



Ministério da Justiça



UnB



Centro de Apoio ao
Desenvolvimento
Tecnológico



latitude

Laboratório de tecnologias da tomada de decisão

Termo de Cooperação/Projeto:

**Acordo de Cooperação Técnica
FUB/CDT e MJ/SE
Registro de Identidade Civil –
Replanejamento e Novo Projeto Piloto**

Documento:

**RT Estudos de Tecnologias para
modelos de artefatos em Papel-
Pesquisa para métodos de detecção de objetos
bidimensionais, detecção facial e reconhecimento
automático de caracteres**

Data de Emissão:

03/06/2015

Elaborado por:

**Universidade de Brasília – UnB
Centro de Apoio ao Desenvolvimento
Tecnológico – CDT
Laboratório de Tecnologias da Tomada
de Decisão – LATITUDE.UnB**

MINISTÉRIO DA JUSTIÇA

José Eduardo Cardozo
Ministro

Marivaldo de Castro Pereira
Secretário Executivo

Helvio Pereira Peixoto
Coordenador Suplente do Comitê Gestor do SINRIC

EQUIPE TÉCNICA

Ana Maria da Consolação Gomes Lindgren
Andréa Benoliel de Lima
Celso Pereira Salgado
Delluiz Simões de Brito
Elaine Fabiano Tocantins
Fernando Saliba Oliveira
Fernando Teodoro Filho
Guilherme Braz Carneiro
Joaquim de Oliveira Machado
José Alberto Sousa Torres
Marcelo Martins Villar
Raphael Fernandes de Magalhães Pimenta
Rodrigo Borges Nogueira
Rodrigo Gurgel Fernandes Távora
Sara Lais Rahal Lenharo

UNIVERSIDADE DE BRASÍLIA

Ivan Marques Toledo Camargo
Reitor

Paulo Anselmo Ziani Suarez
Diretor do Centro de Apoio ao
Desenvolvimento Tecnológico – CDT

Rafael Timóteo de Sousa Júnior
Coordenador do Laboratório de Tecnologias da
Tomada de Decisão – LATITUDE

EQUIPE TÉCNICA

Flávio Elias Gomes de Deus
(Pesquisador Sênior)
William Ferreira Giozza
(Pesquisador Sênior)
Ademir Agostinho de Rezende Lourenço
Adriana Nunes Pinheiro
Alysson Fernandes de Chantal
Amanda Almeida Paiva
Andréia Campos Santana
Antônio Claudio Pimenta Ribeiro
Carolinne Januária de Souza Martins
Daniela Carina Pena Pascual
Danielle Ramos da Silva
Diogenes Ferreira Reis Fustinoni
Fábio Lúcio Lopes Mendonça
Fábio Mesquita Buiati
Flávio de Barros Vidal
Glaudson Menegazzo Verzeletti
Heverson Soares de Brito
Johnatan Santos de Oliveira
José Carneiro da Cunha Oliveira Neto
Kelly Santos de Oliveira Bezerra
Luciano Pereira dos Anjos
Luciene Pereira de Cerqueira Kaipper
Luiz Antônio de Souto Evaristo
Luiz Claudio Ferreira
Marcos Vinicius Vieira da Silva
Marco Schaffer
Pedro Augusto Oliveira de Paula
Roberto Mariano de Oliveira Soares
Sergio Luiz Teixeira Camargo
Soleni Guimarães Alves
Suzane Lais De Freitas
Valério Aymoré Martins
Vera Lopes de Assis
Wladimir Rodrigues da Fonseca

HISTÓRICO DE REVISÕES

Data	Versão	Descrição
05/01/2015	0.1	Versão inicial.
10/01/2015	0.2	Revisão de conteúdo: Rodrigo Távora
05/02/2015	0.3	Revisão de texto e forma
15/03/2015	0.4	Definição de escopo
15/05/2015	0.5	Elaboração de primeira versão
03/06/2015	0.6	Revisão de texto e forma.



Universidade de Brasília – UnB
Campus Universitário Darcy Ribeiro - FT – ENE – Latitude
CEP 70.910-900 – Brasília-DF
Tel.: +55 61 3107-5597 – Fax: +55 61 3107-5590

SUMÁRIO

1	INTRODUÇÃO	5
2	DETECÇÃO DE PADRÕES BIDIMENSIONAIS (2D) CAPTURADOS POR CÂMERAS DIGITAIS PARA DETECÇÃO DE REGIÕES ESPECÍFICAS ESTRUTURADAS	7
2.1	Detecção de Objetos em Imagens bidimensionais (2D).....	7
2.2	A Transformada SIFT	9
2.3	O Algoritmo de Viola e Jones.....	11
2.4	Descritores SURF.....	13
3	DETECÇÃO FACIAL EM IMAGENS	17
3.1	Contextualização	17
3.2	Principais Trabalhos Científicos.....	17
3.3	Método de detecção de faces desenvolvido por Viola e Jones.....	18
4	ESTUDO DE ESPECIFICAÇÕES DE CONJUNTO DE FONTES, ALGORITMOS, ANÁLISE DE PERFORMANCE PARA RECONHECIMENTO AUTOMÁTICO DE TEXTO.....	21
4.1	Estudo das especificações de conjunto de fontes para identidades	21
4.2	Análise de Performance do Reconhecimento Automático de Textos.....	24
4.2.1	Tesseract.....	24
5	IMPLEMENTAÇÕES PARA VALIDAÇÃO DAS TÉCNICAS PROPOSTAS	25
5.1	Detecção da MRZ e extração da Pose a partir de câmeras digitais	26
5.2	Detecção da Região Facial	29
5.3	Reconhecimento dos caracteres contidos na MRZ.....	30
6	CONCLUSÃO	32
7	REFERÊNCIAS BIBLIOGRÁFICAS UTILIZADAS	34

1 INTRODUÇÃO

A Secretaria Executiva (SE/MJ), vinculada ao Ministério da Justiça (MJ), é responsável por viabilizar o desenvolvimento e a implantação do Registro de Identidade Civil, instituído pela Lei nº 9.454, de 7 de abril de 1997, regulamentado pelo Decreto nº 7.166, de 5 de maio de 2010.

Atualmente, a República Federativa do Brasil conta com sistema de identificação de seus cidadãos amparado pela Lei nº 7.116, de 29 de agosto de 1983. Essa lei assegura validade nacional às Carteiras de Identidade, ou Cédulas de Identidade; confere também autonomia gerencial às Unidades Federativas no que concerne à expedição e controle dos números de registros gerais emitidos para cada documento. Essa condição de autonomia, ao contrário do que pode parecer, fragiliza o sistema de identificação, já que dá condições ao cidadão de requerer legalmente até 27 (vinte e sete) cédulas de identidades diferentes. Com essa facilidade legal, inúmeras possibilidades fraudulentas se apresentam de maneira silenciosa, pois, na grande maioria dos casos, os Institutos de Identificação das Unidades Federativas não dispõem de protocolos e aparato tecnológico para identificar as duplicações de registro vindas de outros estados, ou até mesmo do seu próprio arquivo datiloscópico. Consoante aos fatos, os Institutos de Identificação não trabalham interativamente para que haja trocas de informações de dados e geração de conhecimento para manuseio inteligente e seguro para individualização do cidadão em prol da sociedade.

Com foco na busca de soluções para tais problemas, o Projeto RIC prevê a administração central dos dados biográficos e biométricos dos cidadãos no Cadastro Nacional de Registro de Identificação Civil (CANRIC) e ABIS (do inglês *Automated Biometric Identification System*), respectivamente. A previsão desse novo modelo sustenta a não duplicação de registros e a consequente identificação unívoca dos cidadãos brasileiros natos e naturalizados. O Projeto RIC, portanto, visa otimizar o sistema de identificação e individualização do cidadão brasileiro nato e naturalizado com vistas a um perfeito funcionamento da gestão de dados da sociedade, agregando valor à cidadania, à gestão administrativa, à simplificação do acesso aos serviços disponíveis ao cidadão e à segurança pública do país.

Nesse contexto, o termo de cooperação entre MJ/SE e FUB/CDT define um projeto que objetiva identificar, mapear e desenvolver parte dos processos e da infraestrutura

Projeto: MJ/SE-RIC	Emissão: 03/06/2015	Arquivo: 20150603 MJ RIC - RT Estudos de Tecnologias para modelos de artefatos em Papel	Pág.5/36
--------------------	---------------------	---	----------

Confidencial.

tecnológica necessária para viabilizar a implantação do número único de Registro de Identidade Civil – RIC no Brasil.

Neste relatório técnico serão apresentadas as principais metodologias do “estado-da-arte”, contemplando as principais abordagens utilizadas em consenso pela comunidade acadêmica e científica mundial para auxiliarem no desenvolvimento e elaboração de técnicas para identificação de campos e reconhecimento automático de texto a serem incorporados na proposta do novo RIC.

Mesmo utilizando tais metodologias, cabe aqui neste ponto uma ressalva muito importante, pois em se tratando de técnicas envolvendo processamento de imagens e visão computacional, independente de qual referência ou metodologia utilizada, não existe solução completa e fechada. Todas as técnicas são passíveis de falhas, principalmente observando que todos os exaustivos testes realizados por seus autores/proponentes empregaram bases de dados particularizadas e sendo estas controladas para aplicações específicas. Assim, reforça-se a máxima nesta área: em que nenhuma solução pode atender a todos os casos, entretanto mantendo um nível de resultados dentro de valores aceitáveis pela academia e indústria.

Nas seções a seguir serão abordados os seguintes tópicos: a) Detecção de padrões bidimensionais (2D) por câmeras digitais para detecção de regiões específicas em documentos estruturados; b) Pesquisa de trabalhos científicos relevantes sobre detecção facial em imagens de vídeo; c) Estudo de especificações de conjunto de fontes, algoritmos, análise de performance e requisitos básicos de impressão para reconhecimento automático de texto. Estes itens serão utilizados como subsídios para se avaliar a utilização de câmeras digitais de baixa qualidade (por ex.: a dos dispositivos móveis de gerações anteriores), bem como verificar se elas são capazes de, a partir de uma configuração mínima de condições e informações, recuperar as áreas de interesse em um documento de identificação, incluindo as regiões a serem lidas por máquinas (*Machine Readable Zone* - MRZ), como também, neste mesmo padrão de documento, localizar faces humanas e identificar conjuntos com formatação pré-determinadas de caracteres, para processamento futuro destas informações.

Projeto: MJ/SE-RIC	Emissão: 03/06/2015	Arquivo: 20150603 MJ RIC - RT Estudos de Tecnologias para modelos de artefatos em Papel	Pág.6/36
--------------------	---------------------	---	----------

Confidencial.

2 DETECÇÃO DE PADRÕES BIDIMENSIONAIS (2D) CAPTURADOS POR CÂMERAS DIGITAIS PARA DETECÇÃO DE REGIÕES ESPECÍFICAS ESTRUTURADAS

Para se contextualizar este relatório técnico, se faz necessária para a maioria dos casos a definição entre duas situações que são tratadas com distinção para a identificação de códigos/objetos em imagens digitais: o processo de detecção e de reconhecimento. Os dois em conjuntos permitem sua identificação. A detecção (que será abordada neste relatório) consiste em elaborar uma metodologia capaz de localizar na imagem a região espacial (ou conjunto de *pixels*) que compõem o código a ser identificado. Após detectada esta região, realiza-se a aplicação de operações gerando descritores para se reconhecer o código/objeto contido nesta região. Um exemplo desta abordagem pode ser visto em processos de identificação facial automatizada: em um primeiro momento localiza-se a região da face na imagem, em seguida é realizada a extração das minúcias que permitem a identificação do indivíduo.

2.1 Detecção de Objetos em Imagens bidimensionais (2D)

O processo de detecção de objetos em imagens bidimensionais consiste em um conjunto de operações sequenciais, conforme descrito pela [Figura 1](#)~~Figura 1~~~~Figura 1~~, no qual se tem como entrada a imagem digital em que se pretende recuperar a pose¹ do objeto. A partir desta imagem, é extraído/escolhido um conjunto de parâmetros que compõe o conjunto de descritores. Este conjunto de descritores não é obrigatório ser um vetor de características, podendo conter uma imagem do objeto a ser detectado, também conhecido como padrão (ou *template*). Em seguida, um classificador é treinado, seja supervisionado ou não, para responder a este conjunto de descritores, para posterior avaliação da similaridade com o padrão procurado.

¹ Entende-se como pose de um objeto a sua posição no domínio da imagem em coordenadas espaciais horizontais e verticais, e sua orientação em relação ao plano da imagem (rotação e translação).

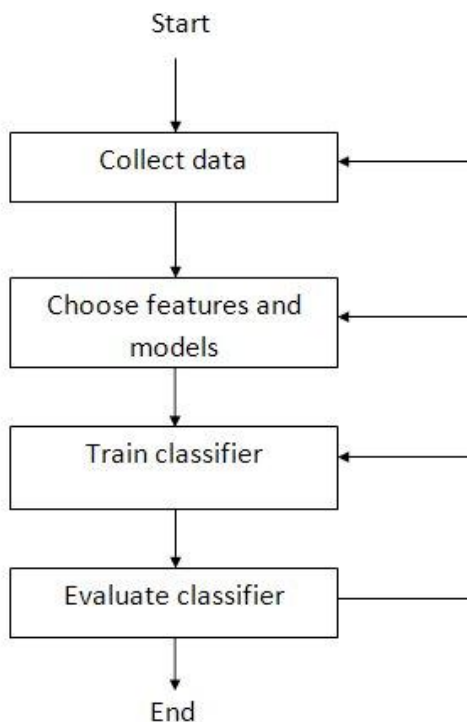


Figura 1 - Fluxograma genérico de um processo de detecção de objetos em imagens. Retirado de [1].

Inicialmente, para o entendimento do processo de detecção de objetos em imagens bidimensionais, deve ser apresentado o conceito de descritores. Por definição, descritores são conjuntos de características discriminantes, extraídas do objeto que se está interessado em detectar, que sejam suficientes para permitir sua aplicabilidade em classificadores automatizados (ex.: reconhecimento de faces) [1]. Alguns descritores exigem a definição de uma métrica própria para calcular a similaridade entre dois vetores de características, como por exemplo a Equação (1), a seguir:

$$d(p, q) = \sum_{i=1}^n |p_i - q_i| \quad (1)$$

Na Equação (1), a função d mede a similaridade entre dois vetores p e q com n elementos. Em termos práticos, o escore definido por esta função será utilizado para a realização da comparação entre os elementos, de forma que serão mais similares os vetores de descritores que produzirem o menor valor da função d .

Dentre os diversos descritores que podem ser encontrados na literatura atual disponível, cada um possui suas peculiaridades e particularidades de aplicação. As principais categorias de descritores [2] são divididas como: Baseados em cor; Baseado em

Projeto: MJ/SE-RIC	Emissão: 03/06/2015	Arquivo: 20150603 MJ RIC - RT Estudos de Tecnologias para modelos de artefatos em Papel	Pág.8/36
--------------------	---------------------	---	----------

Confidencial.

textura; e Baseado em forma.

Os descritores baseados em cores, em sua maioria, utilizam informação provenientes de *gamut* do espectro formado das cores, utilizando informações de acordo com o espaço de cores escolhido [3]. Para os descritores baseados em texturas, são utilizadas informações provenientes dos elementos formadores de texturas, definidos como *Textons*[4]. Utilizando a informação dos *Textons*, descritores de texturas podem ser concebidos a partir da obtenção da Matrizes de Co-ocorrência para Extração de Atributos de Textura (do inglês, GLCM) [5] e explorando suas propriedades de distribuição e orientação destes elementos. Os descritores baseados em forma são formados a partir das informações provenientes das características espaciais, como, por exemplo, os segmentos de retas, círculos e formas geométricas, presentes no domínio da imagem.

Entretanto, mesmo com a gama de tipos e aplicabilidades de descritores, nenhum deles é capaz de realizar sozinho o processo de detecção de objetos. Exemplificando, um descritor baseado em cor não fornece a capacidade de, ao se localizar o objeto desejado, determinar sua pose. Para tal, faz-se necessária a combinação de recursos como determinação de projetivas (transformações espaciais). Neste caso, as informações de localização do objeto no domínio da imagem, juntamente com a utilização de transformações espaciais tornam possível obter uma matriz de transformação estável capaz de mapear no domínio bidimensional da imagem e suas projeções tridimensionais.

Para tal tarefa, faz-se o uso de *frameworks* que utilizam combinações de recursos e características que possam simultaneamente fornecer atributos robustos para os descritores, permitindo, a partir da detecção destes descritores, a determinação de sua pose no domínio da imagem. Dentre esta gama de *frameworks* encontrados na literatura, dois em específicos forneceram grandes avanços nas etapas de detecção de objetos no domínio da imagem bidimensional: o primeiro é ***Transformada de Características Invariantes à Escala*** [6] (do inglês ***Scale-Invariant Features Transform - SIFT***) e o ***algoritmo de detecção de objetos de Viola e Jones*** [7].

2.2 A Transformada SIFT

A Transformada SIFT (como é popularmente conhecida) é uma técnica de processamento de imagens que permite a detecção e extração de descritores locais, razoavelmente invariáveis às mudanças de iluminação, ruído de imagem, rotação, escala

Projeto: MJ/SE-RIC	Emissão: 03/06/2015	Arquivo: 20150603 MJ RIC - RT Estudos de Tecnologias para modelos de artefatos em Papel	Pág.9/36
--------------------	---------------------	---	----------

Confidencial.

e pequenas mudanças de perspectiva dos objetos no domínio da imagem. Esta Transformada proporcionou um passo importante na representação de objetos por meio de suas partes, gerando um grande avanço em diversas áreas como: reconhecimento de objetos, rastreamento e criação de imagens panorâmicas.

Para a construção dos descritores, utiliza-se na Transformada SIFT uma representação piramidal gaussiana de imagens (entenda-se piramidal como uma pilha de imagens) em vários níveis de escala linear, como apresentado na [Figura 2](#) [Figura 2](#) [Figura 2](#).

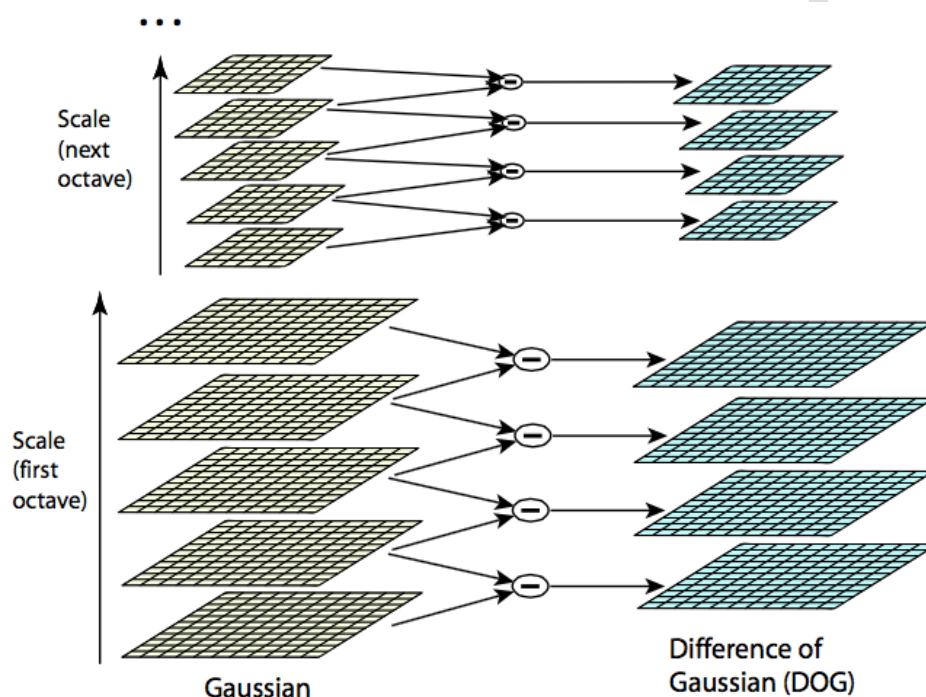


Figura 2 - Representação Piramidal de oitavas gaussianas apresentada por [8]

Esta pirâmide foi definida [8] como *Oitavas de Gaussianas*, em que cada nível f_i da pirâmide contém uma oitava da imagem original obtida a partir de subamostragem sobre a oitava f_{i-1} , localizada no nível imediatamente inferior. Este processo é repetido recursivamente até que se obtenha o nível de representação desejado, contendo blocos de 8×8 *pixels*, associado às maiores escalas observáveis.

A obtenção dos descritores SIFT é realizado a partir dos seguintes passos.

- 1) Detecção dos Máximos e Mínimos: nesta primeira etapa é feita a procura para todas as escalas e localizações da imagem. Neste processo, faz-se o uso da função de *Diferenças de Gaussianas* (DoG) de modo a identificar pontos de interesse invariantes à escala e rotação.

- 2) Localização dos pontos-chaves (*Key Points*): para cada localização em que foi detectado um extremo, um modelo detalhado é ajustado de forma a se determinar a localização da escala. A partir deste modelo, os pontos-chaves são selecionados a partir de medidas obtidas de determinação de estabilidades [9].
- 3) Definição da Orientação: a orientação de cada Ponto-chave é definida a partir dos gradientes locais da imagem. Toda a operação de detecção será realizada a partir dos dados provenientes da imagem transformados em relação à orientação, escala e localização de cada Ponto-chave.
- 4) Descritor dos Pontos-chaves: nesta etapa é realizada a construção dos descritores ao se medir os gradientes locais em uma região vizinha em cada ponto de interesse. Estas medidas são então transformadas para uma representação que permite a interferência de distorções e variações de iluminação. Para cada Ponto-chave, são definidas $n \times n$ regiões, com $k \times k$ pixels cada ao redor da localização x, y do Ponto-chave. Na maioria dos casos $n = k = 4$. Para cada região, extrai-se o histograma em 8 (oito) direções utilizando a magnitude dos pixels pertencentes a cada região, sendo os valores deste histograma os componentes do descritor.

2.3 O Algoritmo de Viola e Jones

O algoritmo de Viola e Jones[7] é o primeiro *framework* de detecção de objetos em imagens que ofereceu a possibilidade da realização desta atividade em intervalo de tempo competitivos, próximos ao tempo real. Embora este algoritmo possa ser treinado para detectar uma variedade de classes de objeto, sua motivação principal foi solucionar o problema de detecção de faces, a ser discutido em Seções futuras deste relatório. O principal recurso utilizado por este *framework* de detecção é o método de otimização da *Imagem Integral* [7].

A *Imagem Integral*, descrita por Viola e Jones [7], permite que a soma em sub-regiões da imagem seja feita rapidamente, permitindo que este cálculo tenha sempre a mesma duração de tempo, independente da dimensão espacial da imagem de entrada. Esta abordagem foi crucial para permitir que o tempo de processamento para as imagens a serem processadas atingisse a marca temporal de execução próxima ao tempo real, o que permitiu sua aplicação a praticamente todos os tipos de sistemas, incluindo dispositivos

Projeto: MJ/SE-RIC	Emissão: 03/06/2015	Arquivo: 20150603 MJ RIC - RT Estudos de Tecnologias para modelos de artefatos em Papel	Pág.11/36
--------------------	---------------------	---	-----------

Confidencial.

móveis.

Estas sub-regiões geralmente possuem a forma retangular, sendo semelhante à Transformada *Haar* [10] só que utilizando várias regiões retangulares combinadas. Estas características das sub-regiões retangulares denotam uma simplicidade quase primitiva, ao se comparar com as técnicas mais modernas de detecção de objetos, no entanto servem como uma representação global da imagem, permitindo sua utilização em tempo real. A [Figura 3](#) apresenta um exemplo das possíveis subimagens retangulares definidas por Viola e Jones.

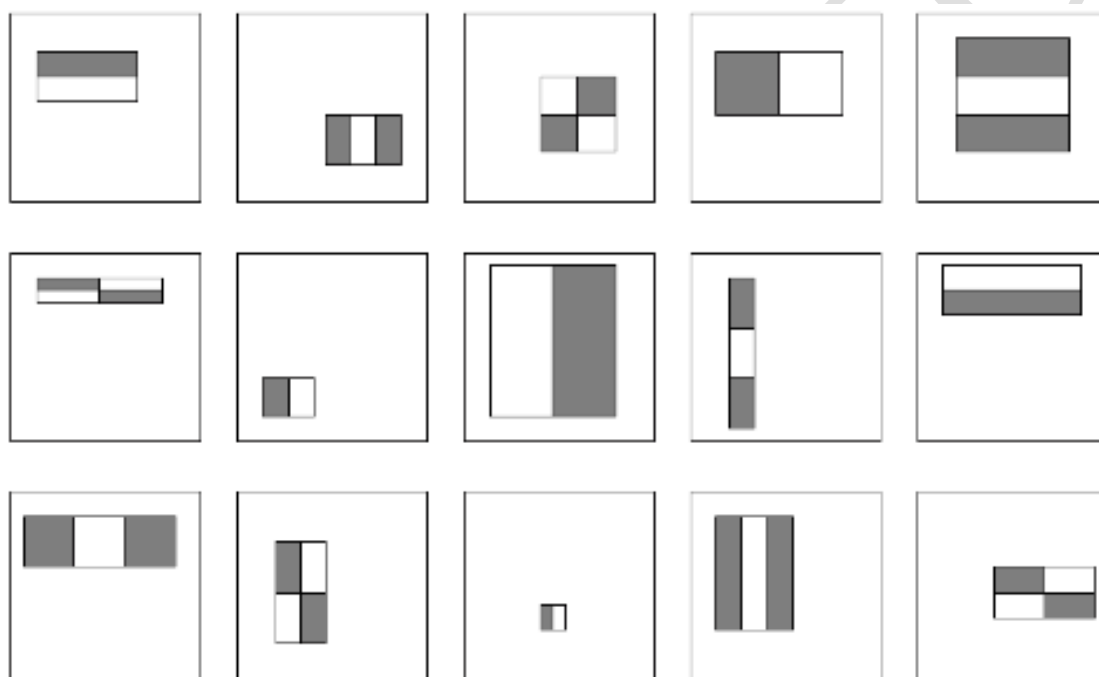


Figura 3 - Subimagens retangulares propostas por Viola e Jones[7].

Neste caso, com a utilização de subimagens para 24×24 pixels é possível extrair até 162.336 características, sendo que todas estas tornam a técnica inviável para a sua realização em tempo real. Para corrigir este problema, Viola e Jones propuseram a utilização de uma variante do algoritmo de aprendizagem *AdaBoost* [11] tanto para selecionar as melhores características, quanto para treinar os classificadores que utilizem este método. Para melhor treinamento, os autores optaram por fazer uso destes classificadores em cascata, conforme [Figura 4](#), em que somente os resultados das características candidatas assimilados como *Verdadeiros Positivos* (do inglês *True Positives*) são propagadas para o classificador seguinte da cascata.

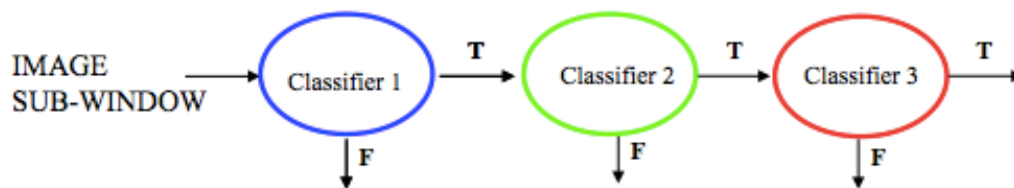


Figura 4 - Proposta de classificação em cascata proposta por Viola e Jones[7].

2.4 Descritores SURF

Entretanto, mesmo sendo estes descritores primitivos eficientes, uma evolução com capacidades melhoradas, estas inspiradas na estrutura da Transformada *SIFT* e com o uso da *Imagem Integral* utilizada por Viola e Jones, permitiu o desenvolvimento do descritor SURF (do acrônimo *Speeded-Up Robust Features*, do inglês).

O descritor SURF [12] é uma versão relaxada do operador de *DoG* da Transformada *SIFT* em que *Wavelets Haar* [10] são utilizadas para calcular a aproximação das derivadas de segunda ordem do núcleo Gaussiano. Esta aproximação é muito próxima à proposta apresentada por Viola e Jones. A detecção dos pontos-chaves do método SURF explora o uso das Imagens Integrais para calcular o operador *DoG* em diferentes escalas, estimando um desempenho 5 (cinco) vezes superior se comparado à Transformada *SIFT*. As posições detectadas na imagem são também refinadas usando interpolação realizada a partir do determinante da matriz Hessiana [13]. Observando que o operador *DoG* apresentava fortes respostas em regiões de borda da imagem, o número de pontos-chaves detectados pelo método SURF é menor, sendo segundo seus idealizadores mais eficientes e robustos que sua versão utilizada na Transformada *SIFT*.

Com intuito de esclarecimentos, as principais diferenças entre o método SURF e a Transformada *SIFT* são: o SURF utiliza um modelo aproximativo do espaço de escala; ambos são baseados em Imagens Integrais; o detector de pontos-chaves no SURF não necessita que o tamanho original da imagem seja alterado; a detecção dos pontos-chaves do SURF é baseado na técnica de supressão de não-máximos do determinante da matriz Hessiana, enquanto a Transformada *SIFT* utiliza a operação Traço desta matriz.

A partir da detecção dos pontos-chaves do objeto de interesse e sua localização na imagem, realiza-se o procedimento de recuperação da pose do objeto detectado em relação

Projeto: MJ/SE-RIC	Emissão: 03/06/2015	Arquivo: 20150603 MJ RIC - RT Estudos de Tecnologias para modelos de artefatos em Papel	Pág.13/36
--------------------	---------------------	---	-----------

Confidencial.

ao plano da imagem (com referência ao dispositivo – câmera – de captura). A pose em imagens bidimensionais, a partir de um conjunto de pontos de referência provenientes do objeto de interesse pode ser facilmente recuperada utilizando técnicas de estimações de transformações rígidas e correção de projeção de perspectiva. Dentre as dezenas de técnicas na literatura, a mais eficiente, ao se tratar de transformações rígidas provenientes de operações de translações e rotações na imagem, é a homografia bidimensional [15], a qual é robusta e eficiente na realização deste processo.

De acordo com o exposto [15], a homografia bidimensional é uma transformação de projetiva que realiza o mapeamento de pontos de um plano para outro plano. A [Figura 5](#) apresenta uma descrição do processo de mapeamento entre planos.

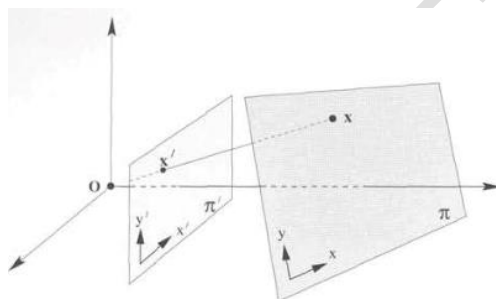


Figura 5 - Mapeamento entre planos. Figura retirada de [15]

Na [Figura 5](#) é apresentado um mapeamento do ponto x no plano π com o seu ponto correspondente x' no plano π' . Esta transformação pode ser definida como um mapeamento linear dos pontos, descrito pela Equação 2.

$$x' = Hx \quad (2)$$

Em que H é a matriz de homografia que mapeia os pontos x no espaço x' .

O processo de estimação da matriz de homografia H pode ser definido a partir de um conjunto de pares de pontos em mapeados por uma transformação em perspectiva fraca, no qual existe um processo de influência direta da profundidade e do alinhamento (rotação e translação) causado por uma transformação rígida entre o plano do objeto de interesse e o da imagem capturada. Em termos práticos, estamos procurando uma matriz que satisfaça a Equação 2, permitindo a recuperação da orientação entre o plano da câmera e do objeto, como exemplificado na [Figura 6](#).

Projeto: MJ/SE-RIC	Emissão: 03/06/2015	Arquivo: 20150603 MJ RIC - RT Estudos de Tecnologias para modelos de artefatos em Papel	Pág.14/36
--------------------	---------------------	---	-----------

Confidencial.



Figura 6 - Exemplo da influência da transformação de perspectiva fraca

Na [Figura 6](#), x', y' e x, y são descritos pela Equação 3, sendo w o fator da projeção dos pontos da imagem da esquerda na imagem da direita.

$$\begin{bmatrix} wx' \\ wy' \\ w \end{bmatrix} = \begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (3)$$

$\mathbf{x}' \quad \quad \mathbf{H} \quad \quad \mathbf{x}$

A obtenção da matriz de Homografia \mathbf{H} é obtida iterativamente a partir do cálculo de um conjunto de pontos correspondentes nas imagens, como demonstrado na [Figura 7](#).

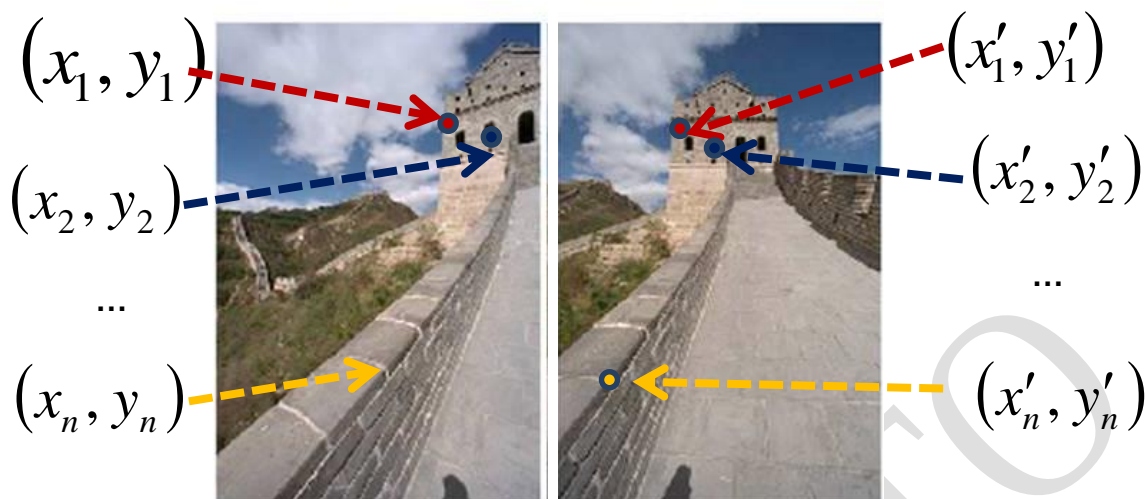


Figura 7 - Conjunto de pontos correspondentes para cálculo da matriz de homografia.

Este conjunto de pontos correspondentes é fornecido pelo algoritmo de detecção de objetos utilizando a transformada SURF, em que os pontos-chaves definidos são utilizados para resolver a Equação 3. O método de solução é iterativo e utiliza decomposição em valores singulares (algoritmo SVD) [16] em que a matriz H é o autovetor da matriz formada pela pré-multiplicação dos conjuntos de pontos x_n, y_n e x'_n e y'_n .

Ambos os descritores e as técnicas de recuperação da pose do objeto de interesse, incluindo também o algoritmo de Viola-Jones, estão disponíveis em bibliotecas *Open-Source* implementados (e otimizados) em linguagens *C/C++*, *Python* e *Java*, disponíveis para os principais sistemas operacionais modernos (*Linux*, *Unix*, *Windows* e *OSX*), incluindo suporte a desenvolvimento em sistemas operacionais móveis (*Android* e *iOS*). A principal biblioteca *Open-Source* que disponibiliza as implementações conjuntas de todos estes descritores é a *OpenCV* [14]. Maiores detalhes (manuais, tutoriais e exemplos) podem ser encontrados no sítio do desenvolvedor em <http://opencv.org/>.

O processo de detecção de regiões estruturadas em imagens, mediante ao domínio de aplicação restrito e estruturado, utilizou uma abordagem definida, tanto pela Transformada *SIFT*, quanto pelo algoritmo de Viola-Jones, combinados na forma melhorada utilizando a partir de detectores SURF [12] Conforme descrito na Seção anterior,, fazendo-se o uso de uma forma combinada destes algoritmos para permitir a detecção destas regiões específicas, bem como a aplicação da Homografia para a recuperação da pose do objeto de interesse na imagem, pode-se verificar que é possível realizar esta detecção solicitada, dentro dos requisitos esperados, que seriam a precisão na detecção e o tempo reduzido de

Projeto: MJ/SE-RIC	Emissão: 03/06/2015	Arquivo: 20150603 MJ RIC - RT Estudos de Tecnologias para modelos de artefatos em Papel	Pág.16/36
--------------------	---------------------	---	-----------

Confidencial.

processamento.

3 DETECÇÃO FACIAL EM IMAGENS

3.1 Contextualização

O processo de detecção automática de faces em imagens digitais é considerado na área de visão computacional como uma das atividades mais elaboradas na academia, despertando grande interesse da indústria e órgãos de segurança mundiais. O problema da detecção facial atualmente é voltado basicamente para o seguinte enunciado: dada uma imagem digital de entrada, na qual contenha pessoas (adultos e crianças) em que a face está visível, delimite a região desta face, permitindo realizar o seu reconhecimento. Neste caso, o reconhecimento está relacionado ao processo de associar a face a um único indivíduo, que pode ser utilizado para fins de identificação forense.

Geralmente em aplicações reais [17] (por exemplo, vídeo-monitoramento de áreas urbanas para fins de segurança), o processo de detecção facial é comprometido, basicamente, pelos seguintes aspectos: pose (orientação) da face na imagem; qualidade da imagem (resolução espacial e de *pixels*); oclusão (total ou parcial) da região da face; disfarces (uso de bonés, chapéus e óculos escuros); e desempenho (tanto no tempo de processamento, quanto da acurácia esperada para tornar a aplicação de detecção facial útil para o processo de reconhecimento).

3.2 Principais Trabalhos Científicos

Por se tratar de um assunto de grande interesse, vários trabalhos podem ser encontrados nos principais repositórios científicos. Contabilizando estes números de artigos e patentes especificamente sobre o tema “detecção facial”, foram encontrados 13.027 artigos no sítio do repositório oficial do *IEEE* (acrônimo em inglês para o Instituto de Engenheiros Elétricos e Eletrônicos) [18], indicando que este tema é ainda de muita importância.

Mesmo com esta grande quantidade de trabalhos científicos abordando o tema, uma unanimidade entre os pesquisadores da área, sendo considerado como o trabalho mais relevante para detecção rápida e com excelente desempenho na detecção facial em

Projeto: MJ/SE-RIC	Emissão: 03/06/2015	Arquivo: 20150603 MJ RIC - RT Estudos de Tecnologias para modelos de artefatos em Papel	Pág.17/36
--------------------	---------------------	---	-----------

Confidencial.

imagens, é a coletânea de artigos publicados por Paul Viola e Michael Jones. Estes são a mesma dupla de autores que desenvolveram o algoritmo Viola-Jones descrito na Seção anterior, em que na verdade o foco do trabalho de pesquisa deles era voltado para a detecção facial em imagens digitais.

Foram apresentadas duas versões para o algoritmo de detecção facial: a primeira versão [19], realizou a detecção robusta da região facial na imagem em tempo real; e a segunda versão realizou a detecção rápida de faces em sistemas em múltiplas vistas (i.e. *multiview systems*).

As técnicas apresentadas por Viola e Jones [19] [20] são consideradas o “estado da arte” no processo detecção facial, em que é capaz de determinar a região na imagem contendo uma face com rotacionadas dentro e fora do plano da imagem. Desde a primeira versão da metodologia proposta pelos autores [19], um bom desempenho computacional é mantido, chegando na época dos testes realizados em valores médios de tempo para a detecção facial, em imagens de 320 x 240 *pixels* de resolução espacial, em torno de 0,12s utilizando uma arquitetura computacional Intel (implementação P4 HT) com ciclo de relógio de 2,8 GHz.

Na Seção a seguir serão descritas as principais características do método proposto por Viola e Jones para a detecção rápida e eficiente de faces em imagens.

3.3 Método de detecção de faces desenvolvido por Viola e Jones

Nos trabalhos apresentados por Viola e Jones [19] [20], uma versão modificada do algoritmo, descrito na Seção 2.3, foi apresentada com uma única alteração no emprego de padrões retangulares que tivessem como prerrogativa a característica unívoca da face humana (independente de cor de pele, descendência e idade) que remete à região formada pelo conjunto de par de olhos. Em síntese, a proposta de Viola e Jones observou que devido à forma do conjunto ósseo, do globo ocular e da base do nariz, foi possível gerar um padrão de tons de cores semelhantes às mascaras retangulares da [Figura 8](#) a seguir.

Projeto: MJ/SE-RIC	Emissão: 03/06/2015	Arquivo: 20150603 MJ RIC - RT Estudos de Tecnologias para modelos de artefatos em Papel	Pág.18/36
--------------------	---------------------	---	-----------

Confidencial.

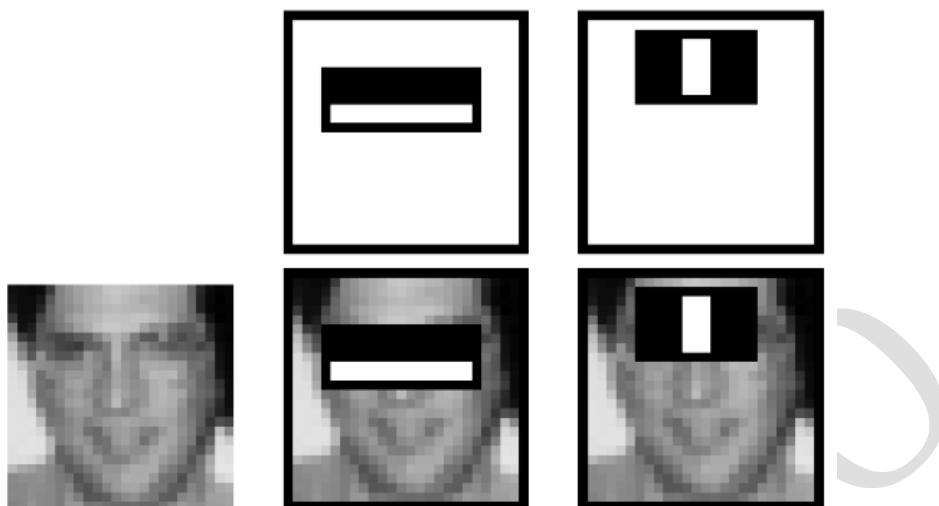


Figura 8 - Máscaras retangulares para detecção facial. Retirado de [19].

Facilmente pode ser observado que estas duas máscaras retangulares extraem a relação da posição do conjunto (par de) olhos na face. A estratégia destas máscaras é subtrair a soma do conjunto de *pixels* cobertos pela parte branca do retângulo pelo somatório dos *pixels* da área da parte preta da máscara. Esta relação permite que quando a máscara estiver sobre a região dos olhos, por ser uma região com menor incidência de iluminação (tendo assim a amplitude dos valores dos *pixels* mais baixa), o resultado desta diferença será maior do que para outras regiões, permitindo assim estimar (via relação de triangulação com a medida do tamanho da região dos olhos) a dimensão área da face na imagem.

Em seguida, utilizando o mesmo algoritmo de treinamento da Seção 2.3, este descritor é melhorado, permitindo assim sua detecção na imagem de entrada e delimitando a região facial.

Para comprovar a eficiência da sua técnica, foi apresentado [19] um conjunto de treinamento formado por 5000 (cinco mil) imagens de faces, todas frontais, re-escaladas para a dimensão espacial de 24 x 24 *pixels* e um conjunto com 300 milhões de imagens de não faces. A classe de imagens não faces é constituída por diversos tipos de imagens, tais como objetos, paisagens, imagens aéreas, entre outros, permitindo uma boa generalização do mecanismo proposto de treinamento. A ~~Figura 1~~ ~~Figura 1~~ ~~Figura 1~~ é um exemplo dos resultados apresentados da técnica implementada para a detecção facial [19].

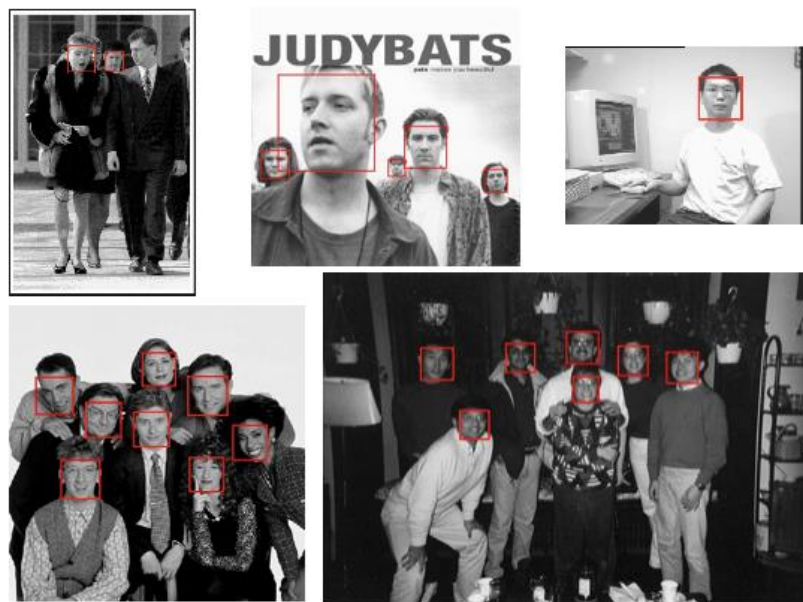


Figura 9 - Resultados da detecção facial. Retirado de [19]

Atualmente todo o processo de reconhecimento facial é disponibilizado [14] para utilização a partir de *templates* para detecção de componentes formadores da face (olhos, nariz e boca, por exemplo), face em perfil e face frontal, em arquivos separados no formato de linguagem de marcação estendida (padrão XML). Todos estes recursos são encapsulados em funções otimizadas para execução em tempo real e sob a licença BSD[14].

E ainda, assumindo questões práticas e funcionais, deve-se atentar à qualidade da imagem capturada para fins de detecção. Esta questão é polêmica e um pouco inconclusiva do ponto de vista de qual exatamente seria a melhor imagem para detecção de face. Entretanto, manuais internacionais recomendam que as imagens, principalmente para documentos oficiais de países utilizem as normas para padronizar e permitir que este tipo de sistema funcione com o maior grau de precisão na detecção. Em resumo, sugere-se o uso dos requisitos de qualidade da imagem facial proposta na Seção IV de [22], no qual os aspectos esperados da imagem facial utilizada para detecção estejam dentro dos atendimentos dos requisitos básicos para identificação forense utilizando características biométricas [22].

Projeto: MJ/SE-RIC	Emissão: 03/06/2015	Arquivo: 20150603 MJ RIC - RT Estudos de Tecnologias para modelos de artefatos em Papel	Pág.20/36
--------------------	---------------------	---	-----------

Confidencial.

4 ESTUDO DE ESPECIFICAÇÕES DE CONJUNTO DE FONTES, ALGORITMOS, ANÁLISE DE PERFORMANCE PARA RECONHECIMENTO AUTOMÁTICO DE TEXTO

O processo de reconhecimento automático de texto é definido seguindo a linha de dispositivos de reconhecimento ótico de caracteres (do inglês *Optical Characters Recognition* - OCR) [21], sendo desenvolvido para o reconhecimento automático de textos a partir de imagens digitalizadas (ou digitais) de textos escritos manualmente, datilografados ou impressos, permitindo sua manipulação por uma tecnologia computacional (por exemplo dispositivos móveis).

Atualmente as principais técnicas desenvolvidas utilizam algum processo de treinamento que permite melhores resultados. Entretanto, a grande maioria das técnicas de OCRs disponíveis no mercado já incorporaram a maioria das fontes textuais existentes, para o caso de imagens com textos gerados por algum sistema computacional que utilizem fontes tipo *True Types*. Na linha de texto manuscrito, ainda poucas técnicas são disponibilizadas a nível de aplicação, sendo restrita a alguns trabalhos científicos para identificação de usuário (quem escreveu) e não do que foi escrito [23].

4.1 Estudo das especificações de conjunto de fontes para identidades

As especificações dos conjuntos de fontes a serem utilizadas para o processo de reconhecimento em dispositivos OCRs, principalmente quando o foco de análise é referente na utilização desta metodologia para a leitura de zonas específicas de leituras automáticas por máquinas (tradução livre do termo em inglês *Machine Readable Zone* – MRZ) a partir de uma imagem digital, implicam a necessidade de que os elementos a serem reconhecidos permitam que as aplicações já existentes (aquelas que usam a leitura via resposta do espectro infravermelho) sejam também aproveitadas. Para o atendimento deste requisito, a melhor abordagem é trazer a padronização da elaboração e determinação da área de MRZ (com o foco específico do modelo de artefato sem chip, com MRZ) para as normas técnicas vigentes.

As normas técnicas que regulamentam explicitamente documentos de identificações (ex.: Passaportes) com informações biométricas ou não, com abrangência internacional entre os países signatários (que adotam) são elaboradas pela Organização Internacional de Aviação Civil (tradução livre do inglês *International Civil Aviation Organization* – ICAO).

Projeto: MJ/SE-RIC	Emissão: 03/06/2015	Arquivo: 20150603 MJ RIC - RT Estudos de Tecnologias para modelos de artefatos em Papel	Pág.21/36
--------------------	---------------------	---	-----------

Confidencial.

A ICAO elabora um conjunto de normas técnicas designadas pela norma ICAO 9303[24], divididas em 3 partes, no qual a primeira parte trata das especificações de passaportes que possuem área de leitura automática por máquina, a segunda parte descreve as especificações a respeito da confecção de vistos para leitura automática por máquina, e a terceira parte descreve as especificações para leitura de dados armazenados por dispositivos reconhecimento ótico de caracteres (OCRs), incluindo documentos da linha com informações biométricas.

Em referência a esta norma vigente, é notório que a padronização exigida tem como foco manter capacidade de intercambiar o uso dos documentos para a identificação oficial entre os diversos países, permitindo sua utilização plena de todas suas funcionalidades com eficiência, incluindo também os aspectos não só organizacionais das informações, como também os aspectos de construção e elaboração de cada um dos itens constantes nestes tipos de documentos. Reforça-se que a descrição que será feita aqui neste relatório contempla basicamente os requisitos indicados na norma ICAO 9303 em relação às especificações necessárias para a elaboração e definição das características da MRZ para permitir a utilização de dispositivos tipo OCRs para o reconhecimento dos caracteres contidos nesta região.

Inicialmente para esta descrição de requisitos/características, foram consultadas as duas partes, a primeira referente aos passaportes e a segunda à emissão de Vistos consulares, ambos possuindo normas relacionadas aos aspectos do reconhecimento da MRZ idênticas, com a exceção das dimensões e o número de linhas, que no caso dos Vistos consulares possuem no máximo 2 (duas) linhas e no caso do passaporte é possível ter até 3 (três) linhas, caso julgue-se necessário.

Para ambos os tipos de documentos que utilizam MRZ, consultado na norma em questão, em relação à fonte utilizada para a impressão dos caracteres, deverá ser utilizado o OCR-B². Na Figura 10 será apresentada uma amostra dos caracteres permitidos na MRZ.

² Arquivos das fontes OCR-B disponíveis gratuitamente para download no sitio: <http://www.font-zone.com/downloads.php?fid=1556>

Projeto: MJ/SE-RIC	Emissão: 03/06/2015	Arquivo: 20150603 MJ RIC - RT Estudos de Tecnologias para modelos de artefatos em Papel	Pág.22/36
--------------------	---------------------	---	-----------

Confidencial.

ABCDEFGHIJKLM
NOPQRSTUVWXYZ
abcdefghijklm
nopqrstuvwxyz
1234567890

Figura 10 - Conjunto de caracteres tipo OCR-B permitidos pela Norma ICAO 9303 para compor a MRZ. Retirado de [24].

Todos os caracteres permitidos devem ser utilizados. A norma proíbe o uso de caracteres acentuados, como ocorrem em alguns caracteres de nomes latinos e em países que utilizam caracteres cirílicos (por ex.: Rússia). O item 12.10 da Parte 1 do Volume 1, da ICAO 9303 apresentada todas as combinações/codificações para cada um dos alfabetos de caracteres acentuados e como deverá ser sua representação impressa na MRZ.

Após a codificação das informações a serem utilizadas na MRZ ser realizada, e em caso de espaço remanescentes, deverá ser obrigatoriamente utilizado o caractere "<<" para o preenchimento (sendo um único "<" para separar o nomes e sobrenomes) na MRZ.

A norma também é clara no quesito da metodologia de leitura, pois a especificação é diretamente direcionada à leitura por dispositivos utilizando infravermelho, de forma a obrigar que o documento que contem a MRZ fique muito próximo ao dispositivo. Esta restrição de proximidade se faz necessária para evitar possíveis fraudes por superposição de informações (por ex.: colagens de informações).

Entretanto, em relação à qualidade de impressão, a ICAO 9303 normatiza que está deverá ser feita para fins de leitura em dispositivos com infravermelho, utilizando o padrão da norma ISO 1831 Range X, na cor Preto (B425 – B680), com inclinação máxima no caractere de 3%. Nesta mesma norma são apresentados os diagramas referentes aos modelos a serem usados na MRZ. E com relação à capacidade, esta fica limitada ao tamanho da zona efetiva, de acordo com os limites de tamanho também especificados na norma.

Projeto: MJ/SE-RIC	Emissão: 03/06/2015	Arquivo: 20150603 MJ RIC - RT Estudos de Tecnologias para modelos de artefatos em Papel	Pág.23/36
--------------------	---------------------	---	-----------

Confidencial.

4.2 Análise de Performance do Reconhecimento Automático de Textos

A análise de performance do processo de reconhecimento automático dos caracteres contidos fica diretamente condicionada ao tipo de técnica (algoritmo) de reconhecimento adotado para este fim. A utilização de dispositivos OCRs para o reconhecimento de caracteres, como todas as aplicações em envolvem imagens, não possui soluções fechadas que atendam a todos os casos. A grande maioria dos OCRs desenvolvidos é designado às situações específicas, sendo estas mais simples ou mais complexas.

No mercado atual existe uma gama de metodologias para OCRs disponíveis, sendo em sua grande parte *software* proprietário, e alguns poucos gratuitos e *open source*. Como um dos pilares deste estudo realizado neste relatório técnico é a análise de técnicas e ferramentas que possuam uma abordagem consagrada na literatura técnica e científica atual, procurando por aquelas que atendam os principais requisitos *open source*, a principal tecnologia OCR, a qual atende esses quesitos, atende pelo nome de *Tesseract* [25]. Nas Subseções a seguir serão apresentadas as principais características do *Tesseract*.

4.2.1 Tesseract

De acordo Smith [25], o *Tesseract* é uma *engine open source* para reconhecimento ótico de caracteres desenvolvida pela HP entre 1984 a 1994. Esta *engine* partiu de um projeto de Doutorado no HP Labs (Bristol), tendo o seu foco de desenvolvimento inicial voltado para dar suporte a linha de digitalizadores de mesa da HP, já que as soluções proprietárias existentes à época não atendiam aos critérios de qualidade exigida pela HP. Durante o seu desenvolvimento e amadurecimento, juntamente com o apoio da divisão de digitalizadores do HP Labs do Colorado, o projeto foi submetido ao anual de acurácia de OCRs, em 1995, em que esta ferramenta provou seu valor, comparado às ferramentas comerciais disponíveis na época. O lançamento oficial ocorreu em 2005, momento em que a HP disponibilizou ao público para *download*³ como uma ferramenta *open source*.

A arquitetura proposta do *Tesseract* para a realização de OCR assume que a imagem de entrada é uma imagem binarizada. O processamento é realizado em um *pipeline* passo-a-passo, no qual o primeiro passo é localização dos componentes conexos da imagem binarizada cujos os pontos desconectados são armazenados em regiões contínuas (i.e.

³ Disponível para download em <http://code.google.com/p/tesseract-ocr>.

Projeto: MJ/SE-RIC	Emissão: 03/06/2015	Arquivo: 20150603 MJ RIC - RT Estudos de Tecnologias para modelos de artefatos em Papel	Pág.24/36
--------------------	---------------------	---	-----------

Confidencial.

blobs). O processo de reconhecimento é realizado em dois passos. No primeiro, tenta-se reconhecer a palavra inteira por tentativa de reconhecimento. Cada palavra é testada em um classificador adaptativo como dado de treinamento. Esta abordagem permite ao classificador obter melhores resultados ao final (parte inferior da página) do texto reconhecimento. A partir deste treinamento do classificador, o segundo passo consiste em melhorar o reconhecimento na parte inicial (parte superior), repetindo o processo de reconhecimento com o classificador agora treinado.

Em sua versão atual (3.02), o *Tesseract* trabalha de forma combinada com a biblioteca de processamento de Imagens Leptonica [26]. A Leptonica é uma biblioteca também *open source* e sua principal característica é na facilidade em lidar com diversos formatos de imagens textuais, permitindo sua conversão em mais de 60 (sessenta) idiomas diferentes.

Devida à sua grande aplicabilidade, em diversas situações, além do forte interesse em tornar a *engine* do *Tesseract* ainda mais completa, na versão 3.0 é possível realizar o treinamento da ferramenta OCR para um conjunto de caracteres exclusivo, como por exemplo o conjunto de caracteres usados na concepção do tipo de fonte OCR-B utilizada na elaboração da MRZ, como pode ser verificado na [Figura 10](#).

Sendo assim, assumindo que a região do MRZ foi corretamente detectada (região efetiva) na imagem de entrada, concebida segundo as normas propostas pela norma ICAO 9303 vigente, e também a partir de processamento simples de binarização (para ajudar o processo de reconhecimento de caracteres), a *engine* do *Tesseract* é uma das três melhores ferramentas para reconhecimento ótico de caracteres [25]. Nestas condições, existem fortes indícios que o elevado desempenho esperado no processo de reconhecimento de caracteres será atendido.

5 IMPLEMENTAÇÕES PARA VALIDAÇÃO DAS TÉCNICAS PROPOSTAS

A partir das considerações e descrições das técnicas apresentadas nas Seções anteriores, foram elaborados conjuntos de implementações computacionais, utilizando recursos e *frameworks open source*, que demonstram a aplicação, utilizando imagens provenientes de uma câmera digital de vídeo, com qualidade muito próxima às utilizadas nos celulares atuais. Assim, foi possível demonstrar a clara possibilidade de realizar a detecção robusta da MRZ e determinar sua pose (utilizando a homografia); detectar a

Projeto: MJ/SE-RIC	Emissão: 03/06/2015	Arquivo: 20150603 MJ RIC - RT Estudos de Tecnologias para modelos de artefatos em Papel	Pág.25/36
--------------------	---------------------	---	-----------

Confidencial.

região que contém a face impressa no novo RIC; e a partir de imagens com região MRZ detectada, aplicando uma técnica de OCR do *Tesseract* em conjunto com *Leptonica*, apresentado na Seção 4.2.1, foi possível realizar o reconhecimento dos caracteres contidos nesta região detectada. As Seções seguintes descrevem a metodologia empregada nestas implementações e os resultados preliminares obtidos, para fins de validação.

O ambiente computacional de desenvolvimento utilizado foi a plataforma *Xcode* 6.3.2 e biblioteca *OpenCV* [14] versão 2.4.10, utilizando a linguagem *C/C++* para sua implementação.

5.1 Detecção da MRZ e extração da Pose a partir de câmeras digitais

Nesta primeira avaliação, montou-se uma base de dados de imagens contendo as principais imagens de protótipos iniciais para a proposta do novo RIC. Todas estas imagens contem a MRZ (completa com caracteres ou somente a região efetiva limitada do MRZ sem caracteres) para serem utilizadas como referência (*Template*) de entrada para o processo de detecção. Nesta implementação, foram utilizados os descritores apresentados na Seção 2.4.

O teste realizado é simples, mas de crucial importância, pois a única informação proveniente será uma imagem de um protótipo do novo RIC, e este procedimento receberá uma imagem da câmera digital com uma imagem de vários outros possíveis padrões do novo RIC, em que contenha ou não caracteres na MRZ, existindo pelo menos a região efetiva da MRZ. Esta condição é necessária, pois é preciso também testar o sistema para a possibilidade de falsos positivos na detecção da MRZ. Um exemplo de falso positivo é apresentar uma imagem para a câmera que não contenha a MRZ e o sistema a localize. A implementação, por questões naturais, realiza a detecção da área delimitada do novo RIC na imagem como um todo. A partir desta região detectada, pode-se facilmente (utilizando estas informações provenientes do padrão geométrico onde a região do MRZ é esperada), recuperá-la pelas proporções geométricas do padrão utilizado para definir a região do RIC.

A [Figura 11](#) ~~Figura 11~~ ~~Figura 11~~ mostra um exemplo das imagens de protótipos anteriores para o novo RIC. Estas imagens foram obtidas a partir de imagens disponíveis da internet em reportagens em que o Novo RIC era apresentado.

Projeto: MJ/SE-RIC	Emissão: 03/06/2015	Arquivo: 20150603 MJ RIC - RT Estudos de Tecnologias para modelos de artefatos em Papel	Pág.26/36
--------------------	---------------------	---	-----------

Confidencial.



Figura 13 – Detectando o novo RIC, utilizando imagens de referência assinadas em imagens com o novo RIC sem assinatura.

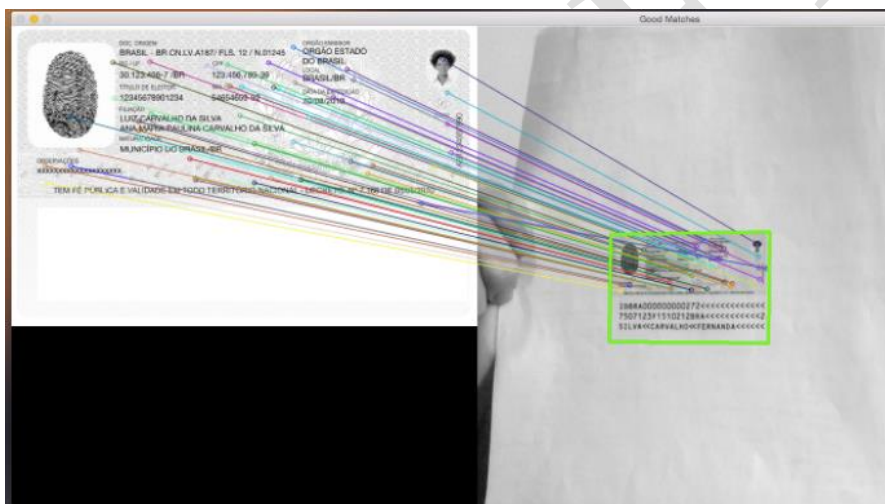


Figura 14 – Detectando o novo RIC utilizando como entrada sem caracteres na MRZ.

Nos exemplos de detecções realizados neste tópico, a pose é recuperada a partir da estimação em tempo real da matriz de homografia (retângulo verde nas imagens das regiões detectadas), seguindo o desenvolvimento da Seção 2.4. A partir desta matriz, aplicando uma transformação afim, é possível recuperar e projetar a imagem detectada diretamente em relação ao plano da imagem. A implementação demonstrou que dado o padrão (contendo ou não caracteres na MRZ) é possível detectar a região e recuperar a sua pose em uma imagem capturada por uma câmera digital.

avaliadas. Para categorias de aplicações que utilizam OCRs como ferramenta de extração de caracteres (por exemplo digitalizar caracteres de uma página de um livro antigo) é aceitável, pois o nível de erro, juntamente com ferramentas de correção textuais (verificador de ortografia) dos demais editores de textos atingem a ordem próxima a 100% de acerto. Entretanto, para a tarefa de reconhecimento dos caracteres contidos na MRZ (que contém em torno de 140 caracteres e estão codificados sem perdas) existe uma grande preocupação que pode inviabilizar a utilização de OCRs no reconhecimento destes caracteres, pois na falta de um único caractere não permitirá que a informação reconhecida pelo OCRs, mesmo os de melhores desempenho, permita o processo de reconstrução/recuperação da informação contida nesta região, impedindo o seu uso em situações reais de aplicação.

INCOMPLETO

Projeto: MJ/SE-RIC	Emissão: 03/06/2015	Arquivo: 20150603 MJ RIC - RT Estudos de Tecnologias para modelos de artefatos em Papel	Pág.31/36
--------------------	---------------------	---	------------------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para a MJ/SE.
É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, da MJ/SE.

6 CONCLUSÃO

Por meio de um trabalho coordenado e interdependente entre as equipes da SE e da Universidade de Brasília, as atividades de elaboração deste RT foram planejadas, discutidas, executadas e documentadas.

Este relatório foi elaborado com o intuito de apresentar uma revisão bibliográfica atual sobre as principais metodologias dos aspectos visando a utilização de técnicas de visão computacional no reconhecimento de regiões específicas (MRZs) em documentos de identificação, colaborando no aprimoramento das funcionalidades a serem aplicadas na proposta do novo RIC. Foram apresentados, nas seções anteriores, estudos sobre os seguintes temas: técnicas de detecção de objetos bidimensionais em imagens digitais; detecção facial em imagens; e análise de desempenho de reconhecimento ótico de caracteres em imagens digitais. Para todos estes temas, foram apresentadas implementações computacionais ajudando na validação e aplicabilidade sobre cada tema.

Com relação ao processo para detecção de objetos em imagens digitais bidimensionais, devido às diversas possibilidades de cenários de aplicação, foi proposto então um cenário em que a região do MRZ fosse detectada a partir da localização e recuperação da pose do documento de identificação. Optou-se então em avaliar o pior caso, no qual a qualidade das imagens utilizadas da câmera e da superfície da imagem (impressão em baixa qualidade) em que constatou a robustez da metodologia proposta utilizando a versão de descritores SURF. Nesta abordagem, os testes iniciais realizados mostraram que, mesmo com poucas variações das imagens utilizadas e daquelas padronizadas com melhor qualidade, os resultados foram satisfatórios, possibilitando a sua utilização para a realização desta tarefa.

Para a metodologia de detecção da região facial em imagens provenientes de câmeras digitais, capturadas a partir de áreas com face impressas em superfícies de documentos de identificação, tanto a metodologia apresentada para a realização desta tarefa quanto dos resultados obtidos nos testes evidenciaram que, mesmo em situações adversas e fora do padrão exigido pela norma (ICAO 9303), foram eficientes e conseguiram detectar a região da face sem maiores complicações. Reforça-se que para se alcançar melhores resultados (neste caso referentes a possíveis dificuldades causadas por situações adversas às testadas), deve-se seguir na elaboração do documento de identificação o que a norma

Projeto: MJ/SE-RIC	Emissão: 03/06/2015	Arquivo: 20150603 MJ RIC - RT Estudos de Tecnologias para modelos de artefatos em Papel	Pág.32/36
--------------------	---------------------	---	-----------

Confidencial.

internacional vigente preconiza, principalmente em relação à qualidade e tamanho da face no documento real impresso.

Já em relação ao desempenho de técnicas de reconhecimento de caracteres das regiões detectadas pelas técnicas anteriores, em que o conteúdo é padronizado (e os caracteres são sempre do mesmo padrão e utilizando fontes adequadas para estes tipos de dispositivos, como a OCR-B) os resultados alcançaram boas taxas de acerto. Entretanto, mesmo utilizando o que há de mais recente na literatura para este fim, o processo de identificação destes caracteres para a aplicação desejada deverá atingir índices de 100% de acerto, o que não é fornecido pela atual configuração.

Contudo, existe a possibilidade deste índice ser alcançado empregando-se restrições na qualidade da imagem de entrada (previamente detectada pelo processo de detecção de objetos bidimensionais) da MRZ, focando tanto na qualidade da impressão na concepção do documento (seguindo o especificado pela norma), quanto nos pré-processamentos necessários para garantir esta precisão, reforçando todo o processo com a aplicação de treinamento para a base de dados de caracteres que será utilizada. Neste caso, recomenda-se a fonte OCR-B, pois em se tratando de dispositivos com MRZ, a utilização de uma padronização a nível internacional pode ser conveniente e permitir maior escalabilidade futura para as novas versões. Neste item, mesmo pela grande abrangência que este assunto propicia, recomenda-se que novos estudos sejam elaborados e que seja realizado um conjunto de testes, utilizando bases de imagens reais (estas disponibilizadas pelo próprio MJ) de forma a permitir que se explore todas estas questões com maiores detalhes.

Vale também ressaltar que todos os testes realizados são iniciais e reforça-se que para resultados mais efetivos, novos e exaustivos testes devem ser realizados, explorando ainda mais as possibilidades que podem ser encontradas em situações que podem ser esperadas com o uso real destas aplicações.

As atividades envolvidas nesta etapa observaram formalmente a execução dos passos da metodologia elencada para gestão do projeto, PMI/PMBok.

A equipe da UnB considera que teve acesso a todas as informações necessárias à boa condução dos trabalhos e que a disponibilização dessas informações pela equipe da SE, assim como as atividades conjuntas de análise e discussão, levou a etapa do projeto a bom termo.

Projeto: MJ/SE-RIC	Emissão: 03/06/2015	Arquivo: 20150603 MJ RIC - RT Estudos de Tecnologias para modelos de artefatos em Papel	Pág.33/36
--------------------	---------------------	---	-----------

Confidencial.

7 REFERÊNCIAS BIBLIOGRÁFICAS UTILIZADAS

- [1] - David A. Forsyth, Jean Ponce, Computer Vision: A Modern Approach, Second Edition, Prentice Hall, 2011;
- [2] - Tinne Tuytelaars and Krystian Mikolajczyk. 2008. Local Invariant Feature Detectors: A Survey. Now Publishers Inc., Hanover, MA, USA.
- [3] - Rafael C. Gonzalez and Richard E. Woods. 2006. Digital Image Processing (3rd Edition). Prentice-Hall, Inc., Upper Saddle River, NJ, USA.
- [4] - Julesz, B. 1981. Textons, the Elements of Texture Perception, and their Interactions. Nature 290 (5802): 91–97. doi:10.1038/290091a0. PMID 7207603.
- [5] - Haralick, R.M., K. Shanmugan, and I. Dinstein, "Textural Features for Image Classification", IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-3, 1973, pp. 610-621.
- [6] - David G. Lowe. 1999. Object Recognition from Local Scale-Invariant Features. In Proceedings of the International Conference on Computer Vision-Volume 2 - Volume 2 (ICCV '99), Vol. 2. IEEE Computer Society, Washington, DC, USA, 1150-.
- [7] - Viola, P.; Jones, M., "Rapid object detection using a boosted cascade of simple features," Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on , vol.1, no., pp.1-511,1-518 vol.1, 2001.
- [8] - Lowe, D. G. Distinctive Image Features from Scale-Invariant Keypoints International Journal of Computer Vision, 2004, 60, 91-110.
- [9] - Shi, J.; Tomasi, C., "Good features to track," Computer Vision and Pattern Recognition, 1994. Proceedings CVPR '94., 1994 IEEE Computer Society Conference on , vol., no., pp.593,600, 21-23 Jun 1994.
- [10] - Ruch, David K.; Van Fleet, Patrick J. (2009). Wavelet Theory: An Elementary Approach with Applications. John Wiley & Sons. ISBN 978-0-470-38840-2.
- [11] - Freund, Y. and Schapire, R. A decision-theoretic generalization of on-line learning and an application to boosting JCSSI, 1997, 55, 119-139.
- [12] - Herbert Bay, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool. 2008. Speeded-Up Robust Features (SURF). Comput. Vis. Image Underst. 110, 3 (June 2008), 346-359.
- [13] - Neudecker, Heinz; Magnus, Jan R. (1988). Matrix Differential Calculus with Applications in Statistics and Econometrics. New York: John Wiley & Sons. p. 136. ISBN 0-471-91516-5.
- [14] - Bradski, G. OpenCV Library - Dr. Dobb's Journal of Software Tools, 2000.

Projeto: MJ/SE-RIC	Emissão: 03/06/2015	Arquivo: 20150603 MJ RIC - RT Estudos de Tecnologias para modelos de artefatos em Papel	Pág.34/36
--------------------	---------------------	---	-----------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para a MJ/SE.
É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, da MJ/SE.

- [15] - Hartley, R. e Zisserman, A. Multiple View Geometry in Computer Vision, second ed. Cambridge University Press, 2004.
- [16] – William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. 2007. Numerical Recipes 3rd Edition: The Art of Scientific Computing (3 ed.). Cambridge University Press, New York, NY, USA.
- [17] – Belhumeur, P.N.; Hespanha, J.P.; Kriegman, D., "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection," Pattern Analysis and Machine Intelligence, IEEE Transactions on , vol.19, no.7, pp.711,720, Jul 1997.
- [18] – Sítio oficial IEEEXplore, <http://ieeexplore.ieee.org> acessado em 19 de Maio de 2015.
- [19] – Viola, P.; Jones, M., "Robust real-time face detection," Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on , vol.2, no., pp.747,747, 2001.
- [20] – Viola, P., Jones, M., Fast Multi-view Face Detection Proc. of Computer Vision and Pattern Recognition, 2003.
- [21] – Schantz Herbert F. (1982) The History of OCR, optical character recognition [Manchester Center, VT.]; Recognition Technologies Users Association.
- [22] – Machine Readable Travel Documents – Part1: Machine Readable Passports; Volume 2 – Specifications for Electronically Enabled Passports with Biometric Identification Capability. ICAO, 6th Edition, 2006. Disponível em: <http://www.icao.int/publications/pages/publication.aspx?docnum=9303>.
- [23] – Guerin Jr., N. D. ; VIDAL, F. B. ; Espinoza, B. L. M. . Handwritten text verification on mobile devices. In: 10th International Conference on Computer Vision Theory and Applications, 2015, Berlim - Alemanha. 10th International Conference on Computer Vision Theory and Applications - VISAPP 2015. Setúbal-Portugal: INSTICC, 2015. v. 1. p. 1-8.
- [24] – Norma ICAO 9303 – Volume 1, Volume 2 e Volume 3, disponíveis em www.icao.int, acessado em 28 de maio de 2015.
- [25] – R. Smith. 2007. An Overview of the Tesseract OCR Engine. In Proceedings of the Ninth International Conference on Document Analysis and Recognition - Volume 02 (ICDAR '07), Vol. 2. IEEE Computer Society, Washington, DC, USA, 629-633.
- [26] – Bloomberg, D. Leptonica Image Processing Library, 2015. <http://leptonica.com>, acessado em 29 de maio de 2015.

Projeto: MJ/SE-RIC	Emissão: 03/06/2015	Arquivo: 20150603 MJ RIC - RT Estudos de Tecnologias para modelos de artefatos em Papel	Pág.35/36
--------------------	---------------------	---	-----------

Confidencial.

Este documento foi elaborado pela Universidade de Brasília (UnB) para a MJ/SE.
É vedada a cópia e a distribuição deste documento ou de suas partes sem o consentimento, por escrito, da MJ/SE.

Universidade de Brasília – UnB

Centro de Apoio ao Desenvolvimento Tecnológico – CDT

Laboratório de Tecnologias da Tomada de Decisão – LATITUDE

www.unb.br – www.cdt.unb.br – www.latitude.eng.br

