

Relatório Técnico

Disciplina: Processamento de imagens 2013.1

Programa de Pós-Graduação em Ciência da Computação

Universidade Federal de Pernambuco

Prof. Carlos Alexandre Barros de Mello

Aluno: JEFFERSON OLIVEIRA ALVES DE ARAUJO

Artigo selecionado:

A combined approach for the binarization of handwritten document images

K. Ntirogiannis, B. Gatos, I. Pratikakis

Pattern Recognition Letters Available online 11 October 2012, ISSN 0167-

8655, <http://dx.doi.org/10.1016/j.patrec.2012.09.026>.

2012

1. Introdução

O trabalho dos autores [1] apresenta uma abordagem de binarização para documentos manuscritos, com especial interesse nos documentos antigos. Utilizando-se da combinação de técnicas não tão boas, para esse tipo de problema, e vários pós-processamentos..

1.1 Motivação

A binarização de documento é um importante pré-processamento, muitas vezes essencial para a boa execução de algoritmos para reconhecimento de escrita manuscrita ou impressa [1].

Quando se trabalha com documentos antigos, existem vários fatores que podem atrapalhar o processo de binarização, dentre outros:

- Papéis delicados que onde existem escrita dos dois lados, fazendo com que a tinta passe para o outro;
- Deterioração
- Umidade

A binarização de documentos antigos ainda é um problema em aberto.

1.2 Objetivos

Assim, o trabalho [1] tem o objetivo de propor uma abordagem combinada para a binarização de documentos antigos. Testando os seus resultados com a base disponível em [4].

Nosso trabalho tem objetivo de implementar, testar e analisar visualmente o método proposto.

2. Metodologia

O método proposto pelos autores contém 12 passos, como lista o fluxograma apresentado na Figura 1.

A primeira fase do algoritmo descrito em [1] usa o algoritmo de binarização Niblack, com os parâmetros $w=60$ e $k=0.2$. O Niblack é um algoritmo de binarização que estima muito bem o caractere, porém traz muito ruído no fundo da imagem, existe a alternativa de utilizar Sauvola que é uma inovação no Niblack para diminuir ruído de fundo, porém o estimação do caractere é depreciada. Por esse motivo os autores preferiram utilizar Niblack [1].

O segundo passo consiste em pegar a imagem resultante de Niblack e aplicar a operação morfológica de dilatação no complemento, com um elemento estruturante 3×3 padrão e depois novamente complementa. Este passo é importante para tentar conter os pixels de transição entre caractere e fundo.

No terceiro passo o fundo da imagem sem os caracteres é estimado. Utilizando-se do método proposto em [1]. Não obteve-se êxito na implementação do método, portanto foi utilizada outra abordagem. Foi aplicado um borramento na imagem original com um filtro Box com o tamanho de 15% da altura da imagem e posteriormente cada pixel correspondente ao primeiro de Niblack dilatado foi substituído na imagem original, desta forma obteve-se a estimação de fundo $BG(x, y)$.

No quarto passo a imagem é normalizada seguindo a equação 1, onde $I(x,y)$ representa a imagem original. E $F(x,y)$ apresentado na equação 2, é a razão entre a imagem de estimação do fundo, $BG(x,y)$, e imagem original $I(x,y)$. Com a normalização da imagem com o fundo os valores de primeiro plano e segundo são bem separados.

$$N(x,y) = \left[(I_{max} - I_{min}) \cdot \frac{F(x,y) - F_{min}}{F_{max} - F_{min}} + I_{min} \right] \quad (1)$$

$$F(x,y) = \frac{I(x,y) + 1}{BG(x,y) + 1} \quad (2)$$

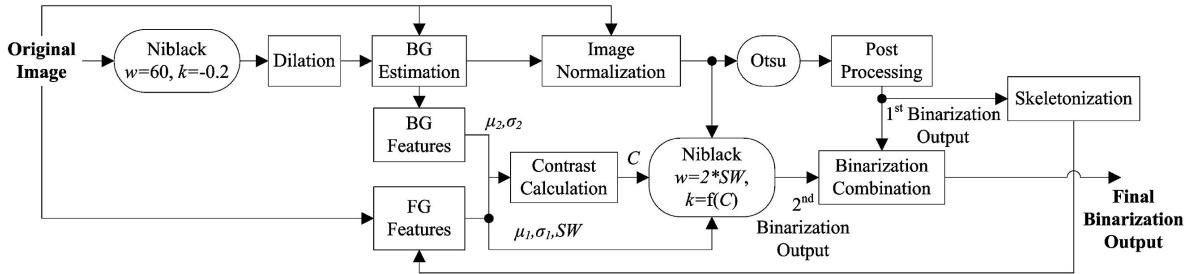


Figura 1 - Etapas do método proposto [1]

Após o passo de normalização, aplica-se o método de Otsu. Este método escolhe o melhor valor de thresholding que tenha a maior variância entre classes [2]. Como a normalização separou bem a escrita do fundo, essa é uma boa estimativa de onde está o primeiro plano. Todavia o método de Otsu costuma ter alguns poucos ruído no plano de fundo, então é aplicado um pós-processamento para tentar removê-lo.

O pós processamento para remoção do ruído da imagem $O(x,y)$, resultante da aplicação de Otsu em $N(x,y)$, é feito somente removendo todos os componentes conectados que não satisfaçam a equação 3. Onde h é a altura do componente conectado, na prática deve-se procurar o valor mínimo de h que satisfaça a equação 3, todos os componentes conectados que tiverem valores menores que esse h devem ser descartados da imagem final Otsu-pós-processado ($OP(x,y)$).

$$\sum_{j=1}^h \frac{RP^j}{RC^j} > 1 \quad (3)$$

$RP^j = \frac{\sum_{x=1}^{lx} \sum_{y=1}^{ly} O^j(x,y)}{\sum_{x=1}^{lx} \sum_{y=1}^{ly} O(x,y)}$, é a razão entre o somatório de todos os pixels dos componentes conectados de tamanho j da imagem $O(x,y)$ ($O^j(x,y)$), com o somatório de todos os pixels da imagem $O(x,y)$.

$RC^j = \frac{NC^j}{NC_o}$, é a razão entre o numero de componentes conectados de tamanho j , com o numero total de componentes conectados na imagem $O(x,y)$. Aplica-se a esqueletização sobre a imagem $OP(x,y)$ para obter os pixels mais centrais possíveis de onde estejam os traços, gerando a imagem $S(x,y)$.

Segundo os autores [1], $S(x,y)$ é uma boa estimativa do centro do traço, e como tal pode ser usado para estimar o tamanho do traço para possibilitar a geração dos melhores valores para uma binarização por Niblack mais precisa. O tamanho médio de traço pode ser obtida a partir da equação 4.

$$SW = \text{Média}(SW_j), \quad j = 1, 2, \dots, \# \text{ esqueleto conectado} \quad (4)$$

onde, $SW_j = \text{Max}(S_j^{sw}(x,y))$, com $S^{sw} = 2D + 1$, onde D é a distancia entre o pixel do esqueleto e o plano de fundo da imagem OP(x,y).

A imagem S(x,y), posteriormente, é usada para calcular o contraste da imagem a partir da equação 5. Onde BG corresponde ao plano de fundo nas coordenadas do esqueleto, e FG corresponde ao primeiro plano nas coordenadas do esqueleto.

$$C = -50 \cdot \log_{10} \left(\frac{FG_{média} + FG_{desvio}}{BG'_{média} - BG'_{desvio}} \right) \quad (5)$$

Apos a extração das características da imagem como contraste entre fundo e escrita e largura do traço. Pode-se estimar os parâmetros para o segundo uso de Niblack onde $k = -0.2 - 0.1 \cdot \left[\frac{c}{10} \right]$ e $w = 2.SW$. Aplica-se o Niblack na imagem normalizada, N(x,y), gerando a imagem NB(x,y).

A NB(x,y) ainda contém alguns ruídos de fundo e alguns caracteres falhos. A imagem binarizada final é resultado da junção da imagem OP(x,y) e NB(x,y) a partir do uso da equação 6, para tentar sanar essas falhas CO(x,y) é o resultado final do método proposto.

$$CO(x,y) = \bigcup_{j=1}^n NB_j(x,y), \quad \forall j : d(j) = \text{verdade} \quad (6)$$

$$\text{onde, } d(j) = \begin{cases} \text{verdade, se } 100 \frac{\sum_{x=1}^{lx} \sum_{y=1}^{ly} OP(x,y) \cdot NB_j}{\sum_{x=1}^{lx} \sum_{y=1}^{ly} NB_j(x,y)} \geq C \\ \text{falso, caso contrário} \end{cases}$$

2.1. Experimentos

O algoritmo proposto em [1] foi implementado em MatLab (no CD que acompanhou este trabalho). O algoritmo foi testado com algumas imagens das bases DIBCO2009 [3] e DIBCO2011 [4].

As imagens do Anexo 1 foram utilizadas para testas o algoritmo proposto. Os resultados constam no Anexo 2.

2.2. Análise

Com os resultados apresentados na seção anterior, foi averiguado que o algoritmo apresenta uma boa binarização mesmo com imagens bastante problemáticas. Porém é possível perceber que alguns artefatos apareceram, principalmente quando eles estão conectados ao texto.

É provável que este problema ocorre pela falha ao tentar implementar o algoritmo de estimação de plano de fundo.

4. Conclusões

Neste relatório, tratamos do problema de binarização de documentos manuscritos antigos. Foi apresentado o método proposto por [1]. Esse método foi implementado em Matlab e testado com as bases [3] e [4].

Apesar dos bons resultados apresentados pelos autores nos testes, obtivemos visualmente uma performance inferior, é provável que este problema tenha ocorrido pelo fato de não ter usado o algoritmo de estimação de plano de fundo proposto pelo autor.

Como extensão deste trabalho sugerimos o teste de outras técnicas de restauração de imagens para estimar o plano de fundo e assim obter melhor performance.

Referências

- [1] K. Ntirogiannis, B. Gatos, I. Pratikakis, A combined approach for the binarization of handwritten document images, Pattern Recognition Letters, Available online 11 October 2012, ISSN 0167-8655, <http://dx.doi.org/10.1016/j.patrec.2012.09.026>
- [2] R.D. GONZALEZ, R.E. WOODS, Processamento Digital de Imagens 3. Ed., Pearson Prentice Hall, 2010
- [3] Site da Internet: <http://users.iit.demokritos.gr/~bgat/DIBCO2009/>, acesso em 10/06/20113
- [4] Site da Internet: <http://utopia.duth.gr/~ipratika/DIBCO2011/>, acesso em 10/06/20113

Anexos

Anexo 1: Imagens de teste

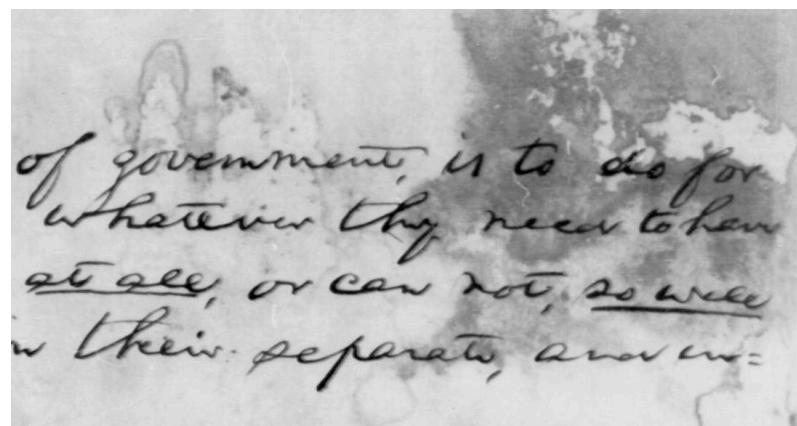


Figura 2 – Primeiro [3]

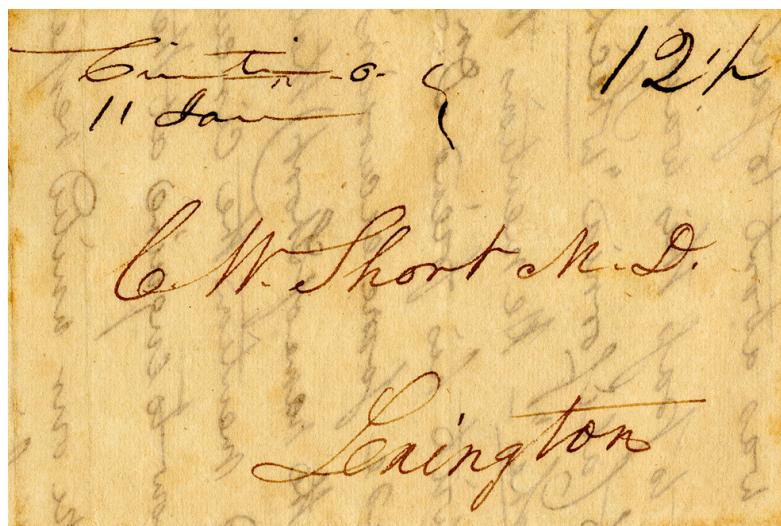


Figura 3 – Segundo [4]

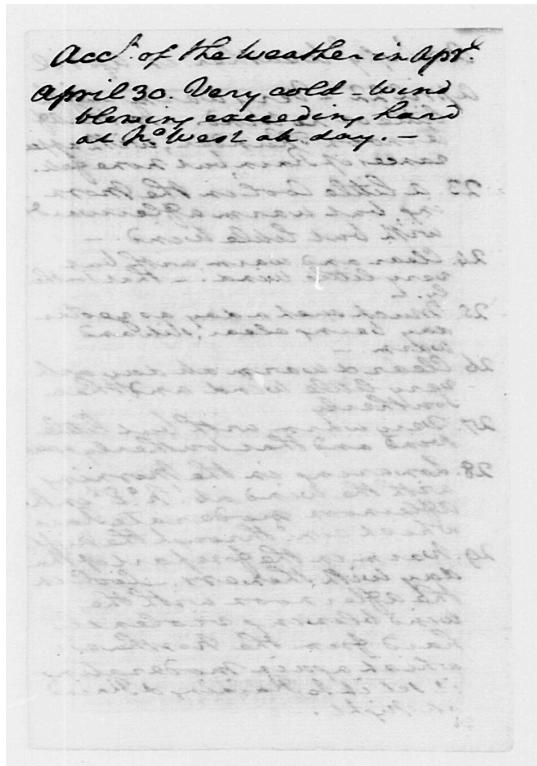


Figura 4 – Quarto [3]

turn me of
h Star saw
very much
y forth saw
has appear
had to command
the same go

Figura 5 – Sexto [4]

^b
of government, is to do for
whatever they need to have
at all, or can do, as well
in their separate, and in

Figura 6 - Resultado Primeiro

~~Garrison~~ 11 days 1/2 h

C. W. Short M.D.

Lafayette

Figura 7 - Resultado Segundo

Acc of the weather - 1808
April 30 Very cold wind
blowing exceeding hard
as he west all day.

Figura 8 - Resultado Terceiro

~~turn me of~~
to the
I have seen
very much
of farther fair
has appear
had to command
the same go

Figura 9 - Resultado Quarto