[[1]](#footnote-1)

Correção de Perspectiva para Imagens de Documentos Capturados por Câmeras

Robson D. Montenegro e Saulo C. S. Machado

*Abstract*—These instructions give you guidelines for preparing papers for IEEE TRANSACTIONS and JOURNALS*.* Use this document as a template if you are using Microsoft *Word* 6.0 or later. Otherwise, use this document as an instruction set. The electronic file of your paper will be formatted further at IEEE. Define all symbols used in the abstract. Do not cite references in the abstract. Do not delete the blank line immediately above the abstract; it sets the footnote at the bottom of this column.

*Index Terms*—image edge detection, image processing

# Introdução

O

volume de informações existente e que ainda é produzido em papel é muito grande. A conversão de papel para mídia eletrônica costumava exigir equipamentos específicos que em muitas situações não podem ser usados por dificuldade de transporte, operacional ou devido aos seus elevados custos. Com o a redução dos preços de câmeras de dispositivos móveis, como celulares, elas se tornaram uma opção muito utilizada para captura de imagens de documentos.

A captura ou digitalização é uma etapa crítica de sistemas de processamento de imagens. Todo o conteúdo da imagem digital é obtido através dessa etapa e, perdas e distorções provenientes dela podem dificultar a interpretação da imagem, seja ela manual ou automática.

Apesar da simplicidade e praticidade de uso das câmeras de dispositivos móveis, suas imagens são diferentes das criadas a partir de equipamentos específicos como escâneres. Uma distorção comum causada pela captura usando câmeras é a distorção de perspectiva.

Mesmo os melhores motores de reconhecimento automático de texto não tratam problemas de distorção de perspectiva e, por isso, apresentam resultados piores para imagens de fotos de documentos.

A transformação de perspectiva é uma tarefa simples, a complexidade está em detectar qual distorção da imagem para que seja aplicada a transformação correta. Neste trabalho a detecção da distorção de perspectiva é baseada em uma abordagem híbrida para detecção de pontos de fuga. A detecção direta de pontos de fuga leva em consideração apenas os pixels da imagem. A abordagem indireta busca o ponto de fuga em um espaço de características obtido para a imagem.

A solução proposta foi aplicada a uma base de cem imagens de celular.

# Conceitos básicos

## Pontos de Fuga

Pontos de fuga são os pontos em que as linhas de profundidade dos objetos descritos na imagem irão convergir. Podemos dividi-los em horizontais e verticais de acordo com a direção em que está a convergência. A Fig. 1 representa os pontos de fuga obtidos para uma imagem através de informações de seus objetos.

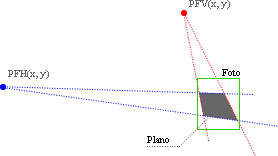


Fig. 1. Pontos de fuga horizontal (PFH) e vertical (PFV) decididos a partir de um plano de um objeto na foto.

## Transformada de Hough

A Transformada de Hough[3] é uma técnica bastante útil na detecção de formas geométricas. Uma imagem *f(x,y)* é representada no espaço de Hough por *H(θ,ρ)* utilizando o mapeamento cartesiano-polar:

*ρ = x cos θ + y sen θ (1)*

Para cada pixel, todas as possíveis linhas que passam pelo ponto são acumuladas em *H(θ,ρ).* Ao fim, é possível encontrar os conjuntos *(θ,ρ)* que obtiveram maior soma, e assim, determinar as linhas retas da imagem. Neste trabalho representamos as linhas através da equação .

# Visão geral

A detecção dos pontos de fuga é dada a partir de informações adquiridas a partir das abordagens direta e indireta. Ambas utilizam informações das linhas encontradas na imagem e suas interseções. A abordagem direta, como o nome sugere, trabalha diretamente com essas informações enquanto a indireta as utiliza na construção de um conjunto de características que será usado para a detecção de pontos de fuga mais precisamente.

O método utilizado parte da imagem do documento com distorção de perspectiva. É feito um pré-processamento que inclui transformação em tons de cinza, limiarização e detecção de bordas e, a partir das bordas são estimadas as linhas presentes na imagem. Essas linhas são usadas para detector os candidatos a pontos de fuga da imagem. Com os candidatos selecionados pela abordagem direta e indireta, é são selecionados os pontos de fuga mais adequados para então realizar uma transformação de perspectiva na imagem original. A transformação cria áreas inválidas na imagem, contendo valores desconhecidos que podem ser interpretados como as regiões que não eram visíveis na imagem. As regiões com pixels inválidos são removidas em uma última etapa de pós processamento. A Fig. 2 descreve esse fluxo.

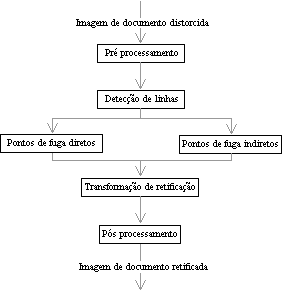


Fig. 2. Etapas de processamento do método utilizado neste trabalho.

O pré-processamento é essencial já que é a partir de seu resultado que são detectadas as linhas da imagem em uma etapa posterior. A imagem é redimensionada utilizando uma escala de 25% de seu tamanho, convertida para tons de cinza e limiarizada utilizando *Block-Otsu*. Também a partir da imagem em tons de cinza são detectadas as bordas da imagem usando o algoritmo de detecção de bordas de *Canny* [2].

# Detecção e seleção de linhas

Após detectadas as bordas, há uma fase de filtragem de componentes. As propriedades de BoudingBox e Eccentricity foram utilizadas para formar uma imagem resultante com apenas componentes adequados. Componentes com 5% de dimensão e excentricidade maior que 95% foram considerados. À imagem resultante foi aplicado a transformada de Hough[3] para detecção da linhas retas mais significativas pelo o seu tamanho.

# Detecção e seleção de pontos de fuga

A partir das linhas detectadas com a abordagem descrita anteriormente, são obtidos os pontos de intersecção. É utilizado o k-médias para o agrupamento dos pontos limitando o número de candidatos a pontos de fuga. Os centroides obtidos são utilizados para seleção dos pontos de fuga a partir dos valores das funções de lucro descritas a seguir.

## Abordagem Direta

A abordagem direta leva em consideração apenas as informações da imagem. A partir dessa premissa, o ponto de fuga mais confiável está entre os centroides obtidos no agrupamento. A escolha do centroide dá-se pela função de lucro que representa o peso do centroide no em relação as demais agrupamentos:

(2)

onde é peso do centroide de , é o número de elementos no cluster , é o total de clusters e é número de elementos no cluster .

## Abordagem Indireta

Esta abordagem utiliza os centroides estimados para o cálculo de *projeções por perspectiva*. Para todos os pontos de fuga candidatos calculamos a projeção da imagem a partir de um ponto de perspectiva , como mostra a Figura 3.



Fig. 3. Cálculo de projeção por perspectiva.

A projeção é dividida em fatias. Para cada pixel na imagem é possível saber qual índice do histograma ele se refere:

(3)

onde é o ângulo formado por e a origem da imagem em relação a H e é determinado por

(4)

onde e são os dois pontos no círculo circunscrito no retângulo formado pela imagem que são tangenciados por retas que passam por H.

O ponto de fuga mais confiável é escolhido pelo comportamento de seu histograma. Como mostra a Figura 4 um histograma construído a partir de um ponto de fuga correto irá gerar um histograma comportado. Isto se deve ao fato que o será realizado na direção da angulação das linhas.

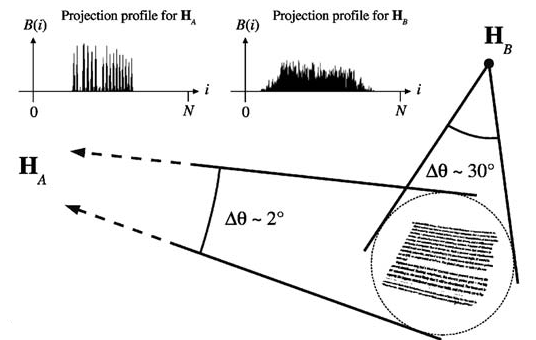


Fig. 4. Histograma de perspectiva.

Podemos perceber que o histograma de apresenta um espaçamento regular entre as barras do histograma, enquanto apresenta um histograma caótico. Para estabelecer qual ponto de fuga é o mais confiável usamos a derivada da soma dos quadrados para um histograma :

(5)

que irá estabelecer o ponto de fuga para qual o histograma é mais bem comportado, sendo assim este sua função de lucro.

## Consolidação dos Pontos de Fuga

Os resultados da abordagem direta e indireta são consolidados em apenas um resultado. Sendo o ponto de fuga estimado pela abordagem direta e pela abordagem indireta. Temos:

(6)

onde é p ponto de fuga mais confiável derivado da média aritmética das duas abordagens.

# Transformação

Como é possível perceber na Figura 5, pode-se determinar um quadrilátero que contém a imagem a partir dos pontos de fuga e do círculo circunscrito. A forma deste quadrilátero é utilizada para determinar a deformação da imagem e a matriz de transformação que será utilizada para realizar o mapeamento de retificação (Figura 6). Para transformações no plano 2D a relação é:

(7)

onde é o pixel corrigido e é o pixel original.

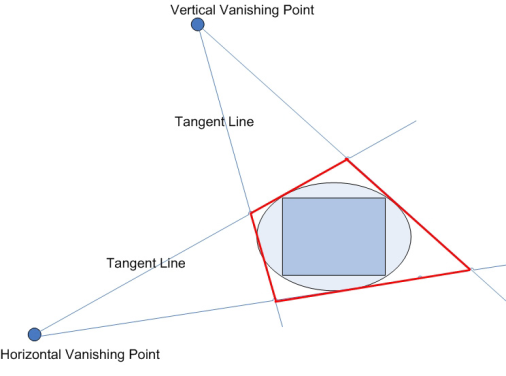
****

Fig. 5. Determinação do quadrilátero a partir dos pontos de fuga.

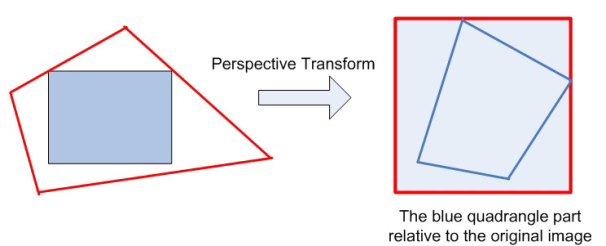
****

Fig. 6. Retificação da imagem.

# Experimentos

# Conclusões

Referências

1. Xu-Cheng Yin, Jun Sun, and Satoshi Naoi, “Perspective rectification for mobile phone camera-based documents using a hybrid approach to vanishing point detection”, International Workshop on Camera-Based Document Analysis and Recognition, 2., 2007, Paraná, Proceedings. Curitiba, 2007, p. 37-44.
2. J. F. Canny, “A computational approach to edge detection”, IEEE Transactions on PAMI, vol. 8, no. 6, 1986, p. 679-698.
3. Duda, R. O. and P. E. Hart, "Use of the Hough Transformation to Detect Lines and Curves in Pictures," Comm. ACM, Vol. 15, pp. 11–15 (January, 1972)

1. Trabalho realizado para a disciplina de Processamento de Imagens da Universidade Federal de Pernambuco, Centro de Informática. Entregue em 01 de Junho de 2012.

   R. D. Montenegro (e-mail: montenegrodr@gmail.com).

   S. C. S. Machado (e-mail: scsm@cin.ufpe.br). [↑](#footnote-ref-1)