

UNIVERSIDADE ESTADUAL PAULISTA

Faculdade de Ciências - Bauru

Bacharelado em Ciência da Computação

Alan Zanoni Peixinho

Android Eye: Sistema para Detecção e Reconhecimento
Facial em Dispositivos Móveis Para Auxiliar Pessoas
com Deficiência Visual



UNESP

2012

Alan Zanoni Peixinho

Android Eye: Sistema para Detecção e Reconhecimento
Facial em Dispositivos Móveis Para Auxiliar Pessoas
com Deficiência Visual

Orientador: Dr. Aparecido Nilceu Marana

Coorientador: André Marcelo Farina

Monografia apresentada junto à disciplina
Projeto e Implementação de Sistemas II, do
curso de Bacharelado em Ciência da
Computação, Faculdade de Ciências, Unesp,
campus de Bauru, como parte do Trabalho de
Conclusão de Curso.

UNESP

2012

Alan Zanoni Peixinho

Android Eye: Sistema para Detecção e Reconhecimento Facial em Dispositivos Móveis Para
Auxiliar Pessoas com Deficiência Visual

Monografia apresentada junto à disciplina
Projeto e Implementação de Sistemas II, do
curso de Bacharelado em Ciência da
Computação, Faculdade de Ciências, Unesp,
campus de Bauru, como parte do Trabalho de
Conclusão de Curso.

BANCA EXAMINADORA

Prof. Adj. Aparecido Nilceu Marana
Professor Adjunto
DCo - FC - UNESP - Bauru
Orientador

Prof. Dr. Simone Domingues Prado
Professora Doutora
UNESP - Bauru

Prof. Adj. Roberta Spolon
Professora Adjunta
UNESP - Bauru

Bauru, 22 de Novembro de 2012.

Resumo

A deficiência visual severa é o tipo de deficiência que mais incide sobre a população brasileira atingindo atualmente cerca de 7 milhões de pessoas, segundo a mais recente pesquisa do IBGE. Esse grande número de deficientes visuais justifica o desenvolvimento de tecnologias que possam auxiliar essas pessoas em suas atividades cotidianas, de forma a melhorar sua integração na sociedade, proporcionando-lhes uma melhor qualidade de vida. Dentre as tecnologias que mais têm se desenvolvido nos últimos anos destacam-se aquelas relacionadas aos dispositivos móveis, como os *smartphones* e os *tablets*. Este projeto de conclusão de curso propõe o desenvolvimento de um sistema de identificação de pessoas por meio do reconhecimento facial em dispositivos móveis, para ser utilizado prioritariamente por pessoas que apresentem deficiência visual. Para tanto, a interface do sistema possui recursos adequados para esse tipo de usuário. O reconhecimento facial foi escolhido, pois é uma das técnicas biométricas mais aceitas pelas pessoas, além de poder ser realizada a distância, e ser coletada pelas câmeras embutidas naturalmente nos dispositivos móveis. Para a implementação desse sistema foram avaliadas algumas técnicas de detecção de faces baseadas na cor da pele e nas características de Haar, conforme proposto por Viola-Jones. Para a representação e descrição das faces foram avaliadas as técnicas baseadas em Modelo de Forma Ativa (ASM - *Active Shape Model*), na análise das componentes principais (PCA - *Principal Component Analysis*), devido à sua ampla utilização, e nos padrões binários locais (LBP - *Local Binary Pattern*), técnica também conhecida como Transformada Census, devido à sua maior invariância às alterações de iluminação. O sistema proposto é uma extensão do sistema BioMobile, desenvolvido recentemente pelo grupo de pesquisa Recogna, do Departamento de Computação, Faculdade de Ciências, UNESP, campus de Bauru. A partir dos resultados experimentais concluiu-se que a técnica de detecção de faces baseada em cor da pele mostrou-se como uma melhor opção quando comparada a técnica Viola-Jones, em termos de tempo e acurácia. Quanto as técnicas de representação da face, esses mesmos resultados, mostraram ainda que a técnica LBP apresenta-se como uma melhor alternativa ao sistema proposto, em termos de acurácia e necessidade de recursos computacionais, quando comparada ao ASM e PCA.

Palavras-chave: Biometria, Reconhecimento da Face, Modelo de Forma Ativa, Análise dos Componentes Principais, Desenvolvimento de Aplicações Móveis.

Abstract

Severe visual impairment is a type of disability that more affects the Brazilian population currently reaching approximately 7 million people, according to the latest statistics of the IBGE (Instituto Brasileiro de Geografia e Estatística). This large number of visually impaired persons justifies the development of technologies that can help these people in their daily activities to improve their integration into society, giving them a better life quality. Among technologies that have developed over the past years include those related to mobile devices such as *smartphones* and tablets. This undergraduate final exam project proposes the development of a system to identify people through facial recognition on mobile devices, for use by persons who are visually impaired. Therefore, the interface system has adequate resources to this type of user. Facial recognition was selected since it is one of the most accepted biometric techniques. Besides, it can perform remotely and the input data (the faces) can be collected by the cameras embedded into the mobile devices. To implement this system were evaluated some techniques for face detection based on skin color and Haar features, as proposed by Viola-Jones. For face representation and description were evaluated techniques based on Active Shape Model (ASM), Principal Component Analysis (PCA), because of its wide use, and Local Binary Pattern (LBP), a technique also known as Census Transform, due to its higher invariance to lighting changes. The proposed system is an extension of the BioMobile system, recently developed by the Recogna research group, Department of Computing, School of Sciences, UNESP, Bauru. From the experimental results we realized that the face detection technique based on skin color proved to be a better option against the Viola-Jones technique, in time and accuracy. To the face recognition, these results also showed that the LBP technique proved to be a better alternative to the proposed system in terms of accuracy and computational resource requirements compared to the ASM and PCA.

Keywords: Biometrics, Face Recognition, Active Shape Model, Principal Component Analysis, Local Binary Pattern, Mobile Development.

Agradecimentos

Antes de tudo gostaria de agradecer imensamente a Deus por guiar meus passos, e me auxiliar, nos momentos em que nada parecia dar certo.

Gostaria de agradecer a toda minha família e amigos, aos quais neguei tantos fins de semana de convívio ao longo da jornada desta graduação. Em especial aos meus pais Anália Maria Zanoni Peixinho e José Peixinho, e meu irmão Alex Zanoni Peixinho.

Ao orientador prof. Dr. Aparecido Nilceu Marana que acreditou em mim, até mais do que eu mesmo, me apoiando e dedicando boa parte de seu, já escasso, tempo em me ajudar, sempre que necessitei.

Ao co-orientador André Marcelo Farina, que tanto me auxiliou no desenvolvimento deste trabalho.

Ao Everton Fernandes da Silva que trabalhou ao meu lado no desenvolvimento deste projeto, mesmo com o advento de seu intercâmbio.

A todos os professores que me ajudaram a crescer, intelectual e pessoalmente, ao longo desses anos.

Aos companheiros do grupo de Pesquisa Recogna que me acolheram e me ajudaram nos momentos em que precisei de sua colaboração.

E por fim, a todos os companheiros, nos cursos de Sistemas de Informação e Ciência da Computação, que lutaram ao meu lado durante todos esses anos de desafios.

Lista de Figuras

1	Modos de operação de sistemas biométricos (JAIN; ROSS; PRABHAKAR, 2004).	20
2	Variações na pose, iluminação, expressão, oclusão, acessórios (óculos), cor, e brilho de uma mesma face (HSU, 2002).	23
3	(1) Imagem contendo face. (2) Imagem com a representação da probabilidades de existência de pele. (3) Binarização da imagem. (4) Segmento correspondente à face com características faciais (CARVALHO; TAVARES, 2005).	24
4	Características de Haar mais utilizadas em detecção de faces (VIOLA; JONES, 2001).	25
5	Relação entre as características de Haar e o contraste natural da face (VIOLA; JONES, 2004).	25
6	Representação da imagem integral (VIOLA; JONES, 2001).	26
7	Região retangular D a ser calculada.	26
8	Classificador em cascata do algoritmo Viola-Jones (VIOLA; JONES, 2004).	27
9	Faces projetadas num espaço n -dimensional (JAIN; ROSS; PRABHAKAR, 2004).	29
10	Representação vetorial de uma imagem.	30
11	Faces (esquerda) e Auto faces (direita) (CDSST, 2012).	31
12	Obtenção do LBP. Em (a) a imagem original. Em (b) uma matriz binária após comparação com pixel central. Em (c) a matriz utilizada para multiplicar as respectivas células dos valores encontrados em (b). E em (d) o resultado dessa multiplicação (Adaptado de (MASCARO; MELLO; SANTOS, 2007)).	32
13	Descritor LBP com janelamento.	33
14	Face média (preto) e variação do primeiro componente principal (cinza) (MILBORROW, 2007).	34
15	Processo de busca do ASM (não são apresentadas todas as formas intermediárias) (MILBORROW, 2007).	35
16	Arquitetura da plataforma Android (ANDROID, 2012).	37

17	Ciclo de vida de uma Activity (ANDROID, 2012).	39
18	Versões da plataforma Android mais utilizadas pelos usuários AppBrain (APP-BRAIN, 2012).	41
19	Teclado físico de dispositivo móvel.	43
20	Utilização do <i>software</i> BrailleTouch (FREY; SOUTHERN; ROMERO, 2011). . . .	44
21	Alfabeto baseado em gestos do Graffiti (ACCESS-COMPANY, 2010-2011).	45
22	Diagrama de casos de uso do Android Eye.	47
23	Processo de reconhecimento do Sistema Android Eye.	49
24	Dispositivo móvel utilizado: Samsung Galaxy 5, com Android 2.2.	50
25	Amostras extraídas de um mesmo indivíduo da base MoBio.	51
26	Faces detectadas pelos algoritmos de detecção Viola-Jones e zonas de pele, respectivamente.	52

Lista de Tabelas

1	Comparação entre tecnologias biométricas (JAIN; ROSS; PRABHAKAR, 2004). .	19
2	Versões correntes da plataforma Android (ANDROID, 2012).	40
3	Resultados do experimento de detecção de faces.	52
4	Resultados do experimento de reconhecimento de faces.	53
5	Resultados do experimento de acurácia do sistema Android Eye.	54
6	Resultados do experimento de distância do Android Eye.	55
7	Resultados do Teste de Usabilidade.	55

Lista de Siglas

API - *Application Programming Interface*

ASM - *Active Shape Model*

DNA - *Deoxyribonucleic Acid*

FN - Falso Negativo

FP - Falso Positivo

IBGE - Instituto Brasileiro de Geografia e Estatística

J2SE - *Java 2 Standard Edition*

JNI - *Java Native Interface*

LBP - *Local Binary Pattern*

NDK - *Native Development Kit*

PC - *Personal Computer*

PCA - *Principal Component Analysis*

RGB - *Red Green Blue*

SDK - *Software Development Kit*

URL - *Uniform Resource Locator*

Sumário

1	Introdução	14
1.1	Problema	15
1.2	Justificativa	16
1.3	Objetivos	16
1.3.1	Objetivo Geral	16
1.3.2	Objetivos Específicos	17
2	Biometria	18
2.1	Biometria	18
2.1.1	Sistema Biométrico de Reconhecimento	20
2.2	Considerações Finais	21
3	Detecção Facial	22
3.1	Detecção de Zonas da Pele	23
3.2	Viola-Jones	24
3.3	Considerações Finais	28
4	Extração de Características	29
4.1	Análise dos Componentes Principais	30
4.2	Padrões Binários Locais	31
4.2.1	LBP com Janelamento	33
4.3	Modelo de Forma Ativa	33
4.3.1	Geração do Modelo	33
4.4	Considerações Finais	35

SUMÁRIO	12
5 Aplicações Móveis	36
5.1 Plataforma Android	36
5.1.1 Android SDK	38
5.1.2 Android NDK	38
5.1.3 Aplicação	38
5.1.4 Android API	39
5.1.5 Ambiente de Desenvolvimento	41
5.2 Considerações Finais	41
6 Acessibilidade	42
6.1 Teclado Físico	42
6.2 Sintetizadores de Voz	43
6.3 Leitores de Tela	43
6.4 Braille	44
6.5 Gestos	44
6.6 Reconhecimento de Fala	45
6.7 Considerações Finais	45
7 Aplicação Proposta: Android Eye	46
7.1 Arquitetura	46
7.2 Interface com o Usuário	46
7.3 Cadastro das Faces	47
7.4 Fase de Treinamento	48
7.5 Reconhecimento das Faces	48
7.6 Considerações Finais	49
8 Resultados Experimentais	50
8.1 Dispositivo Móvel Utilizado	50
8.1.1 Base MoBio	50
8.2 Avaliação das Técnicas de Detecção das Faces	51
8.3 Avaliação das Técnicas de Reconhecimento de Faces	52
8.4 Seleção das Técnicas Utilizadas	54

8.5	Avaliação do Sistema Android Eye	54
8.6	Conclusões Finais	55
9	Conclusão	56
	Referencias	57

Capítulo 1

Introdução

A deficiência visual trata-se de uma situação irreversível de redução da resposta visual. Pessoas que sofrem desse tipo de deficiência encontram diversos obstáculos para interagir com as outras pessoas, além de possuírem uma dificuldade muito maior para realizarem suas atividades do dia a dia. De acordo com a intensidade da deficiência, ela pode ser leve, moderada, profunda, severa ou total, sendo que a deficiência visual do tipo severa é a que possui a maior intensidade, antes da perda total da visão.

Segundo o censo do IBGE realizado em 2010, no Brasil, cerca de 7 milhões de pessoas são afetadas pela deficiência visual severa, sendo o tipo de deficiência que mais incide sobre a população brasileira. Esse grande número de deficientes visuais justifica o desenvolvimento de tecnologias que possam auxiliar essas pessoas em suas atividades cotidianas, de forma a melhorar sua integração na sociedade, proporcionando-lhes uma melhor qualidade de vida.

Dentre as tecnologias que mais têm se desenvolvido nos últimos anos destacam-se aquelas relacionadas aos dispositivos móveis, como os *smartphones* e os *tablets*. Os dispositivos móveis são computadores portáteis que possuem um sistema operacional próprio, com muitas das funcionalidades de um computador propriamente dito, e ainda com a vantagem de terem dimensões e peso reduzidos, permitindo uma mobilidade e praticidade muito maior.

Este trabalho de conclusão de curso propõe o desenvolvimento de um sistema de reconhecimento de pessoas por meio do reconhecimento facial utilizando dispositivos móveis, para ser utilizado prioritariamente por pessoas que apresentem deficiência visual. Para tanto, o sistema foi desenvolvido com uma interface totalmente voltada para que os recursos da mesma sejam adequados a esse tipo de usuário.

Vale salientar que dentre as diversas aplicações de biometria facial Android encontradas, todas implementam um sistema de autenticação de usuários, que indica a autenticidade ou não de um usuário. Contudo, neste trabalho foi desenvolvido um sistema biométrico de reconhecimento, capaz de determinar a identidade de um indivíduo presente numa base de dados.

Dentre as diversas técnicas de biometria, o reconhecimento facial foi escolhido por ser uma das técnicas mais aceitas pelas pessoas, podendo ser realizada a distância, sem a necessidade de

colaboração do indivíduo a ser reconhecido, além de ser coletada pelas câmeras embutidas nos dispositivos móveis, fazendo com que seja necessário que o deficiente visual possua somente um dispositivo móvel com a aplicação desenvolvida instalada no aparelho para que possa usufruir de suas funcionalidades sempre que julgar necessário.

Para a implementação desse sistema foram avaliadas algumas técnicas de detecção de faces baseadas na cor da pele (ZHAO et al., 2003) e nas características de *Haar*, conforme proposto por (VIOLA; JONES, 2001). Para a representação e descrição das faces foram avaliadas as técnicas baseadas na análise das componentes principais (*PCA - Principal Component Analysis*) (TURK; PENTLAND, 1991), devido a sua ampla utilização, nos padrões binários locais (*LBP - Local Binary Pattern*) (AHONEN; HADID; PIETIKAINEN, 2006), técnica também conhecida como Transformada *Census*, devido à sua maior invariância às alterações de iluminação.

O sistema proposto é uma extensão do sistema BioMobile (FARINA, 2011), desenvolvido recentemente pelo grupo de pesquisa Recogna, do Departamento de Computação, da Faculdade de Ciências, UNESP, campus de Bauru.

1.1 Problema

O último censo realizado em 2010 pelo Instituto Brasileiro de Geografia e Estatística (IBGE), aponta que mais de 12 milhões de brasileiros possuem algum tipo de deficiência severa, cerca de 6,7% da população brasileira. Dentre as deficiências apontadas, a deficiência visual destacou-se como a mais incidente, atingindo cerca de 3,5% da população (IBGE, 2011).

A deficiência visual é a perda ou a redução da capacidade visual em ambos os olhos, não havendo a possibilidade de ser melhorada ou corrigida. Ela é dividida em diversos graus, podendo ser leve, moderada, profunda, severa e até total. A deficiência visual severa é a que possui maior intensidade antes da perda total da visão.

Com a ausência do sentido da visão, o deficiente possui muitos obstáculos no seu cotidiano. Como o ato de orientar-se, circular entre as pessoas, desviar de obstáculos presentes pelo caminho, utilizar computadores e a interação com outras pessoas, uma vez que a deficiência visual torna muito difícil a identificação de quem está ao seu redor.

Usualmente, o deficiente visual utiliza na tarefa de identificar as pessoas ao seu redor, o reconhecimento da voz. Contudo, identificar outras pessoas somente através da voz, não é um processo muito confiável, uma vez que a voz é algo possível de ser forjado com uma certa facilidade, além de necessitar da colaboração do indivíduo a ser identificado. Surge então a necessidade de meios mais confiáveis de identificação que possam ser utilizados pelos deficientes visuais.

São diversos os métodos artificiais de identificação de pessoas, como a identificação por meio da digital, íris, retina, entre outros, porém um método eficiente e que não necessita da interação da pessoa a ser identificada, é o reconhecimento facial.

O reconhecimento facial, apesar das vantagens apresentadas, é um método mais amplamente utilizado em máquinas de grande porte e sistemas avançados, sem praticidade alguma para o deficiente visual. Atualmente poucos sistemas de reconhecimento facial foram desenvolvidos para dispositivos portáteis, porém suas interfaces não foram feitas para auxiliar as pessoas afetadas pela

deficiência visual, uma vez que não eram os objetivos das respectivas aplicações.

Para que haja uma melhor interação entre os deficientes visuais e as demais pessoas, é necessário o desenvolvimento de aplicações práticas e eficientes de reconhecimento facial, com interfaces voltadas para assistência dessas pessoas, garantindo assim uma maior acessibilidade para os mesmos.

1.2 Justificativa

Como mencionado um dos principais problemas enfrentados particularmente pelos deficientes visuais consiste na identificação das pessoas que estão ao seu redor, quando não há possibilidade de interação pela voz. Portanto, o desenvolvimento de tecnologias que possam auxiliar pessoas com deficiência visual a identificar automaticamente as pessoas ao seu redor são de grande importância.

Em geral, o desenvolvimento de tecnologias assistivas voltadas a deficientes apresenta duas vertentes (SANTAROSA, 2000):

- **Próteses Físicas:** incluem o conjunto de dispositivos e procedimentos que têm por objetivo o desempenho de funções que o corpo não pode ou tem dificuldades de executar devido a deficiência;
- **Próteses Mentais:** objetivam o desenvolvimento cognitivo, socioafetivo e de comunicação do deficiente.

Em muitos procedimentos é difícil estabelecer limites que categorizem uma prótese como sendo física ou mental, uma vez que elas atuam em ambas as áreas (SANTAROSA, 2000). Um sistema de auxílio ao reconhecimento de pessoas através da face, por exemplo, pode ser categorizado como uma prótese física já que possibilita o reconhecimento facial a distância, impossível ao deficiente visual, e como uma prótese mental já que favorece uma interação social de forma mais natural ao indivíduo.

Por isso, com a grande quantidade de pessoas afetadas pela deficiência visual no país, é necessário que existam cada vez mais próteses físicas e mentais para auxílio dessas pessoas na integração com a sociedade e, com o desenvolvimento contínuo dos *smartphones* e *tablets*, dispositivos que, além de possuírem um pequeno porte, possuem uma facilidade e praticidade em seu uso, além de um preço cada vez mais acessível. Fazendo com que seja possível o desenvolvimento de diversas aplicações para dispositivos móveis com tal finalidade.

1.3 Objetivos

1.3.1 Objetivo Geral

Desenvolver um sistema de identificação de pessoas por meio do reconhecimento facial em dispositivos móveis, para ser utilizado prioritariamente por pessoas que apresentam deficiência visual, através de uma interface totalmente baseada na voz.

1.3.2 Objetivos Específicos

Para o desenvolvimento deste sistema foram estabelecidos alguns objetivos, que foram distribuídos entre os alunos responsáveis pelo desenvolvimento do projeto.

- Conhecer o modo pelo qual é desenvolvido um aplicativo para dispositivos móveis (plataforma Android).
- Implementar a técnica de detecção de faces baseada na cor da pele.
- Implementar a técnica de detecção de faces baseada nas características de *Haar*.
- Implementar a técnica de representação e descrição de faces baseada na análise de componentes de principais (*PCA - Principal Component Analysis*).
- Implementar a técnica de representação e descrição de faces baseada nos padrões binários locais (*LBP - Local Binary Pattern*), técnica também conhecida como Transformada *Census*, devido à sua maior invariância às alterações de iluminação.
- Implementar a técnica de representação e descrição de faces baseada no modelo de forma ativa (*ASM - Active Shape Model*).
- Avaliar quais técnicas são as mais adequadas para dispositivos móveis, levando em consideração o desempenho do dispositivo (visto que dispositivos móveis possuem memória e processamento reduzidos) e a qualidade dos resultados, separando-as por tipo de técnica (qual a melhor das técnicas de detecção de faces e qual a melhor das técnicas de representação e descrição de faces).
- Verificar interfaces para aplicações em dispositivos móveis que sejam voltadas para o auxílio à deficientes visuais.
- Estudar métodos automáticos de pesquisa online (em redes sociais e afins) de imagens de faces para a etapa de cadastramento de pessoas no sistema.
- Projetar um aplicativo para a plataforma Android que tenha a função de identificar pessoas por meio do reconhecimento facial, possuindo uma interface com a função de assistência ao deficiente visual e utilizando as técnicas melhor avaliadas.
- Realizar, testes do aplicativo em condições que simulem a situação para a qual o aplicativo deve funcionar.

Todos os objetivos propostos foram desenvolvidos em conjunto pelos alunos Alan Zanoni Peixinho e Everton Fernandes da Silva. Exceto pela etapa de desenvolvimento, que foi dividida em 4 módulos independentes, divididos da seguinte forma:

1. Detecção de Faces: Desenvolvido pelo aluno Everton Fernandes da Silva;
2. Reconhecimento de Faces: Desenvolvido pelo aluno Alan Zanoni Peixinho;
3. Cadastro de Faces: Desenvolvido pelo aluno Everton Fernandes da Silva;
4. Interface com o Usuário: Desenvolvido pelo aluno Alan Zanoni Peixinho.

Capítulo 2

Biometria

Este capítulo apresenta uma revisão bibliográfica abordando conceitos básicos relacionados a biometria, enfatizando a biometria facial.

2.1 Biometria

A Biometria é definida por (JAIN; ROSS; PRABHAKAR, 2004) como o reconhecimento pessoal baseado em características comportamentais e/ou fisiológicas de um indivíduo.

Diversas são as técnicas biométricas disponíveis na literatura, entre elas: impressões digitais, retina, geometria das mãos, assinatura e face, que foi escolhida para o desenvolvimento deste projeto.

A biometria pode utilizar qualquer tipo de característica fisiológica e/ou comportamental, desde que satisfaça as seguintes condições (JAIN; ROSS; PRABHAKAR, 2004):

- Universalidade: cada pessoa deve possuir a característica;
- Singularidade: quaisquer duas pessoas devem ser suficientemente diferentes em termos da característica;
- Permanência: a característica deve ser suficientemente invariante ao longo do tempo;
- Coletabilidade: a característica deve ser facilmente coletável;
- Desempenho: se refere à precisão e velocidade de reconhecimento;
- Aceitabilidade: indica o grau em que as pessoas estão dispostas a aceitar o uso de um identificador biométrico específico em suas vidas diárias;
- Evasão: reflete a facilidade com que o sistema pode ser enganado através de métodos fraudulentos.

A Tabela 1 apresenta um comparativo entre as principais características biométricas disponíveis na literatura.

Tabela 1: Comparação entre tecnologias biométricas (JAIN; ROSS; PRABHAKAR, 2004).

(A = Alta, M = Média, B = Baixa)

Biometria	Universalidade	Singularidade	Permanência	Coletabilidade	Desempenho	Aceitabilidade	Evasão
Face	A	B	M	A	B	A	B
Impressão Digital	M	A	A	M	A	M	A
Geometria da mão	M	M	M	A	M	M	M
Teclar (senhas)	B	B	B	M	B	M	M
Veias da mão	M	M	M	M	M	M	A
Iris	A	A	A	A	A	B	A
<i>Scanner</i> da Retina	A	A	A	B	A	B	A
Assinatura	B	B	B	A	B	A	B
Voz	M	B	B	M	B	A	B
Termógrafo Facial	A	A	B	A	M	A	A
Odor	A	A	A	B	B	M	B
DNA	A	A	A	B	A	A	A
Padrão de Digitação	M	B	B	A	B	A	M
Cognitiva	M	M	A	M	M	A	M

Entre as principais vantagens da face podemos destacar a alta aceitabilidade em relação as demais características, já que a face é o principal recurso que utilizamos para a identificação de indivíduos. Além disso, ela pode ser coletada a distância sem a interação do indivíduo que se busca reconhecer.

Existem duas abordagens distintas no desenvolvimento de um sistema biométrico. São elas: a autenticação e o reconhecimento (JAIN; ROSS; PRABHAKAR, 2004).

- **Autenticação:** Após extraídas as características biométricas de um indivíduo elas são comparadas com as amostras previamente coletadas deste mesmo indivíduo de forma a estabelecer a sua autenticidade (ou não). Trata-se de uma técnica de comparação do tipo 1 para 1;
- **Reconhecimento:** Após extraídas as características biométricas de um indivíduo elas são comparadas com as amostras previamente coletadas de cada indivíduo presente numa base, apontando sua provável identidade. Trata-se de uma comparação do tipo 1 para n , onde n indica o número de indivíduos cadastrados na base.

A Figura 1 apresenta a diferença entre sistemas biométricos de autenticação e reconhecimento.

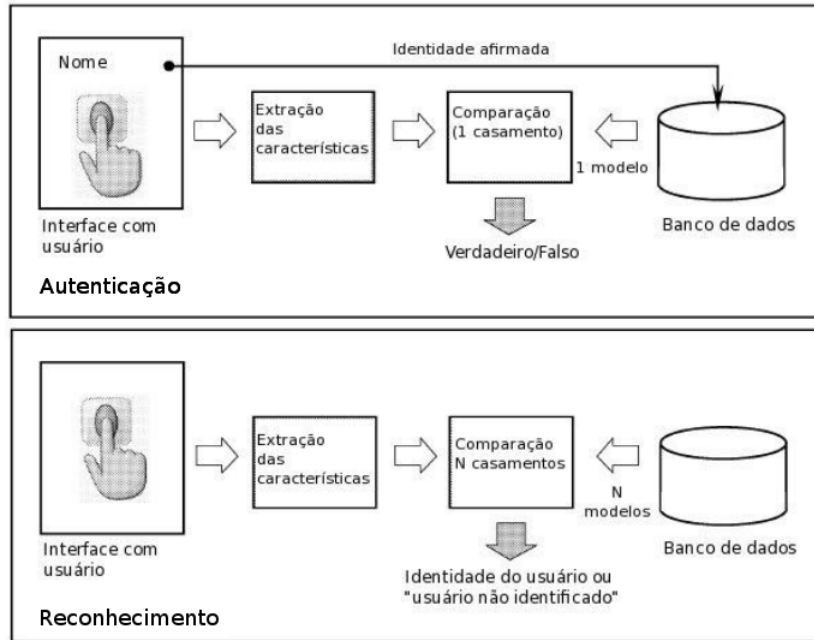


Figura 1: Modos de operação de sistemas biométricos (JAIN; ROSS; PRABHAKAR, 2004).

2.1.1 Sistema Biométrico de Reconhecimento

Nesse trabalho é utilizada uma abordagem de reconhecimento, na qual, dada uma amostra busca-se sua provável identidade. Para isso é calculada a distância da amostra que se deseja reconhecer em relação a todas as amostras presentes numa base rotulada (onde é conhecida a identidade dos indivíduos), o casamento que apresentar a menor distância indica a identidade da amostra submetida ao processo de reconhecimento.

Para determinar o grau de acerto de uma técnica biométrica de reconhecimento pode-se utilizar a taxa de acerto t , dada pela Equação 1:

$$t = \frac{n_{corretas}}{n_{total}} * 100 \quad (1)$$

Onde $n_{corretas}$ representa o número de amostras corretamente reconhecidas e n_{total} representa o número total de amostras submetidas ao processo de reconhecimento.

Pode-se ainda calcular a taxa de acerto considerando-se um Rank- x .

O Rank- x mede a taxa de acerto do algoritmo considerando-se não apenas o indivíduo mais próximo, mas os x indivíduos com a menor distância em relação a amostra, caso qualquer desses x indivíduos represente um casamento correto é considerado um acerto.

O Rank-1 equivale a taxa de acerto simples. Já o Rank- n_{base} , onde n_{base} representa o número de indivíduos que compõem a base, sempre apresenta uma taxa de 100%.

2.2 Considerações Finais

Este capítulo apresentou uma breve introdução ao conceito de biometria, onde foram apresentadas as principais características consideradas na implementação de qualquer sistema biométrico, bem como as diferentes abordagens de utilização da biometria, em especial os sistemas biométricos de reconhecimento, utilizado neste trabalho.

Capítulo 3

Detecção Facial

Este capítulo busca introduzir o leitor aos conceitos básicos a respeito da detecção de faces, bem como apresentar as técnicas utilizadas neste trabalho.

A primeira etapa num sistema de biometria facial consiste da detecção da face numa imagem, seja ela bidimensional ou tridimensional.

Grande é a importância da etapa de detecção de faces, já que todo o sistema depende de uma boa detecção da face a ser autenticada/reconhecida.

Nos anos 90, a maior parte dos estudos em detecção da face buscava a detecção de uma única face na imagem, com vários graus de complexidades de fundo (*background*), utilizando modelos de face inteira, modelos de deformação, cor da pele e redes neurais (VIOLA; JONES, 2001).

Na década seguinte, Viola e Jones (2001) propuseram uma técnica de detecção de faces baseada na ocorrência de determinadas características na imagem denominadas características de Haar, possibilitando a detecção de diversas faces numa mesma imagem.

Contudo ainda existem alguns problemas, apresentados na Figura 2, que dificultam a detecção de faces (YANG; KRIEGMAN; AHUJA, 2002):

- **Pose:** existem variações quanto ao posicionamento da câmera e/ou da face durante a aquisição das imagens.
- **Expressão Facial:** a expressão facial influencia significativamente na aparência da imagem.
- **Presença de Elementos Estruturais:** a presença de alguns elementos como barba, bigode e óculos podem modificar as características em termos de tamanho, luminosidade, entre outros.
- **Ocultação:** em ambientes não controlados as faces podem aparecer parcial, ou totalmente, ocultas, seja por objetos ou por outras faces.
- **Iluminação:** excesso, ou ausência de iluminação também podem dificultar o processo de detecção da face.

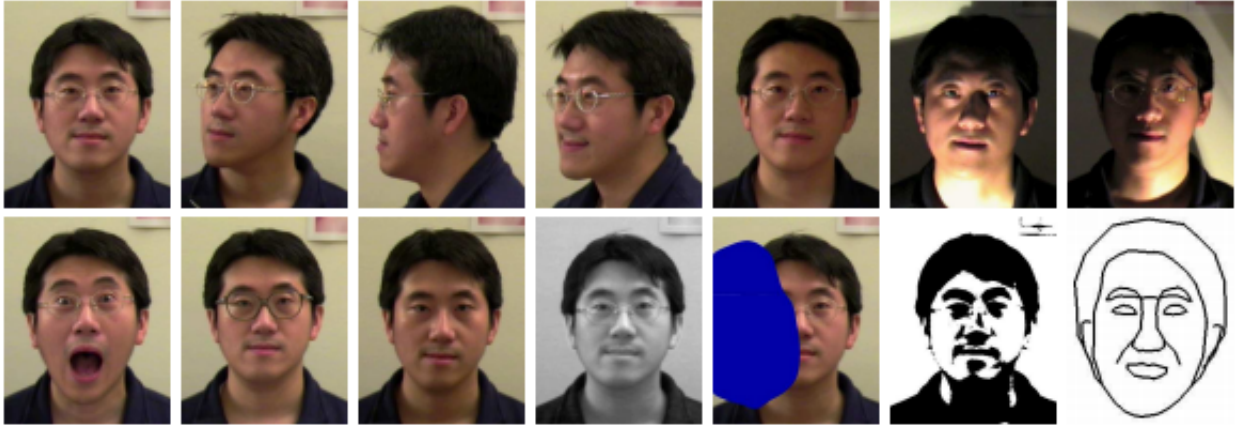


Figura 2: Variações na pose, iluminação, expressão, oclusão, acessórios (óculos), cor, e brilho de uma mesma face (HSU, 2002).

Para o desenvolvimento deste projeto foram avaliadas as técnicas de detecção de faces baseadas nas zonas da pele (JONES; REHG, 1999) e Viola-Jones (VIOLA; JONES, 2001).

3.1 Detecção de Zonas da Pele

A técnica de detecção de zonas da pele (JONES; REHG, 1999; ZHENG; DAOUDI, 2004) utiliza um modelo representativo da pele, construído a partir de um conjunto de imagens de treino contendo amostras de pele retiradas de diversos pontos da face, de vários indivíduos de diferentes idades, etnias e de ambos os sexos (CARVALHO; TAVARES, 2005), por meio do qual é capaz de determinar numa dada imagem a localização de zonas de pele (como a face).

Neste algoritmo é utilizado o espaço de cores YCbCr (ou YUV) em que Y representa a intensidade, Cb representa a componente azul e Cr a componente vermelha.

A conversão do espaço de cores RGB (usualmente utilizado) para o espaço YCbCr é realizado através da Equação 2:

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.144 \\ -0.169 & -0.331 & 0.5 \\ 0.5 & -0.419 & -0.081 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2)$$

O modelo de zona da pele é construído por meio de várias amostras de imagens que apresentam um recorte da pele da face. Calcula-se então a média e a variância de cada uma das componentes (YCbCr), dessa forma considera-se uma distribuição normal (Gaussiana), tornando possível determinar a probabilidade de um dado pixel pertencer a pele.

A partir do modelo construído, é calculada então a probabilidade de cada um dos *pixels* da imagem pertencer a uma zona de pele. Após isso é escolhido um limiar para a binarização da imagem.

A imagem binarizada é então segmentada, onde cada um dos segmentos gerados representa uma potencial face. Durante o processo de segmentação, diversas características podem ser levadas em conta para o refinamento da análise de uma potencial face, sendo as principais:

- Porcentagem de *pixels* detectados como pele;
- Probabilidade média de *pixels* de pele;
- Tamanho em *pixels* do maior componente de pele conectado;
- Número de componentes de pele conectados;

Por fim, é verificada a existência de segmentos de intensidade nula inseridos na provável face, que podem ser representantes de olhos, boca ou sobrancelhas. Aqueles segmentos que não possuam tal requisito são descartados. O processo de detecção é ilustrado pela Figura 3.



Figura 3: (1) Imagem contendo face. (2) Imagem com a representação da probabilidades de existência de pele. (3) Binarização da imagem. (4) Segmento correspondente à face com características faciais (CARVALHO; TAVARES, 2005).

Após essa etapa é realizada uma intersecção do segmento de face com a imagem em tons de cinza, essa imagem é então comparada a uma face canônica utilizando correlação cruzada, que compara uma dada imagem com uma sub imagem adequada, traduzindo numericamente a semelhança existente entre as duas (SCHALKOFF, 1989). De forma que podemos diminuir a ocorrência de detecções errôneas.

3.2 Viola-Jones

O algoritmo Viola-Jones é considerado o estado da arte em termos de detecção de faces, devido a sua eficácia e eficiência, já que apresenta altas taxas de detecção, com um algoritmo rápido, que não necessita de informações de cor ou movimento.

O algoritmo consiste da busca de determinados padrões numa imagem, para a detecção de faces são utilizadas as características de Haar mostradas na Figura 4.

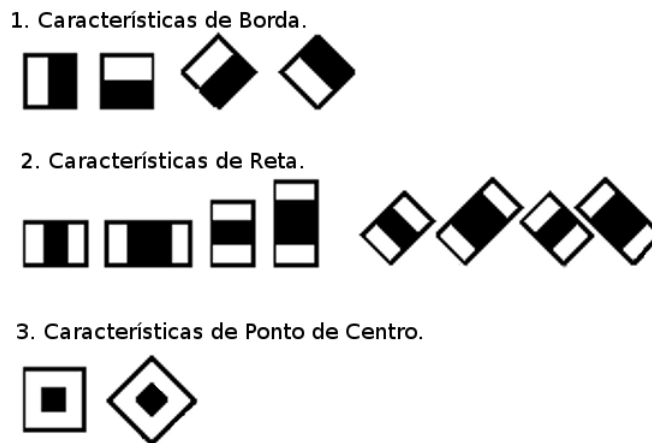


Figura 4: Características de Haar mais utilizadas em detecção de faces (VIOLA; JONES, 2001).

A Figura 5 mostra a comparação das características de Haar com os contrastes naturais da face.

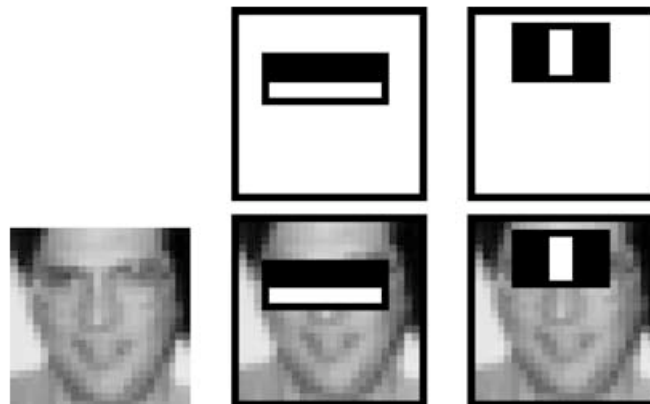


Figura 5: Relação entre as características de Haar e o contraste natural da face (VIOLA; JONES, 2004).

Para determinar o grau de semelhança entre a característica e uma determinada região da imagem utiliza-se a Equação 3.

$$f(w) = \sum p - \sum b \quad (3)$$

Onde $\sum p$ indica o somatório dos *pixels* na região preta e $\sum b$ indica o somatório dos *pixels* na região branca.

De forma a agilizar o cálculo deste cálculo é utilizada uma estrutura auxiliar *ii*, chamada de imagem integral, na qual o valor de $ii(x,y)$ representa o somatório dos *pixels* acima de *y* e a esquerda de *x*. A imagem integral *ii* é obtida através da imagem *i*, definida pela Equação 4:

$$ii(x,y) = \sum_{j=0}^x \sum_{k=0}^y i(j,k) \quad (4)$$

A imagem integral ii permite realizar o cálculo do somatório de *pixels* numa determinada região em tempo constante. Essa imagem pode ser calculada em um passo sobre a imagem original por meio das seguintes equações:

$$s(x,y) = s(x,y-1) + i(x,y) \quad (5)$$

$$ii(x,y) = ii(x-1,y) + s(x,y) \quad (6)$$

Sendo ii a equação que representa a imagem integral, $i(x,y)$ a equação que representa a imagem original, $s(x,y)$ a soma cumulativa da linha, $s(x,-1) = 0$ e $ii(-1,y) = 0$. A imagem integral é ilustrada na Figura 6:

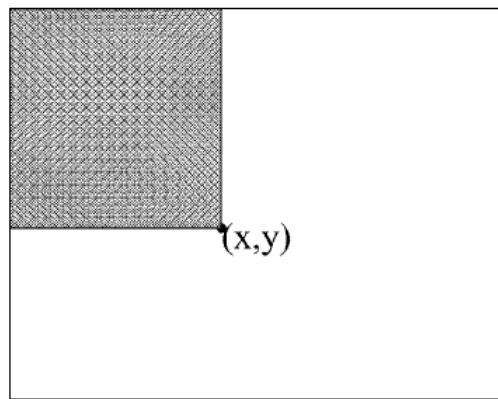


Figura 6: Representação da imagem integral (VIOLA; JONES, 2001).

Com a imagem integral é possível calcular qualquer soma retangular na imagem através dos valores dos pontos da imagem integral que delimita a região retangular a ser calculada.

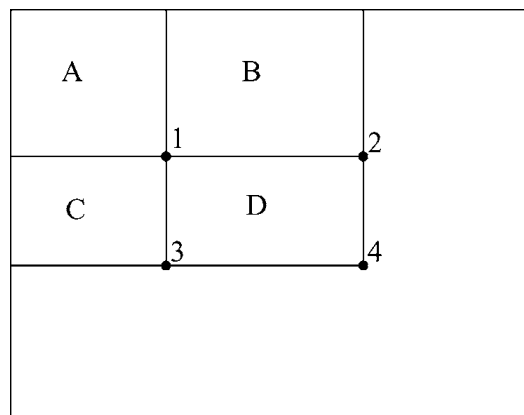


Figura 7: Região retangular D a ser calculada.

Na Figura 7, o vértice 1 corresponde a imagem integral da região A , o vértice 2 corresponde a imagem integral da região $A + B$, o vértice 3 corresponde a imagem integral da região $A + C$ e o vértice 4 corresponde a imagem integral da região $A + B + C + D$. Com essas informações a região D pode ser calculada através das operações de adição e subtração entre os vértices: $4 - 2 - 3 + 1$ ou $ii(4) + ii(1) - ii(2) - ii(3)$.

Cada uma das características de Haar agem como classificadores fracos (já que possuem uma baixa taxa de acerto na detecção de faces). Contudo, utiliza-se o modelo AdaBoost (VIOLA; JONES, 2004), para se obter um classificador forte através da união desses vários classificadores fracos.

De forma a tornar o classificador mais rápido, cada sub-janela considerada da imagem é apresentada a um primeiro classificador, caso este determine que o padrão representa uma face então é invocado o segundo classificador e assim sucessivamente. Se neste processo qualquer um dos classificadores indicar que o padrão não representa uma face, então o procedimento é interrompido e a sub-janela descartada (VIOLA; JONES, 2004). O processo é mostrado na Figura 8.

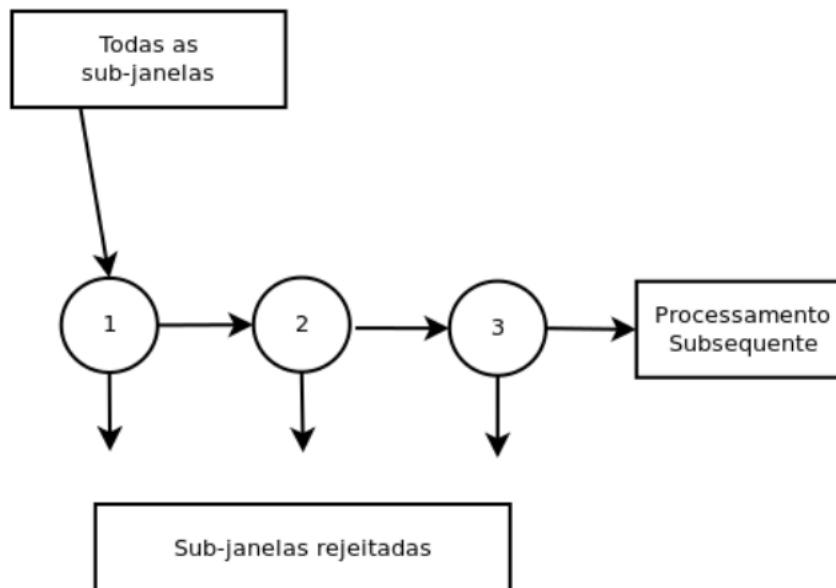


Figura 8: Classificador em cascata do algoritmo Viola-Jones (VIOLA; JONES, 2004).

A escolha dos classificadores que compõem a cascata é feita de tal forma que os primeiros classificadores utilizem um conjunto pequeno de características, e que possuam uma baixa taxa de falsas rejeições (faces reais a serem classificadas como falsas), o segundo nível de classificadores é composto por um maior número de características e uma maior tolerância na taxa de falsas rejeições, além de diminuir a taxa de falsas aceitações. A cada nível os classificadores se tornam mais complexos e mais precisos.

3.3 Considerações Finais

Este capítulo apresentou o leitor aos conceitos básicos de detecção de faces, seus principais problemas e sua importância no contexto de um sistema de reconhecimento baseado na face. Foram ainda apresentados dois algoritmos responsáveis pela detecção de faces em imagens bidimensionais: a detecção de zonas da pele (JONES; REHG, 1999; ZHENG; DAOUDI, 2004) e Viola-Jones (VIOLA; JONES, 2001).

Capítulo 4

Extração de Características

Este capítulo apresenta ao leitor os conceitos da extração de características da face, bem como as técnicas utilizadas neste trabalho.

Após a detecção da face, faz-se necessária a etapa de extração de suas características. Essas características devem representar a face de tal forma que indivíduos distintos tenham valores discrepantes e as características extraídas da face de um mesmo indivíduo tenham valores semelhantes.

Cada uma das faces é representada como um vetor num espaço n -dimensional, onde n indica o número de características extraídas, de tal forma que se possa calcular a distância entre duas faces quaisquer, ou entre uma face e um modelo¹, através de uma função de distância, como mostra a Figura 9.

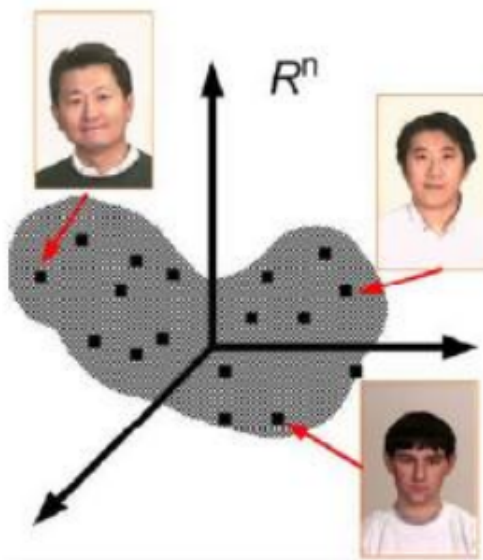


Figura 9: Faces projetadas num espaço n -dimensional (JAIN; ROSS; PRABHAKAR, 2004).

¹Um modelo é construído através da junção de várias amostras.

Existem várias abordagens de obtenção deste vetor de características, como um histograma da imagem, ou simplesmente a intensidade de cada pixel. Contudo existem técnicas mais apropriadas para o reconhecimento de faces.

Neste projeto são estudadas as técnicas baseadas em Análise dos Componentes Principais (*Principal Component Analysis* - PCA) (TURK; PENTLAND, 1991), Padrões Binários Locais (*Local Binary Pattern* - LBP) (AHONEN; HADID; PIETIKAINEN, 2006) e Modelo de Forma Ativa (*Active Shape Model* - ASM) (COOTES et al., 1995).

4.1 Análise dos Componentes Principais

O PCA (*Principal Component Analysis*) (TURK; PENTLAND, 1991) é um algoritmo amplamente utilizado e comumente adotado como base para comparações entre técnicas de reconhecimento de faces. Basicamente, o PCA atua removendo dimensões correlacionadas da imagem, de tal forma que a nova imagem possua apenas as informações mais relevantes, ou seja, aquelas dimensões que apresentam maior variância.

As imagens são representadas como um vetor \mathbf{v}_m , concatenando as linhas da imagem (Figura 10), e arranjadas numa matriz $M_{n \times m}$, onde n representa o número de amostras (imagens) utilizadas, esses dados são centralizados (subtraindo-se a média de cada característica) e então é calculada a matriz de covariância $cov(M)$, segundo a Equação 7:

$$cov(M) = \begin{bmatrix} cov(M_1, M_1) & cov(M_1, M_2) & \cdots & cov(M_1, M_m) \\ cov(M_2, M_1) & cov(M_2, M_2) & \cdots & cov(M_2, M_m) \\ \vdots & \vdots & \ddots & \vdots \\ cov(M_n, M_1) & cov(M_n, M_2) & \cdots & cov(M_n, M_m) \end{bmatrix} \quad (7)$$

Onde M_i representa a i -ésima linha da matriz M .

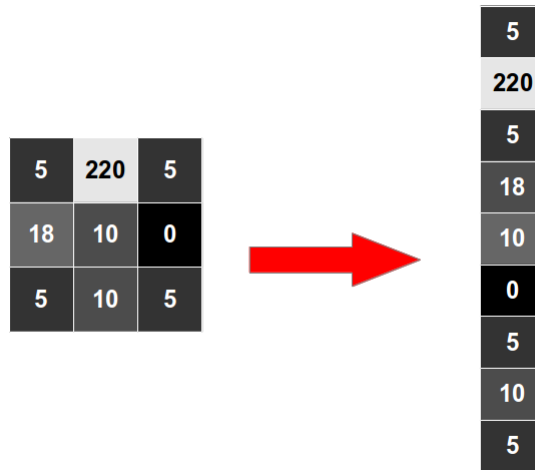


Figura 10: Representação vetorial de uma imagem.

Através da matriz $cov(M)$ são calculados os autovetores correspondentes e gerada uma matriz V tal que o i -ésimo autovetor é associado a i -ésima coluna de V .

A matriz V , matriz de auto faces, configura a matriz de transformação que leva um vetor de imagem de seu espaço de representação habitual ao espaço de representação ótimo, obtido através do PCA. Dessa forma, o vetor obtido da imagem de cada face é multiplicada por V obtendo assim sua representação no novo espaço (Figura 11).



Figura 11: Faces (esquerda) e Auto faces (direita) (CDSST, 2012).

Os autovetores que possuem um baixo autovalor associado (considerados pouco descritivos) podem ser removidos sem grande prejuízo, desconsiderando-se os autovetores a eles relacionados. Isso garante uma redução de dimensionalidade.

Esse subconjunto S de autovetores é escolhido de tal forma que a soma dos autovalores desprezados seja inferior a uma porcentagem p da soma total dos autovalores. Garantindo uma significativa redução dos dados com uma perda de informação inferior a p .

Cada imagem i é representada no espaço definido pela matriz composta pelos autovetores selecionados S , obtendo o descritor u , segundo a Equação 8.

$$u = i * S \quad (8)$$

As faces mapeadas no espaço definido pelo PCA são então comparadas por meio de uma função de distância. Como por exemplo a distância euclidiana definida pela Equação 9.

$$d(u, v) = \sqrt{\sum (u_i - v_i)^2} \quad (9)$$

4.2 Padrões Binários Locais

O LBP (*Local Binary Pattern*), desenvolvido por Ahonen, Hadid e Pietikainen (2006), também chamado Transformada Census, é conhecido pelo seu ótimo desempenho como descritor de tex-

turas e por isso tem sido amplamente adotado em vários tipos de aplicações e, em particular, para a descrição de faces, considerando que uma face é uma composição de micropadrões (AHONEN; HADID; PIETIKAINEN, 2006).

Normalmente o LBP utiliza como entrada uma imagem em tons de cinza, desconsiderando as informações de cor. Contudo existem abordagens que utilizam essas informações de cor.

O LBP é calculado através de uma análise de cada um dos *pixels* da imagem em relação a sua vizinhança. É realizada uma comparação dos valores dos *pixels* vizinhos com o *pixel* central atribuindo 0 aos vizinhos inferiores e 1 aos superiores. Essa matriz binária é multiplicada por valores pré-definidos. A soma dos valores dessa matriz representa o LBP do *pixel* em questão (MÄENPÄÄ, 2003). A Figura 12 ilustra o processo.

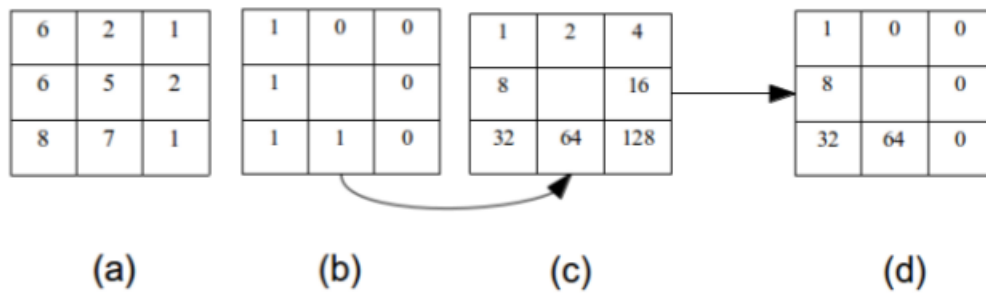


Figura 12: Obtenção do LBP. Em (a) a imagem original. Em (b) uma matriz binária após comparação com pixel central. Em (c) a matriz utilizada para multiplicar as respectivas células dos valores encontrados em (b). E em (d) o resultado dessa multiplicação (Adaptado de (MASCARO; MELLO; SANTOS, 2007)).

Após o cálculo descrito ser realizado para cada pixel da imagem, é gerado um histograma que atua como descritor da imagem, sendo conhecido como descritor LBP (AHONEN; HADID; PIETIKAINEN, 2006).

Por se tratar de um histograma a comparação entre indivíduos é geralmente realizada através da função de distância χ^2 (*Chi-Square*), ao invés da usual distância euclidiana, que é definida pela Equação .

$$\chi^2(u, v) = 0.5 * \sum \frac{(u_i - v_i)^2}{u_i + v_i} \quad (10)$$

Por ser baseado na análise da luminosidade entre pixels vizinhos, o LBP possui a vantagem de ser pouco sensível a mudanças de luminosidade, permitindo a obtenção de imagens em diferentes condições de luminosidade, como ambientes internos e externos.

4.2.1 LBP com Janelamento

A utilização de janelas busca aumentar a eficácia obtida com o LBP. Para tal, a imagem é subdividida em janelas de mesmo tamanho sobre as quais é calculado o histograma LBP, esses histogramas são concatenados gerando-se o descritor da face (AHONEN; HADID; PIETIKÄINEN, 2004), como ilustra a Figura 13.

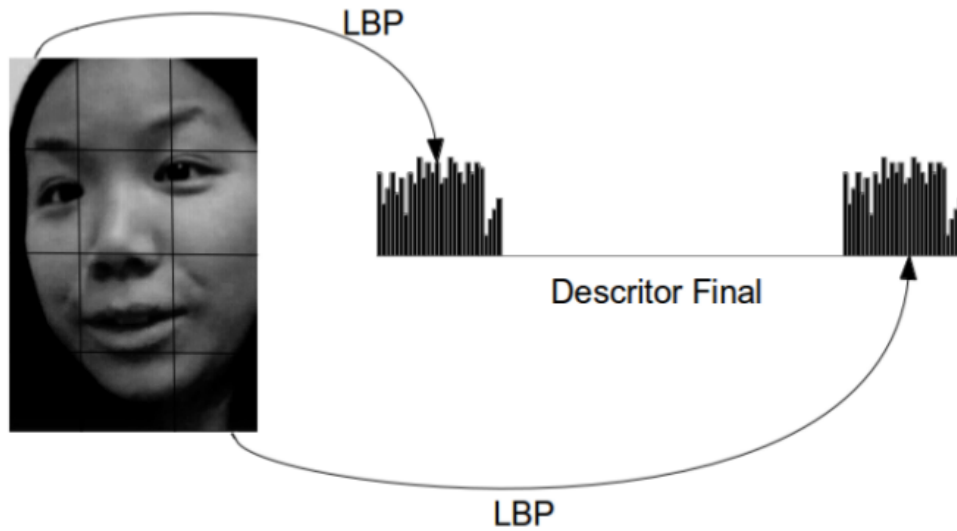


Figura 13: Descritor LBP com janelamento.

4.3 Modelo de Forma Ativa

O *Active Shape Model* (ASM) (COOTES; TAYLOR et al., 2001) tem como objetivo encontrar determinados pontos de referência numa dada imagem (em nosso caso faces) que são utilizados como descritor de forma da imagem em questão.

4.3.1 Geração do Modelo

O primeiro passo para o ASM é desenvolver um modelo com várias faces que tiveram seus pontos de referência marcados manualmente. As amostras desse modelo são alinhadas de tal forma que minimize as distâncias entre as amostras, esse alinhamento é realizado através de rotação (θ), translação (x_t, y_t) e ajuste de escala (s) dos pontos de cada amostra, através da Equação 11.

$$T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_t \\ y_t \end{pmatrix} \begin{pmatrix} s \cos(\theta) & s \sin(\theta) \\ -s \sin(\theta) & s \cos(\theta) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad (11)$$

Após alinhado o modelo, é calculada a média \bar{x} e sua respectiva matriz de covariância S .

Para cada face do modelo são extraídas ainda informações de textura de todos os pontos marcados manualmente, através de uma janela gaussiana, utilizando-se uma abordagem unidimensional ou bidimensional.

Essas informações são utilizadas futuramente durante a busca em novas faces.

De posse desse modelo ele é mapeado num espaço otimizado através da técnica PCA descrita em 4.1, onde é calculada a matriz Φ de autovetores.

Um exemplo de face modelo é apresentado na Figura 14, com sua face média e a variação da primeira componente principal.



Figura 14: Face média (preto) e variação do primeiro componente principal (cinza) (MILBROW, 2007).

Após o PCA a forma da face passa a ser representada segundo a Equação 12:

$$\hat{\mathbf{x}} = \bar{\mathbf{x}} + \Phi \mathbf{b} \quad (12)$$

Para cada nova face apresentada ao método, é posicionada a face média e busca-se minimizar a distância entre a informação de textura de cada ponto e seu correspondente no modelo. A busca da forma da face na imagem é apresentada na Figura 15.



Figura 15: Processo de busca do ASM (não são apresentadas todas as formas intermediárias) (MIL-BORROW, 2007).

Por fim o descritor das faces é formado pela concatenação das componentes x e y , transformando os m pontos contidos no espaço \mathbb{R}^2 num descritor d de dimensão $|d| = 2m$, (Equação 13).

$$d = (x_1, x_2, \dots, x_m, y_1, \dots, y_m) \quad (13)$$

4.4 Considerações Finais

Este capítulo apresentou o leitor aos conceitos básicos a respeito da etapa de extração de características da face, onde foram apresentadas três abordagens distintas de obtenção das informações da face: a Análise dos Componentes Principais (TURK; PENTLAND, 1991), os Padrões Binários Locais (AHONEN; HADID; PIETIKAINEN, 2006) e o Modelo da Forma Ativa (COOTES; TAYLOR et al., 2001).

Capítulo 5

Aplicações Móveis

Este capítulo apresenta uma introdução ao ambiente de desenvolvimento de aplicativos móveis, em especial a plataforma Android.

A partir dos anos 90, houve um grande crescimento no desenvolvimento de tecnologias para comunicação móvel. A popularização dessas tecnologias tem permitido o acesso a informações remotas onde quer que se esteja, abrindo um leque muito grande de facilidades, aplicações e serviços aos usuários (FIGUEIREDO; NAKAMURA, 2003).

Com o crescimento do mercado houve uma grande evolução em termos de *hardware* e *software* para tais plataformas, possibilitando o desenvolvimento de aplicações até então inviáveis.

5.1 Plataforma Android

Neste projeto optou-se pela utilização da plataforma Android, devido ao fato do sistema operacional basear-se num *Kernel Linux*, e portanto tratar-se de um *software* livre e de código aberto, além de estar sendo rapidamente adotado por vários segmentos da sociedade, tais como a comunidade de desenvolvedores de aplicações, a indústria de dispositivos móveis, os fornecedores de *software*, as empresas de telefonia móvel, e muitos outros. Além disso a plataforma adota a linguagem de programação Java que é robusta, de fácil sintaxe, além de contar com um total suporte ao paradigma de orientação a objetos. A Figura 16 ilustra a arquitetura utilizada pela plataforma Android.

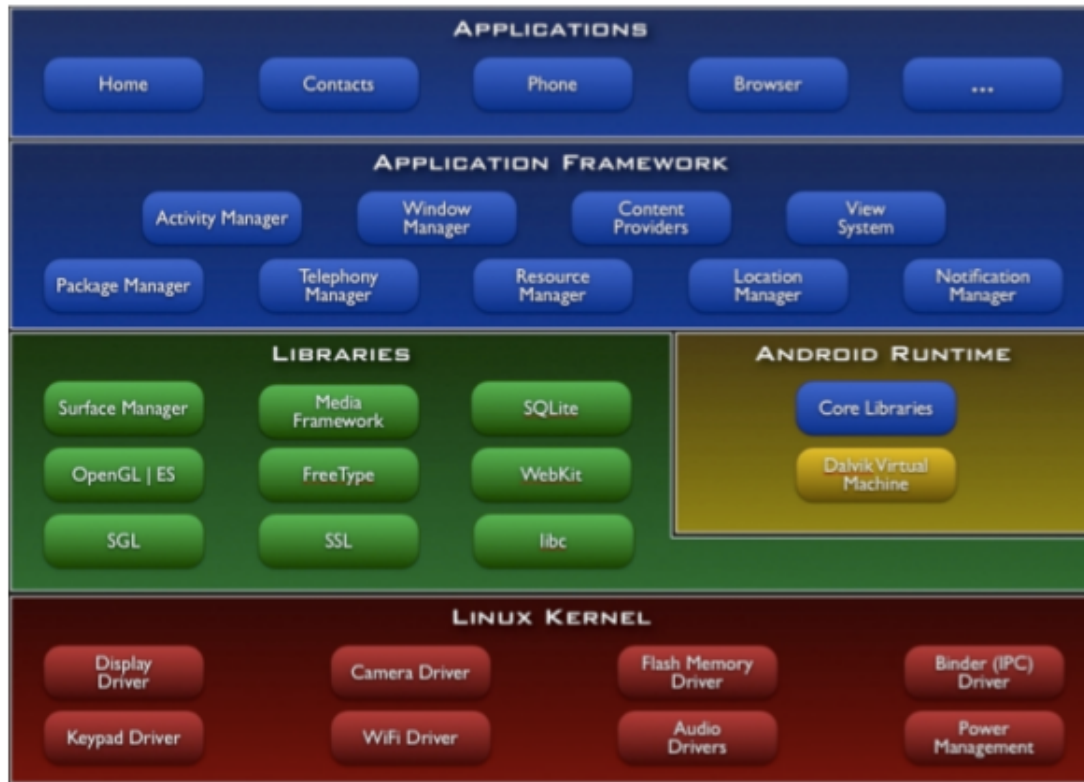


Figura 16: Arquitetura da plataforma Android (ANDROID, 2012).

Essa arquitetura em camadas busca gerar um alto grau de abstração durante o desenvolvimento do *software*, apresentando as seguintes características (ANDROID, 2012):

- **Aplicações:** abriga as aplicações utilizadas pelo usuário, como clientes de e-mail, calendário, jogos, e outras diversas aplicações, que são escritas utilizando a linguagem de programação Java;
- **Framework de Aplicações:** os desenvolvedores são livres para manipular informações do dispositivo de hardware como localização, definir alarmes, adicionar notificações na barra de status, entre outros. É possível utilizar os mesmos recursos que aplicações de núcleo utilizam, e até mesmo substituí-las. Essa camada provê ainda a possibilidade de que qualquer aplicação possa utilizar recursos disponíveis em outras aplicações aumentando significativamente a reutilização de *software*;
- **Bibliotecas:** são disponibilizadas diversas bibliotecas C/C++, utilizadas pelos componentes Android, através da camada de *Framework* da Aplicação. Como por exemplo, bibliotecas de manipulação de mídia, *web* e banco de dados;
- **Ambiente de execução:** inclui a maior parte das funções disponíveis na linguagem Java. Cada aplicação Android roda em seu próprio processo, com sua própria instância da máquina virtual Dalvik, que foi escrita de forma que um dispositivo possa executar várias VMs de forma eficiente, sendo otimizada para consumo de memória mínimo;

- **Kernel Linux:** android utiliza a versão 2.6 do *Kernel Linux* para os serviços centrais do sistema, tais como segurança, gerenciamento de memória, gerenciamento de processos, pilha de rede e modelo de *driver*. O *kernel* também atua como uma camada de abstração entre o hardware e o restante da pilha de *software*.

5.1.1 Android SDK

Como uma forma de facilitar o desenvolvimento de aplicativos para a plataforma, o Android conta com um conjunto de *softwares* de suporte ao desenvolvimento chamado Android SDK (*Software Development Kit*), que conta com bibliotecas, depuradores, emuladores e outras ferramentas que auxiliam na etapa de desenvolvimento.

5.1.2 Android NDK

O desenvolvedor Android conta ainda com o Android NDK (*Native Development Kit*) que disponibiliza uma interface para a utilização de códigos nativos (C/C++) possibilitando sua integração com as bibliotecas Java através da interface JNI (*Java native Interface*).

O NDK possibilita a inclusão de uma gama muito grande de bibliotecas código aberto desenvolvidas nas linguagens de programação C/C++, o que aumenta significativamente o poder da plataforma.

5.1.3 Aplicação

Apesar de utilizar a linguagem de programação Java, as aplicações Android, chamadas *Activities*, tem algumas diferenças em relação ao J2SE (*Java 2 Standard Edition*). O ciclo de execução de uma *Activity*, por exemplo, é semelhante as *applets* Java, contudo possui diferenças que melhor a adaptam a dispositivos móveis. Este ciclo é mostrado na Figura 17.

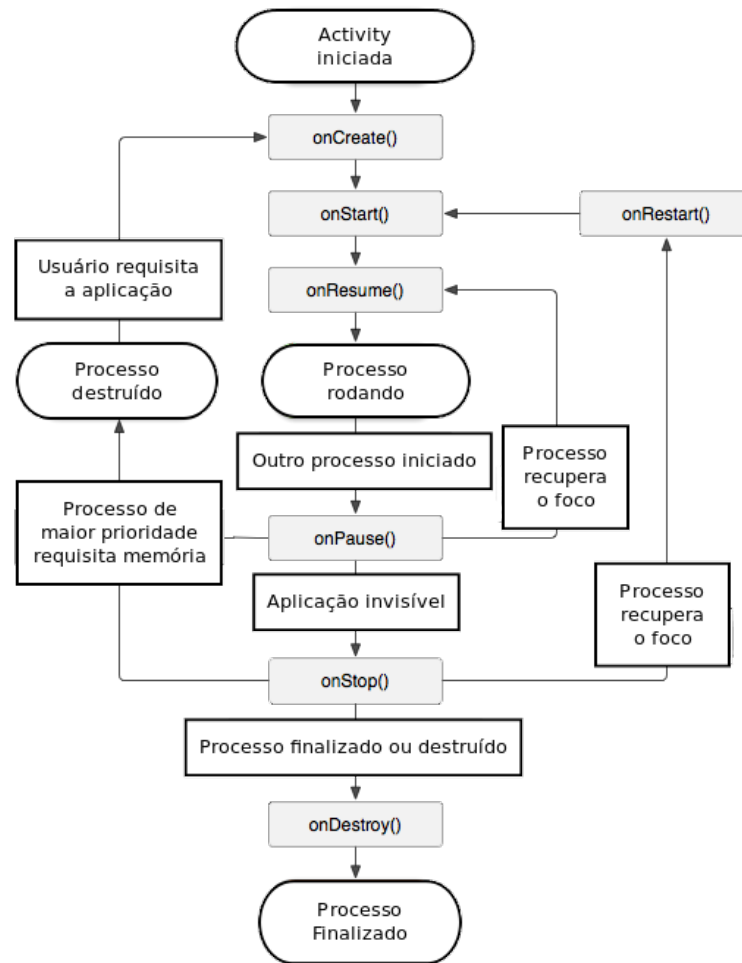


Figura 17: Ciclo de vida de uma Activity (ANDROID, 2012).

Uma característica interessante de se ressaltar é o fato de a plataforma Android não utilizar paginação, ou outro tipo de memória virtual. Logo, quando uma aplicação de maior prioridade (como a aplicação em foco por exemplo) requisita memória, o gerenciador de memória pode matar um processo de menor prioridade para obtê-la. Somado a isso as aplicações contam com uma quantidade bastante limitada de memória, em alguns aparelhos apenas 16Mb estão disponíveis para cada aplicação, o que representa um grande desafio para aplicações mais robustas como a biometria.

5.1.4 Android API

Uma API (*Application Programming Interface*) apresenta um conjunto de rotinas, objetos e estruturas de dados que apresentam uma interface entre o desenvolvedor e detalhes de implementação.

A plataforma Android está em constante desenvolvimento. Dessa forma são realizadas alterações na API de desenvolvimento de forma a prover novas funcionalidades.

As diversas versões da API, chamadas níveis da API, são incrementais, ou seja, um novo nível

contém todas as funções disponíveis nos níveis anteriores, com o acréscimo de novas funcionalidades. A Tabela 2 mostra os diversos níveis da API disponíveis ao desenvolvedor, bem como as versões da plataforma Android que os implementam.

Tabela 2: Versões correntes da plataforma Android (ANDROID, 2012).

Versão da Plataforma	Nível da API	Nome da Versão
Android 4.1, 4.1.1	16	JELLY_BEAN
Android 4.0.3, 4.0.4	15	ICE_CREAM_SANDWICH_MR1
Android 4.0, 4.0.1, 4.0.2	14	ICE_CREAM_SANDWICH
Android 3.2	13	HONEYCOMB_MR2
Android 3.1.x	12	HONEYCOMB_MR1
Android 3.0.x	11	HONEYCOMB
Android 2.3.4	10	GINGERBREAD_MR1
Android 2.3.3		
Android 2.3.2	9	GINGERBREAD
Android 2.3.1		
Android 2.3		
Android 2.2.x	8	FROYO
Android 2.1.x	7	ECLAIR_MR1
Android 2.0.1	6	ECLAIR_0_1
Android 2.0	5	ECLAIR
Android 1.6	4	DONUT
Android 1.5	3	CUPCAKE
Android 1.1	2	BASE_1_1
Android 1.0	1	BASE

Escolher uma API para o desenvolvimento é uma questão importante, pois determina quais recursos poderão ser utilizados durante a etapa de desenvolvimento, além de qual parcela dos usuários Android poderá utilizar o aplicativo a ser desenvolvido. Escolher APIs mais recentes facilita o desenvolvimento, enquanto APIs mais antigas garantem uma maior quantidade de potenciais usuários. A Figura 18 apresenta um estudo desenvolvido pela empresa AppBrain sobre as versões da plataforma Android utilizadas pelos seus usuários.

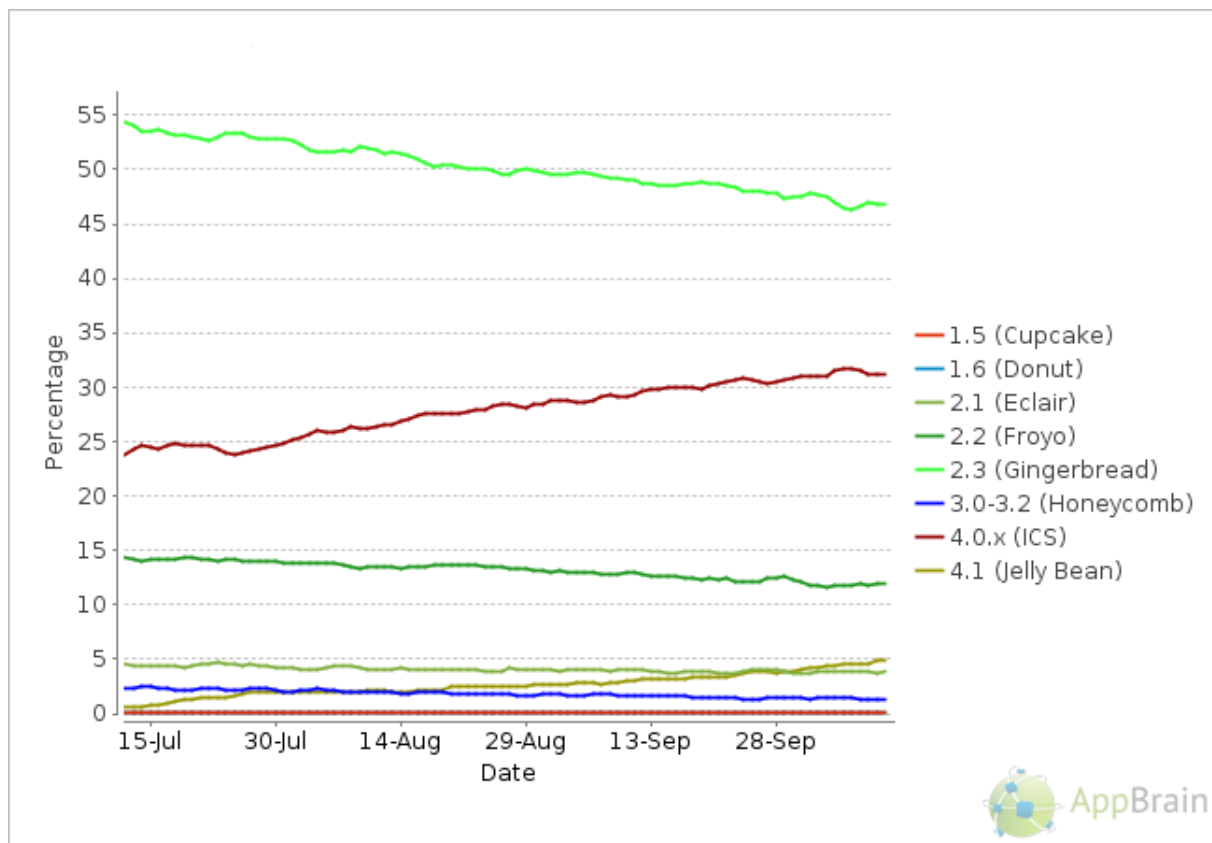


Figura 18: Versões da plataforma Android mais utilizadas pelos usuários AppBrain (APPBRAIN, 2012).

O *software* foi desenvolvido utilizando-se a versão 8 da API Android. Implementada pelos dispositivos que utilizam a versão 2.2 (Froyo) da plataforma Android e superiores.

5.1.5 Ambiente de Desenvolvimento

O projeto foi desenvolvido utilizando-se o ambiente de desenvolvimento multiplataforma MOTODEV da Motorola (MOTOROLA, 2007-2012), disponível gratuitamente para *download* em <http://developer.motorola.com/>, requisitando apenas um cadastro para a utilização.

O MOTODEV é uma adaptação do ambiente de desenvolvimento Eclipse, com ferramentas para o desenvolvimento Android integradas, e um instalador passo a passo que facilita a instalação dos recursos necessários ao Android.

5.2 Considerações Finais

Este capítulo apresentou o leitor ao ambiente de desenvolvimento de aplicações móveis, em especial a plataforma Android. Foram apresentados conceitos importantes como a arquitetura adotada, organização da API e softwares de apoio ao desenvolvimento de aplicações.

Capítulo 6

Acessibilidade

Este capítulo apresenta uma introdução aos conceitos de acessibilidade, bem como algumas soluções que buscam garantir essa acessibilidade ao usuário portador de deficiência visual.

O termo acessibilidade é comumente utilizado como sinônimo de facilidade de uso para deficientes. Contudo seu significado é muito mais amplo.

Segundo Torres e Mazzoni (2004), observar a acessibilidade de um produto consiste em considerar a diversidade de seus possíveis usuários e as peculiaridades da interação dessas pessoas com o produto. Dessa forma a ideia de acessibilidade diz respeito ao quanto um produto atende as necessidades de seus usuários.

Nesse trabalho buscou-se garantir um bom nível de acessibilidade, tendo como público alvo pessoas com deficiência visual, sem contudo excluir possíveis usuários com boa visão. Para tal, foi realizado um estudo a respeito das principais tecnologias que buscam garantir essa acessibilidade aos deficientes visuais no ambiente dos dispositivos móveis.

6.1 Teclado Físico

O teclado físico (Figura 19) presente em boa parte dos dispositivos móveis permite a utilização dos recursos do aparelho através do tato. Contudo, com a popularidade dos dispositivos com telas sensíveis ao toque (*Touch Screen*), o teclado físico já não é mais presente em diversos modelos disponíveis no mercado.



Figura 19: Teclado físico de dispositivo móvel.

Existem ainda modelos de teclados que utilizam o sistema Braille, aumentando o grau de acessibilidade aos deficientes visuais.

6.2 Sintetizadores de Voz

Ferramentas sintetizadoras de voz são uma forma artificial de produzir voz humana, podendo ser implementado em *hardware* ou *software*. Ferramentas de texto para voz (*Text to Speech*) transformam um texto escrito em linguagem falada (ALLEN; HUNNICUTT; KLATT, 1987).

O *software* PicoTTS (SVOX-MOBILE-VOICES, 2008-2012) é uma ferramenta sintetizadora de voz padrão da plataforma Android, capaz de converter qualquer texto em língua inglesa para voz.

Existem ainda outras alternativas para a plataforma como o SVOX Classic TTS Engine (SVOX-MOBILE-VOICES, 2010-2012), que implementa a sintetização de voz para diversas línguas. Contudo, diferente do PicoTTS, trata-se de uma ferramenta não gratuita.

6.3 Leitores de Tela

Os leitores de tela são a forma mais usual de retorno ao deficiente visual na utilização de computadores ou dispositivos móveis. Eles são *softwares* que utilizam ferramentas de sintetização de voz para transmitir através do áudio o conteúdo que é apresentado em forma de texto na tela do dispositivo em uso.

Esse tipo de tecnologia também acaba por guiar o usuário apresentando algumas informações

visuais não escritas, como por exemplo o tipo de elemento de interface (botão, caixa de texto) que está sob o foco do usuário no momento, bem como seu estado corrente, possibilitando, por exemplo, a adaptação de teclados virtuais de dispositivos sensíveis ao toque.

6.4 Braille

O sistema de escrita Braille foi desenvolvido por Louis Braille em 1825, baseado nos trabalhos de Carlos Barbier, que buscava desenvolver um sistema para troca de mensagens militares, utilizando matrizes de pontos. Louis Braille adaptou este sistema para que pudesse ser facilmente utilizado por deficientes visuais, sendo hoje adotado como o principal meio de comunicação escrita entre os deficientes visuais (BAPTISTA, 2000).

O Braille também é utilizado como ferramenta de entrada de informações, mesmo em dispositivos sem um teclado físico. O *software* BrailleTouch (FREY; SOUTHERN; ROMERO, 2011), por exemplo, reconhece caracteres Braille digitados através do toque, como apresentado na Figura 20. Permitindo uma forma bastante rápida para a entrada de informações em forma de texto.



Figura 20: Utilização do *software* BrailleTouch (FREY; SOUTHERN; ROMERO, 2011).

6.5 Gestos

A utilização de gestos ao invés de toques é uma forma alternativa de utilização das telas sensíveis ao toque, que garante uma maior acessibilidade ao usuário, visto que não demanda precisão na localização do toque.

Existem abordagens que mapeiam determinadas funções com gestos específicos como por exemplo, mover o dedo para baixo, ou desenhar um círculo.

O *software* Eyes Free Shell (EYESFREE-PROJECT, 2009-2012), mapeia o menu principal dos dispositivo Android totalmente através de gestos. Um outro *software* interessante nesta categoria é o Graffiti (ACCESS-COMPANY, 2010-2011), também disponível para Android, que proporciona

uma forma de entrada de texto totalmente baseada em gestos, a Figura 21 mostra o alfabeto utilizado pelo Graffiti.

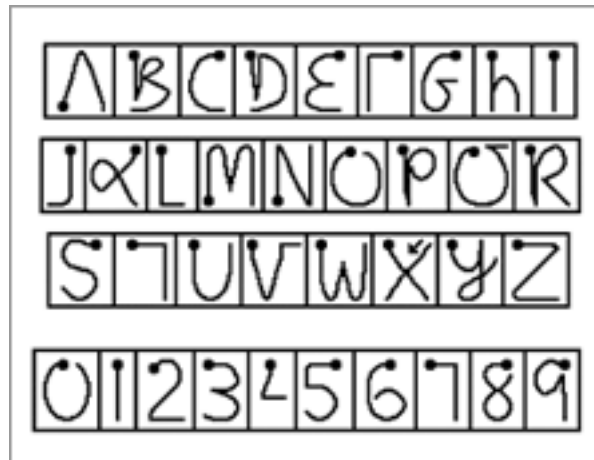


Figura 21: Alfabeto baseado em gestos do Graffiti (ACCESS-COMPANY, 2010-2011).

6.6 Reconhecimento de Fala

A fala é uma forma natural de comunicação humana e é uma boa opção de interação para os deficientes visuais. Contudo diversos problemas como diferenças nos microfones dos aparelhos e a presença de ruído de canal e banda estreita podem afetar o reconhecimento (YNOGUTI, 1999).

Os sistemas mais bem sucedidos são os que se limitam a vocabulários extremamente pequenos, da ordem de 10 a 20 palavras. Para que um sistema seja útil não é necessário um vocabulário muito grande, visto que em alguns sistemas contam com um vocabulário de apenas duas palavras (“sim” e “não”) (YNOGUTI, 1999).

A plataforma Android conta com um recurso de reconhecimento de fala padrão que utiliza o serviço de reconhecimento de fala disponibilizado pelo Google. O reconhecimento é realizado *online* diminuindo a carga de processamento e armazenamento necessárias ao reconhecimento de fala. Contudo esse é um serviço de reconhecimento de fala genérico, contando com um grande vocabulário, o que prejudica sua precisão. Uma alternativa ao reconhecimento de fala para o Android é a biblioteca PocketSphinx (HUGGINS-DAINES et al., 2006) que permite um reconhecimento de fala totalmente *offline* utilizando um vocabulário definido.

6.7 Considerações Finais

Este capítulo apresentou uma breve introdução ao conceito de acessibilidade, onde foram estudadas algumas das soluções que buscam aumentar o grau de acessibilidade de dispositivos móveis aos usuários portadores de deficiências visuais, permitindo escolher aquelas que mais se adequam ao desenvolvimento da aplicação proposta neste trabalho.

Capítulo 7

Aplicação Proposta: Android Eye

Este capítulo apresenta a aplicação desenvolvida neste trabalho, denominada Android Eye, com o intuito de auxiliar os deficientes visuais na tarefa de reconhecer pessoas ao seu redor, chamada Android Eye. Esta aplicação foi desenvolvida tendo como base o projeto BioMobile (FARINA, 2011).

O sistema foi desenvolvido utilizando a plataforma Android 2.2, com o intuito de prover um alto grau de acessibilidade para deficientes visuais, bem como usuários com boa visão.

7.1 Arquitetura

Além do reconhecimento da face, o usuário é capaz de adicionar as fotos de seus contatos à base de faces cadastradas e extrair seus descritores (treinamento) sempre que considerar conveniente.

A aplicação trabalha de forma totalmente *offline*, sem a necessidade de conexão com a *internet* para ser executada, o que garante uma maior mobilidade ao usuário.

7.2 Interface com o Usuário

Toda interação com o usuário é baseada em voz. Visto que esta é a forma mais natural de comunicação utilizada pelas pessoas, provendo uma interface totalmente independente de informações visuais.

Foi utilizada a ferramenta de sintetização de voz PicoTTS disponível no próprio Android. Contudo, devido a sua arquitetura, o Android possibilita a alteração da ferramenta de sintetização de voz para aquela que mais agrada ao usuário.

A entrada de dados utiliza recursos de reconhecimento de fala, permitindo uma interação simples e rápida.

A ferramenta de reconhecimento de voz disponibilizada pelo Android não mostrou-se satisfatória aos objetivos do Android Eye, pois se trata de uma ferramenta *Web*, o que torna sua utilização

restrita a locais onde a conexão à internet é disponível. Além disso, ela utiliza um dicionário de propósitos gerais, com diversas palavras, o que diminui sua acurácia.

Por esse motivo foi adotada a biblioteca PocketSphinx (HUGGINS-DAINES et al., 2006) que permite um reconhecimento de fala totalmente *offline* utilizando um vocabulário definido.

O dicionário utilizado pelo aplicativo possui apenas 6 palavras distintas, que atendem bem as necessidades de uso da aplicação. São elas:

1. *scan*
2. *train*
3. *update*
4. *yes*
5. *no*
6. *exit*

A Figura 22 apresenta o diagrama de casos de uso.

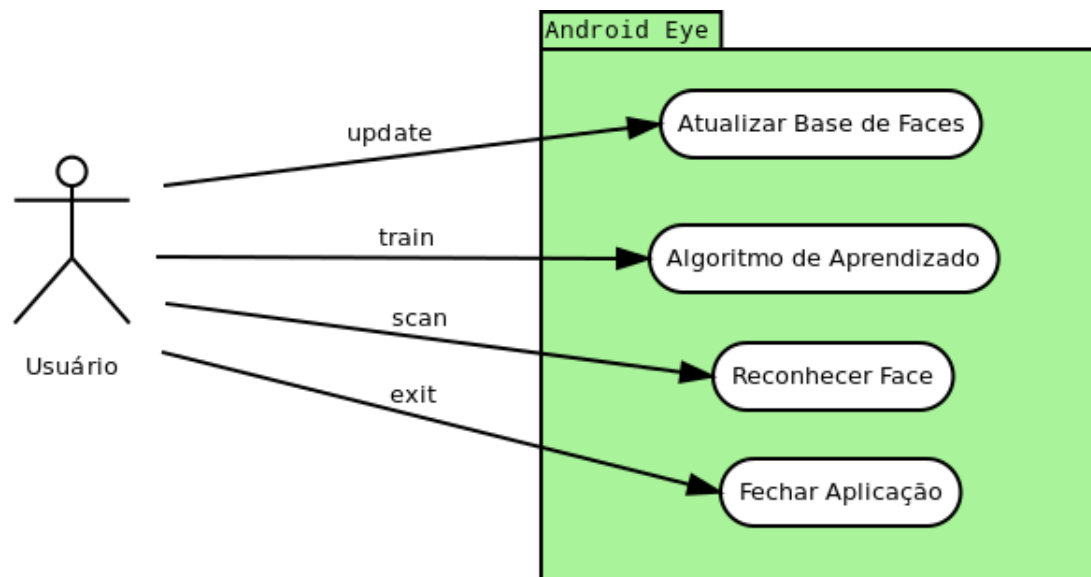


Figura 22: Diagrama de casos de uso do Android Eye.

7.3 Cadastro das Faces

Visto que é uma tarefa difícil ao deficiente visual escolher as imagens necessárias para compor a base de faces faz-se necessária uma forma automática de cadastrar as faces para a utilização do aplicativo.

Buscando resolver este problema foi proposta a criação de um motor de busca capaz de utilizar as imagens de contatos de redes sociais como uma forma de compor a base de faces. Contudo, optou-se por utilizar as imagens dos contatos do próprio celular, visto que existem várias ferramentas capazes de sincronizar os contatos com as mais diversas redes sociais, como o Facebook, Google+ ou o LinkedIn. Permitindo assim, que quaisquer redes sociais, atuais ou futuras, que implementem uma ferramenta de sincronização de contatos com o Android possam ser aproveitadas. Além de possibilitar que o usuário insira novos indivíduos na base manualmente.

Um aspecto importante a ser tratado é a resolução extremamente baixa das imagens de contatos no Android (versões anteriores a 4.1), que é de apenas 96x96 pixels. Isso acontece pelo fato de o Android, por padrão, sincronizar seus contatos com o Gmail, que possui essa limitação.

Este problema pode ser solucionado escolhendo ferramentas de sincronização que trabalhem em maiores resoluções, ou com o uso de aplicações que substituem as imagens sincronizadas de baixa resolução por imagens de maior resolução.

7.4 Fase de Treinamento

De posse da base de imagens adquirida na etapa de cadastro das faces é necessário determinar quais dessas imagens são aptas a compor a base de conhecimento, visto que as imagens utilizadas em perfis nem sempre apresentam a face de seu usuário. Considerando a dificuldade do usuário alvo em determinar quais das imagens são interessantes ao processo de reconhecimento facial, aquelas que contém a face do indivíduo em questão, foi necessário encontrar uma forma automática de selecionar as imagens que devem ser utilizadas na base.

Foi utilizada a seguinte abordagem: através das técnicas de reconhecimento de faces, é realizada uma contagem no número de faces das potenciais imagens. Naquelas imagens onde uma, e apenas uma, face é encontrada, considera-se esta como uma imagem apta a composição da base de conhecimento. Contudo, mesmo com esse cuidado ainda é possível a inserção de faces que não representem realmente o contato.

Após essa etapa são extraídas e armazenadas as características das faces, para utilização na fase posterior: o reconhecimento.

7.5 Reconhecimento das Faces

O sistema proposto captura imagens, através da câmera integrada existente no dispositivo móvel, que são submetidas ao processo de detecção de faces. A partir das faces detectadas, são aplicadas as técnicas de reconhecimento facial buscando as faces encontradas entre as faces previamente cadastradas na base. Ao que é enviada uma resposta sonora ao usuário. A Figura 23 ilustra o processo de reconhecimento do sistema Android Eye.



Figura 23: Processo de reconhecimento do Sistema Android Eye.

7.6 Considerações Finais

Este capítulo apresentou de forma breve a aplicação desenvolvida durante este projeto como solução ao problema proposto do reconhecimento de pessoas através da face.

Foram abordadas a arquitetura do sistema proposto, características de implementação, além do modelo de interface com o usuário adotado pelo Android Eye.

Capítulo 8

Resultados Experimentais

Neste capítulo são apresentados os resultados obtidos através de experimentos realizados com o Android Eye. Buscando mensurar a acurácia das técnicas de detecção e reconhecimento, bem como o grau de usabilidade da interface desenvolvida.

8.1 Dispositivo Móvel Utilizado

Todos os experimentos realizados neste trabalho foram executados no dispositivo móvel Samsung Galaxy 5, que conta com o sistema operacional Android na sua versão 2.2. O aparelho (Figura 24) possui um processador ARM11 de 600 MHz contando com 256 Mb de memória RAM e uma câmera embutida de 2.0 Mpx.



Figura 24: Dispositivo móvel utilizado: Samsung Galaxy 5, com Android 2.2.

8.1.1 Base MoBio

Para a avaliação das técnicas de detecção e reconhecimento facial adotadas nesse trabalho foi necessária a utilização de uma base de faces. Neste projeto foi adotada a base pública MoBio (MCCOOL;

MARCEL, 2009) (*Mobile Biometry*), disponibilizada pelo instituto de pesquisa Idiap e disponível em <https://www.idiap.ch/dataset/mobio>.

A base foi desenvolvida com o intuito de ser uma base multimodal, possibilitando sua utilização em diferentes aplicações, como reconhecimento de fala e facial.

A base MoBio é composta de vídeos de diversas pessoas, que se posicionam em frente a uma câmera enquanto leem textos pré-definidos. Essas amostras são extraídas em diferentes momentos e locais, utilizando-se câmeras de aparelhos móveis. O que a aproxima das condições reais de utilização do sistema Android Eye, proposto neste trabalho.

A Figura 25 apresenta algumas amostras de imagens retiradas da base MoBio.



Figura 25: Amostras extraídas de um mesmo indivíduo da base MoBio.

Pode-se observar que essa base de imagens apresenta desafios, uma vez que há uma grande variação intraclasse, com imagens de um mesmo indivíduo apresentando diferenças significativas.

Os experimentos foram realizados utilizando-se imagens estáticas, extraídas dos vídeos disponibilizados pela base. Foram considerados 75 indivíduos distintos, dos quais foram utilizadas 6 amostras obtidas em diferentes datas, totalizando 450 imagens.

8.2 Avaliação das Técnicas de Detecção das Faces

Para avaliar as técnicas de detecção de faces, as 450 amostras foram submetidas a ambos os algoritmos de detecção.

Para a detecção por zonas da pele foi utilizada a implementação disponível na biblioteca Android, enquanto que para a detecção de faces Viola-Jones foi utilizada a implementação disponibilizada pela biblioteca OpenCV (BRADSKI, 2000).

Essas imagens foram analisadas levando-se em consideração 4 resoluções distintas: 640x480 (resolução original dos vídeos), 480x360, 320x240 e 160x120.

Para cada nível de resolução foi anotado o tempo médio e acurácia de cada uma das técnicas.

A avaliação da acurácia foi realizada através de uma verificação manual das faces marcadas pelos algoritmos. A Figura 26 apresenta amostras submetidas a detecção de faces.

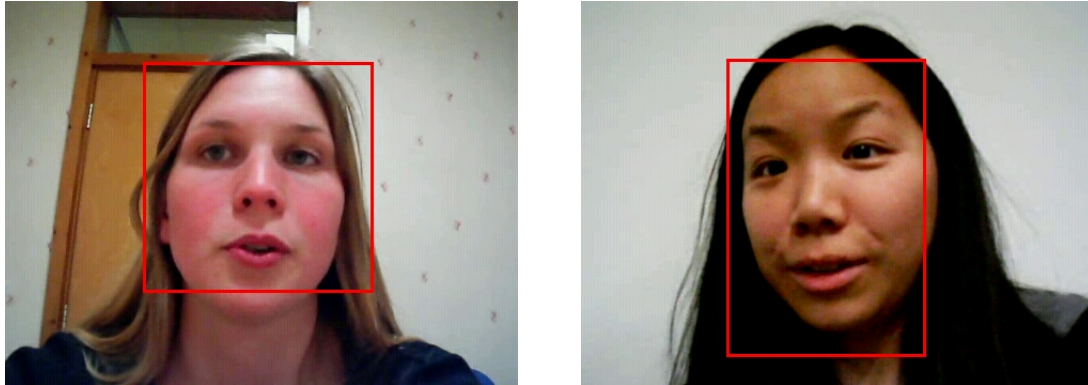


Figura 26: Faces detectadas pelos algoritmos de detecção Viola-Jones e zonas de pele, respectivamente.

Essa acurácia é apresentada através de 2 variáveis:

- FP - Falso positivo: Representa o número de regiões detectadas erroneamente pelo algoritmo.
- FN - Falso negativo: Representa o número de regiões que representam uma face, mas que não foram detectadas pelo algoritmo.

A Tabela 3 mostra os resultados obtidos através dos experimentos.

Tabela 3: Resultados do experimento de detecção de faces.

	Zona da Pele			Viola-Jones		
Resolução	FP	FN	Tempo (s)	FP	FN	Tempo (s)
640x320	7	45	$2,3789 \pm 0,2728$	14	72	$5,5353 \pm 0,3438$
480x360	2	42	$1,2109 \pm 0,0870$	5	58	$3,7918 \pm 0,3007$
320x240	3	41	$0,5175 \pm 0,0345$	2	64	$1,3745 \pm 0,1048$
160x120	0	41	$0,1036 \pm 0,0177$	1	71	$0,2667 \pm 0,0357$

8.3 Avaliação das Técnicas de Reconhecimento de Faces

Foram avaliadas também as técnicas de reconhecimento de faces LBP e ASM.

A técnica LBP, adaptada da implementação disponibilizada pelo BioMobile (FARINA, 2011), foi avaliada utilizando-se 1 (equivalente a abordagem não janelada), 4 e 8 janelas.

A técnica ASM, implementada pela biblioteca STASM (MILBORROW; NICOLLS, 2008), foi avaliada em uma abordagem rápida, utilizando um modelo de 68 pontos com informações de textura unidimensionais, e uma abordagem precisa, que utiliza um modelo de 76 pontos com informações de textura bidimensionais.

A implementação da técnica PCA encontrou barreiras, visto que sua fase de treinamento demanda uma grande porção de memória, $O(n^2)$ em relação a resolução das imagens de entrada. Mesmo para imagens de baixa resolução a técnica supera os limites disponíveis para uma aplicação Android.

A detecção das faces neste experimento foi realizada manualmente, garantindo uma boa detecção das faces.

Foram utilizadas as 450 amostras extraídas da base, que foram divididas em 2 conjuntos de mesmo tamanho para treinamento e teste. Cada conjunto possui 3 imagens distintas de cada um dos 75 indivíduos analisados, num total de 225 imagens.

A avaliação da etapa de reconhecimento foi realizada utilizando-se a medida de acurácia para os casos de Rank-1, Rank-2 e Rank-3.

Além disso, foi mensurado o tempo gasto no processo de extração de descritores (extração de informações da imagem) e o tempo total de reconhecimento¹ que é a soma do tempo de extração de características e do tempo de consulta a base de conhecimento.

A Tabela 4 mostra os resultados do experimento.

Tabela 4: Resultados do experimento de reconhecimento de faces.

Descritor	Rank-1	Rank-2	Rank-3	Tempo Descritor (s)	Tempo Total (s)	Número de Descritores
LBP	22,22%	26,22%	28,89%	0,3866 ± 0,0649	2,1116 ± 0,1139	256
LBP (4 janelas)	36,44%	44,00%	49,33%	0,411418 ± 0,076385	7,073862 ± 0,121046	1024
LBP (8 janelas)	46,22%	54,67%	61,33%	0,485178 ± 0,149978	14,018960 ± 0,512327	2048
ASM (68-1D)	9,33%	15,56%	18,67%	3,696235 ± 0,588743	4,640711 ± 0,670772	136
ASM (76-2D)	8,89%	15,11%	20,00%	23,581487 ± 3,222665	24,704176 ± 3,426079	152

A extração de características baseada em LBP mostrou-se bastante superior ao ASM. Isso se deve ao fato de ASM ser sensível a variações de pose, que são bastante presentes na base analisada.

¹O tempo de reconhecimento varia em função do número de descritores e do número de indivíduos presentes na base. Dessa forma para bases de faces menores, será obtido um menor tempo de reconhecimento. Por outro lado em bases maiores, o tempo gasto no reconhecimento será maior.

A utilização de janelas no LBP mostrou melhoras consideráveis na taxa de acurácia, causando contudo um grande aumento no tempo desprendido no reconhecimento.

8.4 Seleção das Técnicas Utilizadas

Através destes experimentos foram determinadas as técnicas de detecção e extração de características a serem utilizadas no Android Eye.

A técnica de detecção de faces por zona da pele foi adotada, visto que apresentou melhores resultados em tempo de detecção e acurácia, além do fato de ser implementada pela própria biblioteca Android, não necessitando a inclusão de bibliotecas externas.

Para a tarefa de extração de características foi escolhida a técnica LBP utilizando a configuração de 8 janelas, devido aos seus melhores resultados de acurácia.

Para o reconhecimento foi adotada a abordagem de Rank-3, retornando ao usuário final os 3 indivíduos mais próximos encontrados no processo de reconhecimento. De posse da lista de prováveis indivíduos o usuário pode lançar mão de informações de alto nível que o auxiliem no reconhecimento, como por exemplo, qual desses indivíduos frequenta o local em que foi realizado o reconhecimento.

8.5 Avaliação do Sistema Android Eye

Selecionadas as técnicas de detecção de faces e extração de características a serem utilizadas no Android Eye, foi realizado um experimento prático buscando mensurar a acurácia do sistema num ambiente de utilização real.

Para a realização deste experimento foi utilizada bases de faces compostas de 10 indivíduos. A primeira composta de indivíduos cadastrados manualmente nos contatos do dispositivo móvel, e a segunda utilizando indivíduos obtidos da rede social Facebook.

Dos indivíduos cadastrados na base, 3 foram utilizados para uma posterior etapa de testes (dias após a aquisição das imagens), os quais foram submetidos a 10 sessões de reconhecimento, alterando pose e distância da câmera.

A avaliação desta etapa foi realizada utilizando-se a medida de acurácia para os casos de Rank-1, Rank-2 e Rank-3.

A Tabela 5 mostra os resultados obtidos neste experimento.

Tabela 5: Resultados do experimento de acurácia do sistema Android Eye.

Cadastro	Rank-1	Rank-2	Rank-3
Manual	30,0% \pm 26,46	50,0% \pm 20,0	66,67% \pm 11,55
Automático	40,0% \pm 30,0	56,67% \pm 32,15	70,00% \pm 36,06

Um segundo experimento foi realizado com o intuito de determinar o grau de impacto que a distância de aquisição das imagens gera na acurácia do sistema.

Para tal foi utilizada apenas a base de faces construída automaticamente (através do Facebook), visto que obteve resultados semelhantes a base manualmente construída.

A Tabela mostra os resultados obtidos neste experimento.

Tabela 6: Resultados do experimento de distância do Android Eye.

Distância	Rank-1	Rank-2	Rank-3	Erro Detecção
1m	40,0%	70,0%	80,0%	0,0%
2m	30,0%	70,0%	80,0%	10,0%
3m	40,0%	60,0%	60,0%	40,0%

Buscando avaliar a usabilidade da interface desenvolvida foram realizados experimentos nos quais 3 indivíduos tiveram sua visão interrompida, como forma de simular as condições do usuário alvo, e foram submetidos a tarefa de utilizar as funções básicas do Android Eye.

A usabilidade é a medida na qual um produto pode ser usado por usuários específicos para alcançar objetivos específicos com eficácia, eficiência e satisfação em um contexto específico de uso (ISO NBR 9241-11, 2002).

Os testes foram desenvolvidos em ambiente domiciliar, que possuem uma quantidade baixa de ruídos sonoros. Os indivíduos foram submetidos a uma explicação sobre o funcionamento do Android Eye, além do tutorial disponível na própria aplicação. Foram então submetidos a um questionário objetivo (Anexo A) que avalia as grandezas de eficácia, eficiência e satisfação do sistema Android Eye.

Os resultados obtidos são mostrados na Tabela 7.

Tabela 7: Resultados do Teste de Usabilidade.

Objetivos	Ruim	Razoável	Bom
Eficácia	0	0	3
Eficiência	0	0	3
Satisfação	0	1	2

Um dos colaboradores apontou a necessidade de adicionar um comando *help* que invoque o tutorial da aplicação sempre que necessário.

8.6 Conclusões Finais

Neste capítulo foram apresentados resultados obtidos através de experimentos realizados com o software Android Eye. Através destes experimentos foi possível observar a acurácia e tempo obtidos com as técnicas de detecção e reconhecimento, bem como mensurar o grau de acessibilidade obtido com a interface do sistema.

Capítulo 9

Conclusão

Através deste trabalho foi possível ter contato com o crescente ambiente de desenvolvimento de aplicações para dispositivos móveis, mais especificamente a plataforma Android, que a cada dia se torna mais presente.

Foi possível ainda estabelecer contato com o ainda pouco explorado, mundo do desenvolvimento de aplicações voltadas ao deficiente visual. Possibilitando compreender as problemáticas envolvidas e algumas das soluções apresentadas pela literatura e pelo mercado de dispositivos móveis.

Este trabalho também constituiu uma excelente forma de aprender sobre alguns dos conceitos e técnicas de biometria facial, e do processamento de imagens digitais.

Apesar dos problemas encontrados na implementação da técnica de extração de características PCA em ambiente móvel, foi possível obter uma acurácia próxima utilizando o LBP.

Ao longo dos estudos realizados foi possível desenvolver o *software* Android Eye, que implementa técnicas de reconhecimento facial voltadas ao público com deficiência visual, buscando auxiliá-los na difícil tarefa de reconhecer pessoas. Foram incorporadas ao aplicativo as técnicas de detecção de faces por Zona da Pele e de extração de características LBP, devido aos seus bons resultados nos experimentos realizados.

Este *software* utiliza uma interação com o usuário totalmente baseada na fala, através de técnicas de sintetização de voz para a saída e reconhecimento de fala para a entrada. Dessa forma, ela pode ser utilizada tanto por deficientes visuais como por indivíduos com boa visão. A avaliação realizada mostrou uma boa aceitação de usuários em relação ao modelo de interface adotado.

Como trabalho futuro devem ser estudadas formas mais eficientes de comparação de faces capturadas, a partir de diferentes sensores e diferentes resoluções. Pretende-se também utilizar informações de vídeo, ao invés de uma imagem estática, o que pode vir a auxiliar no aumento da acurácia do sistema.

Referências Bibliográficas

ACCESS-COMPANY. *Graffiti Versão 2.0.6*. 2010–2011.

AHONEN, T.; HADID, A.; PIETIKÄINEN, M. *Face recognition with local binary patterns*. [S.l.]: Springer, 2004. 469–481 p.

AHONEN, T.; HADID, A.; PIETIKAINEN, M. *Face description with local binary patterns: Application to face recognition*. [S.l.]: IEEE, 2006. 2037–2041 p.

ALLEN, J.; HUNNICUTT, M.; KLATT, D. *From text to speech: The MITalk system*. [S.l.: s.n.], 1987.

ANDROID. *Android Developer Guide*. 2012. <http://developer.android.com/> Acessado em 20/09/2012.

APPBRAIN. *Top Android SDK versions*. 2012. <http://www.appbrain.com/stats/top-android-sdk-versions> Acessado em 15/10/2012.

BAPTISTA, J. *A invenção do Braille e a sua Importância na Vida dos Cegos*. [S.l.]: Secretariado Nacional para a Reabilitação e Integração das Pessoas com Deficiência, Comissão de Braille, 2000.

BRADSKI, G. *The OpenCV Library*. [S.l.: s.n.], 2000.

CARVALHO, F.; TAVARES, J. da S. *Metodologias para identificação de faces em imagens: introdução e exemplos de resultados*. [S.l.: s.n.], 2005.

CDSST. *Eigenfaces*. [S.l.]: Consortium For The Development Of Specialized Seismic Techniques, 2012. [Http://www.eos.ubc.ca/research/cdsst/Tad_home/eigenfaces.html](http://www.eos.ubc.ca/research/cdsst/Tad_home/eigenfaces.html) Acessado em 22/10/2012.

COOTES, T. et al. Active shape models-their training and application. *Computer vision and image understanding*, San Diego: Academic Press, c1995-, v. 61, n. 1, p. 38–59, 1995.

COOTES, T.; TAYLOR, C. et al. *Statistical models of appearance for computer vision*. [S.l.: s.n.], 2001.

EYESFREE-PROJECT. *Eyes Free Shell Versão 3.1.0*. 2009–2012.

FARINA, A. *BioMobile: Sistema de Identificação de Usuários em Dispositivos Móveis na Plataforma Android Utilizando Reconhecimento de Faces a partir de Vídeo*. [S.l.]: PPGC - Departamento de Computação - Unesp Bauru, 2011.

- FIGUEIREDO, C.; NAKAMURA, E. *Computação móvel: Novas oportunidades e novos desafios. Belo Horizonte: Universidade Federal de Minas Gerais: 28p, 2003.*
- FREY, B.; SOUTHERN, C.; ROMERO, M. *Brailletouch: mobile texting for the visually impaired.* [S.l.]: Springer, 2011. 19–25 p.
- HSU, R. *Face detection and modeling for recognition.* [S.l.: s.n.], 2002.
- HUGGINS-DAINES, D. et al. *Pocketsphinx: A free, real-time continuous speech recognition system for hand-held devices.* [S.l.], 2006. I–I p.
- IBGE, I. B. d. G. e. E. *Pesquisa nacional por amostra de domicílios: Síntese de Indicadores-2010.* [S.l.]: IBGE, 2011.
- ISO NBR 9241-11. Norma, *Requisitos Ergonômicos para Trabalho de Escritórios com Computadores. Parte 11 - Orientações sobre Usabilidade.* [S.l.]: ISO, 2002.
- JAIN, A.; ROSS, A.; PRABHAKAR, S. *An introduction to biometric recognition.* [S.l.]: IEEE, 2004. 4–20 p.
- JONES, M.; REHG, J. *Statistical color models with application to skin detection.* [S.l.], 1999.
- MÄENPÄÄ, T. *The Local Binary Pattern Approach to Texture Analysis: Extensions and Applications.* [S.l.]: Oulun yliopisto, 2003.
- MASCARO, A.; MELLO, C.; SANTOS, W. *Segmentação de Imagens de Mamografias Digitais.* [S.l.: s.n.], 2007.
- MCCOOL, C.; MARCEL, S. *MOBIO Database for the ICPR 2010 Face and Speech Competition.* [S.l.: s.n.], 2009.
- MILBORROW, S. *Locating facial features with active shape models.* Tese (Doutorado) — Faculty of Engineering, University of Cape Town, 2007.
- MILBORROW, S.; NICOLLS, F. *Locating Facial Features with an Extended Active Shape Model.* [S.l.: s.n.], 2008. <http://www.milbo.users.sonic.net/stasm>.
- MOTOROLA. *MOTODEV Versão 4.1.0.* 2007–2012.
- SANTAROSA, L. *Telemática y la inclusión virtual y social de personas con necesidades especiales: un espacio posible en la Internet.* [S.l.]: V Congresso Ibero-Americano de Informática na Educação-RIBIE: Viñadelmar, 2000.
- SCHALKOFF, R. *Digital image processing and computer vision.* [S.l.]: Wiley New York:, 1989.
- SVOX-MOBILE-VOICES. *Pico TTS.* 2008–2012.
- SVOX-MOBILE-VOICES. *Classic Text to Speech Engine.* 2010–2012.
- TORRES, E.; MAZZONI, A. *Conteúdos digitais multimídia: o foco na usabilidade e acessibilidade.* [S.l.]: SciELO Brasil, 2004. 152–160 p.

TURK, M.; PENTLAND, A. *Face recognition using eigenfaces*. [S.l.], 1991. 586–591 p.

VIOLA, P.; JONES, M. *Robust real-time object detection*. [S.l.: s.n.], 2001. 137–154 p.

VIOLA, P.; JONES, M. *Robust Real-Time Face Detection*. [S.l.: s.n.], 2004. 137–154 p.

YANG, M.; KRIEGMAN, D.; AHUJA, N. *Detecting faces in images: A survey*. [S.l.]: IEEE, 2002. 34–58 p.

YNOGUTI, C. *Reconhecimento de Fala Contínua Usando Modelos Ocultos de Markov*. [S.l.: s.n.], 1999.

ZHAO, W. et al. *Face recognition: A literature survey*. [S.l.]: ACM, 2003. 399–458 p.

ZHENG, H.; DAOUDI, M. *Blocking adult images based on statistical skin detection*. [S.l.: s.n.], 2004. 1–14 p.

Anexos

Anexo A

Questionário empregado na avaliação de usabilidade da interface com o usuário do Android Eye.

1. Você conseguiu utilizar todas as funções do aplicativo?

• () Não

• () Algumas

• () Todas

• Sugestões: _____

2. Você sentiu dificuldades na utilização dessas funções?

• () Sem dificuldades

• () Poderia ser mais simples

• () Difícil de se utilizar

• Sugestões: _____

3. O que você achou do Aplicativo?

• () Ruim

• () Mediano

• () Bom

• Sugestões: _____