

Segmentação por Limiarização

Alexandre Xavier Falcão e David Menotti

Instituto de Computação - UNICAMP

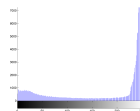
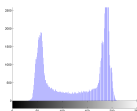
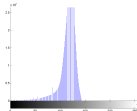
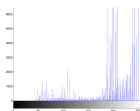
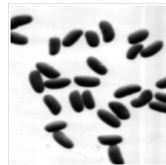
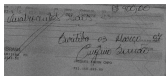
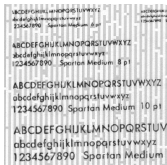
afalcao@ic.unicamp.br

Segmentação por Limiarização

- Para simplificar, vamos considerar problemas envolvendo a segmentação da imagem em duas classes: objeto e fundo.

Segmentação por Limiarização

- Para simplificar, vamos considerar problemas envolvendo a segmentação da imagem em duas classes: objeto e fundo.
- Este é o caso, por exemplo, da segmentação de imagens de texto, imagens de cheques, imagens de impressão digital, imagens de impressão digital, imagens de células, etc.

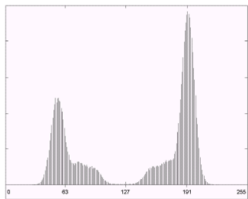


Segmentação por Limiarização

- Seja $\hat{I} = (D_I, \vec{I})$ uma imagem cinza ($\vec{I}(p) = I(p)$).

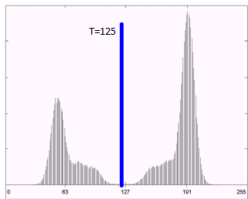
Segmentação por Limiarização

- Seja $\hat{I} = (D_I, \vec{I})$ uma imagem cinza ($\vec{I}(p) = I(p)$).
- A partir do histograma de uma imagem cinza $\hat{I} = (D_I, I)$, percebe-se que é possível separar objeto e fundo a partir de intervalos disjuntos do histograma, *i.e.*, $[l_o, h_o]$ e $[l_f, h_f]$ usando-se um valor (ou limiar/*threshold*) T , tal que $T = h_o = l_f - 1$, por exemplo.



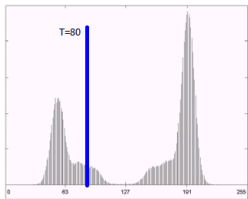
Segmentação por Limiarização

- Seja $\hat{I} = (D_I, \vec{I})$ uma imagem cinza ($\vec{I}(p) = I(p)$).
- A partir do histograma de uma imagem cinza $\hat{I} = (D_I, I)$, percebe-se que é possível separar objeto e fundo a partir de intervalos disjuntos do histograma, *i.e.*, $[l_o, h_o]$ e $[l_f, h_f]$ usando-se um valor (ou limiar/*threshold*) T , tal que $T = h_o = l_f - 1$, por exemplo.



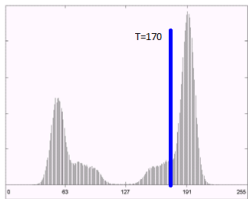
Segmentação por Limiarização

- Seja $\hat{I} = (D_I, \vec{I})$ uma imagem cinza ($\vec{I}(p) = I(p)$).
- A partir do histograma de uma imagem cinza $\hat{I} = (D_I, I)$, percebe-se que é possível separar objeto e fundo a partir de intervalos disjuntos do histograma, *i.e.*, $[l_o, h_o]$ e $[l_f, h_f]$ usando-se um valor (ou limiar/*threshold*) T , tal que $T = h_o = l_f - 1$, por exemplo.



Segmentação por Limiarização

- Seja $\hat{I} = (D_I, \vec{I})$ uma imagem cinza ($\vec{I}(p) = I(p)$).
- A partir do histograma de uma imagem cinza $\hat{I} = (D_I, I)$, percebe-se que é possível separar objeto e fundo a partir de intervalos disjuntos do histograma, *i.e.*, $[l_o, h_o]$ e $[l_f, h_f]$ usando-se um valor (ou limiar/*threshold*) T , tal que $T = h_o = l_f - 1$, por exemplo.



Segmentação por Limiarização - Ótima

- Considere uma imagem cinza $\hat{I} = (D_I, I)$ com $0 \leq I(p) \leq L - 1$ para todo $p \in D_I$, duas classes de interesse, C_1 e C_2 , representando objeto e fundo, e as seguintes definições.
 - P_1 é a probabilidade *a priori* de ocorrer C_1 .
 - P_2 é a probabilidade *a priori* de ocorrer C_2 .
 - $p_1(I)$ é a densidade de probabilidade do nível de cinza I em C_1 .
 - $p_2(I)$ é a densidade de probabilidade do nível de cinza I em C_2 .
 - $p(I)$ é a densidade de probabilidade do nível de cinza I na imagem.

Portanto,

$$p(I) = P_1 p_1(I) + P_2 p_2(I), \quad (1)$$

Segmentação por Limiarização - Ótima

- e pela regra de decisão de Bayes, um pixel p deve ser classificado como classe 1, se $P_1 p_1(I) > P_2 p_2(I)$, e como classe 2, se $P_1 p_1(I) < P_2 p_2(I)$, ficando a igualdade a critério da implementação.
- Esta regra minimiza a probabilidade de erro assumindo como limiar de decisão $P_1 p_1(I) = P_2 p_2(I)$.

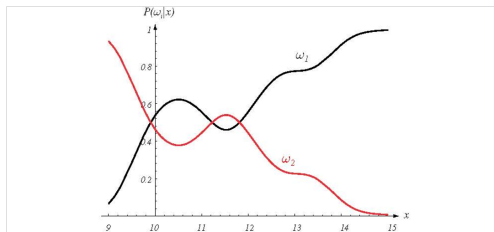


FIGURE 2.2. Posterior probabilities for the particular priors $P(\omega_1) = 2/3$ and $P(\omega_2) = 1/3$ for the class-conditional probability densities shown in Fig. 2.1. Thus in this case, given that a pattern is measured to have feature value $x = 14$, the probability it is in category ω_2 is roughly 0.08, and that it is in ω_1 is 0.92. At every x , the posteriors sum to 1.0. From: Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern Classification*. Copyright © 2001 by John Wiley & Sons, Inc.

- Supondo que $p_1(l)$ e $p_2(l)$ são distribuições Gaussianas com médias μ_1 e μ_2 , e variâncias σ_1^2 e σ_2^2 conhecidas,

$$p_1(l) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left[\frac{-(l - \mu_1)^2}{2\sigma_1^2}\right] \quad (2)$$

$$p_2(l) = \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left[\frac{-(l - \mu_2)^2}{2\sigma_2^2}\right], \quad (3)$$

- e que $0 \leq \mu_1 < \mu_2 \leq L - 1$, os intervalos de brilho que identificam cada classe são definidos a partir de parâmetros.

Segmentação por Limiarização - Ótima

- O problema passa a ser a estimação de parâmetros, *i.e.*,
 $\theta = (P_1, p_1(I), P_2, p_2(I))$.

Segmentação por Limiarização - Ótima

- O problema passa a ser a estimação de parâmetros, *i.e.*, $\theta = (P_1, p_1(l), P_2, p_2(l))$.
- Uma forma de modelar matematicamente este problema é por meio do método *Maximum-likelihood estimation*.

$$\begin{aligned}f(\hat{I}|\theta) &= f(\hat{I}(1), \hat{I}(2), \dots, \hat{I}(p), \dots, \hat{I}(n)|\theta) \\ \mathcal{L}(\theta|\hat{I}) &= f(\hat{I}(1)|\theta) \times \dots \times f(\hat{I}(p)|\theta) \times \dots \times f(\hat{I}(n)|\theta) \\ &= \sum_{j=1}^2 \prod_{i=1}^n P_j p_j(\hat{I}(i)|\theta_j) = \sum_{j=1}^2 \prod_{l=0}^{L-1} P_j p_j(l|\theta_j) h(l) \\ \ln \mathcal{L}(\theta|\hat{I}) &= \sum_{j=1}^2 \left(n \ln P_j + \sum_{i=1}^n \ln p_j(\hat{I}(i)|\theta_j) \right) \\ &= \sum_{j=1}^2 \left(n \ln P_j + \sum_{l=0}^{L-1} \ln p_j(l|\theta_j) h(l) \right)\end{aligned}$$

em que n L são o número de pixels e níveis de cinza na imagem.

Segmentação por Limiarização - Ótima

- Maximizar $\ln \mathcal{L}(\theta|\hat{I})$ (*log-likelihood*) é equivalente a maximizar $\mathcal{L}(\theta|\hat{I})$.

Segmentação por Limiarização - Ótima

- Maximizar $\ln \mathcal{L}(\theta|\hat{I})$ (*log-likelihood*) é equivalente a maximizar $\mathcal{L}(\theta|\hat{I})$.
- É necessário o uso de otimização numérica (sem fórmula fechada).
- O algoritmo *Expectation-Maximization* (EM) é usado para tal fim.
 - Inicialização dos parâmetros do modelo.
 - Enquanto não há convergência (erro/iterações), faça:
 - *Expectation step* - calcula uma função para maximizar a esperança (*expectation*) da função *log-likelihood* com base nos parâmetros estimados.

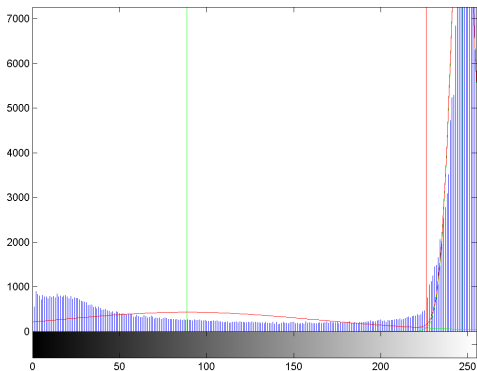
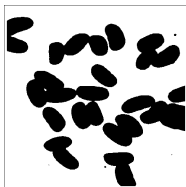
$$\frac{\partial \ln \mathcal{L}(\theta|\hat{I})}{\partial \theta} = 0$$

- *Maximization step* - atualiza os parâmetros do modelo a partir do *Expectation step*.

$$((P_1, \mu_1, \sigma_1), (P_2, \mu_2, \sigma_2))$$

Segmentação por Limiarização - Ótima

- Resultados

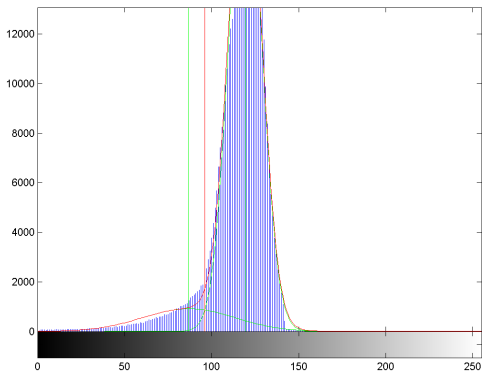
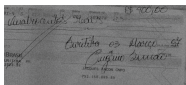


$088.7(\pm 74.4)[000.0 - 226.0]$

$247.2(\pm 06.6)[227.0 - 255.0]$

Segmentação por Limiarização - Ótima

- Resultados

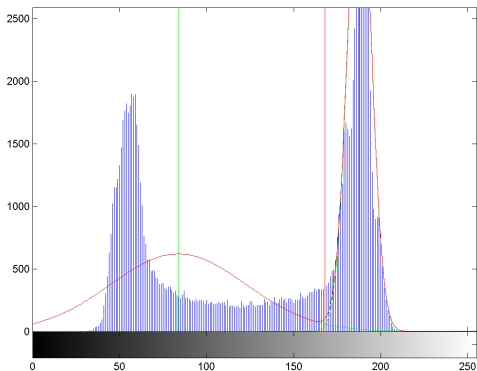


$119.3(\pm 09.3)[097.0 - 151.0]$

$086.8(\pm 26.4)[000.0 - 163.0]$

Segmentação por Limiarização - Ótima

- Resultados

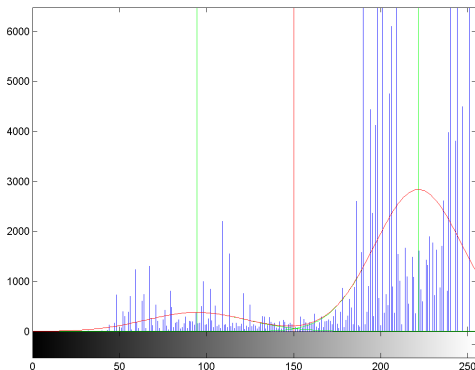
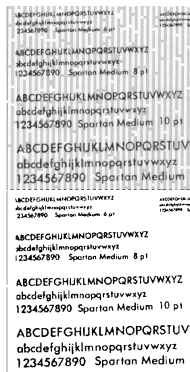


$187.7(\pm 06.6)[169.0 - 213.0]$

$083.7(\pm 38.8)[028.0 - 220.0]$

Segmentação por Limiarização - Ótima

Resultados



221.6(± 25.2)[151.0 – 255.0]

094.3(± 28.9)[032.0 – 150.0]

- Desvantagens:
 - Convergência do EM depende da inicialização

- Desvantagens:
 - Convergência do EM depende da inicialização
 - Método iterativo - razoavelmente complexo computacionalmente para uma tarefa muito simples

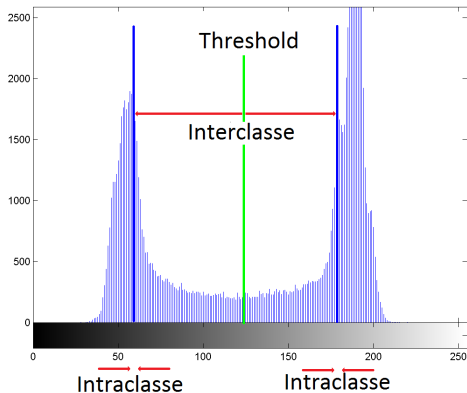
- Desvantagens:
 - Convergência do EM depende da inicialização
 - Método iterativo - razoavelmente complexo computacionalmente para uma tarefa muito simples
 - Suposição de “compartamento” bimodal do histograma

Segmentação por Limiarização - Otsu

- Otsu (1979) propõe uma técnica não paramétrica (não estima parâmetros do modelo) e não supervisionada para a seleção automática do limiar visando a segmentação da imagem.

Segmentação por Limiarização - Otsu

- Otsu (1979) propõe uma técnica não paramétrica (não estima parâmetros do modelo) e não supervisionada para a seleção automática do limiar visando a segmentação da imagem.
- Otimizar a separabilidade das classes (objeto e fundo):
 - maximizar variabilidade interclasse
 - minimizar variabilidade intraclasse



Segmentação por Limiarização - Otsu

- A partir da ideia, três critérios de discriminação são propostos:
 - $\lambda = \sigma_B^2/\sigma_W^2$, $\kappa = \sigma_{Tot}^2/\sigma_W^2$ e $\eta = \sigma_B^2/\sigma_{Tot}^2$

onde

$$\begin{aligned}\sigma_W^2 &= \omega_1\sigma_1^2 + \omega_2\sigma_2^2 \\ \sigma_B^2 &= \omega_1(\mu_1 - \mu_{Tot})^2 + \omega_2(\mu_2 - \mu_{Tot})^2 \\ &= \omega_1\omega_2(\mu_1 - \mu_2)^2 \\ \sigma_{Tot}^2 &= \sum_{l=0}^{L-1} (l - \mu_{Tot})^2 p(l)\end{aligned}$$

$$\begin{aligned}\text{e } \omega_1(T) &= \sum_{l=0}^T h(l), \quad \omega_2(T) = \sum_{l=T+1}^{L-1} h(l), \quad \omega_1(T) + \omega_2(T) = 1 \\ \mu_1(T) &= \sum_{l=0}^T lh(l)/\omega_1(T), \quad \mu_2(T) = \sum_{l=0}^T lh(l)/\omega_2(T), \\ &\omega_1\mu_1 + \omega_2\mu_2 = \mu_{Tot}\end{aligned}$$

- Todavia, os três critérios são equivalentes, *i.e.*

$$\kappa = \lambda + 1, \eta = \lambda/(\lambda + 1)$$

pois

$$\sigma_W^2(T) + \sigma_B^2(T) = \sigma_{Tot}^2$$

para qualquer $T \in [0, L - 1]$.

- Como consequência, temos que σ_{Tot} é independente de T .

- E como η é a medida mais simples de ser calculada, o limiar ótimo de Otsu pode ser calculado maximizando $\sigma_B^2 T$:

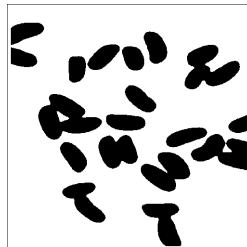
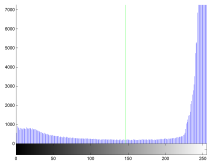
$$\begin{aligned}\sigma_B^2(T) &= \omega_1\omega_2(\mu_2 - \mu_1)^2 \\ &= \omega(T)(1 - \omega(T)) \left[\frac{\mu_{Tot} - \mu(T)}{1 - \omega(T)} - \frac{\mu(T)}{\omega(T)} \right] \\ &= \frac{[\mu_{Tot}\omega(T) - \mu(T)]^2}{\omega(T)[1 - \omega(T)]}\end{aligned}$$

- E o limiar ótimo T^* é

$$\sigma_2^B(T^*) = \max_{0 \leq T \leq L-1} \sigma_B^2(T)$$

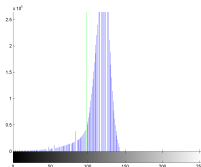
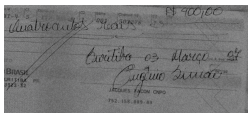
Segmentação por Limiarização - Otsu

- Resultados



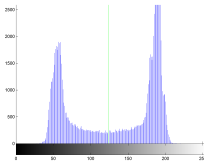
Segmentação por Limiarização - Otsu

- Resultados



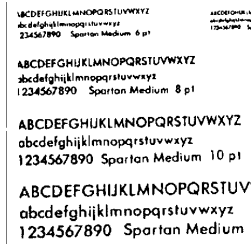
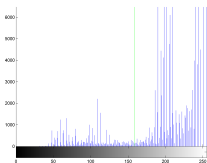
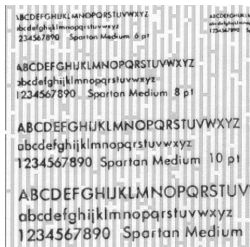
Segmentação por Limiarização - Otsu

- Resultados



Segmentação por Limiarização - Otsu

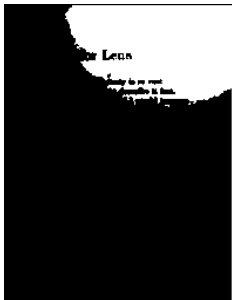
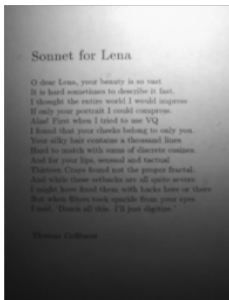
Resultados



- Observe que as técnicas descritas anteriormente são **globais**

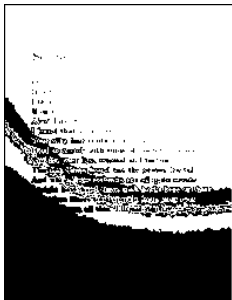
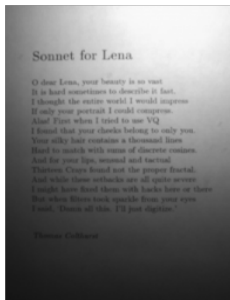
Segmentação por Limiarização - Adaptativa

- Observe que as técnicas descritas anteriormente são **globais**
- Limitação de técnica global.



Segmentação por Limiarização - Adaptativa

- Observe que as técnicas descritas anteriormente são **globais**
- Limitação de técnica global.



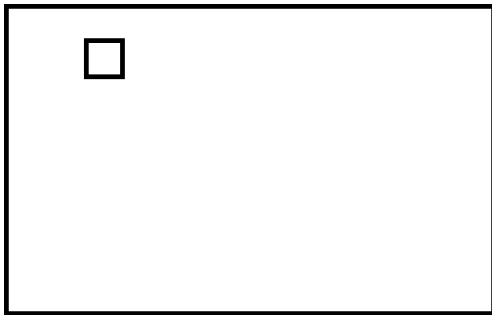
Segmentação por Limiarização - Adaptativa/Local

- A imagem passa a ser considerada como composta de subcomponentes e um processo de limiarização é aplicado a cada uma dessas janelas ($N \times M$ pixels) tornando-o adaptativo e local.



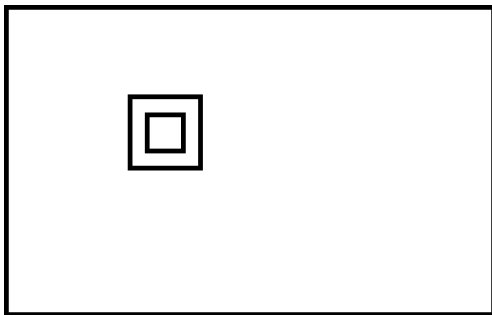
Segmentação por Limiarização - Adaptativa/Local

- A imagem passa a ser considerada como composta de subcomponentes e um processo de limiarização é aplicado a cada uma dessas janelas ($N \times M$ pixels) tornando-o adaptativo e local.



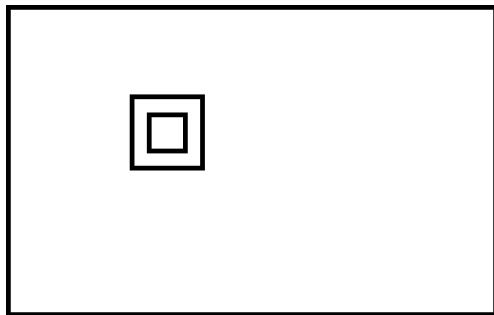
Segmentação por Limiarização - Adaptativa/Local

- A imagem passa a ser considerada como composta de subcomponentes e um processo de limiarização é aplicado a cada uma dessas janelas ($N \times M$ pixels) tornando-o adaptativo e local.



Segmentação por Limiarização - Adaptativa/Local

- A imagem passa a ser considerada como composta de subcomponentes e um processo de limiarização é aplicado a cada uma dessas janelas ($N \times M$ pixels) tornando-o adaptativo e local.



- Existem vários métodos na literatura. Os mais conhecidos são os propostos por [Niblack, 1986] e [Sauvola & Pietikäinen, 2000].

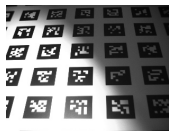
- A ideia de [Niblack, 1986] foi a de estabelecer uma função que levasse em consideração o contraste local da janela (de $N \times M$ pixels) usando medidas estatísticas de primeira e segunda ordem, *i.e.*,

$$T(p) = \mu(p) + k \times \sigma(p)$$

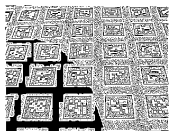
onde $\mu(p)$ e $\sigma(p)$ são a média e o desvio padrão dos pixels dentro da janela considerada (ou da janela centrada em (p)) e k é um parâmetro definido pelo usuário (*default* 0,5).

Segmentação por Limiarização - [Niblack, 1986]

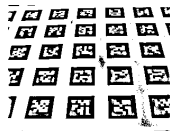
- Resultados



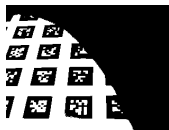
310×240



$N = M = 3$



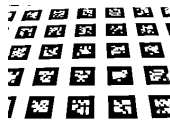
$N = M = 15$



Otsu



$N = M = 21$



$N = M = 31$

Segmentação por Limiarização - [Sauvola & Pietikäinen, 2000]

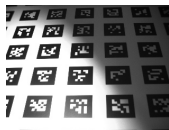
- [Sauvola & Pietikäinen, 2000] propuseram uma nova formulação (pequena modificação) em que o desvio padrão ($\sigma(p)$) pondera diretamente o valor médio da janela ($\mu(p)$) em consideração, *i.e.*,

$$T(p) = \text{round} \left[\mu(p) \left(1 + k \left(\frac{\sigma(p)}{R} - 1 \right) \right) \right]$$

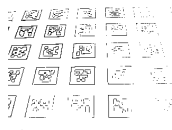
onde R está relacionado com a faixa dinâmica do desvio padrão (*default* 128).

Segmentação por Limiarização - [Sauvola & Pietikäinen, 2000]

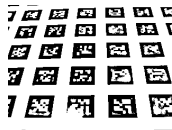
- Resultados



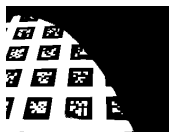
310 × 240



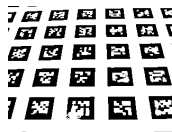
$N = M = 3$



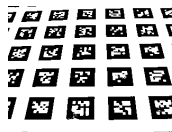
$N = M = 15$



Otsu



$N = M = 21$



$N = M = 31$

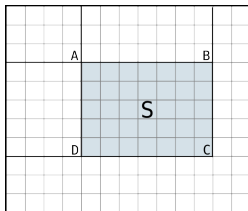
Segmentação por Limiarização - Local/Eficiência

- *Integral image* - evitar o recálculo de somatórios - introduzido por [Viola & Jones, 2001]

1	2	2	4	1
3	4	1	5	2
2	3	3	2	4
4	1	5	4	6
6	3	2	1	3

1	3	5	9	10
4	10	13	22	25
6	15	21	32	39
10	20	31	46	59
16	29	42	58	74

imagem & imagem integral



$$S = C - B - D + A$$

Segmentação por Limiarização - *Moving average*

- Uma outra abordagem local/adaptativa de limiarização, **menos cara** computacionalmente, é aquela baseada em médias em movimento (*moving average*) ao longo de vizinhança linear proposta por [Wellner, 1993], *i.e.*

$$\begin{aligned}\mu(p+1) &= \frac{1}{n} \sum_{i=p+2-n}^{p+1} I(i) \\ &= \mu(p) + \frac{1}{n}(I(p+1) - I(p-n))\end{aligned}$$

em que n define o tamanho da janela.

Segmentação por Limiarização - *Moving average*

- Uma outra abordagem local/adaptativa de limiarização, **menos cara** computacionalmente, é aquela baseada em médias em movimento (*moving average*) ao longo de vizinhança linear proposta por [Wellner, 1993], *i.e.*

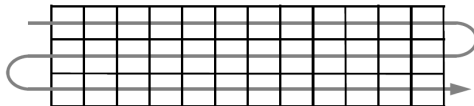
$$\begin{aligned}\mu(p+1) &= \frac{1}{n} \sum_{i=p+2-n}^{p+1} I(i) \\ &= \mu(p) + \frac{1}{n}(I(p+1) - I(p-n))\end{aligned}$$

em que n define o tamanho da janela.

- A janela não pode conter/representar uma região de fundo ou objeto apenas.
- *rule of thumb*: n é 5 vezes o tamanho do objeto a ser segmentado.

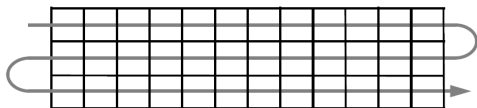
Segmentação por Limiarização - *Moving average*

- A forma em que a janela é deslocada na imagem tem um papel importante no resultado.



Segmentação por Limiarização - *Moving average*

- A forma em que a janela é deslocada na imagem tem um papel importante no resultado.



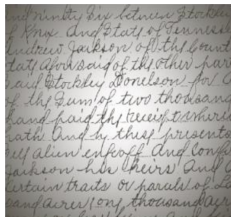
- Em cada pixel $I(p)$, pode-se escolher o limiar com base na média $\mu(p)$ da janela local

$$T(p) = \begin{cases} 1 & \text{se } I(p) > k\mu(p) \\ 0 & \text{caso contrário} \end{cases}$$

em que k é um parâmetro definido pelo usuário (*default* 0,5).

Segmentação por Limiarização - *Moving average*

- Resultados

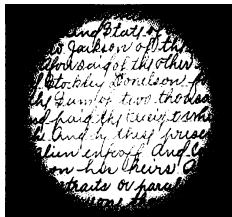


811 × 757

...ninty six between Stokley
Kings and Stats of Gemini
Andrew Jackson of the bond
stat of the said of the other part
and Stokley Donelson for a
of the sum of two thousand
and paid the receipt to their
with and by their presents
of alim enkooff And Conf
Jackson-his theirs And C
ertain traits or parals of L
and aere! on, thousand aere

Niblack

$N = 40$ e $k = 0.5$



Otsu

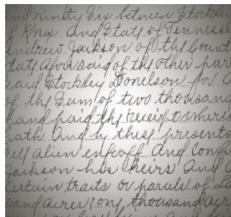
...ninty six between Stokley
Kings and Stats of Gemini
Andrew Jackson of the bond
stat of the said of the other part
and Stokley Donelson for a
of the sum of two thousand
and paid the receipt to their
with and by their presents
of alim enkooff And Conf
Jackson-his theirs And C
ertain traits or parals of L
and aere! on, thousand aere

Wellner

$n = 40$ e $k = 0.95$

Segmentação por Limiarização - *Moving average*

- Resultados



811 × 757

met ninety six between Stockley
Kings And Stat of Tenness
Andrew Jackson of the bound
tats of the said of the other part
and Stockley Donelson for a
of the sum of two thousand
and paid the receipt to their
with and by their presents
of alien enfeoff And Conf
Jackson his theirs And a
ertain traits or parcels of la
and acres one thousand

Niblack

$N = 40$ e $k = 0.5$



Otsu

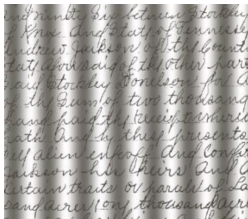
met ninety six between Stockley
Kings And Stat of Tenness
Andrew Jackson of the bound
tats of the said of the other part
and Stockley Donelson for a
of the sum of two thousand
and paid the receipt to their
with and by their presents
of alien enfeoff And Conf
Jackson his theirs And a
ertain traits or parcels of la
and acres one thousand

Wellner

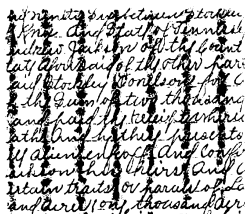
$n = 40$ e $k = 0.95$

Segmentação por Limiarização - *Moving average*

- Resultados



806 × 710

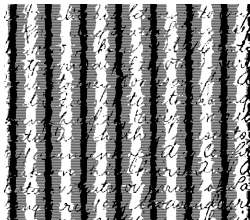


Niblack

$N = 40$ e $k = 0.5$



Otsu

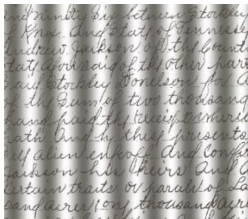


Wellner

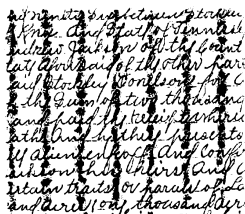
$n = 40$ e $k = 0.95$

Segmentação por Limiarização - *Moving average*

- Resultados



806 × 710

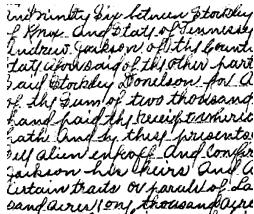


Niblack

$N = 40$ e $k = 0.5$



Otsu



Wellner

$n = 40$ e $k = 0.95$

Referências

[Otsu, 1979] OTSU, N.

A threshold selection method from grey-level histograms.

IEEE Transactions on Systems, Man and Cybernetics. 1979, vol. 9, no. 1, pp. 41–47.

[Niblack, 1986] NIBLACK, W.

An Introduction to Image Processing.

Prentice Hall. 1986, pp 115–116.

[Sauvola & Pietikäinen, 2000] SAUVOLA, J. and PIETIKÄINEN, M.

Adaptive document image binarization.

Pattern Recognition. 2000, vol. 33, pp. 225–236.

[Viola & Jones, 2001] VIOLA, P. & JONES, M.

Rapid object detection using a boosted cascade of simple features

IEEE Int. Conf. on Computer Vision and Pattern Recognition, 2001, vol. 1, pp 511-518.

[Wellner, 1993] WELLNER, P. D.

Adaptive Thresholding for the DigitalDesk.

Rank Xerox Research Centre, 1993, Tech. Rep. EPC-1993-100, pp. 1–17.