

UNIVERSIDADE FEDERAL DA PARAÍBA
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA
DEPARTAMENTO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

ANDRÉ LIRA ROLIM

**UM SISTEMA DE IDENTIFICAÇÃO AUTOMÁTICA DE FACES PARA
AMBIENTES VIRTUAIS DE APRENDIZAGEM**

João Pessoa

2009

ANDRÉ LIRA ROLIM

**UM SISTEMA DE IDENTIFICAÇÃO AUTOMÁTICA DE FACES PARA
AMBIENTES VIRTUAIS DE APRENDIZAGEM**

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Informática, da Universidade Federal da Paraíba, como parte dos requisitos para a obtenção do título de Mestre em Informática.

ORIENTADOR: Prof. Dr. Ed Pôrto Bezerra

João Pessoa

2009

R748s

Rolim, André Lira.

Um Sistema de Identificação Automática de Faces para
Ambientes Virtuais de Aprendizagem / André Lira Rolim.- João
Pessoa: UFPB, 2009.
138f.:il..

Dissertação (Mestrado em Informática) – UFPB/CCEN.
Orientador: Ed Pôrto Bezerra.

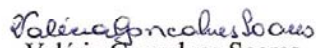
1. Sistema biométrico. 2. Reconhecimento facial. 3. Ambientes
virtuais de aprendizado. 4.Educação a distância I. Título.

UFPB/BC

CDU 004.4(043.2)

Ata da Sessão Pública de Defesa de Dissertação de Mestrado do aluno André Lira Rolim, candidato ao Título de Mestre em Informática na Área de Sistemas de Computação, realizada em 13 de agosto de 2009.

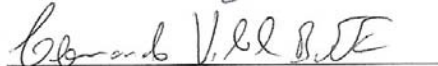
1
2 Aos treze dias do mês de agosto do ano dois mil e nove, às quatorze horas, no Auditório do
3 Centro de Ciências Exatas e da Natureza da Universidade Federal da Paraíba, reuniram-se
4 os membros da Banca Examinadora constituída para examinar o candidato ao grau de
5 Mestre em Informática, na área de “Sistemas de Computação”, na linha de pesquisa
6 “Computação Distribuída”, o Sr. **André Lira Rolim**. A comissão examinadora foi
7 composta pelos professores doutores: Ed Porto Bezerra (DI - UFPB), Orientador e
8 Presidente da Banca Examinadora, Leonardo Vidal Batista (DI-UFPB), como examinador
9 interno e Evandro de Barros Costa (UFAL), como examinador externo. Dando início aos
10 trabalhos, o Prof. Ed Porto Bezerra cumprimentou os presentes, comunicou aos mesmos a
11 finalidade da reunião e passou a palavra ao candidato para que o mesmo fizesse, oralmente,
12 a exposição do trabalho de dissertação intitulado “*UM SISTEMA DE IDENTIFICAÇÃO*
13 *AUTOMÁTICA DE FACES PARA AMBIENTES VIRTUAIS DE APRENDIZAGEM*”.
14 Concluída a exposição, o candidato foi argüido pela Banca Examinadora que emitiu o
15 seguinte parecer: “*aprovado*”. Assim sendo, deve a Universidade Federal da Paraíba
16 expedir o respectivo diploma de Mestre em Informática na forma da lei e, para constar eu,
17 professora Valéria Gonçalves Soares, coordenadora do Programa de Pós-Graduação em
18 Informática, lavrei a presente ata, que vai assinada por mim mesmo e pelos membros
19 membros da Banca Examinadora. João Pessoa, 13 de agosto de 2009.

20
21 
22 Valéria Gonçalves Soares
23

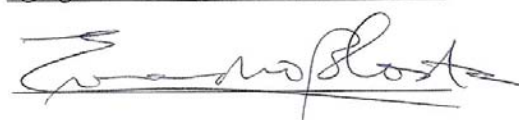
Prof. Dr. Ed Porto Bezerra
Orientador (DI-UFPB)



Prof. Dr. Leonardo Vidal Batista
Examinador Interno (DI-UFPB)



Prof. Dr. Evandro de Barros Costa
Examinador Externo (UFAL)



Dedico este trabalho a minha estimada irmã Ana
Paula, pelo seu incondicional apoio nos momentos
difíceis sem o qual não alcançaria este ideal.

AGRADECIMENTOS

Agradeço,

A Deus, antes de mais nada, pelo dom da vida e capacidade de assimilar conhecimentos com perseverança.

Aos meus pais, José da Silva Rolim e Maria do Socorro Lira Rolim, por acreditarem e contribuírem para realização deste sonho. A minha irmã e amiga Ana Paula Lira Rolim pela dedicação e apoio em todos os momentos de minha vida. Aos meus queridos irmãos Eliomar, Enoque, Emerson e Patrícia, pela força nas horas precisas.

A Salete e Josefa, minhas cunhadas. Aos meus sobrinhos recém chegados à família, Matheus e Vinícius, que me proporcionam muita alegria, descontração e principalmente me ensinam o verdadeiro sentido da vida. Nunca esquecerei de vocês, e sempre os amarei.

A Francisco da Silva Rolim (*in memorian*) e Isaura de Lira do Nascimento, meus avós Maternos, Enoque da Silva Rolim (*in memorian*) e Maria Ana Rolim, meus avós Paternos, pela sabedoria, pelo exemplo e pelas lições de vida que me passam.

Ao meu orientador Ed Porto Bezerra, que pela convivência se tornou um grande amigo, e que a sua orientação, exemplo, e dedicação, tornou possível a realização deste trabalho.

Aos amigos João Teodoro, Francisco Rocha, Ritomar Torquato, Daniel Charles, Angela Cardoso, Nereide Serafin, Michael Schuenck, Saulo Monteiro, Fabrizzia Matos, Graça Nery, Derzu Omaia, Wagner Soares, Francisco Rômulo, Otávio Borges, Alba Lúcia, pela virtude do companheirismo que foi fundamental nesta etapa da minha vida.

A todos que contribuíram direto ou indiretamente para realização deste trabalho de mestrado.

Agradeço ao sol que insistiu em nascer na janela do meu quarto, avisando-me que terminara mais uma noite de trabalho.

RESUMO

O aumento da demanda pela modalidade de Educação a Distância e o surgimento e aperfeiçoamento dos Ambientes Virtuais de Aprendizagem (AVA) apresentam-se num cenário em que se pretende minimizar distâncias e dinamizar estratégias de monitoramento dos alunos. No entanto, os AVA atuais não dispõem de recursos que identifiquem os usuários durante o desenvolvimento das suas atividades. Esta dissertação apresenta um Sistema de Identificação Automática de Faces (SIAF-EAD) para uso em AVA que possui uma arquitetura cliente-servidor e combina estágios de processamento de reconhecimento facial com módulos de um mecanismo que monitora usuários remotos através de uma webcam. Para realizar o reconhecimento facial foi utilizada uma abordagem de seleção de coeficientes da Transformada Discreta do Cosseno (DCT - *Discrete Cosine Transform*) que obteve 97,12% de taxa de acertos em um banco de imagens coletadas sem normalização das faces. Este trabalho descreve, além do projeto arquitetural, um protótipo implementado e apresenta resultados obtidos e ilustra seu funcionamento em um exemplo de execução. Por fim, são discutidos os resultados obtidos e perspectivas futuras para o sistema.

Palavras-chave. Sistema Biométrico, Reconhecimento Facial, Ambientes Virtuais de Aprendizado, Educação à Distância.

ABSTRACT

The increased demand for Distance Education and Virtual Learning Environment (VLE) appearance and development are in a scenario where minimizing distances and boosting tracking of students are aspired. However, current VLE don't have the necessary resources to identify the users in the development of their activities. This dissertation presents an online face identification system for use in VLE that has client-server architecture and combines stages of face recognition processing to modules of a mechanism monitor remote users via a webcam. To accomplish the face recognition was used a approach to select coefficients of Discrete Cosine Transform (DCT), which it obtained 97.12% rate of correct answers in a database of images collected without normalization of faces. This work describes, in addition to architectural design, a prototype implemented and presents results obtained and illustrates its operation in an example implementation. Finally, there are a discuss the results and future prospects for the system.

Keywords: Biometric Systems, Face Recognition, Virtual Learning Environment, Distance Learning.

LISTA DE FIGURAS

Figura 1 – Representação numérica de uma imagem digital monocromática.....	30
Figura 2 - Computação da DCT-II com o processamento em duas etapas.	32
Figura 3 - Aplicação da DCT em imagens monocromáticas e suas visualizações.....	34
Figura 4 – Amostras de três conjuntos de imagens de uma pessoa do banco FERET.	36
Figura 5 – Amostras de Imagens do banco ORL.....	37
Figura 6 – Amostras de imagens do MIT / CMU Data Base	38
Figura 7 – Amostras de imagens do CMU <i>Test Set II</i>	38
Figura 8 - Visão Geral do Processamento de Imagem Facial.	40
Figura 9 - Um sistema genérico de Reconhecimento.	42
Figura 10 – Seleção de Atributos.	43
Figura 11 - Regiões quadradas associadas a seleção de baixas frequências.	45
Figura 12 - Principais conceitos envolvidos no método Viola e Jones.....	52
Figura 13 – Imagem Integral.	53
Figura 14 – Aplicação dos “retângulos de Haar” no método Viola e Jones.	55
Figura 15 – Aplicação em série de classificadores AdaBoost no método Viola e Jones.....	56
Figura 16 - A aplicação do algoritmo <i>Viola e Jones</i> em duas imagens do banco MIT/CMU.	57
Figura 17 - Arquitetura padrão do MOODLE.	61
Figura 18 - Cenário proposto para utilização do MIF.....	63
Figura 19 - Arquitetura do MIF.....	64
Figura 20 - Captura da tela do login, mostrando a execução do Subsistema 1 do MIF.	64
Figura 21 - Redirecionamento feito pelo subsistema 1 para o MOODLE.....	65
Figura 22 – Relatório “Todos os Acessos” do MOODLE mostrando as imagens capturadas pelo MIF.	66
Figura 23 – Cenário de aplicação do SIAF-EAD.	67
Figura 24 – Diagrama de Caso de Uso Geral do SIAF-EAD.....	70
Figura 25 – Diagrama de Estados do Ciclo de vida do SIAF-EAD.	71
Figura 26 – Diagrama de Caso de Uso da primeira estratégia de aquisição das imagens de treinamento do SIAF-EAD.	72
Figura 27 – Diagrama de Caso de Uso da segunda estratégia de aquisição das imagens de treinamento do SIAF-EAD.	73

Figura 28 – Diagrama de Caso de Uso do módulo de treinamento do SIAF-EAD.....	75
Figura 29 – Caso de Uso da Entrada do usuário do sistema.	75
Figura 30 – Diagrama de Caso de Uso da Captura de imagens.	76
Figura 31 – Diagrama de Caso de Uso da Detecção.....	77
Figura 32 – Diagrama de Caso de Uso da face de Reconhecimento.	78
Figura 33 – Caso de Uso da saída do sistema.	78
Figura 34 –Modelo de Domínio do SIAF-EAD.....	79
Figura 35 - – Amostras do Banco de Faces SIAF-EAD, ilustrando diferentes cenários e diferenças de sexo e idade entre as pessoas fotografadas.....	81
Figura 36 – Amostras do Banco de Faces SIAF-EAD, mostrando a presença e ausência de assessorios nas pessoas.	82
Figura 37 – Amostras do Banco de Faces SIAF-EAD, ilustrando a variedade de expressão, pose e iluminação.	83
Figura 38 – Amostras do Banco de Faces SIAF-EAD, ilustrando algumas peculiaridades do banco.....	83
Figura 39 – Representação gráfica da estratégia de testes para a definição da abordagem de detecção.	85
Figura 40 – “Falsos negativos” dos testes de detecção.....	86
Figura 41 – Amostras com a detecção de duas faces na mesma imagem do Banco de Faces SIAF-EAD.	87
Figura 42 - Amostras de Falsos positivos dos testes de detecção.	88
Figura 43 – Amostras de faces detectadas e extraídas corretamente pelo protótipo.	89
Figura 44 – Exemplo de certificação da imagem da face de teste do usuário.....	89
Figura 45 – Estratégia usada para testar e definir o seletor e classificador para a implementação do protótipo do SIAF-EAD.	90
Figura 46 - Classificador Vizinho mais Próximo: seletores de baixas frequências e média das amplitudes.	92
Figura 47 - Classificador de k -vizinhos mais próximo ($k=3$): seletores de baixas frequências e média das amplitudes.	92
Figura 48 – Tela de login do SIAF-EAD. Usuário “edporto” entrando no sistema.	96
Figura 49 – Tela de login com erro de senha incorreta.....	97
Figura 50 – Página de redirecionamento do SIAF-EAD.Usuário “edporto” logado no MOODLE tendo disponível o curso “UFPB – Teste Moodle”.....	98

Figura 51 – Relatório de “Todos os Acessos” mostrando as imagens identificadas e não-identificadas pelo SIAF-EAD.	99
Figura 52 – Ampliação da imagem identificada no relatório.	99
Figura 53 – Gráfico comparativo entre o Banco de Faces SIAF-EAD e o Banco ORL. Classificador NN e seletor de baixas frequências.	113
Figura 54 – Gráfico comparativo entre o Banco de Faces SIAF-EAD e o Banco ORL. Classificador NN e seletor média das amplitudes.	114
Figura 55 – Gráfico comparativo entre o Banco de Faces SIAF-EAD e o Banco ORL. Classificador KNN e seletor de baixas frequências.	115
Figura 56 – Gráfico comparativo entre o Banco de Faces SIAF-EAD e o Banco ORL. Classificador KNN e seletor média das amplitudes.	116

LISTA DE TABELAS

Tabela 1 - Resumo do desempenho dos seletores de baixas frequências e média das amplitudes e classificadores NN e KNN.	45
Tabela 2 - Tabela-resumo com os dados técnicos do Banco de Faces SIAF-EAD.	84
Tabela 3 - Resultados obtidos com os dois seletores e dois classificadores.	91
Tabela 4 – Erros de classificação.	93
Tabela 5 – Tabela detalhada do resultado dos testes de detecção.	111
Tabela 6 – Tabela comparativa entre o Banco de Faces SIAF-EAD e o Banco ORL. Classificador NN e seletor de baixas frequências.	112
Tabela 7 – Comparativo entre o Banco de Faces SIAF-EAD e o Banco ORL. Classificador NN e seletor média das amplitudes.	113
Tabela 8 – Comparativo entre o Banco de Faces SIAF-EAD e o Banco ORL. Classificador KNN e seletor de baixas frequências.	114
Tabela 9 – Comparativo entre o Banco de Faces SIAF-EAD e o Banco ORL. Classificador KNN e seletor média das amplitudes.	115
Tabela 10 – Tabela comparativa de abordagens.	135

LISTA DE QUADROS

Quadro 1 – Requisitos Funcionais do SIAF-AD.....	69
Quadro 2 – Erros de reconhecimento pelo classificador de vizinho mais próximo e seletor de baixas frequências.	94

LISTA DE SIGLAS

AC	Alternate Current
AVA	Ambientes Virtuais de Aprendizagem
CMU	Carnegie Mellon University
DC	Direct Current
DCT	Discrete Cosine Transform
DKNN	Distance k-Nearest Neighbor
DMP	Distância Mínima ao Protótipo
DNN	Distance Nearest Neighbor
DOD	Departamento de Defesa dos Estados Unidos
DWT	Discrete Wavelet Transform
EAD	Educação à Distância
FERET	Facial Recognition Technology
GPL	General Public License
HMM	Hidden Markov Models
ISO	International Organization for Standardization
JTC 1	Joint ISO/IEC Technical Committee
LDA	Linear Discriminant Analysis
KLT	Karhunen-Loève Transform
KNN	k-Nearest Neighbor
MIF	Mecanismo de Identificação Fisionômica
MOODLE	Modular Object-Oriented Dynamic Learning Environment
NN	Nearest Neighbor
OpenCV	Open Source Computer Vision Library
ORL	Olivetti Research Lab
PCA	Principal Component Analysis
PHP	Hypertext Preprocessor
PKI	Public Key Infrastructure
RGB	Red Green Blue
RNA	Redes Neurais Artificiais
SAC	Serviço de Autenticação Central
SC	Sub- Comitê

SGBD	Sistema de Gerenciamento de Banco de Dados
SIAF-EAD	Sistema de Identificação Automática de Faces
UAB	Universidade Aberta do Brasil
UFPB	Universidade Federal da Paraíba
VLE	Virtual Learning Environment

SUMÁRIO

1 INTRODUÇÃO.....	20
1.1 MOTIVAÇÃO.....	20
1.2 TRABALHOS RELACIONADOS	24
1.3 OBJETIVOS	26
1.4 PUBLICAÇÕES RELACIONADAS.....	27
1.5 ESTRUTURA DO TEXTO	28
2 FUNDAMENTAÇÃO TEÓRICA	29
2.1 IMAGEM DIGITAL	30
2.2 TRANSFORMADA DISCRETA DO COSSENO.....	31
2.3 BANCO DE FACES	34
2.4 RECONHECIMENTO DE FACES.....	38
2.5 RECONHECIMENTO FACIAL UTILIZANDO SELEÇÃO DE COEFICIENTES DA DCT	42
2.5.1 Seletor de Baixas Frequências.....	45
2.5.2 Seletor Média das Amplitudes.....	46
2.5.3 Classificador do Vizinho Mais Próximo (NN)	46
2.5.4 Classificador dos K-Vizinhos Mais Próximo (KNN).....	47
2.6 DETECÇÃO FACIAL.....	48
2.6.1 O método Viola e Jones.	52
2.7 AMBIENTES VIRTUAIS DE APRENDIZAGEM	57
2.7.1 Modular Object-Oriented Dynamic Learning Environment (MOODLE).....	58
2.8 MECANISMO DE IDENTIFICAÇÃO FISIONÔMICA	62
2.8.1 Arquitetura do MIF	62
3 O SISTEMA DE IDENTIFICAÇÃO AUTOMÁTICA DE FACES PARA AMBIENTES VIRTUAIS DE APRENDIZADO (SIAF-EAD).....	67
3.1 PROJETO DO SIAF-EAD	68
3.1.1 Requisitos Funcionais do Sistema.....	68
3.1.2 Caso de Uso Geral do SIAF-EAD.....	70
3.1.3 Ciclo de Vida do sistema	71
Fase 1: Cadastro	72
Fase 2: Treinamento	74

Fase 3: Entrada no Sistema.....	75
Fase 4: Captura.....	76
Fase 5: Detecção	76
Fase 6: Reconhecimento	77
Fase 7: Saída do Sistema	78
3.1.4 Modelo de domínio do SIAF-EAD.....	79
3.2 IMPLEMENTAÇÃO DO SIAF-EAD.....	80
3.2.1 Banco de Faces SIAF-EAD.....	80
3.2.2 Protótipo de Detecção Facial e Testes.....	84
Testes do Protótipo de Detecção	85
3.2.3 Protótipo e Testes de Reconhecimento Facial do SIAF-EAD	89
Testes do Protótipo de Reconhecimento	90
3.2.4 Uma Execução do Protótipo	95
4 CONSIDERAÇÕES FINAIS.....	100
4.1 PERSPECTIVAS FUTURAS	102
REFERÊNCIAS.....	104
APÊNDICE A	111
RESULTADO DETALHADO DOS TESTES DE DETECÇÃO FACIAL	111
APÊNDICE B.....	112
COMPARATIVO ENTRE OS RESULTADOS DE RECONHECIMENTO FACIAL OBTIDOS SOBRE O	
BANCO DE FACES SIAF-EAD E O BANCO ORL	112
Vizinho mais Próximo e Baixas Freqüências.....	112
Vizinho Mais Próximo e Média das amplitudes	113
K-vizinhos mais Próximos (3nn) e Baixas Freqüências.....	114
K-vizinhos mais Próximos (3nn) e Média das amplitudes	115
APÊNDICE C	117
BANCO DE FACES DETECTADAS PELO PROTÓTIPO DE DETECÇÃO IMPLEMENTADO	117
PESSOA 1	117
PESSOA 2	117
PESSOA 3	118
PESSOA 4	118

PESSOA 5	118
PESSOA 6	119
PESSOA 7	119
PESSOA 8	119
PESSOA 9	120
PESSOA 10	120
PESSOA 11	120
PESSOA 12	121
PESSOA 13	121
PESSOA 14	121
PESSOA 15	122
PESSOA 16	122
PESSOA 17	122
PESSOA 18	123
PESSOA 19	123
PESSOA 20	123
PESSOA 21	124
PESSOA 22	124
PESSOA 23	124
PESSOA 24	125
PESSOA 25	125
PESSOA 26	125
PESSOA 27	126
PESSOA 28	126
PESSOA 29	126
PESSOA 30	127
PESSOA 31	127
PESSOA 32	127
PESSOA 33	128
PESSOA 34	128
PESSOA 35	128
PESSOA 36	129
PESSOA 37	129
PESSOA 38	129

PESSOA 39	130
PESSOA 40	130
PESSOA 41	130
PESSOA 42	131
PESSOA 43	131
PESSOA 44	131
PESSOA 45	132
PESSOA 46	132
PESSOA 47	132
PESSOA 48	133
PESSOA 49	133
PESSOA 50	133
PESSOA 51	134
PESSOA 52	134
ANEXO A.....	135
RESUMO DOS RESULTADOS DE ALGUNS MÉTODOS DE RECONHECIMENTO FACIAL.....	135
ANEXO B.....	137
ALGORITMO ADABOOST MODIFICADO PARA OS CLASSIFICADORES VIOLA E JONES.	137

1 INTRODUÇÃO

Esta dissertação apresenta um Sistema de Identificação Automática de Faces para Ambientes Virtuais de Aprendizagem (SIAF-EAD). O SIAF-EAD monitora um usuário, capturando imagens através de webcam, bem como detecta a face do usuário na imagem e a identifica dentre outras faces cadastradas no sistema. Um estudo de caso foi realizado com o ambiente virtual de aprendizagem *Modular Object-Oriented Dynamic Learning Environment* (MOODLE) (MOODLE, 2009) que apresenta as imagens das faces identificadas em um dos seus relatórios de participação do usuário.

1.1 Motivação

O papel da educação no mundo de hoje é indiscutível. As evidências científicas sobre suas contribuições para o desenvolvimento econômico e social do mundo moderno têm sido constantemente apontadas (SABÓIA, 1998). A educação, seja ela na modalidade presencial ou à distância, apresenta-se como fator indispensável não apenas à preservação da vida, mas à elevação de sua qualidade. Além disso, a educação, conforme estabelecem os artigos 196, 205 e 206 da Constituição Federal (BRASIL, 2008), está assinalada como direito básico do cidadão e, por consequência, dever do Estado.

A Educação à Distância (EAD), apresenta-se, neste contexto, como uma estratégia para que se possa oferecer educação ao cidadão, favorecendo sua inclusão social e profissional. A utilização de ambientes virtuais como recurso didático para a oferta de cursos na modalidade de EAD enfatiza a necessidade de encontrar um novo conceito que não permita a dicotomia entre a educação presencial e a educação à distância, pois as possibilidades cada vez mais intensas de conectividade e de interação propiciadas pela Internet e pelo desenvolvimento das telecomunicações em geral, tornam a noção de presença e distância bastante discutíveis.

O aumento da demanda pela modalidade de Educação à Distância, os avanços das tecnologias de informação e comunicação, o surgimento e aperfeiçoamento dos Ambientes

Virtuais de Aprendizagem (AVA) apresentam-se num cenário em que se pretende dinamizar estratégias de acompanhamento e monitoramento dos alunos à distância. Todavia, os AVA atuais não dispõem de recursos que garantam a participação dos usuários no desenvolvimento de suas atividades acadêmicas, incluindo seu processo de avaliação *on-line* à distância (RABUZIN *et al.*, 2006; LIN *et al.*, 2004).

Por outro lado, os sistemas biométricos vêm ganhando bastante adesão em aplicações que realizam certificações digitais. Sistemas biométricos são sistemas que realizam a identificação automática de uma pessoa baseados em métodos biométricos que usam suas características fisionômicas ou comportamentais.

Os sistemas biométricos podem utilizar além da face, a íris, assim como realizar varredura de retina, voz, geometria da mão, digital etc. Um sistema pode explorar aspecto comportamental de uma pessoa, um método pode identificar uma pessoa através da forma como ela digita em um teclado de computador, por exemplo, e identificar padrões que são peculiares a cada pessoa como a velocidade de digitação, frequência de palavras utilizadas e estilo gramatical.

Em certas ocasiões os métodos biométricos oferecem mais segurança do que métodos tradicionais que envolvem senhas. Como exemplo, na avaliação realizada por intermédio de um AVA o aluno a ser avaliado deve estar fisicamente presente no ponto de identificação, e para isto, somente a senha e o login não garantem. Além de prevenir contra acessos não autorizados ou uso fraudulento de dados ou equipamentos, os métodos biométricos permitem realizar identificação em tempo-real, isso substitui o uso de documentos de identificação, os quais podem ser perdidos, falsificados e roubados.

O monitoramento também faz parte das aplicações dos sistemas biométricos. No caso do uso do AVA, muitas vezes é importante a confirmação de que o aluno esteve presente durante o tempo de acesso ao sistema, então o monitoramento através de suas características fisionômicas ou comportamentais pode certificar a sua presença desde o momento da sua entrada até o momento de sua saída do ambiente.

Um fator importante no projeto de um sistema real biométrico é determinar como uma pessoa é reconhecida. Dependendo da abordagem, um sistema pode ser tanto um sistema de autenticação ou um sistema de identificação.

Autenticação (Eu sou quem eu digo ser?) implica na confirmação ou rejeição da identidade invocada por uma pessoa, esse tipo de abordagem está sendo bastante utilizado em

aplicações de transações financeiras e em aplicações em dispositivos portáteis (GAO *et al.*, 2003; VANGALA *et al.*, 2004; WATANABE *et al.*, 2004). Para fim de autenticação, os sistemas ainda podem ser baseados em posse, como por exemplo, um cartão magnético ou um login, baseados em algo que somente a pessoa conhece, uma senha, por exemplo, ou em que a pessoa é, ou seja, as suas características físicas ou comportamentais, caracterizando a biometria (MILLER, 2002).

Identificação implica determinar a identidade de uma pessoa (Quem sou eu?) (JAIN *et al.*, 2000), ou seja, a indicação de quem é uma pessoa entre outras.

De maneira geral, para compreender melhor o paradigma identificação versus autenticação, é importante ter a noção da diferença entre ambas. A autenticação consiste em verificar a veracidade da identidade de uma pessoa, para isso, é efetuada uma comparação de 1 para 1 enquanto que a identificação esta associada a descobrir uma pessoa, de entre um conjunto de N pessoas, neste caso a comparação efetuada é de 1 para N.

Os termos utilizados na biometria tiveram padronização quando a ISO (*International Organization for Standardization*) criou um Vocabulário Biométrico Harmonizado (*Harmonized Biometric Vocabulary*) através do seu Comitê Técnico JTC 1 (*Joint ISO/IEC Technical Committee*), sub-comitê SC 37 (ISO/IEC JTC 1, 2009). Em suma, este documento define uma padronização definitiva para os termos “identificação” e “autenticação”:

- **Autenticação:** esse termo já foi usado como sinônimo de aplicação para verificação biométrica, função de verificação biométrica, mas também para sinônimo de aplicação de identificação biométrica e função de identificação biométrica. Usar este termo como sinônimo de verificação biométrica ou identificação biométrica não é mais aceito. O ideal é que se entenda autenticação no ambiente da biometria, como: reconhecimento biométrico.
- **Identificação:** o mais comumente usado hoje é: identificar, que é o processo de busca biométrica em uma base de dados biométricos capturados, para achar e indicar um identificador de referência biométrica atribuído a uma única pessoa.

O barateamento dos equipamentos de coleta biométrica se tornou realidade, como por exemplo, o hardware que mapeia a íris, ou o coletor de digitais, proporcionando as mais diversas possibilidades no desenvolvimento das aplicações biométricas, dentre elas as que

envolvem o reconhecimento facial apresentam a maior conveniência no contexto dos AVA por várias razões, sendo a mais evidente a que possibilita o aluno trabalhar livremente sem ter que interromper seu trabalho posicionando sua digital ou íris no equipamento de coleta, enquanto que a utilização de uma webcam pode realizar a captura da imagem de sua face sem que exista incômodo.

Os sistemas de reconhecimento de face estão sendo cada vez mais utilizados em aplicações que exigem a identificação pessoal, sendo expandido comercialmente e atraindo a atenção de pesquisas científicas que procuram aperfeiçoar as abordagens de reconhecimento (CHELLAPPA *et al.*, 1995; KOHIR *et al.*, 2000; HARANDI *et al.*, 2007) e estão sendo cada vez mais integrados a aplicações que necessitam de segurança (HEMERY *et al.*, 2008; LANG e HONG, 2008).

O desenvolvimento de um sistema de reconhecimento facial pode envolver algumas atividades específicas de construção de software. Essas atividades são padrões da Engenharia de Software¹ só que aplicadas ao projeto de um sistema de reconhecimento facial:

1. Pesquisa, ou criação de abordagens apropriadas de detecção e reconhecimento facial;
2. Projeto de arquitetura;
3. Implementação do projeto;
4. Teste de sistema envolvendo cenários que simulem reais utilizações do sistema.

A pesquisa das abordagens envolve uma avaliação no *estado-da-arte* e a conseqüente escolha de qual método em evidência está trazendo resultados mais satisfatórios, geralmente as pesquisas consideram o percentual de acertos e custo computacional dos algoritmos.

A construção de um banco de imagens pode ser necessária pelo fato das abordagens serem testadas com bancos de faces construídos sob condições controladas, isso é, fator como pose, oclusão, iluminação e escala, dentre outros, são previamente estabelecidos (LI e JAN, 2005), e isso em uma situação real não é possível estabelecer, ou envolve um esforço que

¹ A construção de um software considerado completo envolve muitas outras atividades além destas listadas. A intenção de citar estas específicas é por que elas englobam tarefas que foram importantes e fundamentais na realização do trabalho descrito nesta dissertação. Para mais detalhes sobre Engenharia de Software o leitor pode consultar (SOMMERVILLE, 2008; PFLEEGER, 2007)

acaba atrapalhando o uso do sistema, então para se ter uma comparação dos resultados obtidos sob um banco normalizado e os resultados sob condições em que possivelmente o sistema será aplicado é preciso obter um banco de faces que satisfaça a situação.

Projeto de arquitetura, ou projeto arquitetural, envolve atividades que vão descrever e determinar a estrutura adequada para o sistema. Pfleeger (2007) entende que a arquitetura de um sistema:

Associa as capacidades do sistema identificadas na especificação de requisitos com os componentes do sistema que irão implementá-las. Os componentes são geralmente, módulos, e a arquitetura também descreve as interconexões entre eles. Além disso, a arquitetura define operadores que criam sistemas a partir de subsistemas

O projeto arquitetural tem como objetivo descrever módulos internos que deverão interagir entre si e com outros sistemas, ou subsistemas. A arquitetura descreve como os módulos são dispostos (localmente ou distribuídos), como serão feitas as interações entre eles e, no caso de um sistema de reconhecimento, como estes módulos devem oferecer suporte às abordagens definidas durante a realização da primeira atividade.

A implementação do projeto envolve os aspectos da construção do projeto em si, da programação, utilizando a linguagem adequada. Nessa etapa, é de bom senso a busca de funcionalidades já implementadas. A atividade seguinte consiste em testar o sistema já integrado, com todos os módulos construídos, e a determinação de possíveis correções de *software* quando necessárias.

1.2 Trabalhos Relacionados

A maioria dos AVA utiliza apenas o sistema de verificação de acesso por identificação pessoal de login e senha. Mas, algumas instituições preocupadas em fornecer maior segurança na identificação do usuário desenvolveram outras técnicas de identificação pessoal, além da utilização do tradicional login e senha.

A Universidade de Zaragoza (Espanha) desenvolveu um sistema de certificação digital, baseado na tecnologia de chave pública PKI (*Public Key Infrastructure*), que recebeu o nome de MoodlePKI (VASQUEZ, 2006). Ele é um mecanismo de assinatura digital capaz

de identificar o usuário através do certificado digital fornecido pela Universidade de Zaragoza, que é uma organização reconhecida e credenciada pelo órgão certificador espanhol.

A Universidade de Vigo, também na Espanha, desenvolveu um mecanismo de identificação biométrica (MURAS *et al.*, 2007; AGULLA *et al.*, 2008) integrado ao Serviço de Autenticação Central (SAC) originalmente desenvolvido pela Universidade de Yale (Estados Unidos). O mecanismo aproveita as funcionalidades de segurança do SAC para fornecer um único serviço básico de login e senha de acesso a todas as aplicações web do usuário, incorporando a verificação biométrica.

Nestes dois mecanismos, a segurança digital restringe-se apenas ao acesso do usuário, ou seja, eles não garantem que é mesmo o usuário autenticado que está realizando as atividades no MOODLE.

Outros trabalhos enfocam a necessidade de adicionar aos AVA sistemas de autenticação ou identificação biométrica, como trabalho descrito em Rabuzin *et al.*, (2006) que explora as possibilidades de utilizar dados biométricos para adicionar aos AVA sistemas de segurança que realizem a identificação biométrica adequando-se a uma “sociedade em rede”. Apesar de Rabuzin *et al.* (2006) mostrar a importância dos diversos tipos de dados biométricos para a segurança, não descreve uma solução específica que mostre a integração destes sistemas com os AVA.

Conseqüentemente, surgiram trabalhos que buscam adicionar os sistemas biométricos aos AVA. Penteado *et al.* (2008) descrevem um sistema, ainda em andamento, que busca autenticar usuários por meios biométricos, principalmente através da face. Esse trabalho descreve alguns resultados parciais em relação à autenticação facial realizada através do vídeo de uma webcam, porém não apresenta a solução integrada ao AVA, se limitando basicamente aos métodos de reconhecimento facial utilizados.

Esta dissertação utiliza algumas idéias de trabalhos já realizados que serviram de base ou complementaram o propósito de construir um sistema que faça a identificação automática de um usuário durante o seu acesso ao AVA, são eles :

1. Na Universidade Federal da Paraíba - UFPB, desenvolveu-se um Mecanismo de Identificação Fisionômica (MIF), descrito em Silva (2008), para incorporar a imagem de alunos no relatório de atividades do AVA MOODLE. O MIF captura imagens de alunos obtidas de uma webcam, e as exibe nos relatórios, objetivando o

acompanhamento de suas atividades neste AVA. Este mecanismo, apesar de apresentar uma solução interessante, não adiciona ao AVA uma certificação digital automática do usuário. Apesar de utilizar uma característica fisionômica, a certificação é feita a olho humano, não caracterizando assim um sistema biométrico de segurança (ISO/IEC JTC 1, 2009).

2. Foi desenvolvido na UFPB um estudo sobre o reconhecimento facial utilizando a seleção de coeficientes da Transformada Discreta do Cosseno, ou DCT (*Discrete Cosine Transform*), que conseguiu 99,25% de acertos nos testes realizados (MATOS *et al.*, 2008). Esse trabalho compara os resultados obtidos com a DCT em relação a outros métodos existentes e mostra algumas vantagens de se utilizar essa transformada matemática no reconhecimento facial. O uso da DCT está sendo evidenciado em vários trabalhos científicos na área de reconhecimento facial, sendo também apontado como um método que funciona relativamente independente de fatores como iluminação, posição e escala da face (KOHIR e DESAI, 2000; FAÚNDEZ, 2003).

Esta dissertação descreve um Sistema de Identificação Automática de Faces (SIAF-EAD) que adiciona ao MIF a funcionalidade de identificar por computador a real identidade de um usuário dentre outros usuários cadastrados. O SIAF-EAD é um sistema baseado na web composto de vários módulos que implementam o reconhecimento de faces utilizando a seleção de coeficientes da Transformada Discreta do Cosseno.

Pretende-se que, com a utilização do SIAF-EAD, administradores, coordenadores, professores e tutores tenham uma certificação da participação de um aluno durante a realização de suas atividades e avaliações em um AVA.

1.3 Objetivos

O objetivo geral dessa dissertação é a concepção, projeto e implementação de um sistema de identificação automática de faces que contenha uma arquitetura cliente-servidor e que aproveite a estrutura criada no Mecanismo de Identificação Fisionômica descrito em Silva, (2008) e que adicione a informação ao AVA sobre a identificação processada utilizando

a abordagem de reconhecimento com base na seleção de coeficientes da DCT (MATOS, 2008).

Os objetivos específicos dessa dissertação são os seguintes:

1. **Cadastro de imagens de treinamento.** O sistema proposto deve implementar uma estratégia de aquisição das imagens faciais de usuário do sistema;
2. **Treinamento.** Desenvolver um módulo para o treinamento das faces cadastradas dos usuários do AVA;
3. **Entrada no Sistema.** O sistema proposto deve carregar informações da face dos usuários cadastrados no sistema ao efetuar entrada no mesmo;
4. **Captura.** Realização da captura do usuário em intervalo de tempo determinado durante a utilização do mesmo no AVA.
5. **Deteção.** A deteção facial na imagem capturada do usuário é necessária para realizar a identificação do mesmo.
6. **Reconhecimento.** Realizar o reconhecimento com as faces detectadas do usuário durante o seu acesso no AVA.
7. **Testes.** Realizar testes nas abordagens de deteção e reconhecimento facial escolhidas, assim como analisar e documentar os mesmos;
8. **Coleta de um Banco de Faces.** Para testar as abordagens escolhidas na implementação do sistema é preciso coletar um banco de imagens capturadas de uma webcam simulando um cenário o em que o sistema será utilizado.

1.4 Publicações Relacionadas

Os resultados das pesquisas desenvolvidas durante este trabalho, até o momento, resultaram na seguinte publicação:

- Rolim, A. L.; Bezerra, E. P. Um sistema de identificação automática de faces para um ambiente virtual de ensino e aprendizagem. XIV Simpósio Brasileiro de Sistemas Multimídia e Web (*WebMedia* 2008). Vila Velha - ES 2008.

Foi também submetido para a publicação o seguinte artigo:

- Rolim, A. L.; Omaia, D.; Bezerra, E. P.; Batista, L. An architecture for a facial identification system in a teaching and learning virtual environment. IEEE Transactions on Learning Technology. Link da revista: < <http://www.ewh.ieee.org/soc/es/ilt.html> >. Acessado em 01/10/2009.

1.5 Estrutura do Texto

O trabalho aqui desenvolvido considera desde a contextualização dos principais conceitos envolvidos na pesquisa, até os resultados adotados na concepção do SIAF-EAD e os detalhes técnicos referentes ao desenvolvimento. Sendo assim a dissertação está organizada nos seguintes capítulos:

- **Capítulo 2 – Fundamentação Teórica.** Neste capítulo são descritos os principais conceitos que envolvem o trabalho realizado, com um estudo sobre o estado-da-arte de abordagens de detecção, reconhecimento, normalização facial, mostrando alguns resultados obtidos em trabalhos relacionados às suas respectivas áreas. Também é descrito um trabalho que deu origem ao SIAF-EAD, o Mecanismo de Identificação Fisionômica (MIF).
- **Capítulo 3 – Sistema de Identificação Automática de Faces.** Descreve a arquitetura do SIAF-EAD, descrevendo seus requisitos funcionais, casos de usos, ciclo de vida e modelo conceitual do projeto. Neste capítulo foram definidos os principais módulos constituintes do sistema. Também descreve os resultados obtidos e testes dos protótipos gerados durante as atividades de desenvolvimento assim como um exemplo de execução do SIAF-EAD.
- **Capítulo 4 – Considerações Finais.** São apresentadas as contribuições e conclusões das atividades desenvolvidas durante o projeto, bem como perspectivas de trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

Este capítulo aborda uma visão geral sobre os principais conceitos utilizados no desenvolvimento do trabalho descrito nesta dissertação. São eles :

- Imagem Digital. Incluindo sua descrição matemática e distinção entre uma imagem monocromática e RGB;
- DCT. A definição e propriedades de uma variante desta transformada matemática aplicada a imagens digitais bidimensionais, a DCT II – 2D, que além de outras aplicações está sendo utilizada em reconhecimento facial;
- Banco de Faces. Os bancos de faces mais utilizados em projetos de abordagens faciais são descritos. Entender como esses bancos são construídos foi importante para construir um banco próprio para os testes do SIAF-EAD;
- Reconhecimento de Faces. Um estudo sobre Reconhecimento Facial foi realizado com a intenção de contextualizar o trabalho realizado;
- Reconhecimento Facial Utilizando Seleção de Coeficientes da DCT. Uma abordagem que aproveita os coeficientes gerados na aplicação da transformada em uma imagem facial para realizar o reconhecimento;
- Detecção Facial. Um estudo sobre Detecção Facial foi realizado com a intenção de contextualizar o trabalho realizado. A abordagem que foi utilizada é a Viola e Jones;
- Ambientes Virtuais de Aprendizagem. Definição, características e aplicação na Educação à Distância desses ambientes. Também é descrito o AVA MOODLE;
- Mecanismo de Identificação Fisionômica. Este mecanismo foi utilizado como base para construção SIAF-EAD.

Todos esses temas foram objetos de estudo deste trabalho, visando contribuir na compreensão desses elementos para o desenvolvimento da pesquisa. Essa fundamentação foi realizada através da revisão de literatura existente.

2.1 Imagem Digital

Uma imagem é uma função de intensidade luminosa bidimensional denotada por $f(x,y)$, em que o valor de f nas coordenadas espaciais é a intensidade luminosa (brilho) da imagem naquele ponto.

A digitalização de uma imagem faz-se necessária para processá-la em um computador, e para isso é preciso realizar uma amostragem das coordenadas espaciais junto com uma quantização da intensidade luminosa. Discretizando uma função contínua $f(x, y)$ com amostras igualmente espaçadas, e arranjadas na forma de matriz $N \times M$, com cada elemento sendo uma quantidade discreta, temos a equação (1), onde o lado direito da equação representa uma imagem digital, e cada elemento dessa matriz é chamado de pixel (abreviatura de “*picture element*”) (GONZALEZ *et al.*, 2003).

$$f(x,y) = \begin{bmatrix} f(0,0) & f(0,1) & \dots & f(0, M-1) \\ f(1,0) & f(1,1) & \dots & f(1, M-1) \\ \dots & \dots & \dots & \dots \\ f(N-1,0) & f(N-1,1) & \dots & f(N-1, M-1) \end{bmatrix} \quad (1)$$

Uma imagem digital monocromática, ou em escala de cinza, pode ser considerada como uma matriz onde seus elementos representam apenas a intensidade luminosa naquele ponto. A Figura 1 mostra uma representação numérica de uma imagem digital monocromática de tamanho 10x10 pixels, onde se utilizou uma escala de 255 tons de cinza, quantificado de 0, ausência total da luminosidade, a 254 maior intensidade considerada.

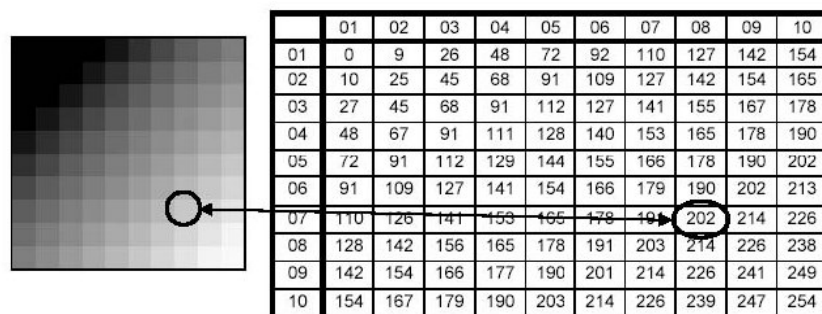


Figura 1 – Representação numérica de uma imagem digital monocromática.

Em princípio, quanto mais denso for o espaço da imagem, ou seja, menor o espaço entre os pixels, maior o grau de detalhamento possível na imagem, e quanto maiores os valores possíveis assumidos pelos pixels, maior a qualidade da imagem. Entretanto outros fatores como a quantidade de níveis de cinza, ruído e sistema óptico do equipamento de captura interferem na qualidade da imagem (BATISTA, 2002).

No modelo usado em imagens coloridas, chamado de RGB (Iniciais de *Red*, *Green* e *Blue*), por exemplo, a imagem é decomposta nas cores vermelho, verde e azul, transformando a imagem em uma matriz tri-dimensional, estabelecendo nos pixels uma intensidade para cada uma das três cores, a soma dos três valores forma a cor final, logo quanto maiores os valores das intensidades, maior será a quantidade de cores da imagem e melhor a qualidade da imagem e conseqüentemente maior será o seu tamanho.

2.2 Transformada Discreta do Cosseno

A teoria das transformadas matemáticas ganhou importância em processamento de sinais. As transformadas são inversíveis e geram um conjunto de coeficientes a partir dos quais é possível restaurar as amostras originais da imagem (BATISTA, 2002).

Uma transformada adequada à compressão de dados produz coeficientes estatisticamente descorrelacionados e concentra a energia do sinal em poucos coeficientes de valor absoluto elevado.

A Transformada Discreta do Cosseno (DCT) vem sendo intensamente utilizada para compressão de imagem, sendo o método usado nos padrões JPEG e MPEG (SALOMON, 2004).

A DCT é uma função linear e inversível, $R \rightarrow R$, que expressa os dados como uma soma de funções cosseno. O sinal original é convertido para o domínio da frequência e é possível converter o sinal de volta para o domínio do tempo ou espaço aplicando-se a DCT inversa. Existem vários tipos de DCT, sendo a DCT-II bidimensional, ou DCT-II 2D, a mais utilizada para imagens, sendo inclusive a implementada no padrão JPEG (RAO, 1990).

Uma propriedade importante da DCT-II 2D é a separabilidade, o que permite a transformada ser processada em duas etapas, através do processamento da DCT-II unidimensional primeiramente nas linhas e depois nas colunas, conforme mostra a Figura 2.

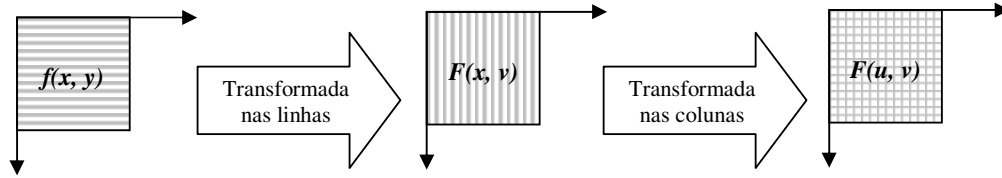


Figura 2 - Computação da DCT-II com o processamento em duas etapas. Adaptado de (KHAYAM, 2003).

Seja uma função $f(x)$, com $x = 0, 1, \dots, N-1$, a DCT-II unidimensional de $f(x)$ gera uma transformada $F(u)$ cujos coeficientes $u = 0, 1, \dots, N-1$, são dados por:

$$F(u) = \left(\frac{2}{N}\right)^{1/2} \alpha(u) \sum_{n=0}^{N-1} f(x) \cos \left[\frac{(2n+1)u\pi}{2N} \right], \quad u = 0, 1, \dots, N-1$$

onde

$$\alpha(u) = \begin{cases} (1/2)^{1/2} & \text{para } u = 0 \\ 1 & \text{para } u = 1, 2, \dots, N-1 \end{cases}$$

Esta operação leva vetores de um determinado domínio (comumente, do domínio do tempo ou do espaço) para o chamado domínio da frequência. A função $f(x)$ pode ser recuperada aplicando-se a $F(u)$ a DCT inversa:

$$f(x) = \left(\frac{2}{N}\right)^{1/2} \sum_{k=0}^{N-1} \alpha(x) F(u) \cos \left[\frac{(2n+1)k\pi}{2N} \right], \quad x = 0, 1$$

A definição da DCT-II 2D é apresentada na equação (4), considerando que a imagem original corresponde à matriz de escala de cinza $f(x,y)$, de dimensões $a \times b$, com $a = b = N$. A aplicação da DCT-II 2D produz então a matriz $F(u,v)$, também de dimensão $a \times b$. As variáveis x e y são as coordenadas no domínio espacial e u e v são as coordenadas no domínio da frequência.

$$F(u,v) = \alpha(u)\alpha(v) \sum_{x=0}^{a-1} \sum_{y=0}^{b-1} f(x,y) \cos\left[\frac{(2x+1)u\pi}{2N}\right] \cos\left[\frac{(2y+1)v\pi}{2N}\right] \quad (4)$$

em que,

$$\alpha(u), \alpha(v) = \begin{cases} \sqrt{\frac{1}{2}}, & u, v = 1 \\ \sqrt{\frac{1}{2N}}, & \text{demais casos} \end{cases}$$

A função original $f(x,y)$ pode ser recuperada aplicando-se a DCT bidimensional inversa em $F(u,v)$:

$$f(x,y) = \frac{1}{\sqrt{2N}} \sum_{u=0}^{a-1} \sum_{v=0}^{b-1} \alpha(u)\alpha(v) F(u,v) \cos\left[\frac{(2x+1)u\pi}{2a}\right] \cos\left[\frac{(2y+1)v\pi}{2b}\right] \quad (5)$$

O primeiro coeficiente, $F(0,0)$, é referenciado como coeficiente DC (*Direct Current*) e depende apenas do brilho da imagem. Os demais coeficientes de $F(u,v)$ indicam a amplitude correspondente do componente de frequência de $f(x,y)$ e são referenciados como coeficientes AC (*Alternate Current*) (SALOMON, 2004). Esta nomenclatura é oriunda da análise de sinais elétricos, onde o sinal DC representa a corrente constante e o sinal AC representa a corrente variando no tempo.

A Figura 3 exibe uma visualização do resultado da DCT aplicada em duas imagens. Claramente, as imagens têm a sua energia concentrada, ou seja, antes da aplicação da DCT, as intensidades se encontram espalhadas, quando aplicada a DCT, a energia da imagem é concentrada para a região de baixa frequência (isto é, região superior esquerdo da imagem).

Um olhar mais cuidadoso na primeira imagem da Figura 3 (a), indica que ela contém grandes áreas de suaves diferenças de intensidades. Estas imagens podem ser classificadas como de baixa frequência com poucos detalhes espaciais. A DCT operando nestas imagens fornece uma excelente compactação de energia na região de baixa frequência da imagem transformada. Isto significa que, com poucos coeficientes de baixas frequências, é possível restaurar a imagem com pouca perda de qualidade. No entanto, a Figura 3 (b) é uma imagem

de conteúdo espacial de alta frequência. Conseqüentemente, os coeficientes da transformada estão distribuídos por baixas e altas frequências (KHAYAM, 2003).

Além de ser uma poderosa ferramenta na compressão de imagens, a Transformada Discreta do Cosseno vem mostrando excelentes resultados em metodologias de reconhecimento facial, não só no requisito precisão, assim como no custo computacional (KOHIR *et al.*, 2000; ZANG *et al.*, 2006; ER *et al.*, 2005).

Na Seção 2.3 é descrita uma abordagem baseada em seleção de coeficientes da DCT que obteve um bom resultado em reconhecimento facial, e que foi a abordagem utilizada no trabalho desta dissertação.

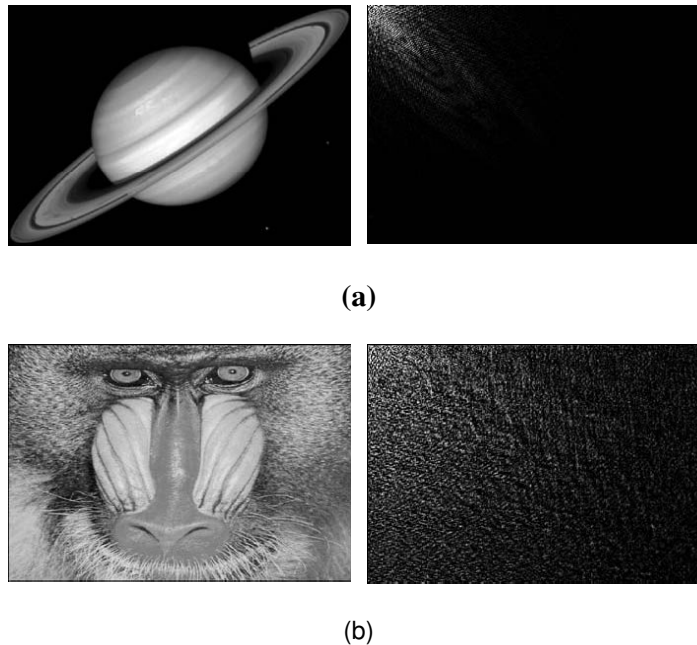


Figura 3 - Aplicação da DCT em imagens monocromáticas e suas visualizações. Fonte: Khayam, (2003).

2.3 Banco de Faces

Em pesquisas sobre abordagens de processamento facial (detecção, extração reconhecimento entre outros) são usados bancos de imagens faciais. Imagens faciais, são

imagens que contêm uma ou mais região que representa uma face humana, onde a classificação do que é uma face é feita por intuição humana, já que atributos visuais da face podem ser facilmente distintos de outros objetos na cena por olho humano. Esta distinção é muitas vezes mais complexa por métodos de detecção facial, como veremos mais adiante nesta dissertação.

Os bancos de imagens faciais existentes trazem algumas variações de aspectos que são fundamentais nos testes das abordagens propostas, pois estas variações têm influências na aparência da face, por exemplo, iluminação, pose, expressão facial, idade, oclusão, adornos, etc. Entretanto, em relação aos aspectos modificadores da face humana, temos bancos com maior controle desses aspectos enquanto outros o fazem de forma mais reduzida. Uma relação e caracterização de bancos de faces 2D referenciados na literatura estão disponíveis em (LI e JAIN, 2005).

Um banco bastante referenciado na literatura é o FERET (*Facial Recognition Technology*) (PHILLIPS *et al.* 2000) , que surgiu diante de repetidos pedidos de órgãos governamentais sobre o estado atual da tecnologia de reconhecimento facial, o que levou o Departamento de Defesa dos Estados Unidos (DOD) a desenvolver um programa oficial para estabelecer um novo conjunto de avaliações. As imagens do banco foram coletadas em um ambiente semi-controlado.

Para manter certa coerência ao longo de todo o banco de dados, a mesma configuração física foi utilizada em cada sessão de fotografia. Por causa da necessidade de reagrupar o equipamento para cada sessão, houve alguma pequena variação em imagens coletadas em datas diferentes. As imagens do FERET foram coletadas em 15 sessões entre agosto de 1993 e julho de 1996. O banco de dados contém 1.564 conjuntos de imagens para um total de 14.126 imagens de 256 X 384 pixels de resolução, que inclui 1.199 pessoas e 365 conjuntos de imagens duplicadas.

O controle nas imagens do banco FERET pode ser visto na Figura 4 que mostra três conjuntos de imagens.

O primeiro conjunto, Figura 4 (a), mostra cinco categorias de imagens frontais do banco, onde as duas primeiras, *fa* e *fb*, foram capturadas com diferentes expressões faciais. As mudanças são tipicamente sutis, muitas vezes comutadas entre "neutro" e "sorrindo". As imagens da categoria representada pela imagem *fc* foram gravadas com câmeras diferentes e em diferentes condições de iluminação sendo que um certo número de pessoas voltaram a ser

fotografadas em uma data. A categoria indicada pela imagem *D.I* são imagens duplicadas onde as pessoas foram fotografadas de 0 a 1.031 dias entre as sessões de fotografia (mediana de 72 dias, média de 251 dias), a outra categoria de “duplicados”, representada pela imagem *D.II*, 18 meses, pelo menos, separaram as sessões (mediana de 569 dias, média de 627 dias).

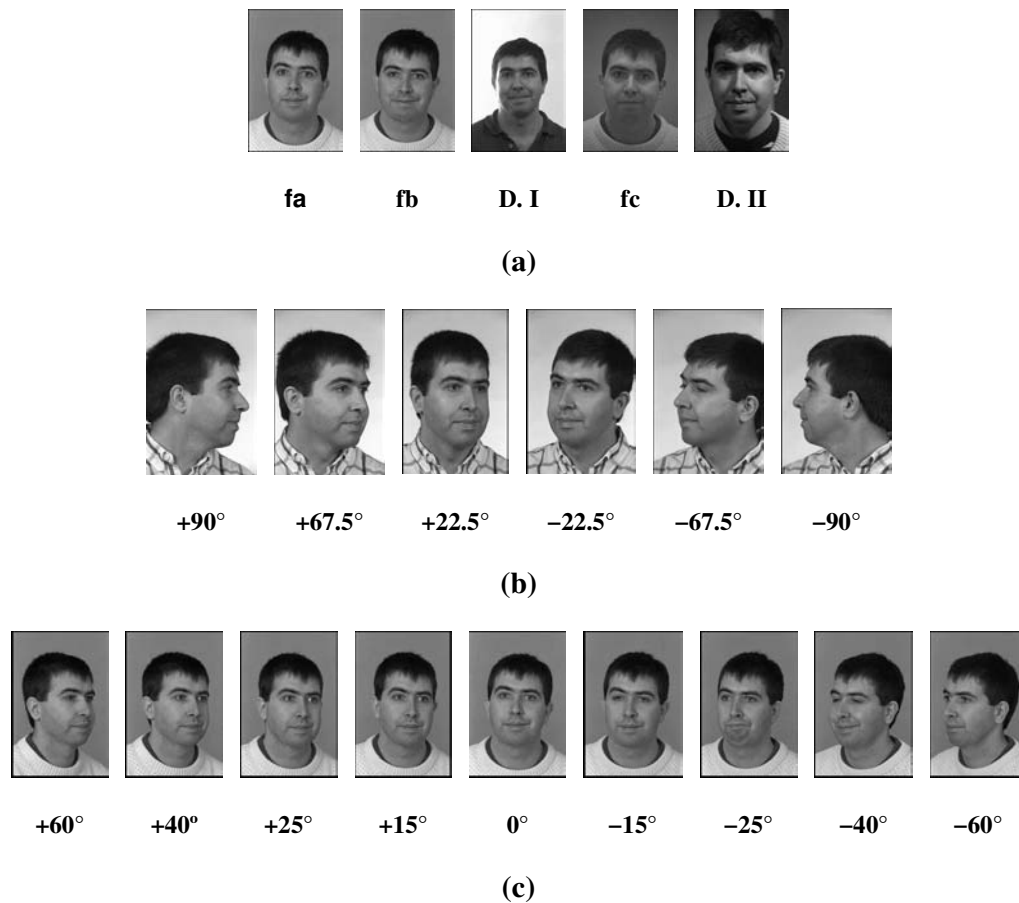


Figura 4 – Amostras de três conjuntos de imagens de uma pessoa do banco FERET. Fonte: Li e Jain, (2005).

No segundo conjunto, Figura 4 (b), foram coletadas as imagens considerando o perfil da cabeça e certa variação de ângulo: grande, 90° , médio, $67,5^\circ$ e pequeno $22,5^\circ$, tanto para esquerda (ângulos positivos) quanto para a direita (ângulos negativos).

O último conjunto, Figura 4 (c), representa a categoria em que a pessoa rotaciona a cabeça e o corpo em um intervalo de $+60^\circ$ a -60° , essas imagens foram coletadas para 200 pessoas (LI e JAIN, 2005).

Um outro banco também bastante referenciado na literatura é o banco ORL (*Olivetti*

Research Lab) *Face Database* (AT&T, 2009) que contém 400 imagens de 112 x 92 pixels, referentes a 40 pessoas, cada uma em 10 poses.

No banco há homens e mulheres e todas as imagens foram tomadas contra um fundo escuro homogêneo com as pessoas em uma posição vertical, frontal (com tolerância de alguns movimentos laterais). Para alguns temas, as imagens foram tomadas em momentos diferentes, variando as condições de iluminação, expressões faciais (olhos abertos / fechados, sorrindo / não sorrindo) e acessórios (óculos / sem óculos). Ele está disponível na WEB para download sem necessidade de pedido ou de senha, sendo um dos mais referenciados nos trabalhos de reconhecimento facial. A Figura 5 ilustra amostras de imagens do banco de faces ORL apresentando as pessoas de número 6, 7, e 8 respectivamente, cada uma em suas 10 poses.

Algoritmos de detecção facial normalmente devem ser treinados com imagens de faces e não-faces para criar uma representação interna de uma face humana, e para isto os bancos FERET e ORL são bastante referenciados e usados em pesquisas recentes (YANG *et al*, 2002).



Figura 5 – Amostras de Imagens do banco ORL.

Para avaliar comparativamente o desempenho dos algoritmos de detecção facial, conjuntos de imagens comuns são necessários. Estes conjuntos de imagens devem ser imagens representativas do mundo real contendo faces em diferentes orientações contra um fundo complexo. Nos últimos anos dois bancos de imagens públicos emergiram como quase padrão na avaliação de abordagens: a combinação MIT / CMU *DataBase* para testes em detecção de faces frontais (ROWLEY *et al.*, 1998); e CMU *Test Set II* para testes de detecção de faces frontais e não frontais (SCHNEIDERMAN *et al*, 2000).

Esses bancos contêm imagens em escala de cinza de diversos tamanhos, contendo cenas variadas. As Figuras 6 e 7 mostram algumas imagens do banco MIT / CMU *DataBase* e CMU *Test Set II*, respectivamente.



Figura 6 – Amostras de imagens do MIT / CMU Data Base



Figura 7 – Amostras de imagens do CMU Test Set II.

2.4 Reconhecimento de Faces

Numerosas abordagens vêm mostrando um considerável avanço na área de reconhecimento facial (CHELLAPPA *et al.*, 1995; ZHAO *et al.*, 2003), sendo esta área um

campo especializado da Visão Computacional que trata de imagens faciais e que envolve tarefas que são naturais de um ser humano, como detectar e reconhecer uma face humana em uma imagem (CHAI *et al.*, 2004).

As abordagens de reconhecimento facial geralmente comparam uma imagem de uma face com outras armazenadas em um banco de imagens, sendo que algumas também utilizam vídeos como fonte de aquisição de imagens faciais, sendo observadas três abordagens (PHILLIPS *et al.*, 2000):

1. **Imagem-para-imagem** onde tanto a consulta como a bases de dados é constituída por imagens estáticas;
2. **Vídeo-para-vídeo** onde tanto a consulta como a base de dados é constituída de vídeos;
3. **Imagem-para-vídeo** em que a consulta é feita por um vídeo comparado com imagens estáticas armazenadas na base de dados.

Aplicações que envolvem vídeos estão sendo utilizadas em situações que envolvem o monitoramento realizado por computador (CHIN *et al.*, 2005; PENTEADO *et al.*, 2008), enquanto que as abordagens que consideram as imagens estáticas também possuem um campo de aplicação abrangente, principalmente em aplicações relacionadas a segurança que vão desde a autorização de acesso físico e lógico até as mais sofisticadas aplicações de reconhecimento de suspeitos (SRISUK *et al.*, 2003; HEMERY *et al.*, 2008; LANG e HONG, 2008).

A Figura 8 mostra três escopos aplicados em processamento de imagem facial (CHAI *et al.*, 2004): captura; análise e compressão.

A captura é o procedimento onde se adquire a imagem contento uma face, a compressão se refere à compactação da imagem e a análise de imagem facial é o escopo mais abrangente se dividindo em três subáreas: detecção, extração e reconhecimento.

Detecção facial consiste em detectar uma face em uma imagem dada, isso pode ser feito observando atributos visuais que distingue uma face humana de outros objetos na cena (YANG *et al.*, 2002). A detecção facial desempenha um papel muito importante em um sistema de reconhecimento, sendo que a detecção eficiente ajuda no desempenho global do sistema.

Extração envolve a separação da região facial do fundo da cena. E por fim o reconhecimento consiste em identificar uma pessoa comparando sua face com a de outras pessoas armazenadas em um banco de imagens.

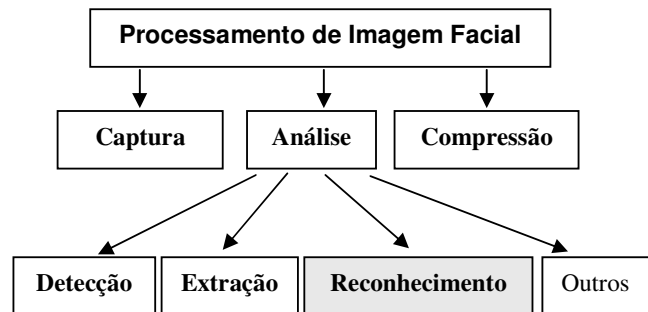


Figura 8 - Visão Geral do Processamento de Imagem Facial. Adaptado de CHAI *et al.* (2004).

Outras áreas de pesquisas relacionadas a faces, que são referenciadas na literatura, correspondem à detecção de faces em vídeo e análise de expressões faciais (YANG *et al.*, 2002; ZHAO *et al.*, 2003; CHAI e WRONG, 2004;).

Pesquisas apontam a importância de detectar expressões faciais em alguns tipos de aplicações, segundo Pantic *et al.* (2000): “... um sistema que realize essas operações com precisão e em tempo real proporcionaria um grande passo para alcançar um ser humano”.

Apesar desta atividade ser para o homem fácil de se processar, para o computador não é, pois envolve várias etapas inclusive a classificação da expressão (por exemplo, nas categorias de emoção).

Também existem pesquisas relacionadas à normalização de faces, as quais estudam métodos que são aplicados as imagens faciais a fim de obter-se uma face que potencialize a eficiência do reconhecimento. Na Seção 2.4 são descritos algumas destas abordagens.

O cérebro humano é capaz de realizar rapidamente o reconhecimento de uma face, essa tarefa é fácil para o ser humano que pode identificar corretamente uma pessoa a partir de sua imagem facial mesmo sobre as mais diversas condições, como variações de iluminação, observando apenas uma de suas características ou partes, e até mesmo com distorções ou deformações, o que é uma tarefa complexa em máquinas, que se faz com métodos sofisticados e robustos (ADLER e SCHUCKERS, 2007).

Observa-se na literatura a existência dos métodos: holísticos, estruturais e híbridos (ZHAO *et al.*, 2003):

1. Os **métodos holísticos** são os que usam toda a região da face sem identificar características físicas individuais, eles estão em grande evidência na literatura por sua alta precisão e baixo custo computacional.
2. **Métodos estruturais**, baseados em características, são técnicas mais recentes que utilizam medidas geométricas (ângulos e distâncias) relativas entre diversos pontos notáveis da face, como olhos, nariz, boca e bochechas.
3. **Métodos híbridos** tentam oferecer o melhor dos dois métodos, na tentativa de se aproximar do sistema de percepção humano, que se utiliza tanto da aparência global da face quanto das características locais. Estes métodos possuem em comum a dificuldade de comparação quando a aparência das características muda de forma significativa, como por exemplo, olhos fechados, olhos com óculos ou boca aberta.

Alguns métodos holísticos que consideram abordagens estatísticas estão conseguindo obter altas taxas de reconhecimento na literatura, como é o caso do modelo de Markov e da Análise de Componentes Principais (*Principal Componentes Analysis* – PCA), conhecido como *Eigenfaces* quando aplicado ao reconhecimento facial, assim como a Análise de Discriminantes Lineares de Fisher (*Linear Discriminant Analysis* - LDA) conhecida como *Fisherfaces* quando aplicados ao reconhecimento facial (TURK E PENTLAND, 1991; KOHIR *et al.*, 2000) outros métodos se baseiam nas transformadas matemáticas, como DCT, DWT (*Discrete Wavelet Transform*) e KLT (*Karhunen-Loève Transform*) (YU *et al.*, 2006; KOHIR *et al.*, 2000; FAÚNDEZ-ZANUY, 2003), esses métodos também são combinados com outros algoritmos, como por exemplo, redes neurais (ZHANG *et al.*, 2006; ER *et al.*, 2005).

Em condições de laboratório, os algoritmos de reconhecimento de face podem apresentar taxas de erros bastante aceitáveis. A solução encontrada tem sido restringir os problemas de captura de imagens pelo fornecimento de condições controladas, como por exemplo, a normalização das imagens faciais detectadas o qual compensa variações que possam existir em uma face, tratando em especial aspectos como iluminação, aproximação e posição, produzindo ao final uma imagem o mais próxima possível do padrão do banco de dados da comparação (MATOS, 2008).

O Anexo A apresenta um resumo com os resultados obtidos em um estudo comparativo (Tabela 10), na qual é mostrado o desempenho de métodos existentes. Verificando a Tabela 10, os trabalhos envolvendo a DCT produzem resultados com taxa de acertos mais elevada que outros métodos, sendo também relativamente independente de fatores como iluminação, posição e escala (KOHIR e DESAI, 2000; FAÚNDEZ-ZANUY, 2003; ZHONG et al., 2004).

2.5 Reconhecimento Facial utilizando Seleção de Coeficientes da DCT

O trabalho descrito em MATOS (2008) apresenta uma abordagem baseada em seleção de coeficientes da DCT, tratando o problema de reconhecimento de faces como um problema genérico de classificação, em que se tem uma etapa de treinamento e a classificação de uma amostra.

A Figura 9 ilustra o método de reconhecimento de faces, visto de forma genérica, dividido em duas etapas que são padrões no processamento do reconhecimento facial, a etapa de Treinamento e a de Classificação.

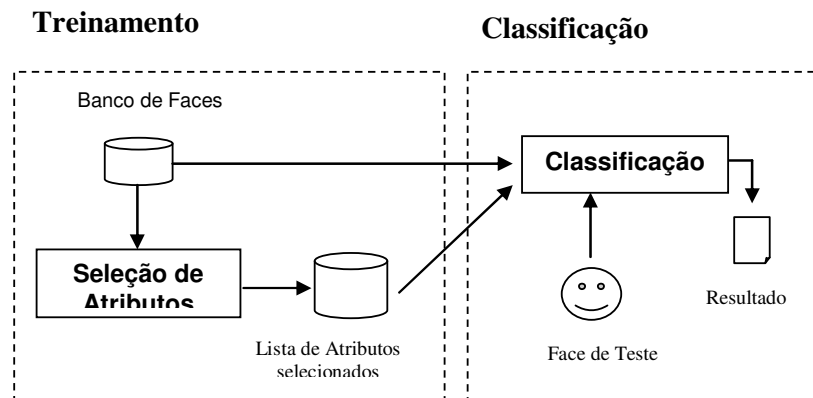


Figura 9 - Um sistema genérico de Reconhecimento. Adaptado de MATOS et al. (2008).

Na etapa de Treinamento é realizada a extração de características das imagens faciais armazenadas no banco de faces, essa etapa considera que as faces já foram previamente detectadas e normalizadas. A normalização tem o objetivo de eliminar dados desnecessários ou distorcidos fazendo com que a entrada apresente apenas dados relevantes para o

reconhecimento da face em análise.

A seleção de atributos gera a lista de atributos que melhor representa a face da pessoa analisada utilizando as várias poses armazenadas no banco de faces. O objetivo principal da etapa de seleção de atributos é caracterizar classes distintas de objetos por atributos que sejam adequados para a classificação.

A extração de características é uma etapa importante para todo o processo já que realiza a análise dos dados de entrada a fim de extrair e derivar informações úteis para o processo de reconhecimento.

Embora exista uma distinção entre seleção e extração de atributos, na literatura verifica-se de forma generalizada o uso do termo seleção para expressar redução de dimensionalidade de imagens de faces, que consiste no número de atributos utilizados para representar a imagem de face. Portanto, após a aplicação da DCT sobre a imagem original seguido da seleção de coeficientes, teremos a representação da face com a sua dimensionalidade reduzida.

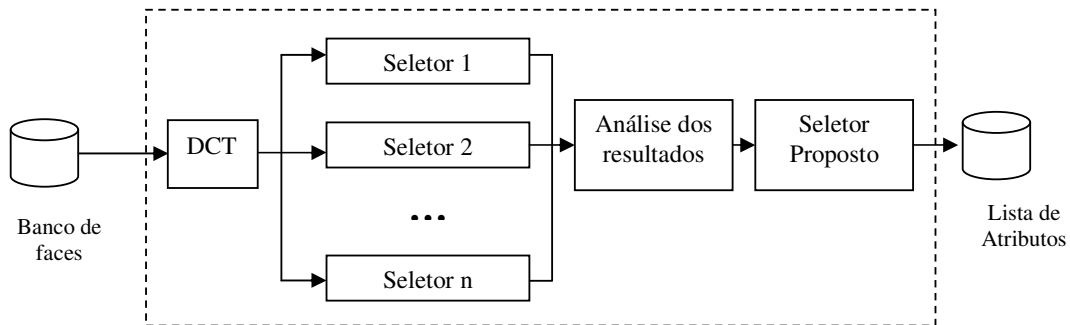


Figura 10 – Seleção de Atributos. Adaptado de Matos, 2008.

Vários seletores podem ser utilizados para a extração de características, como mostra a Figura 10, permitindo a análise dos resultados obtidos para posterior escolha do seletor mais eficiente ao reconhecimento. As abordagens de seleção podem ser agrupadas em duas categorias (MATOS, 2008):

1. Seletores que identificam um conjunto de coeficientes por pessoa. Estes seletores escolhem um conjunto de coeficientes para cada pessoa. Fazem parte deste grupo os

seguintes seletores: média das amplitudes, média excluindo extremos, mediana, contagem de ocorrências, posição média dos coeficientes e outros.

2. Seletores que identificam um único conjunto de coeficientes para todas as pessoas. Neste grupo, os seletores selecionam um único conjunto de coeficientes para representar todas as pessoas do banco. Fazem parte deste grupo os seguintes seletores: distância normalizada pela variância, baixas frequências e outros.

A etapa de classificação é onde acontece a comparação entre a face de teste e as faces armazenadas no banco de faces. Para isso, considera-se a lista de atributos selecionados para indicar a quem pertence a face de teste.

A classificação é feita baseada em probabilidade, pois uma classificação computacional perfeita é considerada impossível (DUDA *et al.*, 2000). Sendo assim, o classificador calcula um valor que indica o grau de semelhança da face de teste em relação a uma no banco de treinamento. O grau de semelhança é computado como uma função dos atributos selecionados e é utilizado para indicar a classe que mais se aproxima do objeto de teste.

O conjunto de atributos selecionados para distinguir as faces influencia diretamente no resultado do classificador e normalmente verifica-se que uma seleção apropriada conduz a uma maior taxa de acertos.

Algumas das abordagens mais utilizadas, para a classificação, são: distância mínima ao protótipo (DMP), vizinho mais próximo (*Nearest Neighbor* - NN) e *k*-vizinhos mais próximo (*k-Nearest Neighbor* - KNN) (MATOS, 2008).

Na subseção 2.4.3 é descrito o classificador vizinho mais próximo e na subseção 2.4.4 o classificador *k*-vizinhos mais próximos. A subseção 2.4.1 descreve o seletor de baixas frequências e a subseção 2.4.2 o seletor média das amplitudes, visto que estes foram os pares seletor/classificador utilizados na implementação, e testes, do trabalho descrito nesta dissertação, bem como foram as abordagens que demonstraram melhores resultados nos testes realizados em MATOS (2008) sobre o banco de faces ORL, como mostra a Tabela 1, que relaciona o desempenho dos demais seletores (no total de oito seletores) usando os três classificadores : DMP, NN, KNN.

Os melhores resultados dos testes realizados foram os que utilizaram o seletor de baixas frequências com o classificador de vizinho mais próximo (99,25% de taxa de acerto com 36 coeficientes). A Tabela 1 também mostra o número de coeficientes (entre parênteses)

que obteve o melhor resultado. Os testes foram feitos utilizando a abordagem de testes chamada de *leave-one-out* (deixe um de fora) descrito na seção 3.2.3.

Tabela 1 - Resumo do desempenho dos seletores de baixas frequências e média das amplitudes e classificadores NN e KNN.

	NN	KNN (K=3)
Baixas frequências	99,25 (36)	97,75 (400)
Média das amplitudes	98,75 (36)	97,00 (121)

Fonte: Matos (2008).

2.5.1 Seletor de Baixas Frequências

Segundo MATOS (2008), essa abordagem de seleção é simples, não avalia valores dos coeficientes DCT das imagens de treinamento, não realiza cálculos e nem qualquer tipo de comparação. Consiste simplesmente na seleção dos coeficientes DCT de mais baixa frequência, considerando uma região quadrada com início no primeiro coeficiente da matriz DCT, utilizando os coeficientes que concentram maiores energia.

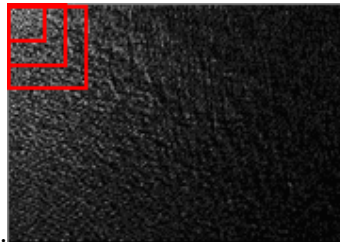


Figura 11 - Regiões quadradas associadas a seleção de baixas frequências.

A Figura 11 ilustra a abordagem de seleção por baixas frequências indicando três regiões quadradas de seleção, começando pelo elemento (0,0) até o elemento (x,x), sendo x menor que cada uma das dimensões da matriz de coeficientes.

2.5.2 Seletor Média das Amplitudes

A média aritmética é uma medida estatística de tendência central largamente utilizada, que resume informações de um conjunto de dados em torno do seu ponto central de distribuição. Ela é definida como a razão entre o somatório dos valores e o número de valores.

Considerando que a amplitude representa o valor do coeficiente em módulo, então a média aritmética da amplitude do i -ésimo coeficiente da pessoa j é dada por:

$$\bar{\mu}_{ij} = \frac{1}{q} \sum_{k=1}^q |x_{ijk}| \quad (6)$$

Sendo q o número de pose de treinamento de cada pessoa e x_{ijk} o valor do i -ésimo coeficiente da pessoa j na pose k . O conjunto de coeficientes selecionados para identificar a pessoa j corresponde aos m coeficientes que apresentem maior valor de $\bar{\mu}_{ij}$.

O seletor baseado na média de amplitude é uma abordagem de seleção simples e fácil de implementar. Ele se concentra apenas nos valores dos coeficientes das poses de uma dada pessoa, sem considerar dispersão dos dados nem qualquer relacionamento de semelhanças ou diferenças representativas que possam ocorrer entre pessoas distintas (MATOS, 2008).

2.5.3 Classificador do Vizinho Mais Próximo (NN)

Sejam $\{y_{1j}, y_{2j}, \dots, y_{mj}\}$ os m coeficientes selecionados para representar a pessoa j e sejam $\{w_{1jk}, w_{2jk}, \dots, w_{mjk}\}$ as amplitudes dos coeficientes de treinamento da pessoa j na pose k , com w_{ijk} correspondendo ao coeficiente de mesma posição que y_{ij} .

Seja f a pessoa a ser classificada e sejam $\{v_{1f}, v_{2f}, \dots, v_{mf}\}$ as amplitudes dos coeficientes DCT da pessoa f , com v_{if} correspondendo ao coeficiente de mesma posição

que y_{ij} .

A distância entre a face f e a pessoa treinamento j na pose k , com $j=1, 2, \dots, p$ e $k=1, 2, \dots, q$, é dada por:

$$DNN_{fjk} = \sum_{i=1}^m |w_{ijk} - v_{if}| \quad (7)$$

A face de teste f é classificada como a pessoa j quando:

$$DNN_{fjk} \leq DNN_{fgh}, \forall j \neq g, \forall k \neq h \quad (8)$$

2.5.4 Classificador dos K-Vizinhos Mais Próximo (KNN)

A abordagem de classificação KNN é semelhante à abordagem NN. Entretanto, ao invés de classificar a face de teste como a primeira face de treinamento mais próxima, ele identifica as k faces mais próximas e classifica a face de teste considerando a classe que for mais freqüente entre as k faces de treinamento mais próximas (MATOS, 2008).

Sejam $\{y_{1j}, y_{2j}, \dots, y_{mj}\}$ os m coeficientes selecionados para representar a pessoa j e sejam $\{w_{1jk}, w_{2jk}, \dots, w_{mjk}\}$ as amplitudes dos coeficientes de treinamento da pessoa j na pose k , com w_{ijk} correspondendo ao coeficiente de mesma posição que y_{ij} .

Seja f a pessoa a ser classificada e sejam $\{v_{1f}, v_{2f}, \dots, v_{mf}\}$ as amplitudes dos coeficientes DCT da pessoa f , com v_{if} correspondendo ao coeficiente de mesma posição que y_{ij} .

O classificador KNN classifica a face f baseado nos seguintes passos:

1. Calcula-se a distância entre a face f e a pessoa de treinamento j na pose k , com $j=1, 2, \dots, p$ e $k=1, 2, \dots, q$, dada por:

$$DKNN_{fjk} = \sum_{i=1}^m |w_{ijk} - v_{if}| \quad (9)$$

2. identificam-se os k menores valores $DKNN_{fjk}$ e define-se o i -ésimo vizinho mais próximo como a pessoa j associada ao i -menor valor de $DKNN_{fjk}$, com $i=1, 2, \dots, k$;
3. A classificação da pessoa f corresponde então à pessoa j mais frequente entre os k -vizinhos identificados;
4. Em caso de empate, foi definido nesse trabalho se adotar a abordagem NN para desempate.

2.6 Detecção Facial

A detecção facial em uma imagem representa uma etapa de considerável importância em um sistema de reconhecimento facial. Geralmente em uma imagem existem outros objetos no fundo além de uma ou mais faces. A localização precisa da face em uma cena é uma tarefa importante, pois influenciará no desempenho geral do sistema visto que isso permitirá a extração futura de características que estão relacionadas somente à face em questão, desprezando assim ruídos, ou seja, dados que não representam características diretamente relacionadas a face.

A grande dificuldade para desenvolver um detector de faces robusto é que faces humanas não seguem um padrão rígido, e variam bastante em relação a forma, cor, tamanho e textura. A detecção facial em uma imagem é uma tarefa desafiadora devido à variabilidade na dimensão, localização, orientação (para cima/direita, rotação), e pose (frontal perfil). Expressão facial, oclusão, e condições de iluminação também alteraram a aparência geral da face (YANG *et al.*, 2002).

Os métodos de detecção facial procuram extrair características que são comuns nas faces humanas, considerando tanto a face como um todo quanto as partes que a constituem, por exemplo, boca, nariz e olhos. Sem considerar características locais, a detecção de faces é

considerada bem sucedida se a presença de uma ou mais faces for indicada corretamente assim como a sua localização. Sem a correta detecção e localização das características faciais é observada uma queda de desempenho no reconhecimento.

Até meados da década de 1990, a maioria dos trabalhos sobre a detecção foi centrada em uma única face em fundos simples ou complexos. Esta é uma versão simplificada do problema de detecção, pois se assume que existe somente uma face na imagem, o que resume a solução do problema em detectar a existência e localização de características, tais como os olhos, nariz, narinas, sobrancelhas, boca, lábios, orelhas, etc. Existem abordagens que consideram o rosto por completo e modelos baseados em características deformáveis, como a cor da pele, e abordagens que usam rede neural (ROWLEY *et al.* 1998).

Avanços significativos têm sido feitos nos últimos anos na realização da detecção automática em condições diversas. Comparados a métodos baseados em extração de características e modelos que combinam métodos diferentes (modelos híbridos), trabalhos baseados em aparência ou métodos baseados em treinamento de imagens com grande número de amostras alcançaram os melhores resultados. Isto pode não ser surpreendente, uma vez que objetos faciais são complexos, muito semelhantes entre si, e diferentes de objetos não-faces. Através de treinamento extensivo, os computadores podem ser muito bons em detectar faces (ZHAO *et al.* 2003).

Apesar de existirem diferentes abordagens, alguns problemas são comuns a todas elas, como (YANG *et al.*, 2002):

- **Pose:** a imagem de uma face varia de acordo com a posição da câmera que captura a imagem e alguns elementos faciais, tais como um olho ou o nariz pode aparecer parcial ou totalmente ocluso;
- **Expressão Facial:** a expressão da face influencia diretamente na aparência da imagem de face;
- **Presença de Elementos Estruturais:** a presença de elementos como barba, bigode e óculos que podem modificar as características em termos de tamanho, luminosidade, etc;
- **Ocultação:** no caso de imagens feitas em ambientes não controlados as faces podem aparecer parcial ou totalmente sobrepostas, por objetos ou até mesmo por outras faces;

- **Orientação da Imagem:** imagens faciais podem variar diretamente sobre o eixo óptico da câmera, possibilitando, por exemplo, a captura de uma imagem de uma face de cabeça para baixo proporcionando uma disposição não regular dos elementos faciais (olhos, boca, nariz etc);
- **Condições da Imagem:** quando a imagem é capturada, fatores como a iluminação (aspectos, fonte de distribuição e intensidade), características da câmera (sensor de captura, lentes) afetam a aparência da face.

Resumidamente, as abordagens de detecção facial estão organizadas em quatro categorias:

1. **Métodos baseados em conhecimento:** utiliza conhecimento humano para codificar o que significa uma face, ou seja, utilizam alguma base de regras estabelecida a partir do conhecimento prévio sobre o problema, ou seja, métodos que possuem regras que definem o que é uma face, de acordo com o conhecimento de uma pessoa. Por exemplo, sabemos que faces humanas possuem determinadas características válidas para a maioria de suas instâncias, como: todas as faces em geral possuem dois olhos, um nariz e uma boca. Além disso, esses elementos se encontram distribuídos de maneira específica sobre a face. Com este conhecimento é possível estabelecer regras que identificam uma face humana (KOTROPOULOS e PITAS, 1997; YANG *et al.*, 2002).
2. **Técnicas baseadas em características invariantes:** essas técnicas procuram utilizar características estruturais que existem mesmo quando a pose, o ponto de vista ou as condições de iluminação variam. Particularmente, estes métodos são inspirados na capacidade que os seres humanos possuem de identificar objetos independentemente do ponto de vista (SABER e TEKALP, 2003).
3. **Métodos baseados na comparação de templates:** a face é descrita como um template e então a detecção é feita com base na correlação entre o template e a entrada. Uma técnica clássica de detectar objetos é procurar pelo mesmo dentro da imagem e testar se ele corresponde a um modelo prévio da sua forma. Uma das formas mais comuns de modelar a forma de um objeto é descrevê-lo através de seus componentes geométricos básicos, como círculos, quadrados ou triângulos, esta técnica é

denominada *template*. A detecção do objeto, portanto, consistirá em achar a melhor correspondência, definida através de uma função de energia, entre o objeto presente na imagem e o seu molde (*template*). No caso de detecção de faces a *template* mais utilizada é aquela que trata a face como uma elipse (ALATTAR e RAJALA, 1999).

4. **Métodos baseados na aparência:** utiliza modelos de faces treinados de uma base de treinamento para realizar a detecção de faces. Esses métodos recebem este nome devido ao fato de não utilizarem nenhum conhecimento *a priori* sobre o objeto ou características a ser detectada. Nesta classe de algoritmos surgem os conceitos de aprendizado e treinamento, uma vez que as informações necessárias para realizar a tarefa de detecção são retiradas do próprio conjunto de imagens sem intervenção externa. Alguns métodos bastante referenciados na literatura fazem parte dessa categoria: *eigenfaces* baseado na transformada de *Karhunen-Loève* (KLT), ou PCA (*Principal Component Analysis*) (TURK e PENTLAND, 1991); Redes Neurais Artificiais (RNA's) são modelos matemáticos biologicamente inspirados, com o intuito de emularem os processos cognitivos do cérebro humano (ROWLEY *et al.* 1998); os Modelos Ocultos de Markov (*Hidden Markov Models - HMM*) (ARA, 1998).

Um problema relacionado e importante é como avaliar o desempenho dos métodos propostos. Muitos documentos recentes de detecção de face comparam o desempenho de vários métodos, normalmente em termos de taxa de detecção e falsos alarmes. Também vale notar que muitas métricas foram adotadas para avaliar algoritmos, como tempo de treinamento, tempo de execução, o número de amostras requeridas no treinamento, e a relação entre a taxa de detecção e falsos alarmes, ou mesmo a relação entre o número de faces detectadas corretamente e o número de faces determinado por um humano.

Em geral, detectores podem cometer dois tipos de erros: *falsos negativos* em quais as faces são perdidas, o que resulta em baixas taxas de detecção e *falsos positivos* nos quais uma região de imagem é considerada como uma face, mas não é. Uma avaliação justa deveria levar estes fatores em consideração (YANG *et al.*, 2002).

Na subseção 2.5.1, é apresentada uma breve descrição de um método proposto para a detecção de faces utilizando um modelo baseado nas características invariantes que realiza a detecção de faces frontais de forma rápida e robusta em tempo real.

2.6.1 O método Viola e Jones.

O método Viola e Jones (VIOLA e JONES, 2004) é utilizado em imagens simples em tons de cinza e com isso se distingue de outros modelos, pois não é baseada em informações auxiliares como, por exemplo, a diferença de sequência de vídeos e o uso de cores. No entanto, mesmo com maiores limitações de informações utilizadas, foram obtidas altas taxas de detecção. Este método baseia-se em três conceitos básicos, o primeiro é o uso de características de *Haar*.

Características invariantes são originados no trabalho (PAPAGEORGIOU *et al.*, 1998), que evidencia o uso de *wavelets Haar* para a identificação de diferenças de intensidade entre regiões diferentes. Essas características são avaliadas rapidamente por meio de uma nova forma de representação da imagem chamada "Imagem Integral".

O segundo conceito é a aplicação do algoritmo para gerar um conjunto de características baseadas no conceito da Imagem Integral. Esse algoritmo de aprendizado (*boosting*) é chamado de *AdaBoost*, o qual filtra o conjunto de característica no intuito de diminuir o número de características geradas.

O terceiro conceito é a geração de uma estrutura de árvore dos classificadores fracos *boosted*, para que gere inferências rápidas e robustas na construção de um classificador forte (VIOLA E JONES, 2004). A Figura 12 ilustra a aplicação desses conceitos.

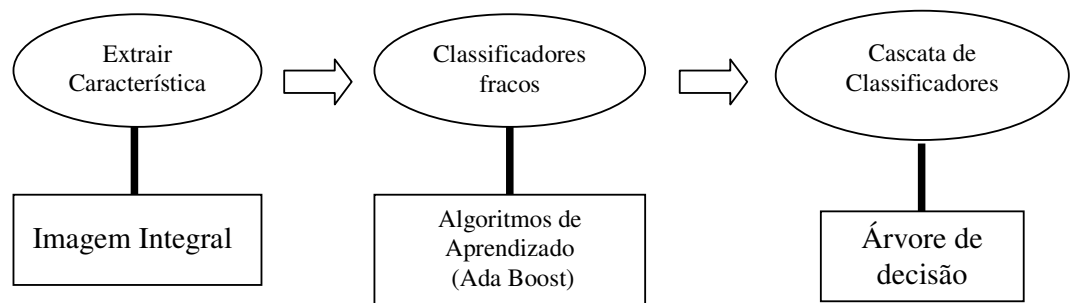


Figura 12 - Principais conceitos envolvidos no método Viola e Jones.

O conceito da Imagem Integral serve para determinar eficientemente a presença ou ausência de características de *Haar* em cada imagem em diversas escalas e locais. Em geral, "integrar" significa acrescentar pequenas unidades juntas. Neste caso, as pequenas unidades

são valores dos pixels. O valor integral para cada pixel é a soma de todos os pixels acima dele e à sua esquerda. Começando no canto superior esquerdo e atravessando para a direita e para baixo, toda a imagem pode ser integrada com um pequeno número inteiro de operações por pixel. A definição formal dessa operação é dada pela equação 10:

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y') \quad (10)$$

Onde $ii(x, y)$ é a imagem Integral de uma imagem $i(x, y)$.

Conforme a Figura 13 (a) após a integração, o valor em cada pixel de localização (x, y) , contém a soma de todos os valores dentro de uma região retangular do canto superior esquerdo da imagem até (x, y) . Para encontrar o valor médio dos pixels neste retângulo, é preciso dividir o valor em (x, y) pela área do retângulo.

A Figura 13 (b) mostra uma imagem dividida em quatro regiões distintas (retângulos A, B, C e D) e quatro pontos (1, 2, 3 e 4).

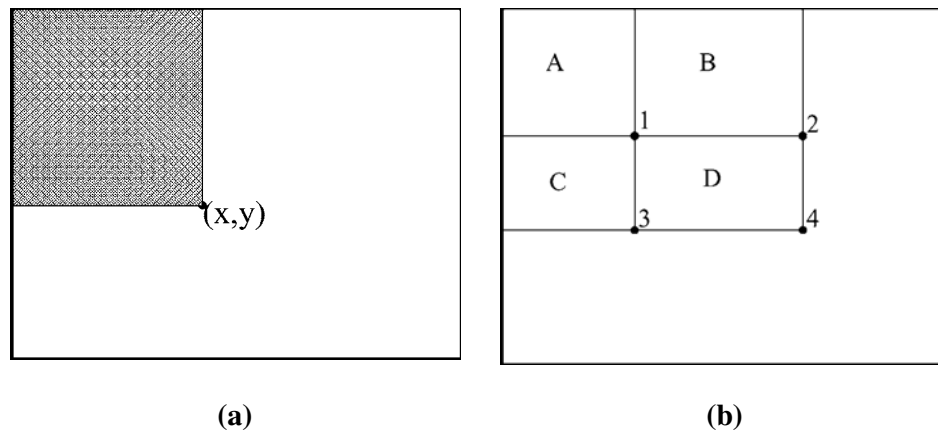


Figura 13 – Imagem Integral. Fonte: Viola E Jones, 2004.

Supondo que se deseja saber o valor da Imagem Integral no retângulo D. A soma dos pixels dentro do retângulo D pode ser calculada da seguinte forma: o valor integral da imagem na posição indicada no ponto 1, que é a soma dos pixels no retângulo A. O valor da localização 2, que é a soma dos pixels nos retângulos A e B, no ponto 3 é a soma dos pixels nos retângulos A e C, e na posição 4 é a soma dos pixels nos quatro retângulos A, B, C e D. A

soma dentro de D pode ser computada como sendo a diferença do valor no ponto 4 e a soma das regiões A, B e C, sendo assim temos :

$$D = I_4 - (A + B + C) \quad (11)$$

Sendo que $B = I_2 - I_1$, $C = I_3 - I_1$ e $A = I_1$ então temos:

$$D = I_4 + I_1 - I_2 - I_3 \quad (12)$$

Wavelets de Haar são ondas quadradas de comprimentos únicos, em que são representadas por um par de retângulos adjacentes - um claro e outro escuro. O método Viola e Jones considera combinações de retângulos (quatro retângulos) mais adequadas para a tarefa de detecção facial, portanto não usa o conjunto original completo desses retângulos.

A Figura 14 mostra um exemplo de aplicação de dois retângulos desse conjunto adaptados ao método Viola e Jones. Apesar de o método considerar quatro retângulos, a Figura 14 mostra um caso em que apenas dois retângulos são necessários para detectar uma face com eficácia, já que avaliar todo o conjunto de retângulos em cada sub-imagem se torna computacionalmente dispendioso. Aqui vale a nota que o algoritmo procura as faces percorrendo a imagem inteira considerando varias sub-imagens de tamanhos diferentes a partir de 24 x 24 pixels, então quanto mais diminuir a quantidade de resoluções de sub-imagens mais eficiente será o algoritmo.

As características nada mais são do que avaliações do conjunto das informações de intensidade de um ponto (pixel). Esse processo consiste na soma da intensidade dos pixels nas regiões brancas das características, subtraído da soma da intensidade dos pixels da imagem, nas regiões pretas das características. Os resultados são usados como valor da característica em uma determinada localização e podem ser combinados para a formação de hipóteses fracas nas imagens.

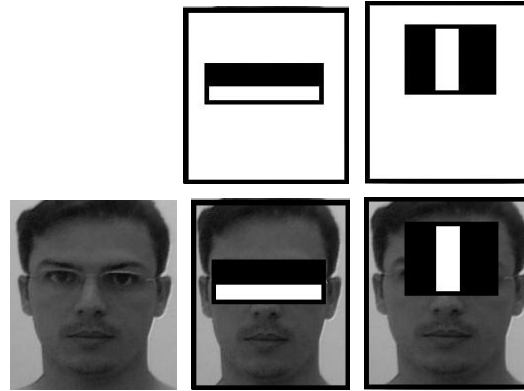


Figura 14 – Aplicação dos “retângulos de Haar” no método Viola e Jones.

As características ilustradas na Figura 14 são a primeira e segunda características selecionadas pelo algoritmo *AdaBoost*. As duas características são mostradas na linha superior e, em seguida, aplicadas em uma imagem de face típica na linha inferior. A primeira característica mede a diferença em intensidade entre a região dos olhos e região abaixo dos olhos. A característica baseia-se na observação que a região de olho é frequentemente mais escura que a região abaixo dos olhos. A segunda característica compara as intensidades nas regiões de olho e a intensidade na ponte do nariz. Essa característica se baseia na propriedade que a região dos olhos é geralmente mais escura do que da ponte do nariz.

Para selecionar características *Haar* específicas, e para definir limiares, o método Viola e Jones usa uma máquina de aprendizagem chamado *AdaBoost*. O *AdaBoost* combina muitos classificadores "fracos" para criar um classificador "forte". "Fracos" aqui, significa que o classificador só obtém a resposta certa um pouco mais frequentemente do que o faria se adivinhasse aleatoriamente. Porém, se existir um conjunto inteiro desses classificadores fracos, cada um obtendo resposta na direção certa, pode-se ter um classificador forte, combinando as forças para chegar à solução correta. *AdaBoost* combina um conjunto de classificadores fracos e atribui um peso para cada. Esta combinação ponderada é o classificador forte.

O método Viola e Jones combina uma série de classificadores *AdaBoost* em cadeia, como filtros, mostrados na Figura 15. Cada filtro C_i , $i = 1, 2, \dots, N$, é um distinto classificador *AdaBoost* com um número relativamente pequeno de classificadores fracos.

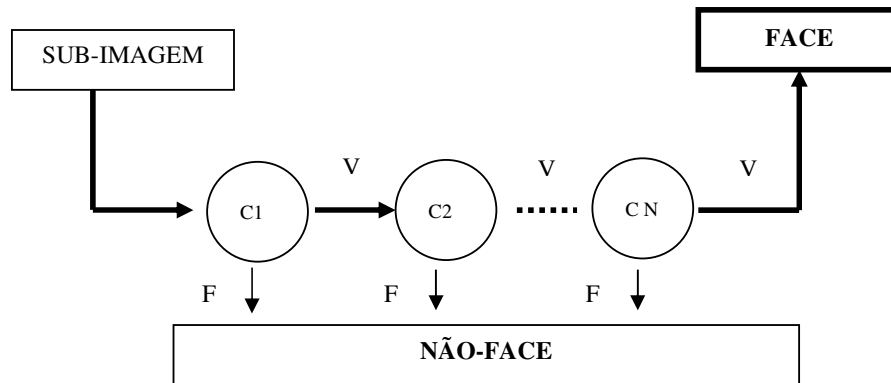


Figura 15 – Aplicação em série de classificadores AdaBoost no método Viola e Jones.

O limiar de aceitação em cada nível é fixado como baixo o suficiente para passar todas, ou quase todas, as faces de exemplo no conjunto treinado. Os filtros são treinados em cada nível para classificar imagens de treinamento que passaram por todas as etapas anteriores.

O conjunto de treinamento é um banco de faces, com cerca de mil ou mais imagens. Durante a utilização, se qualquer um destes filtros para não classificar uma região de imagem, essa região é imediatamente classificada como "não-face". Quando um filtro passa uma região de imagem, ela vai para o próximo filtro na cadeia. Regiões de imagem que passam por todos os filtros na cadeia são classificadas como "face". Esta filtragem em cadeia é chamada de cascata.

No Anexo B é mostrado o algoritmo *AdaBoost* modificado para os classificadores Viola e Jones, assim como o algoritmo de Treinamento utilizado na construção da cascata.

A Figura 16 ilustra a aplicação do algoritmo em duas imagens do banco MIT/CMU, demonstrando a sua eficiência na detecção de faces frontais e em diversas escalas considerando diversos cenários e número de faces existentes.



Figura 16 - A aplicação do algoritmo Viola e Jones em duas imagens do banco MIT/CMU.
Fonte: Viola e Jones (2004).

2.7 Ambientes Virtuais de Aprendizagem

A EAD tem sido uma alternativa de ensino/aprendizagem em um cenário marcado pelas dificuldades de acesso da nossa população ao ensino formal de modalidade presencial. Sendo assim, a EAD surgiu para superar as distâncias geográficas, no entanto várias outras distâncias nos afastam em virtude das diferenças culturais, sociais, afetivas ou econômicas.

E-learning é basicamente a Educação via Internet (E-LEARNING BRASIL, 2009), rede local ou mesmo através de um único computador e está relacionado ao uso de sistemas eletrônicos, aplicativos e processos para o aprendizado. Aplicativos e processos incluem ensino baseado na Internet, ensino baseado em computadores, salas de aula virtuais e colaboração digital. O conteúdo pode ser disponibilizado via Internet, rede local, CD/DVD, fitas de áudio e vídeo, transmissão de TV, e até mesmo por celulares.

Nesse contexto, os AVA, também conhecidos como Sistema de Gerenciamento da Aprendizagem (*Learning Management System*) ou Sistema de Gerenciamento de Cursos (*Course Management System*), vêm ganhando espaço como uma ferramenta poderosa de

auxílio ao ensino à distância, disponibilizando várias funcionalidades que facilita ao tutor, ou professor, na comunicação com os seus alunos, fóruns, *chats*, *wiks*, murais, listas de discussão, ajudam na elaboração de atividades didáticas.

Os AVA são sistemas computacionais (softwares) com interfaces gráficas amigáveis que facilitam a disponibilização de conteúdos, permitindo a colaboração e a comunicação entre o corpo docente e discente de uma instituição de ensino. Consiste, na maioria dos casos, de um conjunto de mecanismos que apresenta informações de maneira organizada, realiza troca de mensagens (envio e recebimento), permite a inserção e acesso a conteúdos educacionais e o gerenciamento de informações, além de possibilitar o gerenciamento de tarefas pedagógicas e de atividades administrativas, eles foram desenvolvidos com focos educacionais para serem utilizados como suporte para cursos à distância *on-line*, assim como apoio nas atividades presenciais, “permitindo expandir as interações da aula para além do espaço-tempo, do encontro face-a-face, ou para suporte às atividades de formação semi-presenciais onde o ambiente digital poderá ser utilizado tanto nas ações presenciais como nas atividades à distância” (ALMEIDA, 2008).

Existe no mercado uma enorme quantidade de AVA, alguns gratuitos, como MOODLE (2008), e outros pagos, como Blackboard/WebCT (WEBCT, 2008). Segundo Mariño (2007) e Ono *et al* (2007), o MOODLE é um dos mais populares, melhores e mais bem sucedidos AVA gratuito do mercado. Em 2006, algumas avaliações apontaram essas ferramentas de comunicação e a criação e administração de elementos de aprendizado, além dos amplos conceitos didáticos, como os pontos fortes do MOODLE, sendo considerado o melhor ambiente virtual de aprendizado *Open Source* avaliado (GRAF, 2006).

2.7.1 Modular Object-Oriented Dynamic Learning Environment (MOODLE)

O MOODLE é uma plataforma computacional, formada por diferentes ferramentas de comunicação, interação, inserção de documentos e gerenciamento de informações. Ele facilita a disponibilização de conteúdos, permitindo a colaboração e a comunicação entre o corpo docente e discente. Os cursos regulares, geralmente, se encontram em espaços fechados

em que os usuários se cadastram ou são previamente cadastrados e tem acesso ao conteúdo do curso apenas com autorização através de login e senha.

Durante o acesso, o ambiente rastreia e monitora, automaticamente, as ações do aluno no curso, registrando as interações deste com o ambiente e apresentando estas informações de forma que o professor possa acompanhar a frequência, a participação e o progresso dos alunos. Além disso, é registrado cada passo dado pelo aluno, mostrando assim que tipo de acesso o aluno tem feito (unidades visitadas, ferramentas utilizadas etc.), bem como o tempo gasto com o acesso desses (TAROUCO *et al*, 2007).

No entanto, o sistema de acompanhamento é realizado através de relatórios contendo informações como número de acessos, data e hora do último acesso, e frequência de acesso do aluno durante um determinado período. Nele é considerado o registro de quem informou o login e a senha de acesso, mas não há como garantir a identidade real (OTSUKA, 2002) de quem está acessando e realizando as interações e intervenções no processo de ensino-aprendizagem.

O MOODLE foi desenvolvido em código aberto e distribuído sob a Licença Pública Geral para software livre, mais conhecida como *General Public License* ou GPL. É utilizado em mais de 196 países, existem mais de 35.314 sites registrados, já foi traduzido para mais de 70 idiomas e tem uma grande comunidade de usuários com grande participação na manutenção da distribuição, sugerindo sempre modificações, novas habilidades e reportando eventuais defeitos. Embora a maioria das discussões seja da comunidade internacional, existe também a comunidade brasileira atuando no aperfeiçoamento da plataforma e contribuindo com discussões relevantes e trabalhos colaborativos, auxiliando na construção do aprendizado da comunidade sobre o ambiente (MUZINATTI, 2006).

O AVA MOODLE surgiu nos anos 90, a partir da necessidade de um *webmaster* e administrador de sistema de desenvolver uma ferramenta de software baseada na web que fornecesse, além de recursos tecnológicos integrados de compartilhamento de informação, interação e comunicação, a possibilidade de professores desenvolverem uma abordagem de ensino e aprendizagem online que realçasse o papel do aluno na aprendizagem.

Uma importante característica do projeto MOODLE é a disponibilização de uma página na Internet que está em constante atualização, buscando atender e adaptar-se às necessidades da comunidade, podendo ser acessada gratuitamente por qualquer pessoa. A

principal função da página é prover um ponto central para informação, discussão e colaboração entre os usuários.

A popularização do MOODLE no Brasil deve-se à expansão da EAD, através, por exemplo, de projetos como Universidade Aberta do Brasil (UAB, 2008) e Escola Aberta do Brasil (E-TEC BRASIL, 2008). Atualmente, existem mais de 2.000 sites registrados oficialmente utilizando o Moodle no Brasil. Tal sucesso deve-se, dentre outros motivos, aos seguintes fatores: baixo custo de implantação e disponibilização do sistema, sem royalties para uso nem custos de aquisição; fácil utilização por usuários leigos em programação visual, permitindo que a maior parte das áreas de texto (fontes, recursos, fóruns, notícias) possa ser editada, usando um editor HTML (HTML, 2008); seus conteúdos podem ser facilmente adaptados e disponibilizados para os alunos; dispõe de interfaces gráficas para a língua portuguesa; e permite suporte gratuito através da comunidade internacional de usuários e atualização gratuita de recursos via Internet.

O ambiente do MOODLE agrega várias ferramentas para o apoio ao ensino e aprendizado à distância. Essas ferramentas são classificadas de acordo com seus objetivos pedagógicos, podendo ser classificadas como Ferramentas Síncronas, que são aquelas em que os interlocutores estão conectados no ambiente ao mesmo tempo, como uma conversa em um *chat*, videoconferência, etc, ou Ferramentas Assíncronas onde os interlocutores acessam a mesma ferramenta sendo que em tempo diferente, como listas de discussão, mural e *wiks*. Entre elas, é importante mencionar os Fóruns de discussão on-line nos quais, apesar de ser uma ferramenta de comunicação assíncrona, os alunos podem prosseguir suas discussões acerca dos conteúdos estudados. Uma outra ferramenta é a aquela em que os alunos podem realizar exercícios dos mais variados tipos – com questões de múltipla escolha, discursivas e associativas.

Em suma, o MOODLE constitui-se em um conjunto sofisticado de ferramentas síncronas e assíncronas, além de possuir alguns recursos para acompanhamento das atividades dos seus usuários. No entanto, esses recursos ainda são insuficientes, pois os professores continuam tendo dificuldades tanto para acompanhar os alunos, como para garantir a identidade do usuário durante a realização das atividades.

A implementação do MOODLE é feita na linguagem PHP (um acrônimo recursivo para "*PHP: Hypertext Preprocessor*") e permite o desenvolvimento de extensões, o que é fundamental para a realização de pesquisas e a construções de módulos que complementam as funcionalidades já existentes assim como ferramentas novas. Junto ao PHP a construção do

banco de dados é feito com o MYSQL que é um verdadeiro servidor de banco de dados multiusuário e multitarefa. Ele apresenta uma arquitetura cliente-servidor multiplataforma com suporte a uma série de linguagens (CONVERSE, 2004; GREENSPAN, 2001).

A arquitetura padrão do MOODLE é a cliente-servidor, onde nos computadores Clientes rodam os programas que requisitam serviço de um ou mais computadores Servidores e estes respondem as solicitações. A Figura 17 mostra a existência de um servidor Web denominado Servidor EAD, onde está disponível o AVA (páginas em PHP) com o sistema de gerenciamento de banco de dados (SGBD) que é o MYSQL, onde ficam armazenados todos os dados gerados durante o desenvolvimento dos cursos envolvidos.

O Servidor EAD disponibiliza o AVA para a utilização de Administradores, Coordenadores, Professores, Tutores e Alunos, que utilizam o AVA em computadores Clientes, trocando mensagens (Pedido/Resposta) através da Internet.

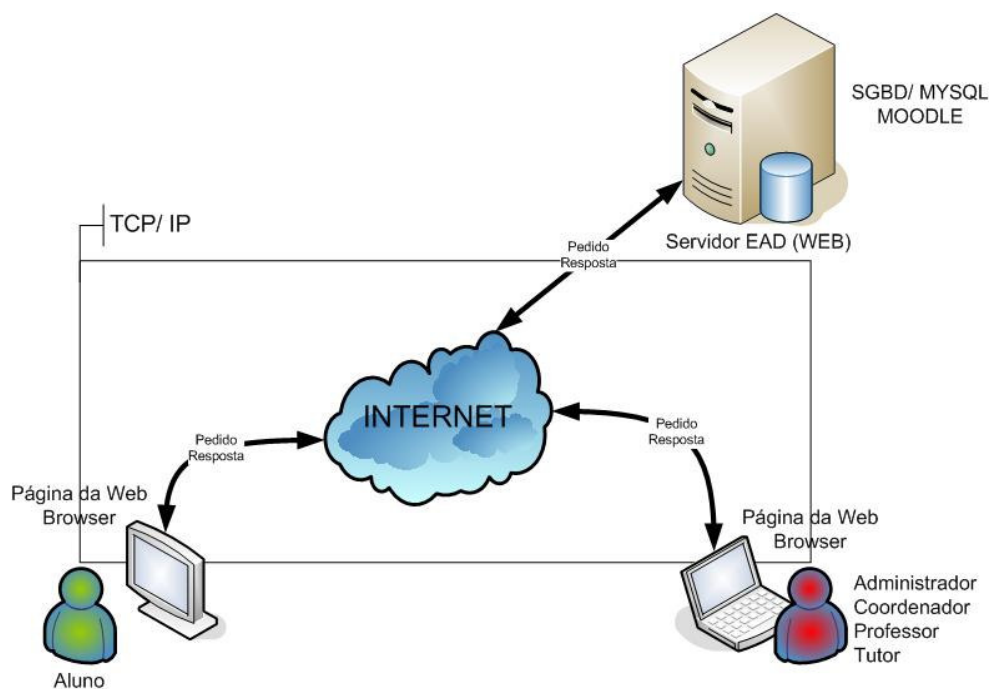


Figura 17 - Arquitetura padrão do MOODLE. Fonte: Silva, 2008.

2.8 Mecanismo de Identificação Fisionômica

O Mecanismo de Identificação Fisionômica, o MIF, nasceu da necessidade de certificar a utilização dos usuários do MOODLE durante a sua permanência desde a entrada até a saída do ambiente. O MOODLE permite o acesso ao sistema através de login e senha o que não garante a legitimidade do aluno. A certificação real deste aluno é importante em EAD, pois evita fraudes o que é importante principalmente para a sua avaliação. Em outras situações a comprovação do tempo de estudo na modalidade de educação à distância é também de suma importância. Silva (2008) apresenta um estudo de caso que corresponde a essa situação. Alunos apenados que estudam em cursos em EAD precisam comprovar o tempo de estudos para que a justiça o conceda a remissão penal que lhe é de direito, sendo assim, precisa-se garantir a identidade do aluno. Para a realização dessa certificação é necessário o desenvolvimento de um mecanismo integrado ao AVA que realize de alguma forma essa identificação, já que o mesmo não a realiza sozinho.

O MIF é um mecanismo que monitora o uso do AVA por parte do aluno, capturando progressivamente em tempos predefinidos a sua imagem facial, com o auxílio de uma webcam, permitindo assim, o registro de quem está realizando as atividades pedagógicas no MOODLE, reforçando a garantia da participação do aluno matriculado no curso.

O MIF, além de capturar as imagens, as envia e armazena em um servidor, e também as anexa em um dos relatórios de participação do MOODLE, possibilitando a análise, através da face do aluno, por parte dos professores ou quaisquer administradores do ensino à distância.

2.8.1 Arquitetura do MIF

A Figura 18 mostra um cenário proposto para o MIF em (SILVA, 2008), onde se ilustram interações entre as entidades envolvidas. O “Aluno” é o usuário do MOODLE que pode ou não ser monitorado pelo MIF através de uma webcam que captura suas imagens em intervalos estabelecidos. Este monitoramento é feito no computador Cliente e as imagens

enviadas ao Servidor EAD através da Internet. As imagens enviadas são armazenadas no Banco de Imagens naquele Servidor, que serviram para serem anexadas aos relatórios de acesso do MOODLE.

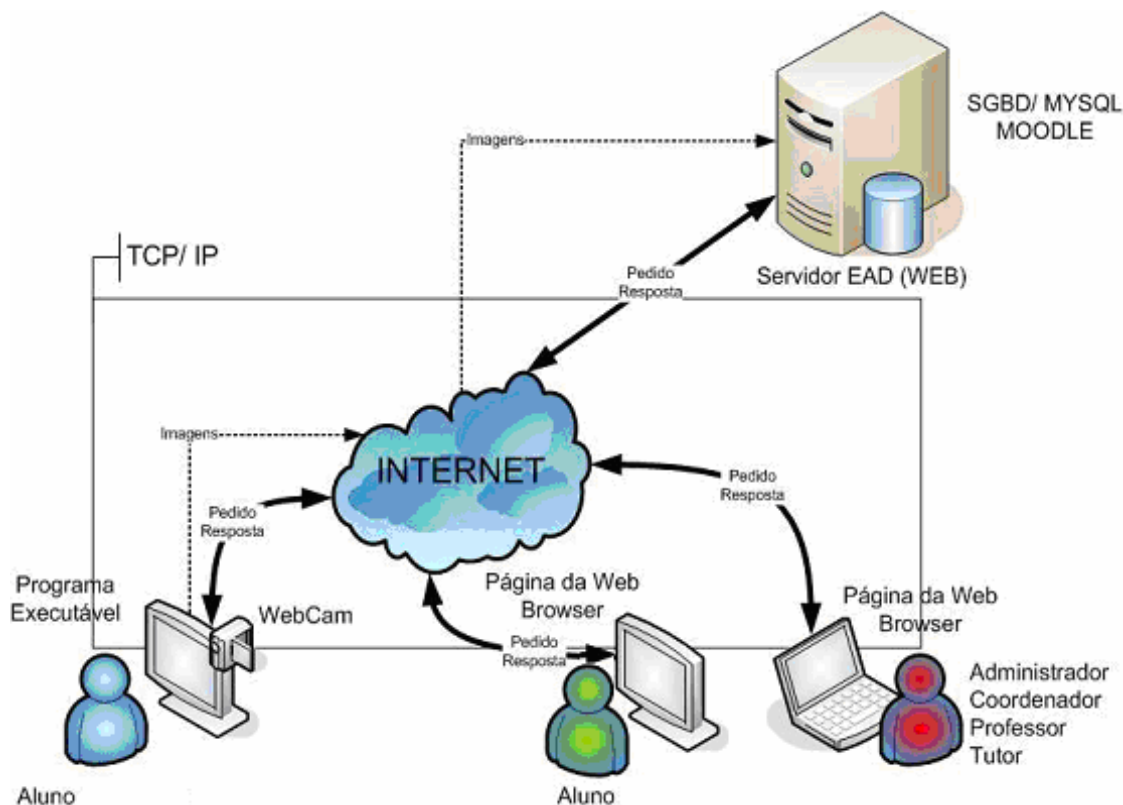


Figura 18 - Cenário proposto para utilização do MIF. Adaptado de Silva (2008).

O MIF é composto por dois subsistemas, Figura 19. O Subsistema 1 permite ao usuário entrar no MOODLE através de login e da senha. Ele abre o *browser* na página inicial do MOODLE (1); captura (2) as imagens do usuário através da webcam; envia (3) estas imagens para armazenamento no Servidor EAD. A captura é feita a cada instante predefinido e a imagem capturada é armazenada no banco de imagens. O subsistema 1 é um programa executável com a implementação na linguagem C++, o que permite a utilização do MIF em qualquer computador Cliente que tenha o Subsistema 1 instalado.

O Subsistema 2 é uma extensão do MOODLE, portanto implementado em PHP, que recupera imagens (4) no banco de imagens e as apresenta (5) nos relatórios do MOODLE.

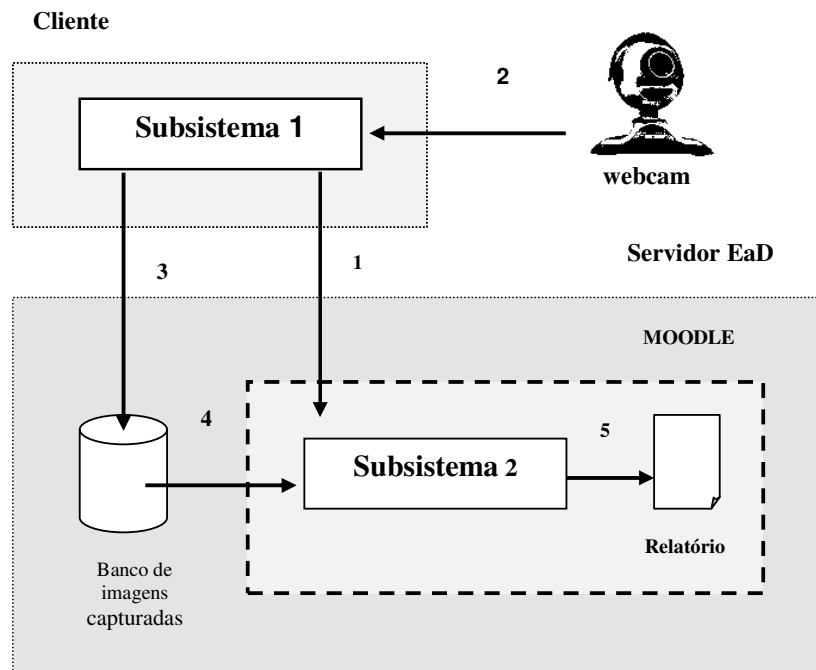


Figura 19 - Arquitetura do MIF.

A Figura 20 mostra a tela principal deste sistema. Durante a exibição desta tela, o usuário tem a oportunidade de ajustar a webcam e efetuar a entrada no MOODLE.

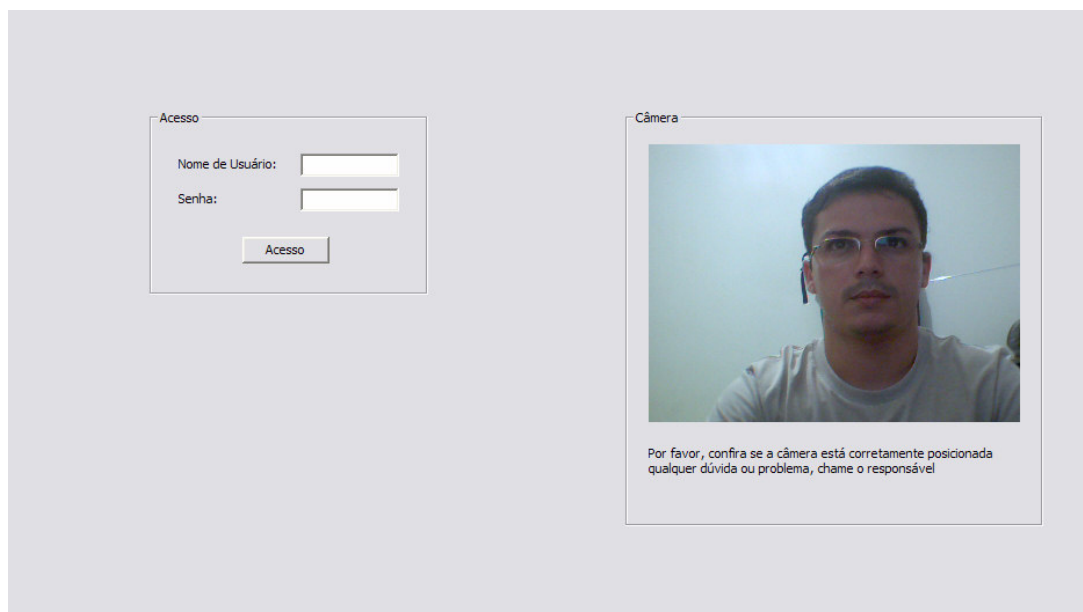


Figura 20 - Captura da tela do login, mostrando a execução do Subsistema 1 do MIF.

A Figura 21 mostra um *browser* aberto com o MOODLE, após ter entrado com um usuário no ambiente através do Sistema 1 com login e senha.

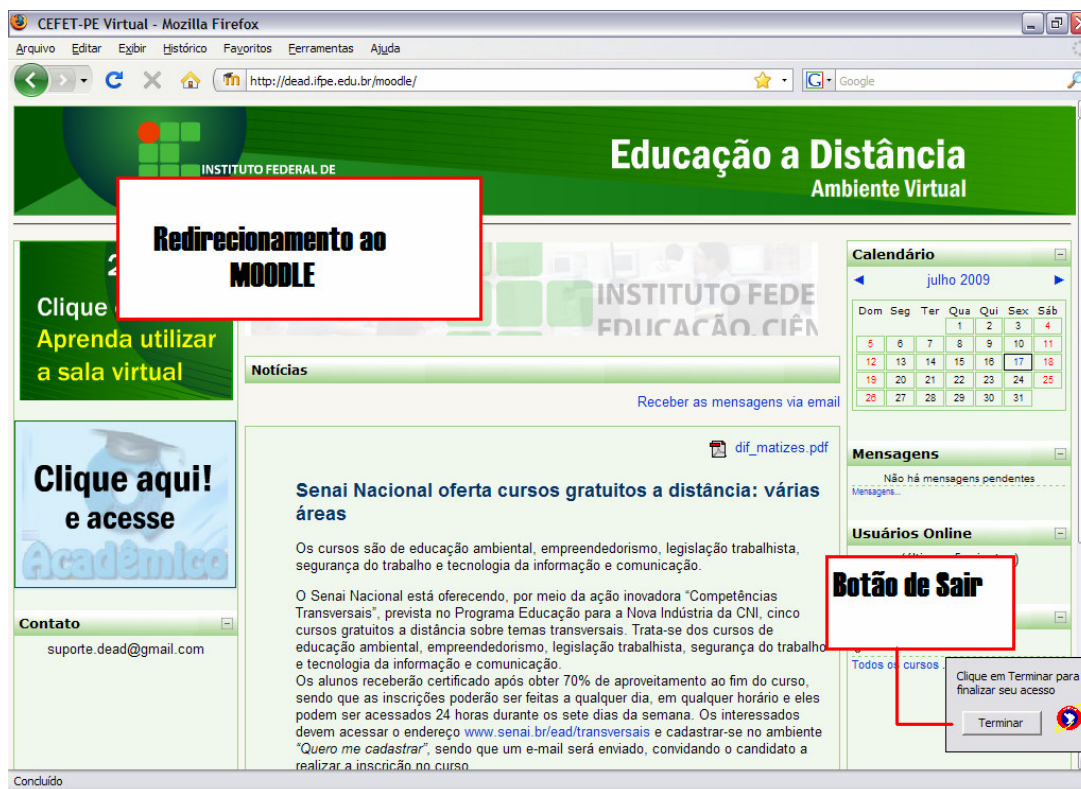


Figura 21 - Redirecionamento feito pelo subsistema 1 para o MOODLE.

O Subsistema 2 é uma extensão do MOODLE que realiza a distinção de qual usuário vai ou não apresentar a imagem anexada ao relatório. Os usuários que terão suas imagens capturadas e adicionadas ao relatório do MOODLE são distintos no ato de cadastro dos mesmos e o único relatório do MOODLE que apresenta as imagens é o de “Todos os Acessos” (SILVA, 2008).

A Figura 22 mostra um relatório de “Todos os Acessos” modificado de um usuário fictício chamado de “edporto” que entrou no MOODLE através do MIF e que teve suas imagens captadas e enviadas ao Servidor EAD, sendo que essas imagens foram anexadas ao relatório através do Subsistema 2.

Mostrando 123 registros
Página: 1 2 (Próximo)

Ed Porto	
edporto	

Imagem	Hora	Ação	Informação
	Sex 24 julho 2009, 05:18	course user report	Ed Porto
	Sex 24 julho 2009, 05:18	course user report	Ed Porto
	Sex 24 julho 2009, 05:18	user view	Ed Porto
	Sex 24 julho 2009, 05:18	user view all	Ed Porto
	Sex 24 julho 2009, 05:18	course view	UFPB -Teste Moodle
	Qui 23 julho 2009, 00:24	course user report	Ed Porto

Figura 22 – Relatório “Todos os Acessos” do MOODLE mostrando as imagens capturadas pelo MIF.

3 O SISTEMA DE IDENTIFICAÇÃO AUTOMÁTICA DE FACES PARA AMBIENTES VIRTUAIS DE APRENDIZADO (SIAF-EAD)

Este capítulo aborda a descrição do projeto, implementação e testes do SIAF-EAD, que implementa o cenário proposto na Figura 23.

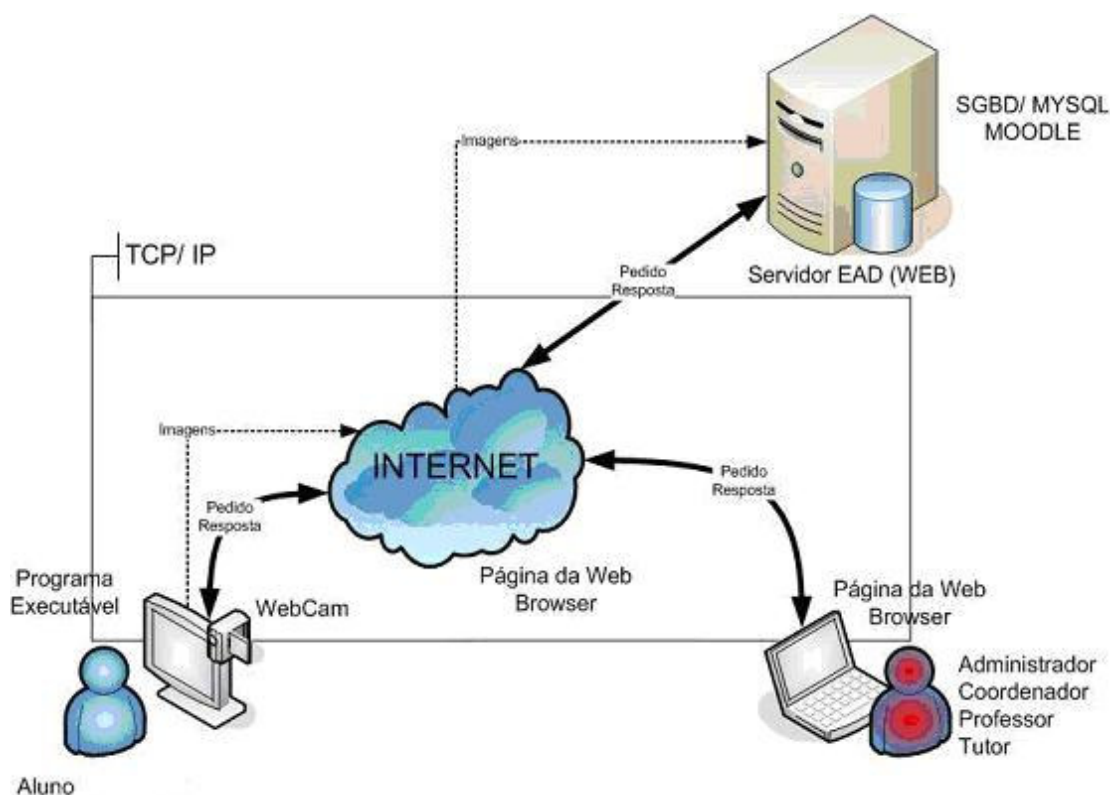


Figura 23 – Cenário de aplicação do SIAF-EAD.

A Figura 23 ilustra um usuário chamado de “Aluno” que utiliza um AVA que é executado em um *Browser*, conectando-se a um Servidor EAD através da Internet, e outros tipos de usuários, “Administrador”, “Coordenador”, “Professor” e “Tutor” que são os responsáveis por administrar o uso do AVA por parte dos usuários “Alunos”. Os usuários são monitorados através de um programa executável que captura a imagem do usuário através de uma webcam e as envia ao Servidor EAD. Este cenário é bem parecido com o proposto para o MIF, só que engloba apenas os usuários “Alunos” monitorados pela webcam, entretanto neste mecanismo não existe a garantia de que a imagem capturada seja realmente do aluno cadastrado com o login e senha.

Considerando o cenário proposto, assim como o levantamento do *estado-da-arte* no desenvolvimento de abordagens e sistemas computacionais na área de processamento facial, é proposto um sistema de arquitetura distribuída que identifica automaticamente um usuário que se submete ao uso do AVA através do reconhecimento de sua face em imagens digitais capturadas por uma webcam. A certificação da identidade do usuário será imprimida na imagem facial detectada e identificada dentre um conjunto de imagens de treinamento cadastradas no sistema, e será anexada aos relatórios do AVA para posterior consulta por parte das pessoas responsáveis.

3.1 Projeto do SIAF-EAD

O projeto do SIAF-EAD foi feito através do levantamento dos requisitos funcionais do sistema observando as necessidades envolvidas no cenário proposto. O projeto arquitetural utiliza alguns módulos já implementados no MIF, e outros módulos exclusivos para realizar a identificação automática.

3.1.1 Requisitos Funcionais do Sistema

Durante o desenvolvimento do SIAF-EAD foram levantados alguns questionamentos que originaram uma lista de requisitos funcionais, servindo para especificar o Sistema. Requisitos de um sistema descrevem os serviços oferecidos e suas restrições (SOMMERVILLE, 2008). O Quadro 1 descreve os principais requisitos do SIAF-EAD.

ID	Requisito	Descrição
(RF01)	Cadastro das imagens de treinamento dos usuários.	Cadastrar um número determinado de imagens que contenha a face do usuário do AVA Afim de determinar um conjunto para a fase de treinamento. As imagens devem ser escolhidas por uma pessoa responsável e devem seguir alguns padrões: *A cada usuário do sistema será atribuído um número de identificação única do banco de imagens (ID); *Cada imagem deve conter uma face humana que seja detectada pelo algoritmo de detecção utilizado; *As faces devem ser extraídas da cena e armazenadas no servidor; *O conjunto de poses para cada usuário obtido deve satisfazer o número adotado para o treinamento.
(RF02)	Treinamento do conjunto de faces cadastradas.	Realizar o treinamento no conjunto de faces detectadas e armazenadas, conforme (RF01), seguindo os seguintes passos: *Aplicação da DCT em todas as faces; *Realizar a seleção de atributos, utilizando o seletor que apresentar melhor resultados nos testes descritos no Capítulo 4 desta dissertação; *Armazenar a lista de atributos selecionados . *Para que essas informações sejam preservadas, e atualizadas com frequência, é importante que a operação descrita aqui seja realizada no Servidor EAD.
(RF03)	Detectar a face na imagem capturada	Detecção da face nas imagens armazenadas, conforme o (RF02), e nas imagens capturadas em uma sessão de usuário.
(RF04)	Realizar o reconhecimento da face detectada.	Reconhecer a face detectada, realizando a identificação da mesma dentre o conjunto de faces armazenadas, seguindo os seguintes passos: *Utilização de um classificador que obtenha melhor resultado nos testes descrito no Capítulo 4 desta dissertação; *Aplicação da DCT na face de teste, para a comparação com as faces do conjunto de treinamento; *Para que haja uma melhor distribuição de carga de processamento, é importante que essa operação seja realizada no lado Cliente do Sistema.
(RF05)	Certificar a imagem como identificada ou não.	Adicionar alguma informação na própria imagem, de forma que possa certificar a identidade do usuário naquele momento de utilização do AVA. *Esta certificação pode ser feita imprimindo na imagem de teste o resultado da identificação realizada em (RF05).
(RF06)	Enviar imagem certificada ao Servidor EAD.	Por fim, o envio da mensagem ao Servidor EAD, para que a mesma possa ser anexada aos relatórios do AVA, através do Subsistema 2, implementado no MIF.

Quadro1 – Requisitos Funcionais do SIAF-EAD.

3.1.2 Caso de Uso Geral do SIAF-EAD

O caso de uso mostrado na Figura 24 ilustra de um modo geral as funcionalidades e algumas dependências entre elas, essas funcionalidades serão exploradas com mais detalhes na seção 3.2.3 de acordo com o ciclo de vida do sistema.

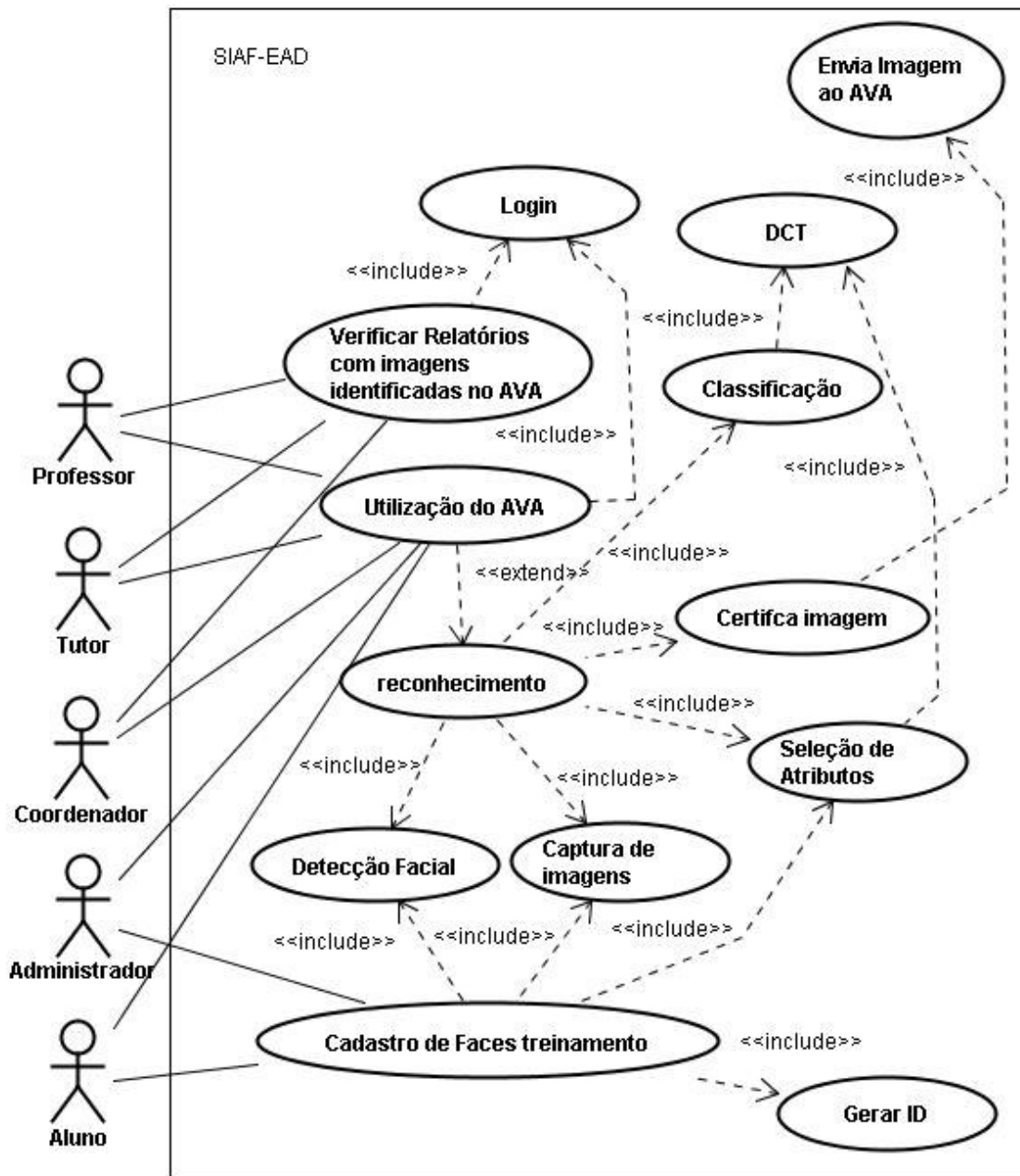


Figura 24 – Diagrama de Caso de Uso Geral do SIAF-EAD.

3.1.3 Ciclo de Vida do sistema

A Figura 25 ilustra o ciclo de vida do SIAF-EAD, mostrando seus principais estados possíveis. Nela é notada a existência de sete fases distintas: Cadastro, Treinamento, Entrada no Sistema, Captura, Detecção, Reconhecimento e Saída do Sistema.

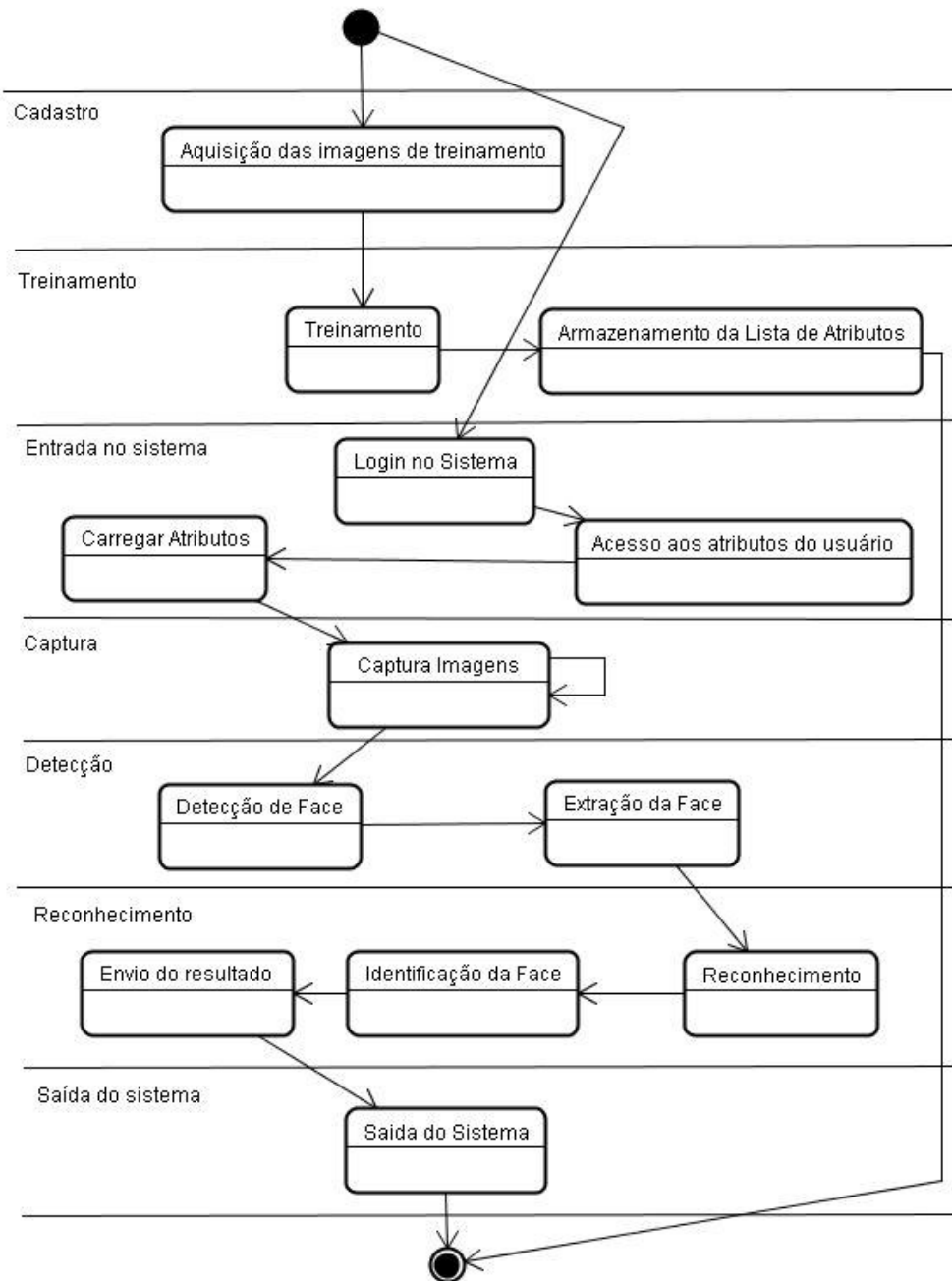


Figura 25 – Diagrama de Estados do Ciclo de vida do SIAF-EAD.

Fase 1: Cadastro

O cadastro de usuários que terão os seus acessos ao AVA monitorados pelo SIAF-EAD terá que disponibilizar algumas imagens faciais para o treinamento. A arquitetura do SIAF-EAD é distribuída baseada na arquitetura atual do MIF (cliente-servidor), portanto alguns módulos são implantados do lado Servidor e outros módulos no lado Cliente, sendo que esse conjunto de imagens deve ser armazenado no Servidor com o propósito de manter o conjunto atualizado sempre que for preciso, assim como aumentar a privacidade e segurança dos dados do usuário, já que no treinamento é gerada a lista de atributos importante na identificação do usuário. Então, baseado nessa estrutura foram criadas duas estratégias para adquirir o conjunto de imagens.

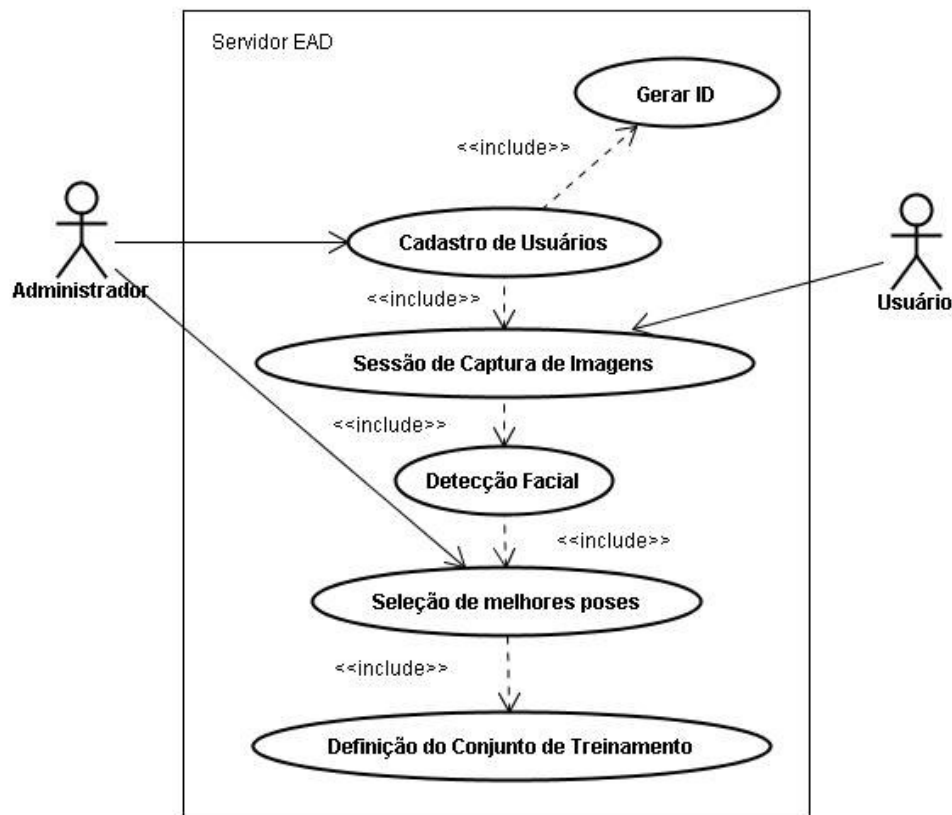


Figura 26 – Diagrama de Caso de Uso da primeira estratégia de aquisição das imagens de treinamento do SIAF-EAD.

1. A primeira estratégia é a coleta das imagens na localidade onde fica instalado a estrutura do Servidor EAD, como mostra o Caso de Uso ilustrado na Figura 26. A captura das imagens deve ser feita através de uma webcam, utilizando um módulo de software que além de realizar as capturas das imagens, detecta e extrai a imagem da face do usuário armazenando-a no Servidor EAD e também gera o ID do usuário. O conjunto das faces mais adequadas para o treinamento é escolhido pelo especialista (Administrador), já que algumas faces extraídas são detectadas de forma errônea (discutido na seção 2.3), ou de forma distorcida (desalinhada, alto grau de rotação, variação de expressão, oclusa etc);
2. A segunda estratégia é utilizar tanto o lado Cliente quanto o lado Servidor da arquitetura como mostra o Diagrama de Caso de Uso ilustrado na Figura 27. Nesta estratégia a sessão de captura das imagens e a detecção das faces do usuário são feitas no computador Cliente e em seguida enviadas ao Servidor. O Administrador será notificado que o usuário está utilizando o sistema pela primeira vez, isto significa que ele ainda não tem as poses de treinamento armazenadas. Então durante os primeiros acessos ao sistema o Administrador gera o ID do usuário e seleciona as poses mais apropriadas para o treinamento, que ficarão armazenadas no Servidor EAD.

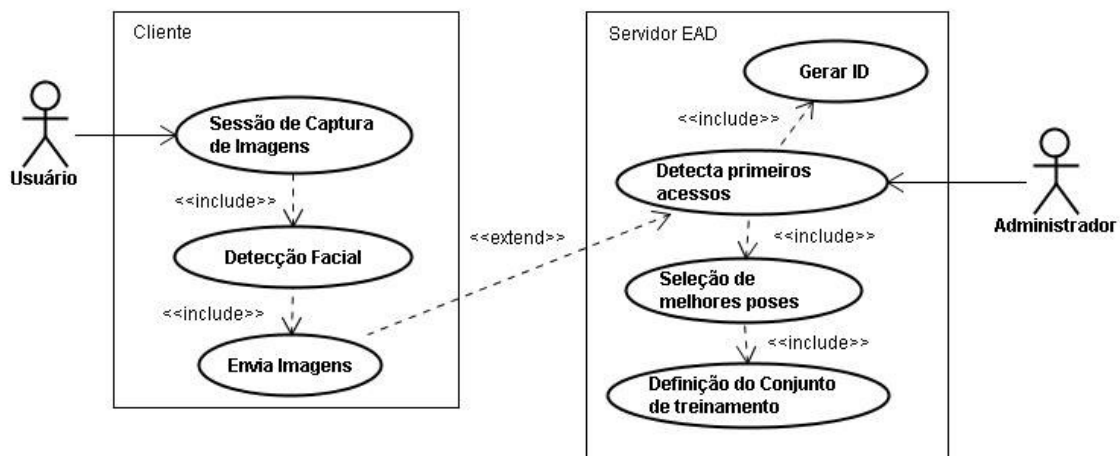


Figura 27 – Diagrama de Caso de Uso da segunda estratégia de aquisição das imagens de treinamento do SIAF-EAD.

Por enquanto, o SIAF só possui um módulo que implementa a primeira estratégia, sendo a segunda, alvo de implementações futuras. A opção da captura das imagens no computador Cliente oferece algumas vantagens, a mais evidente é a que fornece maior conveniência ao usuário já que não precisará se deslocar até a localidade de implantação do Servidor EAD.

Fase 2: Treinamento

A fase de treinamento é executada no Servidor EAD onde está disponibilizado o banco de imagens das faces detectas e selecionadas no cadastro dos usuários. O fato da fase de treinamento ser realizada no Servidor EAD é importante pois o treinamento geralmente é realizado somente uma vez, disponibilizando a lista de atributos selecionada para ser utilizada no reconhecimento.

A DCT concentra energias em poucos atributos, como visto na seção 2.1, então a seleção de poucos atributos é o suficiente para que haja uma boa taxa de acerto no reconhecimento. O Capítulo 4 mostra os resultados dos testes de reconhecimento realizados neste trabalho, e como se pode notar, foi alcançada maior taxa de acerto utilizando apenas 64 atributos da seleção de baixas frequências.

Para ilustrar um exemplo desta redução de dimensionalidade considere-se que tem-se 52 usuários do sistema e são cadastradas 8 poses de cada um. Cada imagem tem a resolução 220x240, o que resulta em 52800 pixels, a quantificação luminosa dada pelo intervalo de 0 a 255 é representada por 1 *byte*, então cada imagem é representada por 52,8 KB que multiplicado pelo total de imagens, 416, é aproximadamente 220 MB. No entanto, a representação de 64 atributos selecionados é de 26,7 KB, o que reduz em muito o tamanho da informação, sendo aceitável a transmissão da lista de atributos em cada acesso de usuário em computadores Clientes.

O treinamento é composto por um módulo que utiliza a seleção de atributos da Transformada Discreta do Cosseno e armazena os atributos selecionados, como mostra o Caso de Uso ilustrado na Figura 28.

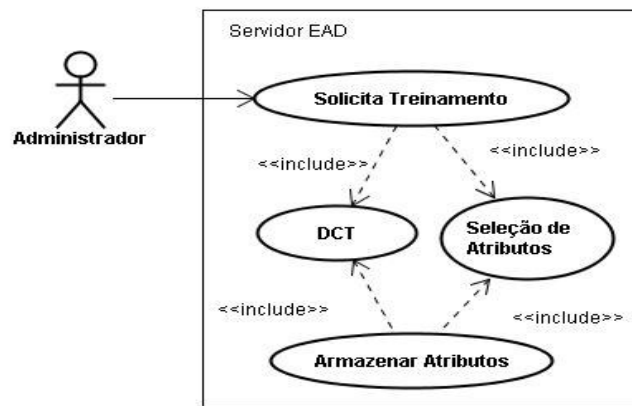


Figura 28 – Diagrama de Caso de Uso do módulo de treinamento do SIAF-EAD.

Fase 3: Entrada no Sistema

A entrada do usuário no sistema representa a fase em que ele efetua o login. O SIAF-EAD utiliza os recursos existentes no MIF para efetuar o login e direcionar o usuário para a página inicial do AVA, conforme descrito no Capítulo anterior, a diferença é que o SIAF-EAD adiciona a funcionalidade de solicitação dos atributos armazenados no Servidor EAD, carregando-os para posterior utilização. A Figura 29 ilustra o Caso de Uso desta fase.

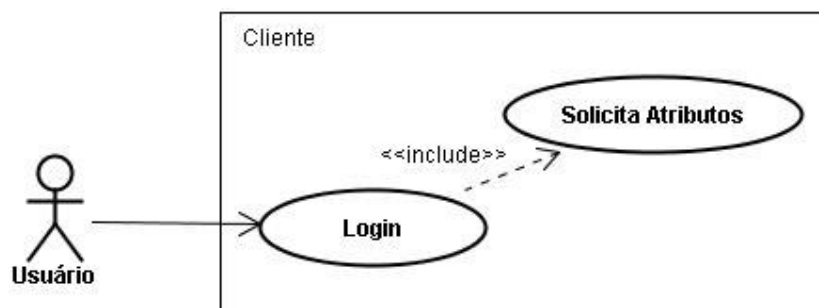


Figura 29 – Caso de Uso da Entrada do usuário do sistema.

Fase 4: Captura

A captura corresponde à tarefa realizada logo após a entrada do usuário no sistema, onde são capturadas as imagens do usuário por uma webcam em intervalos de tempo pré-definidos durante as sessões dos usuários, esta fase é executada no lado Cliente da arquitetura.

A webcam captura um vídeo (um certo número de imagens por segundo, chamadas de *frames*), sendo que o SIAF-EAD determina um intervalo de tempo durante o qual o usuário do AVA estará sendo monitorado. As imagens capturadas fora desse intervalo de tempo são descartadas, as que se enquadram são encaminhadas para a detecção. A Figura 30 mostra o Caso de Uso da captura de imagens.

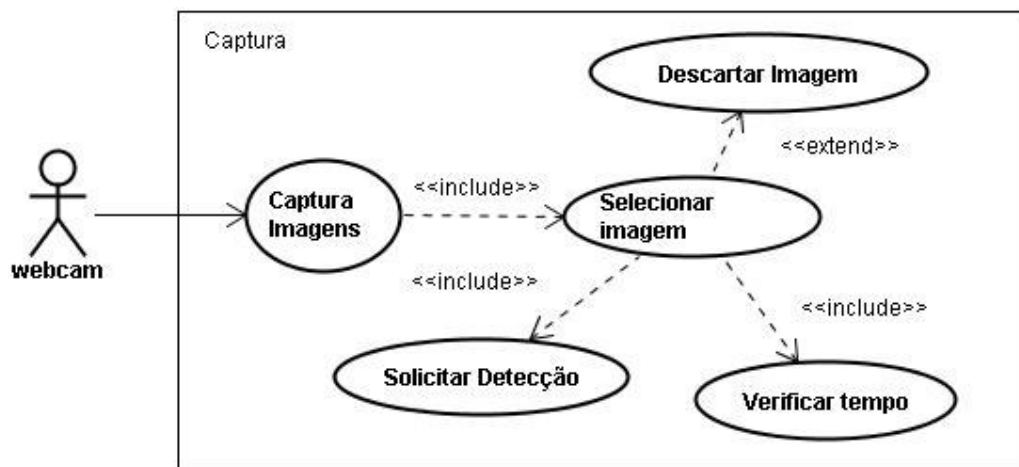


Figura 30 – Diagrama de Caso de Uso da Captura de imagens.

Fase 5: Detecção

Após a captura correta da imagem é efetuada a solicitação de detecção facial na imagem. Na fase de detecção é realizada uma verificação da face de interesse, pois na cena da imagem capturada pode existir mais de uma face humana, ou o algoritmo implementado pode cometer o erro de “falsos positivos” descrito na seção 2.5.

Se não houver faces, ou se o algoritmo de detecção cometer o erro de “falsos negativos” para todas as faces existentes na imagem, a imagem capturada será descartada. Se uma face de interesse for detectada a região que representa essa face é extraída da imagem original a qual será utilizada como a face de teste na fase de reconhecimento. A Figura 31 ilustra o Caso de Uso da fase de Detecção.

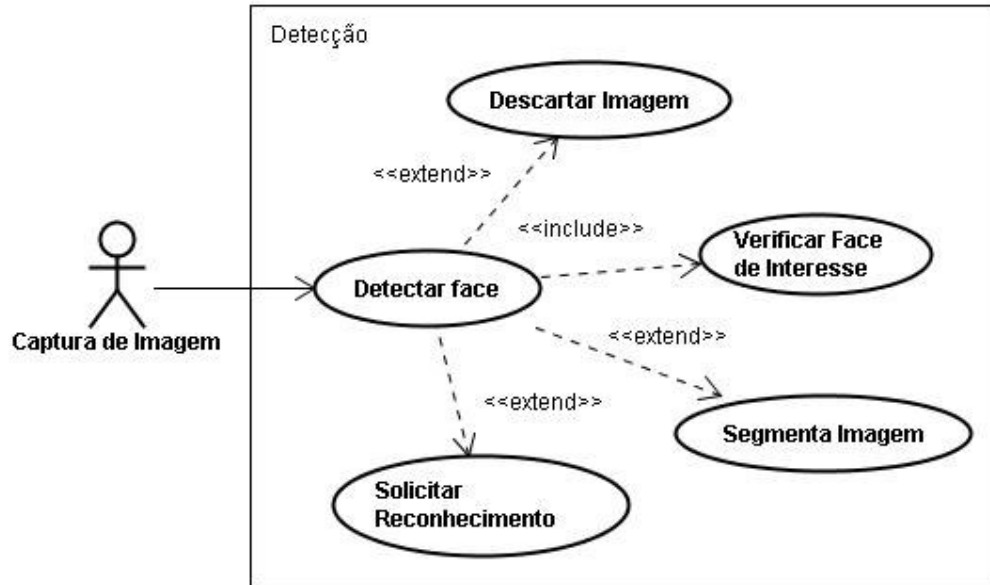


Figura 31 – Diagrama de Caso de Uso da Detecção.

Fase 6: Reconhecimento

A fase de reconhecimento é executada no lado Cliente do sistema. De posse da imagem facial detectada no momento, assim como da lista de atributos dos usuários do sistema o próximo passo é realizar a identificação do usuário.

Primeiramente é aplicada a DCT na face de teste, então um classificador identificará aquela face dentre as existentes no banco de treinamento. O classificador indicará a face no conjunto de treinamento que tem os seus atributos mais parecidos com a face de teste.

O SIAF-EAD desenha uma certificação na face de teste, indicando se ela foi ou não identificada pelo sistema, e por fim a envia ao Servidor EAD, onde será anexada aos

relatórios do AVA através do Subsistema 2 integrante do MIF. A Figura 32 ilustra o Caso do Uso da fase de reconhecimento.

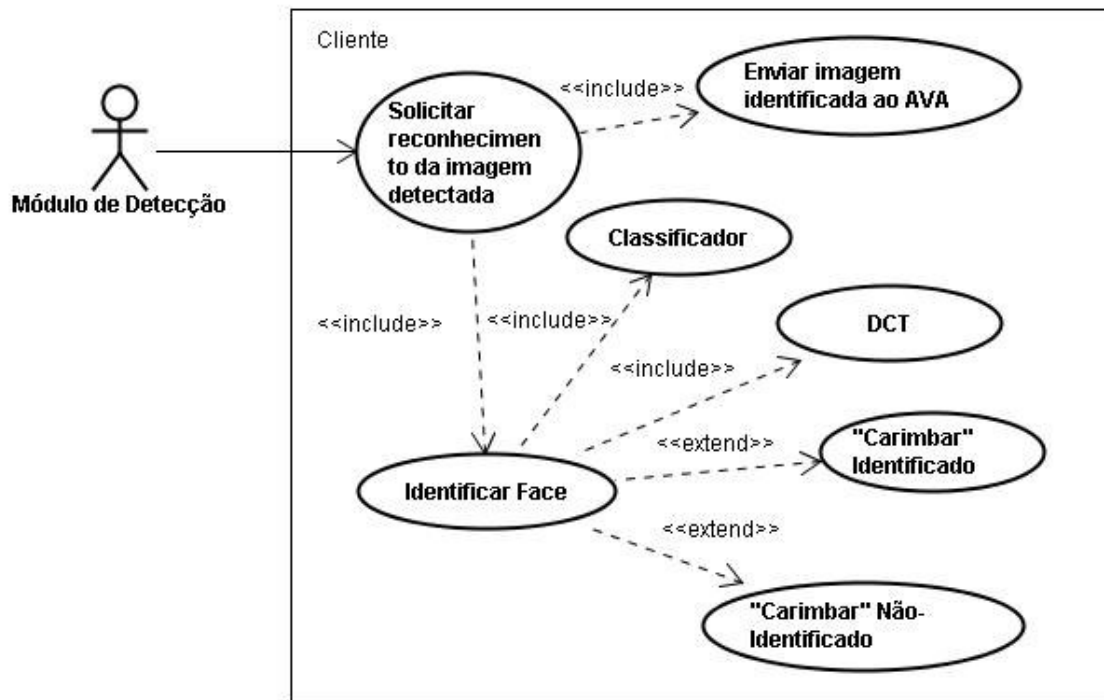


Figura 32 – Diagrama de Caso de Uso da fase de Reconhecimento.

Fase 7: Saída do Sistema

A última fase é quando o usuário efetua a saída do sistema. Nesta etapa é efetuada a saída do AVA e liberação de recursos. A Figura 34 ilustra o Caso de Uso da saída do sistema.

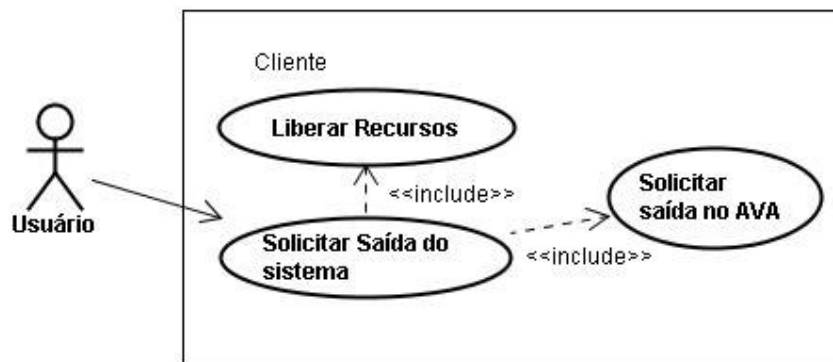


Figura 33 – Caso de Uso da saída do sistema.

3.1.4 Modelo de domínio do SIAF-EAD

Para Larman (2003) o modelo de domínio é o resultado da identificação de conceitos, atributos e associações que são de interesse para a descrição do domínio do problema a partir da perspectiva de uma classificação por objetos. Ou seja, o modelo de domínio é utilizado para denotar ou agrupar um conjunto de sistemas ou de áreas funcionais, dentro dos sistemas, que exibem funcionalidades similares, sendo que o artefato gerado nos mostra uma abstração de alto nível do problema.

Considerando que o SIAF-EAD reúne conceitos envolvidos na área de processamento facial, assim como utiliza Subsistemas já existentes do MIF com a organização oferecida pela arquitetura cliente-servidor, é possível agrupar as classes existentes no projeto em módulos conceituais, como descrito no modelo de domínio ilustrado na Figura 34.

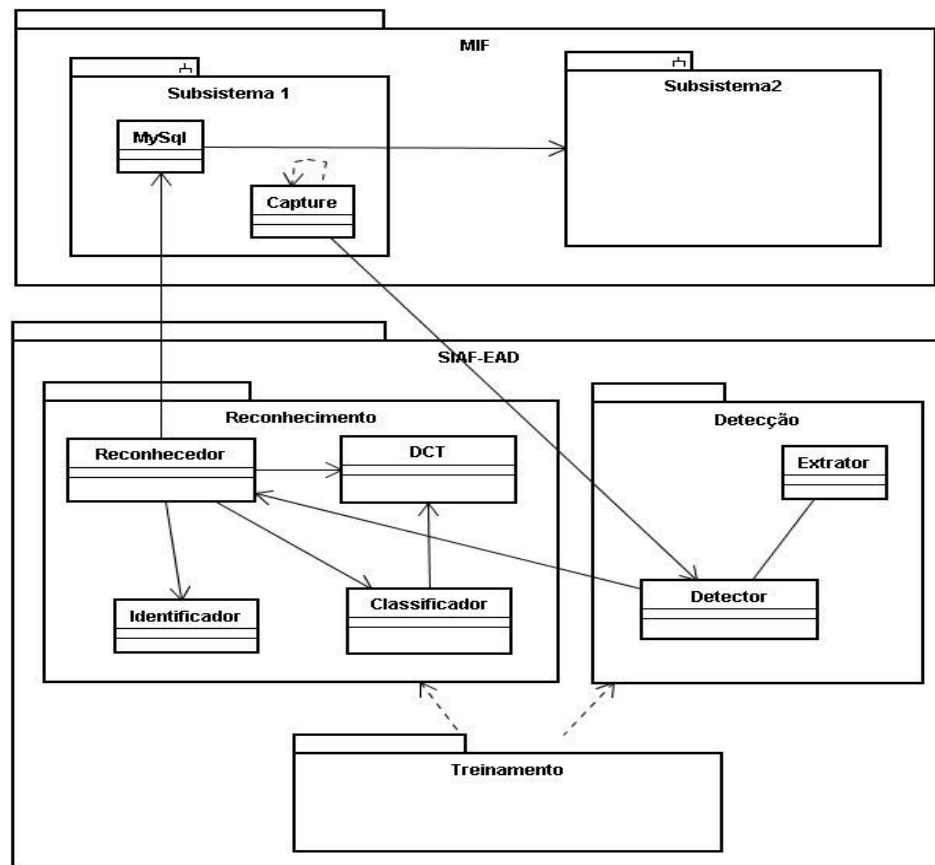


Figura 34 –Modelo de Domínio do SIAF-EAD.

3.2 Implementação do SIAF-EAD

Nesta seção é descrito um banco de imagens faciais coletado para testar as abordagens de detecção e reconhecimento facial utilizado pelo SIAF-EAD, também é descrito o protótipo desenvolvido que implementa a estrutura apresentada na seção 3.1.

Foram realizados testes com o propósito de se obter algoritmos satisfatórios de detecção e reconhecimento facial para a construção do SIAF-EAD.

Para que fossem realizados os testes descritos nesta seção foram desenvolvidos módulos ou, quando possível, modificados alguns já existentes em trabalhos correlacionados, que contêm a implementação dos algoritmos testados, sendo que as abordagens que obtiveram os resultados mais relevantes foram incorporadas ao protótipo final do SIAF-EAD. Nesta seção estão descritos como foram conduzidos esses testes e seus resultados.

3.2.1 Banco de Faces SIAF-EAD

As abordagens de detecção e reconhecimento facial utilizam bancos de faces normalizados para a realização de testes, todavia as situações reais exigem a aplicação de tais algoritmos em condições diversas, fator que diminui a eficiência das abordagens comprometendo a taxa de acerto, entretanto nestes casos a realização de uma etapa extra de normalização para a tentativa de melhorar os resultados pode ser efetuada.

Um sistema que realize reconhecimento facial integrado a um AVA tem que se adaptar a nenhum ou pouco controle das condições de captura das imagens faciais durante sua execução. Sendo assim é necessário testar as abordagens envolvidas em bancos de imagens que retratem as reais condições de aplicação do sistema, o que justifica a necessidade da obtenção de um banco de imagens que se ajuste a situação para que os testes realizados retratem melhor o cenário de aplicação do sistema.

Neste contexto, o Banco de Faces SIAF-EAD surgiu da necessidade de se obter um conjunto de imagens que melhor representasse o cenário proposto em (SILVA, 2008) para

serem testadas as abordagens que venham a compor os módulos do SIAF-EAD descrito nesta dissertação.

A construção do Banco de Faces SIAF-EAD foi realizada com a colaboração de 52 voluntários de diferentes idades, raça e sexo, que permitiram que fossem submetidos a uma sessão de captura de imagens por uma webcam, com cenários (fundos) diversos, como um quarto, escritório e laboratório de informática.



Figura 35 - -- Amostras do Banco de Faces SIAF-EAD, ilustrando diferentes cenários e diferenças de sexo e idade entre as pessoas fotografadas.

A Figura 35 ilustra algumas amostras de imagens do Banco de Faces SIAF-EAD, ilustrando uma pose das pessoas 3 (pose 2), 37 (pose 2) e 43 (pose 1) em diferentes cenários (escritório, quarto e laboratório de informática respectivamente) com diferenças de sexo e idade (adulto, jovem e adolescente respectivamente).

Foram capturadas 8 poses de cada pessoa em uma mesma sessão, isto é, todas as poses de cada pessoa foram fotografadas na mesma hora do dia. Também foi permitido o uso alternado de acessórios como boné Figura 36 (a), pessoa 45 (poses 2 e 5), e óculos Figura 36 (b), pessoa 12 (poses 6 e 5). Foram capturadas poses da mesma pessoa com e sem acessórios.

Uma das principais características do Banco de Faces SIAF-EAD é o pouco ou nenhum controle sobre as condições citadas nos trabalhos de reconhecimento facial, como iluminação, pose e expressão.



(a)



(b)

Figura 36 – Amostras do Banco de Faces SIAF-EAD, mostrando a presença e ausência de acessórios nas pessoas.

A Figura 37 ilustra algumas amostras que variam em expressão Figura 37 (a), na pessoa 34 (poses 1, 4 e 8), em pose Figura 37 (b), na pessoa 1 (poses 7, 6 e 2) e em iluminação, Figura 37 (c), na pessoa 39 (poses 3, 4 e 5).

Também foi permitido que os voluntários ficassem à vontade na frente da webcam, sendo assim, aconteceu que algumas amostras foram capturadas quando eles estavam com a mão no cabelo ou ajustando um boné na cabeça, também foi permitida a presença de outra pessoa ao fundo com sua face à amostra.



(a)



(b)



(c)

Figura 37 – Amostras do Banco de Faces SIAF-EAD, ilustrando a variedade de expressão, pose e iluminação.

A Figura 38 apresenta algumas amostras ilustrando algumas peculiaridades no banco: da esquerda à direita, pessoa 52 (pose 7) com as mãos nos cabelos, pessoa 35 (pose 8) com outra face de outra pessoa ao fundo e o pessoa 45 (pose 3) ajusta o boné durante a captura da imagem.



Figura 38 – Amostras do Banco de Faces SIAF-EAD, ilustrando algumas peculiaridades do banco.

A webcam foi posicionada frontalmente às pessoas fotografadas a uma distância entre 1 a 1,4 metros com a resolução de 640 X 480 das imagens capturadas em RGB. A Tabela 2 resume todos esses dados técnicos.

Tabela 2 - Tabela-resumo com os dados técnicos do Banco de Faces SIAF-EAD.

Numero de pessoas	Poses	Numero de imagens	Resolução das imagens	Distância da webcam para a face	Posição da webcam
52	8	416	640X480	De 1m a 1,4m	Frontal

3.2.2 Protótipo de Detecção Facial e Testes.

A implementação do protótipo na linguagem C++ utilizou a biblioteca OpenCV (*Open Source Computer Vision Library*) (OPEN CV, 2009) que implementa o algoritmo de detecção Viola e Jones.

OpenCV é uma biblioteca multi-plataforma criada pelo grupo de desenvolvimento da Intel em 2000. Possui uma coleção de ferramentas poderosas para o desenvolvimento de aplicativos na área de Visão Computacional, possuindo módulos de processamento de imagens e vídeo, estrutura de dados, álgebra linear, interface básica com sistema de janelas independentes, controle de mouse e teclado, além de mais de 350 algoritmos utilizados em visão computacional como: filtros de imagem, calibração de câmera, reconhecimento de objetos, análise estrutural, entre outros.

Ao realizar a detecção, o protótipo extrai a região da face da imagem, gerando uma outra imagem que representa apenas a região da face detectada, e para contornar os problemas com os erros de detecção discutidos na seção 2.5, e implementar a funcionalidade “verifica face de interesse” ilustrado no Diagrama de Caso de Uso da Figura 31, foi adotado um procedimento simples: o protótipo considera a primeira face encontrada com a largura maior que 100 pixels. Este procedimento leva em consideração que a face de interesse deve estar a

uma distancia mais próxima da webcam, desprezando as faces ao fundo que obviamente tem um tamanho menor.

O protótipo extrai a face considerando um tamanho padrão de 220 pixels de largura e 240 pixels de altura e converte-a para monocromática.

Considerando que o projeto do sistema prevê que a detecção é utilizada em duas fases do ciclo de vida (Treinamento e Detecção), após a extração da face o protótipo pode armazená-la, com o intuito de formar o banco de treinamento, ou a mesma será utilizada como face de teste a ser identificada.

Testes do Protótipo de Detecção

Os testes de detecção facial foram realizados através da implementação de um protótipo que utiliza o método Viola e Jones. O protótipo usa como entrada todas as imagens do Banco de Faces SIAF-EAD e indica as faces detectadas desenhando um retângulo na região das mesmas, gerando um conjunto de imagens do banco original com as faces detectadas para posterior análise dos resultados obtidos, conforme mostra a Figura 39.

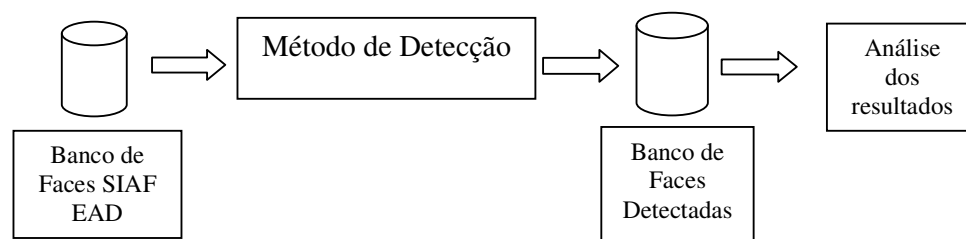


Figura 39 – Representação gráfica da estratégia de testes para a definição da abordagem de detecção.

Durante os testes, foram observados alguns erros de detecção, “falsos negativos” (Figura 40) e “falsos positivos” (Figura 42).

A abordagem utilizada apresentou o erro de falso negativo apenas em cinco imagens do Banco de Faces SIAF-EAD, reproduzidas na Figura 40, onde três poses (1,7,8,

respectivamente) da pessoa 1 (a) e duas poses (6 e 7 respectivamente) da pessoa 9 (b) não tiveram suas faces detectadas.



(a)



(b)

Figura 40 – “Falsos negativos” dos testes de detecção.

Algumas imagens do banco apresentam duas faces humanas, sendo uma mais próxima da webcam, Figura 41 (a) , que é a face de interesse para o sistema e outra ao fundo. Estes casos são considerados de sucesso, já que a detecção é feita corretamente, porém, o sistema apresenta uma solução para indicar a face de interesse.

Outro caso é quando detecta regiões aninhadas de tamanhos distintos na mesma face ou não-face, Figura 41 (b) (c).

Na análise do resultado, para o caso mostrado na Figura 41 (b) foi considerado que o algoritmo funcionou com sucesso, ou seja, a face foi encontrada, apesar deste caso representar um erro, pois a região maior não delimita uma face apenas. O fato de considerar este caso de sucesso é por causa que o algoritmo do sistema considerou a face menor que representa corretamente a face da pessoa na imagem.

Já para o caso mostrado na Figura 41 (c) foram considerado dois erros de “falsos positivos”, além da detecção correta da face real.

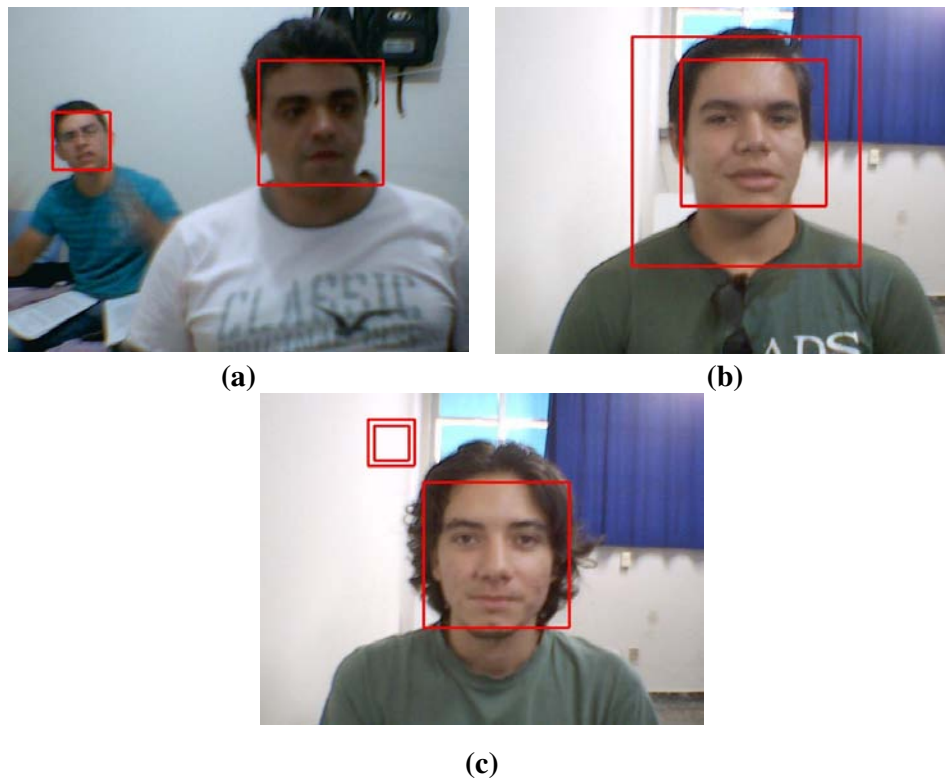


Figura 41 – Amostras com a detecção de duas faces na mesma imagem do Banco de Faces SIAF-EAD.

Os “falsos positivos” ocorrem de várias formas, sendo três as mais comuns, como mostra a Figura 42. A primeira ocorrência é quando o algoritmo indica uma face em uma região dentro da face real, Figura 42 (a), ou seja, a região indicada é menor que a face real, a segunda e mais freqüente é quando ele indica a existência de uma face na região ao fundo da imagem, Figura 42 (b) e a última é quando há intersecção de regiões indicada como face, Figura 42 (c).

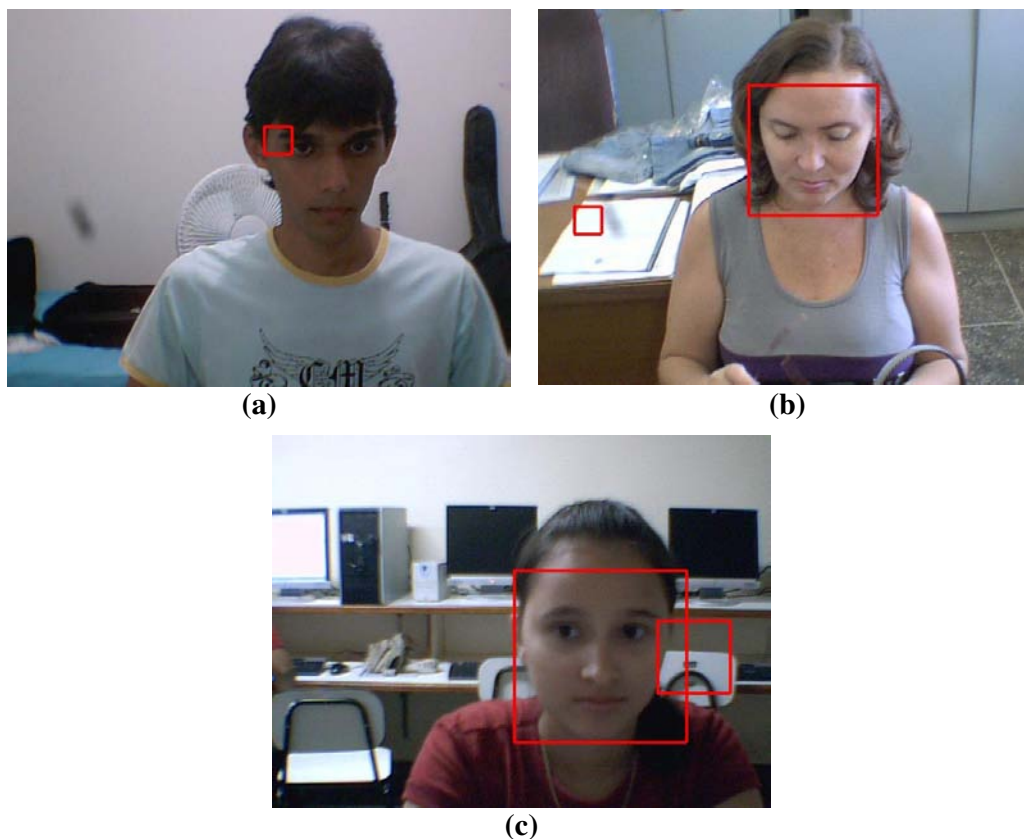


Figura 42 - Amostras de Falsos positivos dos testes de detecção.

O protótipo implementado obteve 63 “falsos positivos” e 5 “falsos negativos” em 62 imagens das 416 existentes no banco de faces. O Apêndice A mostra uma tabela detalhada dos testes realizados, mostrando por pessoa a quantidade e tipo de erro ocorrido.

É importante notar que, para serem realizados os testes de reconhecimento descritos na próxima seção, as cinco imagens em que não foi possível a detecção automática tiveram suas faces reais extraídas manualmente usando um editor de imagens, com o intuito de deixar constante o número de poses de cada pessoa.

A Figura 43 ilustra algumas amostras das faces detectadas nas imagens mostradas nesta seção. Foi gerado um banco com as 411 faces detectadas no Banco de Faces SIAF-EAD, sendo que as cinco faces restantes correspondem às imagens em que foram obtidos erros de “falso negativo”. O Apêndice C mostra o banco completo das faces detectadas neste protótipo.

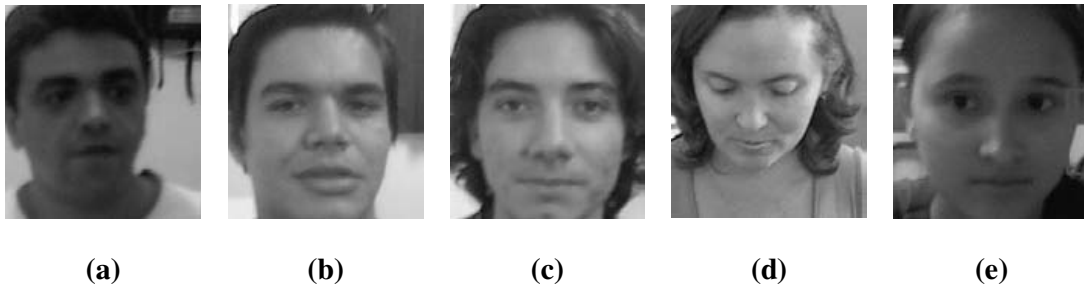


Figura 43 – Amostras de faces detectadas e extraídas corretamente pelo protótipo.

3.2.3 Protótipo e Testes de Reconhecimento Facial do SIAF-EAD

O protótipo de reconhecimento consiste em um módulo que realiza o reconhecimento facial, implementando o classificador de vizinho mais próximo e o seletor de baixas frequências. A escolha desse classificador e seletor foi devido as maiores taxa de acertos apresentadas nos testes, e a facilidade de implementação do seletor de baixas frequências.

A identificação da face de teste é realizada e é impressa uma indicação na imagem como mostra a Figura 44. Quando a imagem é reconhecida pelo protótipo, ele imprime na face a palavra “IDENTIFICADO” da cor verde, quando não-reconhecida, imprime um símbolo em vermelho.



Figura 44 – Exemplo de certificação da imagem da face de teste do usuário.