1. Revisão Bibliográfica

Neste capitulo serão abordados todos os conhecimentos necessários para o desenvolvimento do projeto, como as características da Carteira Nacional de Habilitação, as bibliotecas e linguagens utilizadas para o desenvolvimento e principais algoritmos utilizados na implementação.

* 1. Carteira Nacional de Habilitação

A CNH, Carteira Nacional de Habilitação, que também é conhecida como carteira de motorista é um documento de identificação obrigatório para qualquer cidadão que pretenda conduzir um veiculo automotor. Atualmente o código brasileiro divide a CNH em cinco categorias de acordo com o tipo de veículos que o condutor está habilitado a conduzir, sendo elas (DETRAN PR):

A – condutor de veículo motorizado de duas ou três rodas, com ou sem carro lateral (motos);

B – condutor de veículo motorizado não abrangido pela categoria A, com peso bruto total inferior a 3.500 quilos e lotação máxima de oito lugares, além do motorista (automóveis);

C – condutor de veículo motorizado usado para transporte de carga, com peso bruto superior a 3.500 quilos (como caminhões);

D - condutor de veículo motorizado usado no transporte de passageiros, com lotação superior a oito lugares além do motorista (ônibus e vans, por exemplo);

E - condutor de combinação de veículos em que a unidade conduzida se enquadre nas categorias B, C ou D e cuja unidade acoplada ou rebocada tenha peso bruto de 6 mil quilos ou mais; ou cuja lotação seja superior a oito lugares; ou, ainda, que seja enquadrado na categoria trailer.

A primeira CNH só pode ser retirada nas categorias A ou B e ela deve participar de cursos teóricos preparatórios, médico e psicotécnico. Após a primeira habilitação existem algumas regras para mudança de categoria (DETRAN MG):

* Categoria B: ter mais de 18 anos completos;
* Categoria C: ter, no mínimo, um ano na categoria “B”;
* Categoria D: ter 21 anos completos, estar habilidade no mínimo a 2 anos na categoria B ou 1 ano na categoria “C”;
* Categoria E: ter 21 anos completos, estar habilitado, há um ano nas categorias “C” ou “D”;

Atualmente a CNH possui, além dos dados acerca da habilitação, fotografia, número da carteira de identidade (RG) e do Cadastro de Pessoa Física (CPF). Assim a CNH pode ser utilizada como um documento de identificação pessoal em todo território nacional (Art. 159 do CTB). Além disto a CNH tem data de validade, ou seja, motoristas com menos de 65 anos devem renovar a carteira a cada 5 anos, após os 65 anos ou para alguns casos especiais o prazo de renovação é de 3 anos. Dessa forma, garante-se que o documento apresentado é sempre recente.

* 1. Imagem Digital

As imagens digitais podem ser produzidas por meio de diversos aparelhos, como câmeras de vídeo, câmeras fotográficas, scanners, aparelhos de raio-X, entre outros. A tecnologia da imagem digital permite a codificação digital documentos analógicos, em forma de imagem digital para armazenagem, transmissão em sistemas computadorizados.

Uma imagem digital é a representação de uma imagem bidimensional como uma sequencia finita de valores digitais. Cada um destes valores representa um pixel, que é o menor elemento que compõe uma imagem digital ou uma componente de cor de um pixel, no caso de imagens coloridas.

O valor de cada pixel é obtido pela quantização do sinal analógico (carga elétrica) obtido pelo dispositivo de captura, citados anteriormente. Os três tipos principais de imagens são:

* Imagens Binárias: Também conhecida como imagens preto e branco. Nestas imagens cada pixel pode assumir valores 0 ou 1. Este tipo de imagem normalmente surge no processamento de uma imagem digital como resultado de uma mascara ou certas operações como segmentação, threshold e ruído (estes conceitos serão discutidos nas próximas sessões).
* Imagens em escala de cinza: Nestas imagens cada pixel é representado por apenas um valor, que pode variar de 0 que indica a cor preta até 255 que indica a cor branca, totalizando 256 tons de cinza diferentes.
* Imagens coloridas: Nestas imagens, cada pixel é representado por pelo menos três informações, uma para cada canal de cor, como por exemplo o RGB que será um dos itens discutidos na próxima sessão. O valor de cada canal e composição da imagem colorida depende do espaço de cores utilizado.
  1. Visão computacional

A visão computacional é um campo que inclui métodos para a aquisição, processamento, análise e compreensão de imagens. A finalidade dessas análises e processamento é a obtenção de informações numéricas ou simbólicas de imagem, como por exemplo, na forma de decisões (SHAPIRO, 2001). Dessa forma, o objetivo de um sistema de visão computacional é extrair informações úteis a partir de dados de imagem usando modelos construídos com a ajuda da geometria, física, estatística e teoria de aprendizagem.

O sistema cognitivo do ser humano processa e extrai informações de imagens a todo momento e utilizar de sistemas computacionais para realizar a mesma atividade é algo extremamente complexo e que exige diversos conhecimentos e abstrações. Para isso a organização de cada sistema depende muito da sua aplicação. No entanto existem algumas funções típicas, que são encontradas em muitos sistemas de visão computacional e normalmente são executadas conforme a imagem XX:

[Imagem processamento]

**Aquisição de imagens**

A primeira etapa consiste em adquirir a imagem digital a partir de uma câmera digital, webcam ou scanner.

**Processamento**

Uma imagem digital pode possuir muitas informações que podem influenciar negativamente na tomada de decisão do sistema. O processamento permite que estas informações indesejadas sejam ignoradas ou retiradas (ZUECH, 1988). O processamento de imagens envolve técnicas de transformação visando melhorar, segmentar e extrair as características desejadas da imagem. As principais etapas deste processo são:

* Pré processamento: Nessa etapa a imagem é submetida a um processo de realce dos contrastes, deixando-a mais nítida para facilitar as etapas seguintes.
* Segmentação: A segmentação é uma etapa em que será delimitada a os objetos de interesse.
* Detecção: Após a segmentação é necessário determinar quais das áreas delimitadas são realmente necessárias para extrair a informação desejada, no sistema em questão, será possível determinar e classificar o dado da CNH.
* Interpretação: A interpretação dos dados a interpretação e classificação dos dados detectados.

Nas subseções seguintes serão descritas algumas técnicas e métodos utilizados nesta aplicação.

* + 1. Conversão para Escala de Cinza

Como já foi discutido nas seções anteriores, as imagens digitais possuem diversas representações. A mais comum é a chamada RGB (Red, Green, Blue) e ela produz várias cores através das três cores básicas: vermelha, verde e azul. Nesta representação cada pixel possui três bytes, um para cada cor, que podem possuir valores entre 0 e 255. Para reproduzir a cor preta as três dimensões devem possuir valores iguais a 0. Para obter a cor vermelha basta colocar 255 na componente R do pixel e 0 nas demais componentes. Dessa forma, é possível criar a cor desejada a partir da combinação ponderada de cada um dos componentes.

Porém algumas operações do processamento de imagens são facilitados quando os pixels têm apenas uma dimensão, ou seja, quando a imagem está em escala de cinza ou binária. Para converter uma imagem colorida em escala de cinza é feita uma média simples entre os valores de RGB do pixel e esta média é atribuída ao novo valor do mesmo.

* + 1. Binarização

A identificação de partes de interesse em uma imagem é uma das etapas mais criticas no processamento de imagens. A segmentação de uma imagem consistem em dividir a imagem em regiões de interesse, isolando aqueles pixeis que não fazem parte do mesmo. O

O threshold é um valor limite que é usado em um critério de seleção. Todos os pixels de uma imagem são comparados a esse critério e são alterados confome a necessidade. Um exemplo da aplicação de threshold é comparar todos os pixels a um valor limite e caso sejam maiores ou iguais a esse limiar ele é transformado em pretos (valor 0) e transformando em brancos (valor 255), caso contrário.

* + 1. Segmentação baseada em bordas

Na detecção de bordas são analisadas as descontinuidades nos níveis de cinza. Neste trabalho ela é empregada para delimitar os campos dos dados no documento.

Uma borda é o limite entre duas regiões com propriedades relativamente distintas de nível de cinza. As bordas na imagem caracterizam os contornos dos objetos e são bastante úteis para segmentação e identificação dos pontos de interesse. Porém, quando a imagem é diferenciada, todas as variações dos níveis de cinza são detectadas e, por consequência, detectam-se também bordas indesejadas. Para que as bordas indesejadas, provenientes de ruído ou textura da imagem, não sejam detectadas, deve-se tratar a imagem a fim de suaviza-la. Contudo, existem efeitos negativos ligados à suavização, como a perda de informação e o deslocamento de estruturas de feições relevantes na imagem.

Historicamente, uma variedade de detectores de bordas tem sido desenvolvida visando diferentes propósitos, com formulações matemáticas diferenciadas e com propriedades algorítmicas distintas. Em 1986, Canny desenvolveu um processo de detecção de bordas a partir de critérios de quantificação de desempenho de operadores de bordas conhecidos como os critérios de detecção e de localização.

* + 1. Transformada de Hough

A transformada é uma importante técnica para extrair características de uma imagem . Foi desenvolvida por Paul Hough em 1962 e se transformou popular no processamento de imagens em 1981, na publicação do artigo “Generalizing the Hough transform to detect arbitrary shapes”. A proposta desta técnica é basicamente a identificação qualquer forma que possa ser representada matematicamente, porém seu uso mais comum é para detecção de linhas, círculos e eclipses em imagens binárias, normalmente a imagem em que a transformada é aplicada passou anteriormente pela operação de reconhecimento de bordas.

* 1. Sistemas OCR

Segundo (Bhaskar, 2010), o algoritmo de reconhecimento óptico de caracteres (Optical Character Recognition – OCR) é a tradução , eletrônica ou mecânica, de imagens escaneadas de documentos escritos à mão, datilografados em máquinas de escrever ou impressos, para textos codificados em máquinas. As aplicações dos sistemas OCR são as mais variadas, indo desde o reconhecimento de placas de automóveis ou de outros tipos de placas e avisos, passando pela digitalização de documentos para o formato texto e finalmente reconhecimento de caligrafias manuais.

Segundo (Mori, 1992) os primeiros avanços destas tecnologias começaram nos anos 50, inicialmente apenas conceitualmente, mas gradualmente avançando para ferramentas mais concretas. As primeiras versões eram bastante sujeitas à falhas e se baseavam em comparação de documentos ou em imagens como modelos. Com o passar do tempo, concluiu-se que a comparação de modelos não era suficiente para obter bons resultados e começaram os estudos para utilização da analise estrutural dos caracteres para identificação dos mesmo. Esta mudança de abordagem deu origem aos sistemas atuais.

Atualmente, com o barateamento dos computação, aumento da capacidade de processamento [RUMELT,2002 ] e do grau de maturidade da tecnologia OCR, os sistemas de leitura de documentos estão cada vez mais acessíveis às pessoas e novos algoritmos, bem como técnicas para pré-processamento, extração e classificação de dados foram criados [CHERIET, 2007].

Muito além de reconhecer texto em imagens, leitura automática de documentos refere-se a todo e qualquer processo ou mecanismo que permita extrair, de forma automatizada, a informação textual existente nas imagens dos documentos. Neste projeto pretende-se não somente extrair o texto dos documentos, mas também atribuir sentido à essa informação. O objetivo deste trabalho é justamente reconhecer e atribuir sentido aos dados extraídos de um documento de identificação, no caso a CNH.

Existem alguns sistemas de leitura automatizada de documentos, mas a maioria deles utiliza um scanner próprio, como o que pode ser visto na imagem XXX, para obter a imagem de entrada. Dessa forma, o sistema fica limitado e restrito. Alguns fornecedores, como Mitek e TrustID, possuem ferramentas mobile e web para leitura automática de documentos, porém o fato de não serem soluções nacionais torna o suporte um pouco mais lento e as atualizações quanto à legislação e validade de documentos mais demorada.

[imagem scanner]

\section{Resumo do Caítulo}

Esse capítulo teve como objetivo uma revisão bibliográfica da teoria de visão computacional e etapas de processamento de imagens contextualizadas no sistema de leitura automática de documento. Com base nas informações apresentadas, no próximo capítulo será detalhado como essas técnicas foram utilizadas para a elaboração do projeto realizado.

<http://www.detran.pr.gov.br/modules/catasg/servicos-detalhes.php?tema=motorista&id=130>

EMGU. EmguCV. Disponível em: . Acesso em: 11 jun. 2014.

OPENCV. (2013). OpenCV: About OpenCV. Disponivel em: Acessado em: 17 Ago 2013.

<http://opencv.org/about.htm>

COUTO, Leandro Nogueira . Sistema para localização robótica de veículos autônomos baseado em visão computacional por pontos de referência . Biblioteca Digital da USP. São Carlos . 2012. Disponível em: . Acesso em: 18 Novembro 2012.

33. WATERS, D. J. Do microfilme à imagem digital: projeto conservação

preventiva em bibliotecas e arquivos. Disponível em:

<http://www.cpba.net>. Acesso em: 29 abr. 2003.

<http://www.prefeitura.sp.gov.br/cidade/upload/cpba_49_1253284217.pdf>

SHAPIRO, L. G. (2001). *Computer Vision* (1 ed., Vol. 1). Prentice Hall.

ZUECH, N. (1988). *Understanding and Applying Machihe Vision* (2 ed.). Yardley, Pennsylvania: Marcel Dekker INC.

DAVIES, R. (2005). *Machine Vision: Theory, Algorithms, Practicalities* 1 ed., Vol. 1. Morgan Kaufmann.

Bhaskar, implementing Optical Character Recognition on the Android Operating System for business cards

3.3.1 Aquisicão das Imagens

Para realizar a aquisição das imagens, utilizou-se um código disponível na *Internet* do programa *C# Webcam Capture* desenvolvido por Philip Pierce (PIERCE, 2003) que foi adicionado ao projeto do sistema biométrico no *Visual Studio* e adaptado conforme a necessidade. Esse programa captura a imagem de uma *webcam* conectada ao computador, mostrando a imagem a uma certa taxa de tempo, definida pelo usuário.

As imagens são adquiridas em uma resolução de 640 x 480 *pixels*.