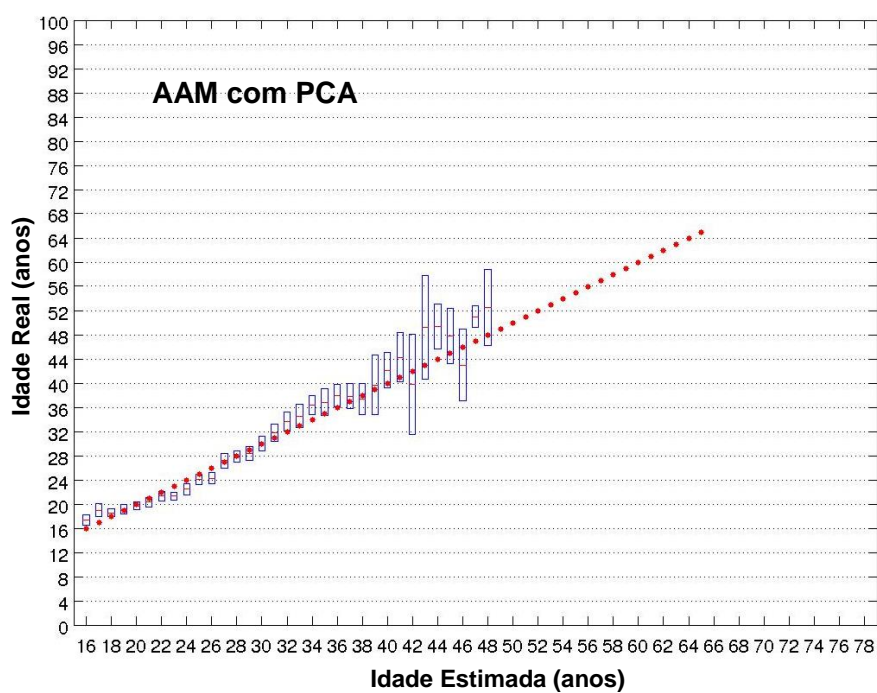


**Figura 4.17** Intervalos de confiança para o PCA da raça Africana no bando de dados MORPH. O nível de confiança é 95% e os atributos foram extraídos unicamente através de AAM.





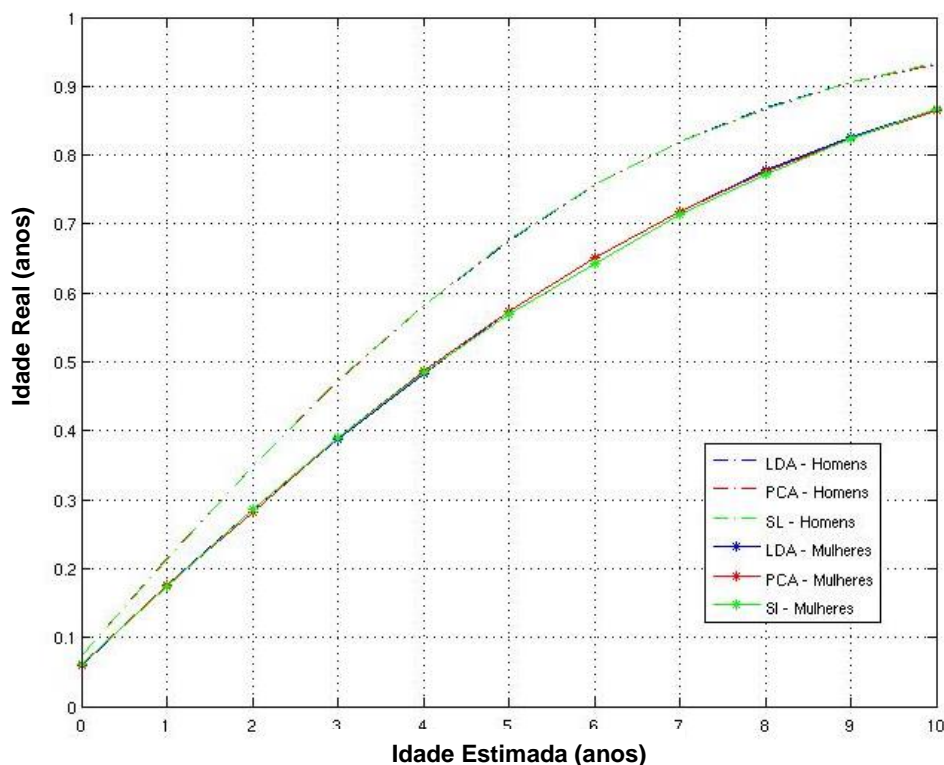
**Figura 4.18** Resultado dos intervalos de confiança para o PCA da raça Branca - bando de dados MORPH. O nível de confiança é 95% e os atributos foram extraídos unicamente através de AAM.



**Figura 4.19** Intervalos de confiança para o PCA da raça Hispânica no bando de dados MORPH. O nível de confiança é de 95% e os atributos foram extraídos unicamente através de AAM.

Finalmente, a Figura 4.20 apresenta as curvas da acurácia contra o nível de erro, correspondentes às avaliações feitas sob a influência do gênero no desempenho do estimador de idade. Observa-se que em todos os casos, as curvas da acurácia dos homens superam em desempenho as curvas da acurácia das mulheres. Este resultado pode estar relacionado ao uso de maquiagem por parte das mulheres, o qual geralmente altera a informação da textura nas faces, e, por conseguinte a estimativa da idade. Além disso, o número de imagens de mulheres é menor em uma relação de 3/10 com respeito aos homens. Houve, portanto, nestes experimentos menos imagens de mulheres do que de homens para o treinamento dos respectivos modelos.

Naturalmente, desenvolver um sistema de estimação de idades particular para cada gênero, ajudaria consideravelmente na acurácia das estimações. O problema de discernir entre sexo não é uma tarefa complexa para um sistema de reconhecimento facial (Gehrig, Steiner and Ekenel 2011), e facilmente pode ser incluído como etapa previa para o sistema de estimação de idade. Não obstante, esta tarefa não foi objeto desta dissertação.



**Figura 4.20** Erro Acumulado (CS). Comparam-se as curvas do erro acumulado dos métodos de redução de dimensionalidade e gênero das pessoas.

○ **Banco de Dados FG-NET**

A seguir se apresentam os resultados reunidos de experimentos realizados com o banco de dados FG-NET.

Nas Tabela 4.9 e Tabela 4.10 se resumem os resultados do MAE e MAE/D das experimentações realizadas com o banco de dados FG-NET. Do mesmo modo que com o MORPH, comparam-se as técnicas de extração de atributos de textura e de redução de dimensionalidade. Note-se que com este banco de dados é mais evidente a diferença de desempenho entre o sistema que usa AAM ou aquele que emprega AAM misturado com o LBP. Nas tabelas se observa que a diferença no erro médio absoluto está na ordem de 30%, uma diferença considerável. Adicionalmente, na Figura 4.21 se apresentam os intervalos de confiança para cada uma das idades estimadas por meio do sistema que utiliza AAM misturado com LBP e PCA como método para reduzir dimensionalidade. Os resultados para AAM e o PCA, são mostrados na Figura 4.22. Ambos os grupos de resultados apontam para a mesma tendência.

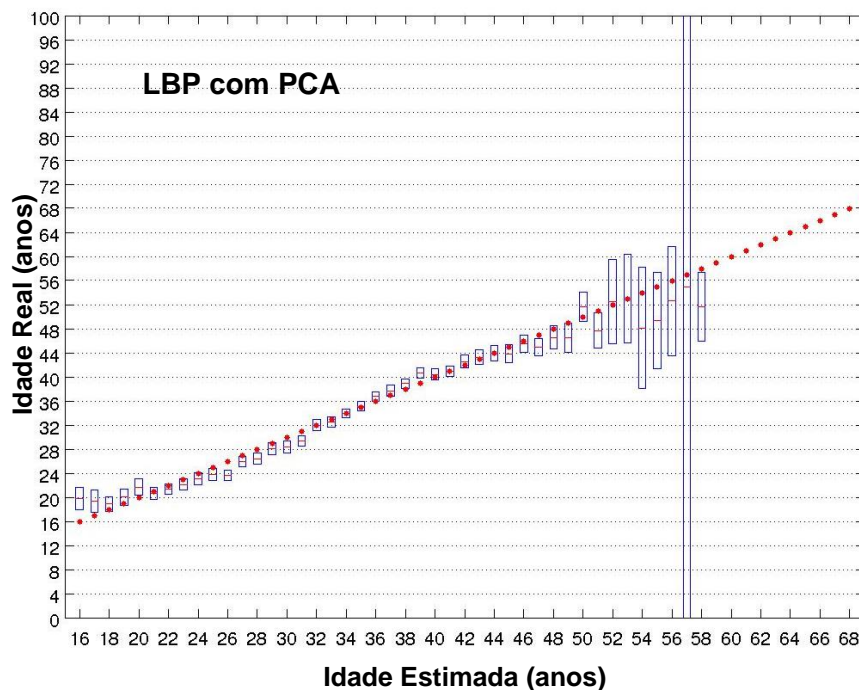
<b>Método de Extração de Atributos</b>	<b>MAE</b>
AAM	6,25
AAM + LBP	9.17

**Tabela 4.9** Erro Médio Absoluto (MAE) para estimativa da idade no banco de dados FG-NET.

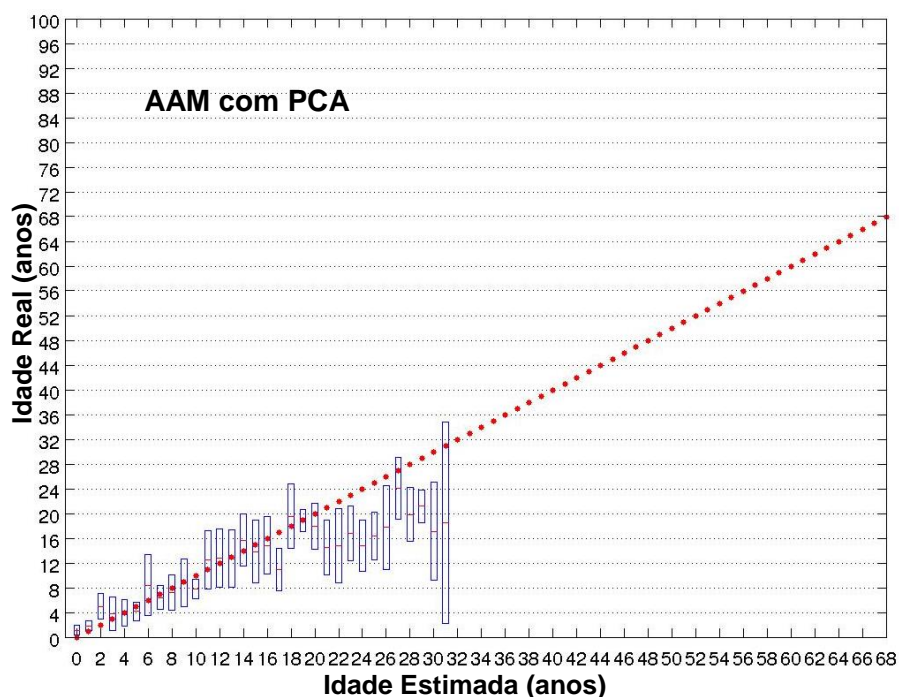
<b>Década</b>	<b>AAM</b>	<b>AAM + LBP</b>
00 – 09	4,23	8,19
10 – 19	3,75	3,61
20 – 29	4,61	7,78
30 – 39	11,30	18,46
40 - 49	19,50	28,90
50 – 59	29,53	30,66
60 - 69	39,28	40,71

**Tabela 4.10** Erro Médio Absoluto por Década (MAE/D) para estimativa da idade do banco de dados FG-NET. Utilizou-se o PCA na etapa de redução de dimensionalidade.





**Figura 4.21** Resultado dos intervalos de confiança para o PCA no banco de dados FG-NET. O nível de confiança é de 95% e os atributos foram extraídos através de AAM misturado com o LBP.



**Figura 4.22** Resultado dos intervalos de confiança para o PCA no banco de dados FG-NET. O nível de confiança é de 95% e os atributos foram extraídos através de AAM.

### 4.3.1.3. Experimento 3.

**Objetivo:** Comparar os métodos de regressão.

Neste experimento se avalia o desempenho dos métodos de estimação de idade ante as técnicas de regressão, ou seja, SVR e Regressão Linear. Ambos os bancos de dados foram usados nos experimentos, em conjunto com as três técnicas de redução de dimensionalidade utilizadas em todo este trabalho. Adotou-se o melhor método de extração de atributos obtido no experimento 2, o AAM.

Seguindo a mesma estratégia experimental para cada banco de dados, os experimentos foram realizados, primeiro com SVR, e depois com regressão linear. Para SVR foi selecionada a melhor configuração de parâmetros através de validação cruzada.

**Resultados:** os resultados são apresentados a seguir separados por banco de dados.

- **Banco de Dados MORPH**

Como nas seções anteriores, se resumem os resultados do MAE e MAE/D de todos os testes realizados em duas tabelas, especificamente, na Tabela 4.11 e na Tabela 4.12, respectivamente. Compara-se o desempenho das duas técnicas de regressão ante os três métodos de redução de dimensionalidade. Observa-se que, em todos os casos, a regressão realizada por meio das SVRs supera em desempenho a regressão linear. Este comportamento também se nota na Figura 4.23, onde se ilustram os resultados do MAE/y produzidos a partir de SVR e regressão linear, utilizando PCA.

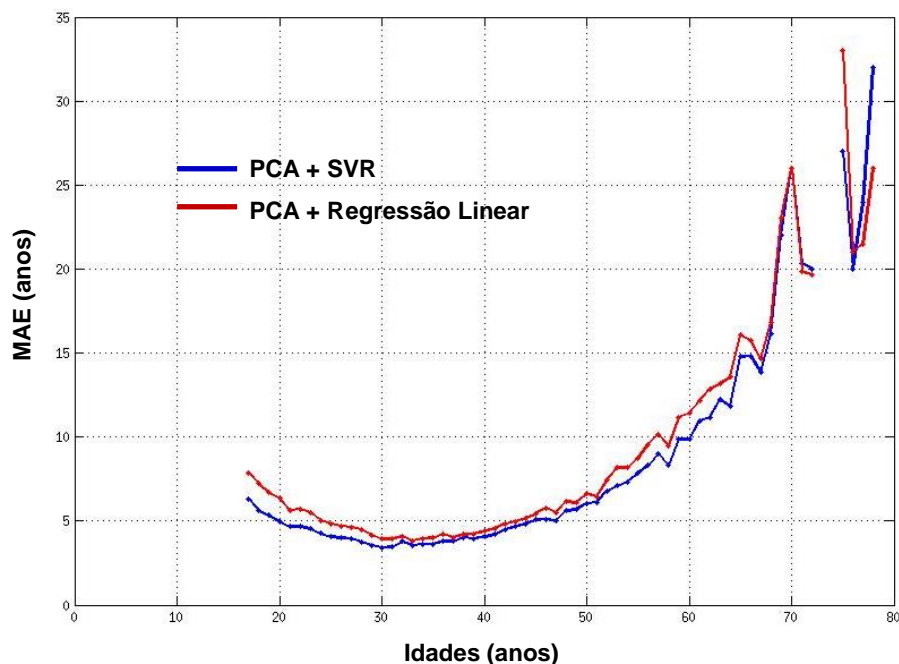
Técnica de Regressão	Método para Reduzir Dimensionalidade		
	PCA	LDA	Subspace L.
SVR	4.57	4.58	4.90
Regressão Linear	5.26	5.26	5.35

**Tabela 4.11.** Erro Médio Absoluto (MAE) para estimativa da idade no banco de dados MORPH, Album2. Comparam-se as técnicas de regressão e redução de dimensionalidade.

Observa-se que, especificamente na faixa de idades compreendida entre 16 e 30 anos, o erro associado à regressão linear é substancialmente maior do que o erro associado à SVR, embora não exorbitante.

Década	SVR			Regressão Linear		
	PCA	LDA	SL	PCA	LDA	SL
10 – 19	5,48	5,48	6,18	6,95	6,94	7,16
20 – 29	4,14	4,14	4,49	4,94	4,94	5,12
30 – 39	3,78	3,78	3,93	4,09	4,09	4,19
40 – 49	4,95	4,95	5,24	5,38	5,38	5,43
50 – 59	7,42	7,44	7,85	8,29	8,28	8,12
60 – 69	12,94	12,97	13,22	14,16	14,16	14,18
70 – 79	22,07	22,07	20,61	21,80	21,80	27,06

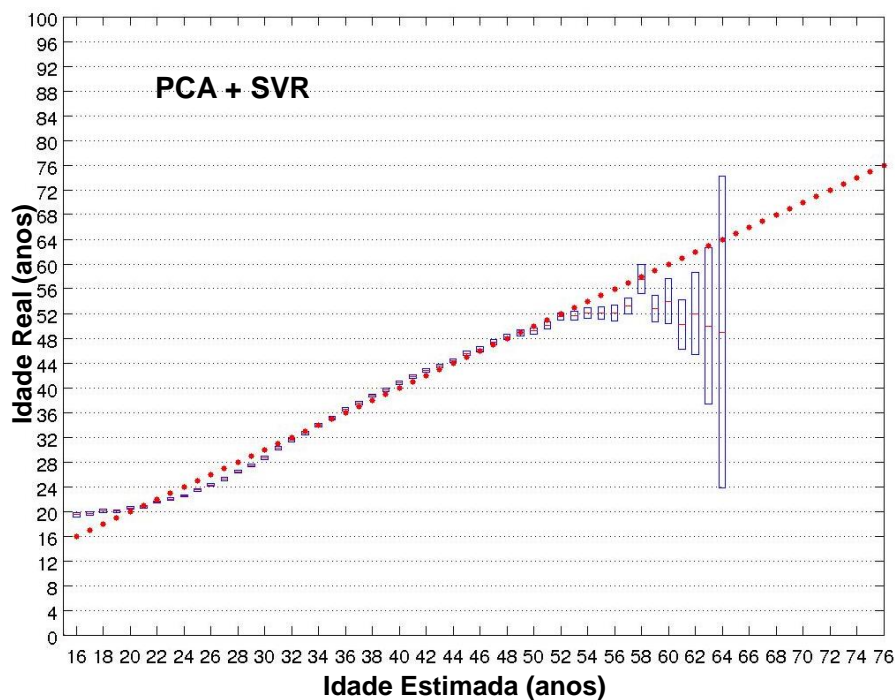
**Tabela 4.12** Erro Médio Absoluto por Década (MAE/D) para estimativa da idade do banco de dados MORPH, Album2. Comparam-se as técnicas de regressão e redução de dimensionalidade.



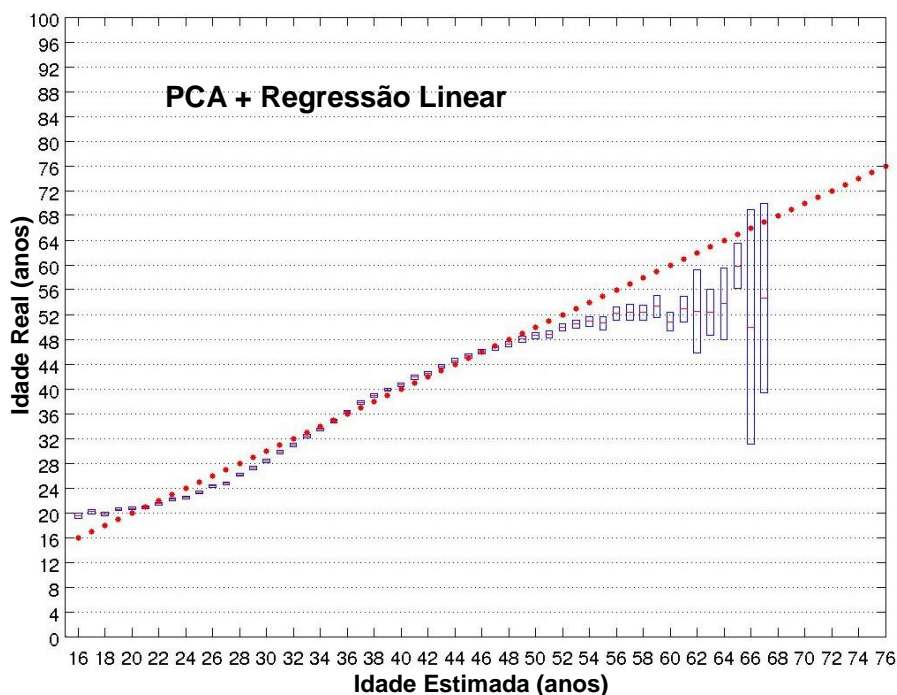
**Figura 4.23** Erro Médio Absoluto por Idade (MAE/y) da Estimativa da Idade do álbum 2, no banco de dados MORPH. O PCA é usado para redução de dimensionalidade, e se comparam dois métodos para regressão; Regressão Linear e SVRs.

Finalmente, na Figura 4.24 e Figura 4.25, se avaliam os intervalos de confiança. Na Figura 4.24 se apresentam os resultados para o SVR, e na Figura

4.25 os resultados para a regressão linear. Em ambos os casos, foi utilizado o PCA como método de redução de dimensionalidade. Observa-se que os intervalos de confiança a partir de 48 anos (eixo horizontal) ficam abaixo da curva pontilhada, no caso da regressão linear, e a partir dos 55 anos de idade, para SVR.



**Figura 4.24** Intervalos de confiança para PCA + SVR no banco de dados MORPH. O nível de confiança é de 95%.



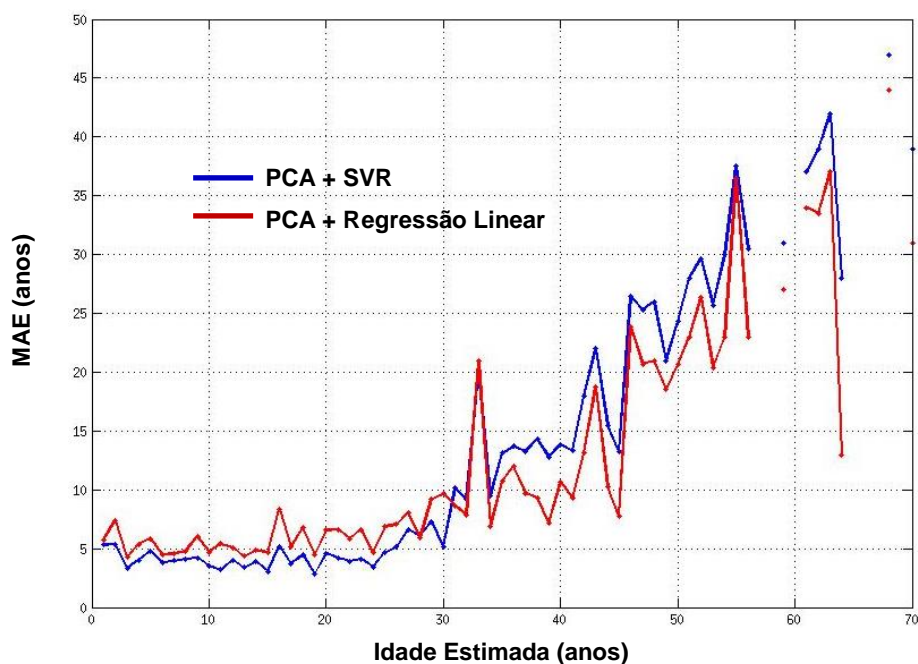
**Figura 4.25** Intervalos de confiança ao usar PCA +Regressão linear no banco de dados MORPH. O nível de confiança é de 95%.

○ Banco de dados FG-NET

Técnica de Regressão	Método para Reduzir Dimensionalidade		
	PCA	LDA	Subspace L.
SVR	6,25	6,23	6,80
Linear	6.88	7.03	7.04

**Tabela 4.13** Erro Médio Absoluto (MAE) para estimativa da idade no banco de dados FG-NET. Comparam-se as técnicas de regressão e de redução de dimensionalidade.

Os resultados obtidos com o banco de dados FG-NET, apresentam a mesma tendência revelada nos experimentos realizados com o banco MORPH. Neste sentido, o desempenho do estimador de idade que utiliza SVR, supera o obtido por regressão linear, conforme mostra a Tabela 4.13. Não obstante, a curva de MAE/y referente à regressão linear mostrado na Figura 4.26 apresenta erros mais baixos na faixa de idade entre 35 e 60 anos do que os que amostra a curva relacionada a SVR. Do mesmo modo, observa-se na Figura 4.27 e na Figura 4.28 que, na faixa compreendida entre 30 e 45 anos, a regressão linear em muitos casos teve melhor desempenho.



**Figura 4.26** Erro Médio Absoluto por Idade (MAExA) da Estimativa da Idade do banco de dados FG-NET. O PCA é usado para redução de dimensionalidade, e se comparam dois métodos para regressão; Regressão Linear e SVRs.

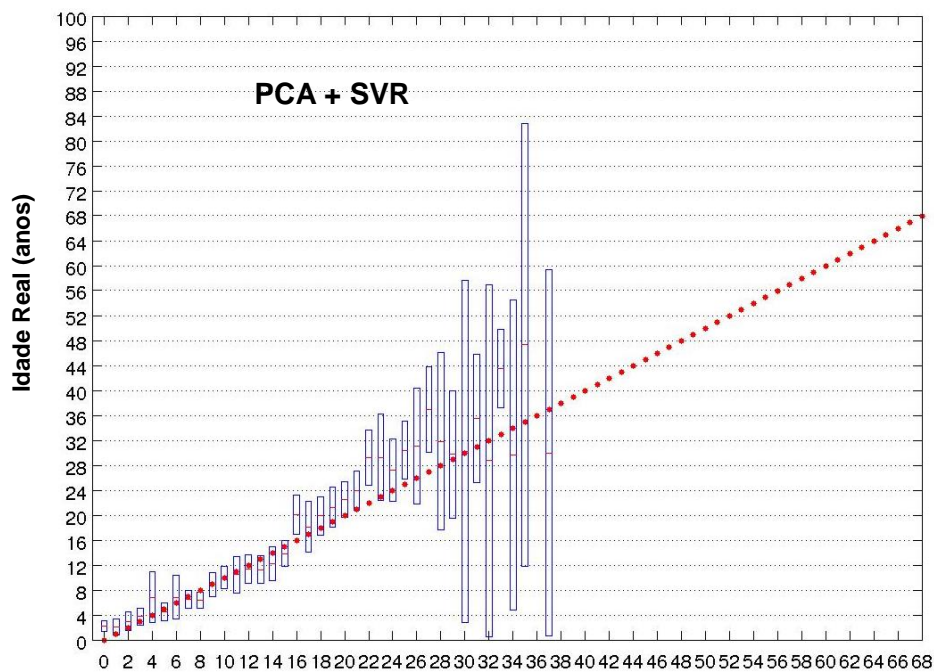
Em resumo, nas faixas etárias mais jovens, os intervalos de confiança referentes à SVR, apresentaram melhor desempenho, mas entre 35 e 60 anos, os resultados da regressão linear superaram os da SVR.

#### **4.3.2. Sumário dos Experimentos**

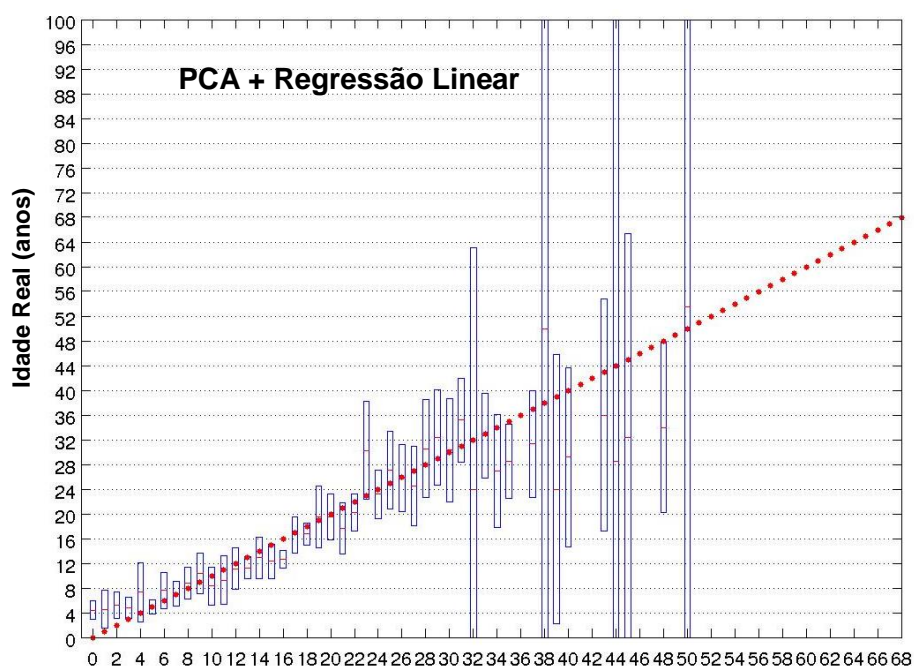
Nos três experimentos levados a cabo, se conseguiu determinar a influência de vários fatores, tais como acurácia na localização dos pontos fiduciais, atributos, raça, métodos de redução de dimensionalidade e de regressão, no desempenho do sistema de estimação de idade.

Surpreende o fato que técnicas como o LDA, em muitos casos, apresentou pior desempenho do que o PCA, e que os resultados obtidos através de regressão linear, não foram amplamente superados pelos de SVR.

Embora não tenha sido uma métrica avaliada, cabe mencionar que o tempo de processamento no treinamento é substancialmente mais longo para SVR em comparação com os tempos de treinamento para regressão linear. Por exemplo, num experimento com todo o banco de dados MORPH, o treinamento do SVR tomou um dia aproximadamente, enquanto que a regressão linear, menos de 15 min.. Não obstante, terem sido utilizadas implementações distintas, esta diferença é substancial.



**Figura 4.27** Intervalos de confiança utilizando o PCA + SVR no banco de dados FG-NET. O nível de confiança é de 95%.



**Figura 4.28** Intervalos de confiança utilizando o PCA + regressão linear no banco de dados FG-NET. O nível de confiança é de 95%.



## 5 Conclusões

Neste capítulo se apresentam as principais conclusões desta dissertação. Discutem-se os resultados coletados nas experiências realizadas, assim como também os trabalhos futuros a serem desenvolvidos na sequência desta pesquisa.

### 5.1. Discussão

Neste trabalho investigaram-se métodos alternativos para estimação de idade a partir de imagens faciais. O estudo baseou-se em cinco métricas de desempenho.

Nos experimentos realizados, primeiro pôde-se determinar que o impacto da acurácia da detecção dos pontos fiduciais na performance da estimativa da idade é moderado, mas importante. O erro associado à detecção automática foi cerca de 10% maior do que quando se utilizam pontos fiduciais detectados manualmente.

Segundo, a representação de textura através do algoritmo LBP foi consistentemente inferior à representação de textura utilizada no método AAM

Em terceiro, o PCA foi a técnica de redução de dimensionalidade que apresentou o melhor desempenho na maioria das experiências, ainda que não muito superior ao LDA e ao *Subspace Learning* (SL). Isto é significativo tomando-se em conta que é, entre os métodos de redução de dimensionalidade analisados, o mais simples.

Finalmente, verificou-se que as funções de regressão conseguidas por meio das SVRs superaram consistentemente as dos modelos de regressão linear.

Por outro lado, é importante ressaltar que a distribuição das idades nos bancos de dados afetou de maneira direta nos resultados obtidos. A escassez de



imagens de treinamento em certas faixas etárias causou os erros nestas idades, conforme se observou nas idades acima de 60 anos.

Por último, os experimentos realizados com os conjuntos de testes discriminados por raça ou por gênero, permitiram determinar que a acurácia do estimador da idade depende da raça e do gênero, tendo sido mais exato para a raça hispânica e para homens. Deste modo, é de se acreditar que a estimativa de idade se beneficie de uma pré-classificação das imagens faciais de entrada em termos de raça e gênero.

## **5.2. Trabalhos Futuros**

A estimativa da idade a partir de imagens faciais é um problema que tem atraído grande interesse nos últimos anos. Foram propostas muitas abordagens que procuram dar uma solução a este problema, como está detalhado no capítulo segundo capítulo desta dissertação. Mesmo assim, uma solução definitiva continua sendo ainda um desafio, dada a complexidade do problema. Em particular, os resultados medidos neste trabalho foram satisfatórios para aplicações com exigência moderada, mas não superlativos em muitos casos. Ainda assim, os possíveis estudos que podem se desenvolvidos baseados na abordagem adotada sugerem resultados mais prometedores.

Como se detalhou no capítulo sobre a Análise Experimental, a análise dos intervalos de confiança em muitos casos mostrou um *bias* entre a idade ideal e o centro do intervalo de confiança estimado. Uma análise simples sugere que é possível corrigir este erro, mudando a resposta da função de regressão de maneira adequada para as idades em que se apresenta o comportamento anteriormente descrito. Maiores testes serão necessários para se confirmar esta suspeita.

Por outro lado, a pouca disponibilidade de bancos de dados de faces, rotuladas com a idade, continua sendo uma dificuldade. Uma solução possível seria aproveitar o grande volume de imagens faciais disponíveis na internet, por exemplo, para realizar experimentações com técnicas de aprendizado semi-

supervisionado, como se aventa em (Ueki et al. 2010) e (Chen and Hsu 2013). Deste modo, se poderia suprir a escassez de imagens rotuladas em determinadas faixas etárias.

Finalmente, um outro desdobramento futuro nesta linha de pesquisa diz respeito à estimação de idade a partir de imagens em movimento e em ambientes não controlados.

## 6 Referências Bibliográficas

AHONEN, T.; HADID, A.; PIETIKAINEN, M. Face description with local binary patterns: Application to face recognition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, v. 28, n. 12, p. 2037–2041, 2006. ISSN 0162-8828.

ALLEY, T. R. Social and applied aspects of face perception: An introduction. In: *Social and applied aspects of perceiving faces*. [S.l.]: Lawrence Erlbaum Assoc Hillsdale, NJ, 1988, p. 1–8.

ASTHANA, A. et al. Robust discriminative response map fitting with constrained local models. In: *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*. [S.l.]: IEEE, 2013, p. 3444–3451. ISBN 1063-6919.

AYMA, V. H. Estudo de Métodos Automáticos de Reconhecimento Facial para Vídeo Monitoramento. *Dissertação de Mestrado – Pontifícia Universidade Católica do Rio de Janeiro – PUC-RIO*, 2014.

CHANG, C. -C.; LIN, C.-J. Libsvm: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, v. 2, n. 3, p. 27, 2011. ISSN 2157-6904.

CHEN, Y.; HSU, C. Subspace learning for facial age estimation via pairwise age ranking. 2013. ISSN 1556-6013.

CHRISTOU, N.; DINOV, I. D. Confidence interval based parameter estimation—a new socr applet and activity. *PloS one*, v. 6, n. 5, p. e19178, 2011. ISSN 1932-6203.

COOTES, T. F.; EDWARDS, G. J.; TAYLOR, C. J. Active appearance models. *IEEE Transactions on pattern analysis and machine intelligence*, v. 23, n. 6, p. 681–685, 2001. ISSN 0162-8828.

CORTES, C.; VAPNIK, V. Support-vector networks. *Machine learning*, v. 20, n. 3, p. 273–297, 1995. ISSN 0885-6125.

CRISTINACCE, D.; COOTES, T. F. Feature detection and tracking with constrained local models. In: *BMVC*. [S.l.: s.n.], 2006. v. 2, p. 6.

EASTON, V. J.; MCCOLL, J. H. Book. *Statistics glossary*. [S.l.]: Steps, 1997.

FARKAS, L. G. Book. *Anthropometry of the Head and Face*. [S.l.]: Raven Pr, 1994. ISBN 0781701597.

FISHER, R. A. The use of multiple measurements in taxonomic problems. *Annals of eugenics*, v. 7, n. 2, p. 179–188, 1936. ISSN 2050-1439.

FU, Y.; XU, Y.; HUANG, T. S. Estimating human age by manifold analysis of face pictures and regression on aging features. In: *Multimedia and Expo, 2007 IEEE International Conference on*. [S.l.]: IEEE, 2007. p. 1383–1386. ISBN 1424410169.

GEHRIG, T.; STEINER, M.; EKENEL, H. K. Generic, Draft: Evaluation Guidelines for Gender Classification and Age Estimation. 2011.

GENG, X.; YIN, C.; ZHOU, Z.-H. Facial age estimation by learning from label distributions. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, v. 35, n. 10, p. 2401–2412, 2013. ISSN 0162-8828.

GENG, X.; ZHOU, Z.-H.; SMITH-MILES, K. Automatic age estimation based on facial aging patterns. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, v. 29, n. 12, p. 2234–2240, 2007. ISSN 0162-8828.

GENG, X. et al. Learning from facial aging patterns for automatic age estimation. In: *Proceedings of the 14th annual ACM international conference on Multimedia*. [S.l.]: ACM, 2006. p. 307–316. ISBN 1595934472.

GUO, G. et al. A study on automatic age estimation using a large database. In: *Computer Vision, 2009 IEEE 12th International Conference on*. [S.l.]: IEEE, 2009. p. 1986–1991. ISBN 1424444209.

HAYASHI, J. et al. Method for estimating and modeling age and gender using facial image processing. In: *Virtual Systems and Multimedia, 2001. Proceedings. Seventh International Conference on*. [S.l.]: IEEE, 2001. p. 439–448. ISBN 0769514022.

HSU, C.-W.; CHANG, C.-C.; LIN, C.-J. Generic, A practical guide to support vector classification. 2003.

JOHNSON, R. A.; WICHERN, D. W. Book. Applied multivariate statistical analysis. [S.l.]: *Prentice hall Englewood Cliffs, NJ*, 1992. v. 4.

JOLLIFFE, I. Book. Principal component analysis. [S.l.]: *Wiley Online Library*, 2005. ISBN 0470013192.

KWON, Y. H.; LOBO, N. da V. Age classification from facial images. In: *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR'94., 1994 IEEE Computer Society Conference on*. [S.l.]: IEEE, 1994. p. 762–767. ISBN 0818658258.

LANITIS, A.; TAYLOR, C. J.; COOTES, T. F. Toward automatic simulation of aging effects on face images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, v. 24, n. 4, p. 442–455, 2002. ISSN 0162-8828.

LEWIS, M. B.; ELLIS, H. D. How we detect a face: A survey of psychological evidence. *International Journal of Imaging Systems and Technology*, v. 13, n. 1, p. 3–7, 2003. ISSN 1098-1098.

LIU, C.; WECHSLER, H. Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition. *Image processing, IEEE Transactions on*, v. 11, n. 4, p. 467–476, 2002. ISSN 1057-7149.

LOBO, N.; KWON, Y. Generic, Automatic feature detection and age classification of human faces in digital images. [S.l.]: *Google Patents*, 1998.

LU, J.; TAN, Y.-P. Ordinary preserving manifold analysis for human age and head pose estimation. *Human-Machine Systems, IEEE Transactions on*, v. 43, n. 2, p. 249–258, 2013. ISSN 2168-2291.

MARQUES, J. M.; MARQUES, M. A. M. As componentes principais no descarte de variáveis em um modelo de regressão múltipla. *Revista FAE*, p. 93–101, 2005.

MATTHEWS, I.; BAKER, S. Active appearance models revisited. *International Journal of Computer Vision*, v. 60, n. 2, p. 135–164, 2004. ISSN 0920-5691.

MCLACHLAN, G. Book. Discriminant analysis and statistical pattern recognition. [S.l.]: *John Wiley & Sons*, 2004. v. 544. ISBN 0471691151.

MONTGOMERY, D. C. Book. Applied Statistics and Probability for Engineers 6th edition. [S.l.]: *Wiley*, 2013.

MUSTI, U.; ZHOU, Z.; PIETIKAINEN, M. Facial 3d shape estimation from images for visual speech animation. In: *Pattern Recognition (ICPR), 2014 22nd International Conference on*. [S.l.]: *IEEE*, 2014. p. 40–45. ISBN 1051-4651.

OJALA, T.; PIETIKÄINEN, M.; HARWOOD, D. A comparative study of texture measures with classification based on featured distributions. *Pattern recognition*, v. 29, n. 1, p. 51–59, 1996. ISSN 0031-3203.

PHILLIPS, P. J. et al. The feret evaluation methodology for face-recognition algorithms. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, v. 22, n. 10, p. 1090–1104, 2000. ISSN 0162-8828.

RAMANATHAN, N.; CHELLAPPA, R. Modeling age progression in young faces. In: *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*. [S.l.]: *IEEE*, 2006. v. 1, p. 387–394. ISBN 0769525970.

RICANEK, K.; TESAFAYE, T. Morph: A longitudinal image database of normal adult age-progression. In: *Automatic Face and Gesture Recognition, 2006. FGR*

2006. *7th International Conference on. [S.l.]: IEEE*, 2006. p. 341–345. ISBN 0769525032.

RIESENHUBER, M.; POGGIO, T. Hierarchical models of object recognition in cortex. *Nature neuroscience*, v. 2, n. 11, p. 1019–1025, 1999.

SAGONAS, C. et al. A semi-automatic methodology for facial landmark annotation. In: *Computer Vision and Pattern Recognition Workshops (CVPRW), 2013 IEEE Conference on. [S.l.]: IEEE*, 2013. p. 896–903.

SERRE, T. et al. Robust object recognition with cortex-like mechanisms. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, v. 29, n. 3, p. 411–426, 2007. ISSN 0162-8828.

SEUNG, H. S.; LEE, D. D. The manifold ways of perception. *Science*, v. 290, n. 5500, p. 2268–2269, 2000.

SMOLA, A. J.; SCHÖLKOPF, B. A tutorial on support vector regression. *Statistics and computing*, v. 14, n. 3, p. 199–222, 2004. ISSN 0960-3174.

FG-NET:2011 Web Page, THE FG-NET Aging Database [Online]. 2011. Disponível em: <<http://www.fgnet.rsunit.com/>>.

TURK, M.; PENTLAND, A. Eigenfaces for recognition. *Journal of cognitive neuroscience*, v. 3, n. 1, p. 71–86, 1991.

UEKI, K.; SUGIYAMA, M.; IHARA, Y. Semi-supervised estimation of perceived age from face images. In: *VISAPP (2)*. [S.l.]: Citeseer, 2010. p. 319–324.

VEZZARO, L. Generic, ICAAM - inverse compositional active appearance models, 2010. Disponível em: <<http://sourceforge.net/projects/icaam/>>.

VIOLA, P.; JONES, M. Rapid object detection using a boosted cascade of simple features. In: *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on. [S.l.]: IEEE*, 2001. v. 1, p. I–511–I–518 vol. 1. ISBN 0769512720.

WANG, Y.; LUCEY, S.; COHN, J. F. Enforcing convexity for improved alignment with constrained local models. In: *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*. [S.l.]: IEEE, 2008. p. 1–8. ISBN 1424422426.

Wenwen, L.; Xiaoxue, X.; Fu, L., & Yu, Z. Application of Improved Grid Search Algorithm on SVM for Classification of Tumor Gene. *International Journal of Multimedia and Ubiquitous Engineering*, 2014. p. 181-188.

YAN, S.; LIU, M.; HUANG, T. S. Extracting age information from local spatially flexible patches. In: *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*. [S.l.]: IEEE, 2008. p. 737–740. ISBN 1424414830.