Correção de Perspectiva para Imagens de Documentos Capturados por Câmeras

Robson D. Montenegro e Saulo C. S. Machado

**[[1]](#footnote-1) *Abstract*—There’s a large amount of data that exists and still is produced in paper. Creating digital documents using images captured from mobile cameras is even a greater challenge due to perspective distortions. The proposed method uses geometric concepts to estimate vanishing points and common image processing algorithms to detect useful information from the image.**

*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.

O método apresentado foi aplicado a imagens de câmeras de celular e os resultados foram comparados a resultados estimados visualmente.

# 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.

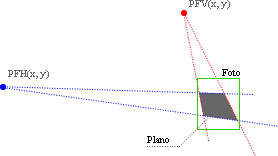


Fig. 1. Pontos de fuga horizontal (PFH) e vertical (PFV) decididos a partir de um plano de um objeto na foto.

## 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 as linhas são representadas pela equação .

# Visão geral

A detecção dos pontos de fuga é dada a partir de informações adquiridas a partir das abordagens indireta e direta. Ambas utilizam informações das linhas encontradas na imagem e suas interseções. A abordagem indireta é mais simples e leva em consideração apenas a quantidade de interseções para escolher o ponto de fuga. Já a direta utiliza essa informação para limitar o espaço de observações que são realizadas para a escolha dos pontos de fuga.

O método utilizado parte da imagem do documento com distorção de perspectiva. É feito um pré-processamento que inclui transformação em tons de cinza, binarização 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. Os candidatos selecionados pela abordagem indireta e direta, são consolidados em um resultado final 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.

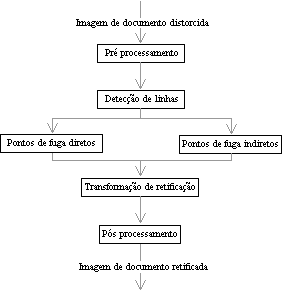


Fig. 2. Etapas de processamento do método utilizado neste trabalho.

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 reduzida utilizando a 25% de seu tamanho original, convertida para tons de cinza e binarizada utilizando *Block-Otsu*. Também a partir da imagem em tons de cinza são detectadas as bordas com o algoritmo de *Canny* [2].

# Detecção e seleção de linhas

A partir das bordas é realizada uma filtragem dos componentes. As propriedades de *BoudingBox* e *Eccentricity* foram utilizadas para formar uma imagem resultante com apenas componentes adequados. Apenas componentes com 5% de dimensão e excentricidade maior que 95% foram considerados. A imagem resultante foi transformada para o espaço de Hough [3] para detecção das linhas retas mais significativas de acordo com o tamanho.

# Detecção e seleção de pontos de fuga

As equações das linhas obtidas são utilizadas para encontrar os pontos de interseção entre elas. A Fig. 1 deixa claro por que esses pontos são escolhidos para compor os candidatos a pontos de fuga. Mas em uma imagem real certamente haverão mais linhas detectadas que as bordas do documento que alvo da correção. Para limitar o número de candidatos é feito um agrupamento desses pontos a partir do algoritmo K-Médias e apenas os centroides dos grupos são considerados candidatos a pontos de fuga.

Cada abordagem detalhada a seguir define uma função de lucro para a escolha de pontos de fuga. Essas funções são consolidadas em um terceiro momento para a escolha dos pontos de fuga finais.

## Abordagem Indireta

A abordagem indireta é baseada apenas nos centros dos agrupamentos. A partir de uma observação superficial da imagem é possível inferir que os pontos de fuga mais adequados onde mais linhas convergiram, isto é, o centroide do grupo que contém mais pontos. A função de lucro indireta pode então ser escrita como:

(2)

onde é o centroide de um grupo, seu peso sobre os centroides, é o número de elementos no grupo de , é o total de grupos e é número de elementos no grupo .

## Abordagem Direta

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 Fig. 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.

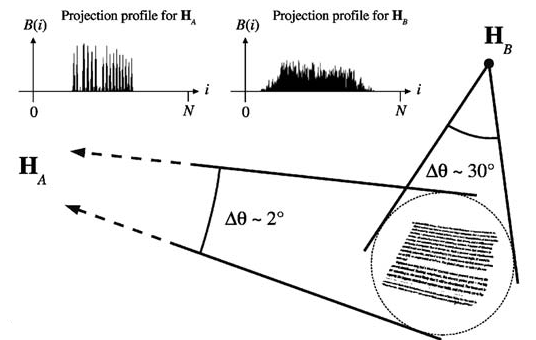


Fig. 4. Histograma de perspectiva.

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 obtido a partir do centroide :

(5)

que irá estabelecer o ponto de fuga para qual o histograma é mais bem comportado. Esta é a função de lucro que por conveniência é convertida para um coeficiente:

(6)

onde é a quantidade de centroides e é um centroide em .

## Consolidação dos Pontos de Fuga

Após a obtenção dos resultados pelas duas abordagens é feita a consolidação para definir o ponto de fuga final. Sendo o ponto de fuga estimado pela abordagem direta e pela abordagem indireta. Temos:

(7)

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

# Transformação

Como é possível perceber na Fig. 5, pode-se determinar um quadrilátero que contém a imagem a partir dos pontos de fuga e do círculo que o circunscreve. 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 (Fig. 6). Para transformações no plano 2D a relação é:

(8)

onde é o pixel corrigido e é o pixel original.

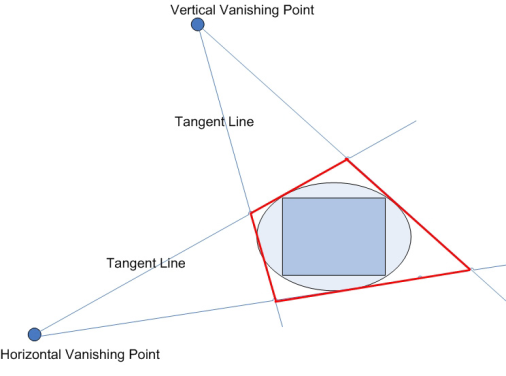
****

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

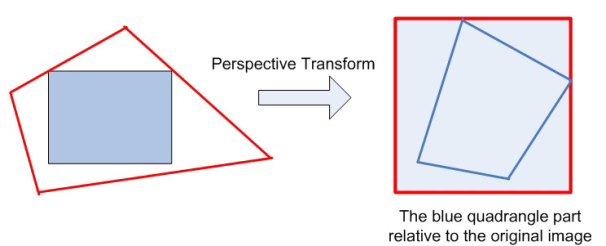
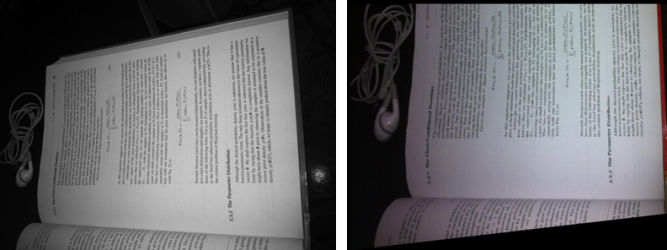
****

Fig. 6. Retificação da imagem.

# Resultados

Para avaliar a qualidade do método foram estimados visualmente resultados ideais para as imagens com distorção de perspectiva. Após a correção automática o resultado foi comparado e, como pode ser observado nas Fig. 7 e 8, foram satisfatórios.

Na Fig. 7 (b) é possível observar que, com a correção de perspectiva, é criada uma região inválida (em preto) na imagem. Isso acontece porque com a transformação o conteúdo disponível da imagem foi ajustado, mas a área visível continua a mesma. Para remover essas áreas é feito um corte considerando as menores dimensões visíveis para: topo, esquerda, baixo e direita. A Fig. 8 (b) já apresenta esse corte.



(a) (b)

Fig. 7. Imagem de livro antes (a) e depois (b) da retificação.



(a) (b)

Fig. 8. Imagem de quadro antes (a) e depois (b) da retificação.

# Conclusões

A solução apresentada visa superar os desafios da correção de perspectiva. A detecção automática do ponto de fuga é realizada com base nos conceitos geométricos. O método detecta as linhas e a partir delas encontra candidatos a pontos de fuga que são analisados usando duas abordagens diferentes para que, escolhendo os melhores, sejam consolidados os pontos de fuga finais para a retificação da imagem.

Apesar de ter observado bons resultados, é possível melhorar o método aprofundando o estudo na influência dos parâmetros dos algoritmos usados.

Além disso, tendo como objetivo dispositivos móveis, é necessário buscar alternativas para alguns algoritmos usados que sejam menos custosas computacionalmente.

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)