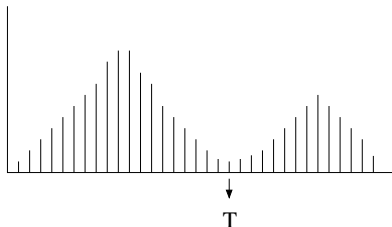


Limiarização

- A *limiarização* é uma das técnicas mais simples de segmentação e consiste na classificação dos pixels de uma imagem de acordo com a especificação de um ou mais limiares.
- Seja o histograma de níveis de cinza, mostrado na figura a seguir, correspondente a uma imagem $f(x, y)$, de maneira que os pixels dos objetos e os do fundo tenham seus níveis de cinza separados em dois grupos dominantes.



- Uma maneira óbvia de extrair os objetos do fundo é por meio da seleção de um limiar T que separe os dois grupos.

Limiarização

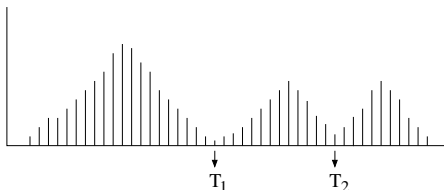
- Então, cada ponto (x, y) tal que $f(x, y) > T$ é denominado um ponto do objeto; caso contrário, o ponto é denominado um ponto do fundo.
- A imagem limiarizada $g(x, y)$ pode ser definida como

$$g(x, y) = \begin{cases} 0, & \text{se } f(x, y) \leq T \\ 1, & \text{se } f(x, y) > T \end{cases} \quad (34)$$

- Portanto, os pixels rotulados como 1 (ou qualquer outro nível de cinza conveniente) correspondem aos objetos, enquanto que aqueles rotulados como 0 correspondem ao fundo.
- Nesse caso, a limiarização é denominada *binarização*, pois a imagem resultante possui apenas dois valores de intensidade, 0 (preto) ou 1 (branco).

Limiarização

- A figura a seguir mostra um caso ligeiramente mais geral dessa abordagem, em que três grupos dominantes caracterizam o histograma da imagem (por exemplo, dois tipos de objetos iluminados sobre um fundo escuro).



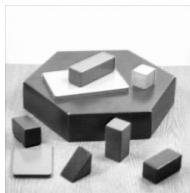
- A limiarização da imagem pode ser definida como

$$g(x, y) = \begin{cases} l_1, & \text{se } f(x, y) \leq T_1 \\ l_2, & \text{se } T_1 < f(x, y) \leq T_2 \\ l_3, & \text{se } f(x, y) > T_2 \end{cases} \quad (35)$$

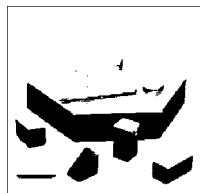
tal que, para cada intervalo, é especificado um nível de cinza correspondente.

Limiarização

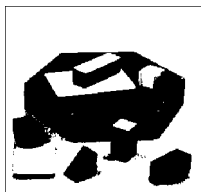
- A seleção correta do valor de limiar é crucial para que o processo de segmentação baseada na limiarização produza bons resultados.
- A figura a seguir ilustra a segmentação de uma imagem com o uso de diferentes valores de limiar.



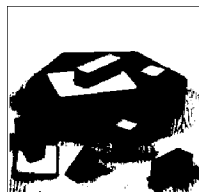
(a) imagem original



(b) $T = 108$



(c) $T = 179$



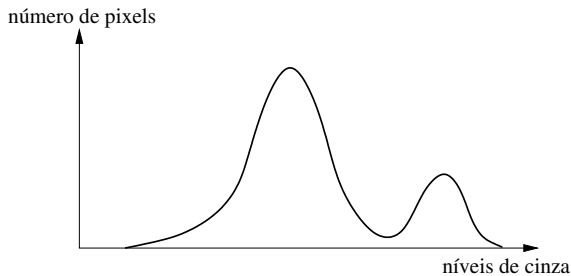
(d) $T = 213$

Limiarização

- *Limiarização global*: a utilização de um único valor de limiar para segmentar toda a imagem não é em geral adequada, pois as imagens podem conter variações nos níveis de cinza dos objetos e do fundo por causa da iluminação não-uniforme, ruído, parâmetros do dispositivo de aquisição não-uniformes ou outros fatores.
- *Limiarização local*: valores de limiar podem variar sobre a imagem como uma função de suas características locais.

Limiarização Global

- Como mencionado anteriormente, a maneira mais direta de selecionar um valor global de limiar é a partir da distribuição das intensidades dos pixels na imagem.
- Para uma imagem consistindo de objetos com níveis de cinza aproximadamente iguais, diferindo do nível de cinza do fundo, o histograma possuirá dois picos distintos, um formado pelos pixels dos objetos e outro pelos pixels do fundo.
- Um histograma com tal distribuição é chamado *bimodal*.



(a) histograma bimodal

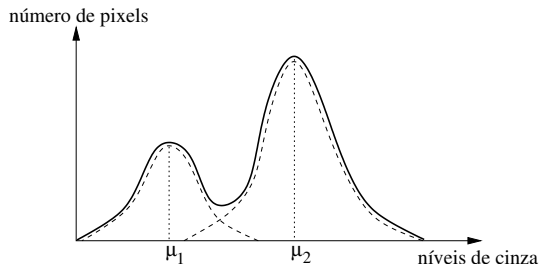
- Intuitivamente, o valor do limiar corresponde ao nível de cinza que melhor separa os dois picos.
- Vários métodos têm sido propostos para encontrar os pontos de máximo (picos) e, então, selecionar o valor de limiar como o ponto de mínimo (vale) entre eles.
- Há várias desvantagens em selecionar o limiar global por meio da distribuição das intensidades dos pixels.
 - ▶ Nem sempre as intensidades dos objetos e do fundo da imagem são bem distintas em virtude, por exemplo, da ocorrência de baixo contraste ou ruído.
 - ▶ Além disso, a determinação de picos e vales não é um problema trivial. Por exemplo, muitos pontos de mínimo ou máximo locais podem existir na imagem.

- Uma abordagem proposta por Mason et al. (1975) consiste em modificar o histograma de acordo com uma função do valor de contraste local.
- O histograma das intensidades dos pixels é ponderado pelo inverso dos seus valores de gradiente.
- As intensidades de regiões de pixels com valores de gradiente baixos são mais influenciadas do que as intensidades com valores de gradiente mais altos, desse modo, os picos do histograma causados por tais regiões podem ser removidos, facilitando a determinação do limiar entre eles.

- Uma técnica similar é ponderar as intensidades de pixels com valores de gradiente e, então, escolher o limiar no pico do histograma, correspondendo à transição entre as regiões.
- Esse método requer a detecção de um ponto de máximo, o qual é normalmente mais fácil do que a detecção de vales entre dois picos. Outra abordagem utiliza o operador Laplaciano, que é um operador de diferenciação de segunda ordem.
- O Laplaciano resulta em valores altos para as transições das regiões em bordas.
- Caso alguma propriedade da imagem a ser segmentada seja previamente conhecida, a seleção do limiar pode ser simplificada, já que o valor do limiar é escolhido para assegurar que essa propriedade seja satisfeita.
- Conhecendo-se a percentagem p de pixels do objeto na imagem, essa informação pode ser utilizada para escolher o limiar T , tal que o valor p da área da imagem tenha níveis de cinza menores que T e o restante dos níveis de cinza maiores que T .

Limiarização Global

- Algumas técnicas determinam o valor de limiar por meio da otimização de certas medidas de separação entre as classes de objetos na imagem.
- As distribuições de intensidade dos pixels podem ser modeladas como funções de densidade de probabilidades, por exemplo, a função Gaussiana.



- Para cada distribuição, a média e a variância são calculadas, e o limiar é escolhido tal que o erro definido por essas funções de probabilidade seja minimizado.

- Métodos para escolha de limiar que resultam em segmentação com erro mínimo entre os picos de duas ou mais funções de probabilidade com distribuição normal foram propostos por vários autores.
- Uma dificuldade com esses métodos é a estimativa de parâmetros das distribuições normais, bem como a incerteza inerente ao se considerar a função de probabilidade como uma distribuição normal.
- Um método para selecionar iterativamente o valor de limiar baseado na combinação de modelos de duas distribuições Gaussianas foi proposto por Ridler e Calvard (1978).
 - ▶ Após uma estimativa inicial para o limiar, o método refina o cálculo do limiar baseado na média das distribuições dos níveis de cinza do objeto (μ_1) e do fundo (μ_2), até que esses valores não sofram alterações em iterações sucessivas.
 - ▶ Na prática, as iterações terminam quando a diferença $|T_{i+1} - T_i|$ entre os limiares T nas iterações i e $i + 1$ torna-se suficientemente pequena.

- O algoritmo a seguir apresenta os principais passos do método de Ridler e Calvard (1978).

Algoritmo 5 Limiarização global iterativa

- 1: Estimar um valor inicial, $T = T_0$, para o limiar, por exemplo, igual à intensidade média da imagem.
 - 2: Particionar a imagem em duas regiões, R_1 e R_2 , com o limiar T , tal que todos os pixels com intensidade menor que T sejam atribuídos à região R_1 e todos os outros sejam atribuídos à região R_2 .
 - 3: Na iteração i , calcular valores de intensidade média μ_1 e μ_2 associados às partições R_1 e R_2 .
 - 4: Calcular um novo limiar como $T = T_{i+1} = \frac{\mu_1^i + \mu_2^i}{2}$.
 - 5: Repetir os passos (2) a (4) até que o limiar T não sofra alteração significativa em iterações sucessivas, ou seja, $T_{i+1} = T_i$.
-

- Trussel (1979), baseando-se no algoritmo de Ridler e Calvard, propôs que o novo valor de limiar T_{i+1} na iteração $i + 1$ fosse calculado como

$$T_{i+1} = \frac{\mu_1^i + \mu_2^i}{2} = \frac{1}{2} \left(\frac{\sum_{i=0}^{T_i} i p_i}{\sum_{i=0}^{T_i} p_i} + \frac{\sum_{i=T_{i+1}}^{L-1} i p_i}{\sum_{i=T_{i+1}}^{L-1} p_i} \right)$$

em que μ_1^i e μ_2^i são as intensidades médias dos níveis de cinza abaixo e acima, respectivamente, do limiar T_i e p_i é a probabilidade de se encontrar um número de pixel n_i com nível de cinza i em uma imagem com n pixels, ou seja, $p_i = n_i/n$.

- Lloyd (1985) propôs uma modificação no método de Ridler e Calvard para seleção iterativa do limiar.
- O passo (4) do Algoritmo 5 é substituído por

$$T = \frac{1}{2} (\mu_1 + \mu_2) + \frac{\sigma^2}{(\mu_1 - \mu_2)} \log \frac{n_2}{n_1}$$

em que μ_1 e μ_2 são os níveis de cinza médios associados aos números de pixels n_1 e n_2 das regiões R_1 e R_2 , respectivamente.

- A variância σ^2 pode ser estimada a partir dos dados, em particular, $\sigma^2 = \sigma_1^2 = \sigma_2^2$.

- O método de Otsu (1979) é baseado no fato de que o histograma de uma imagem é composto de duas classes, os pixels do objeto e os pixels do fundo.
 - ▶ Cada classe possui suas próprias características, ou seja, sua média e desvio padrão.
 - ▶ A variância σ_T^2 e a média μ_T globais da imagem são calculadas.
 - ▶ O método procura maximizar a razão $\eta(T)$ da variância entre as classes σ_B^2 em relação à variância total, para todos os valores possíveis de limiar T , sendo

$$\eta(T) = \frac{\sigma_B^2}{\sigma_T^2} \quad (36)$$

tal que

$$\begin{aligned} \sigma_T^2 &= \sum_{i=0}^{L-1} (i - \mu_T)^2 p_i & \mu_T &= \sum_{i=0}^{L-1} i p_i & \sigma_B^2 &= \omega_1 \omega_2 (\mu_1 - \mu_2)^2 & \omega_1 &= \sum_{i=0}^T p_i \\ \omega_2 &= 1 - \omega_1 & \mu_1 &= \frac{\mu_T}{\omega_1} & \mu_2 &= \frac{\mu_T - \mu_S}{\omega_2} & \mu_S &= \sum_{i=0}^T i p_i & \sum_{i=0}^{L-1} p_i &= 1, & p_i &= \frac{n_i}{n} \end{aligned}$$

em que n_i é o número de pixels com nível de cinza i , n é o número total de pixels da imagem, p_i representa a probabilidade do nível de cinza i ser encontrado na imagem e L é o número de níveis de cinza da imagem.

- Conforme equação 36, a razão $\eta(T)$ é calculada para todos os valores possíveis de T .
- O limiar ótimo pode ser determinado como

$$T = \arg \max \eta(T)$$

- O método de Otsu possui bom desempenho em imagens com maior variância de intensidade.
- Uma desvantagem é que o método assume que o histograma da imagem seja bimodal.

- Um método proposto por Yanni e Horne (1994) assume que dois picos distintos g_{pico1} e g_{pico2} nos níveis de cinza são identificáveis na distribuição de probabilidade.
- Um ponto intermediário é inicialmente estabelecido como $\hat{g}^* = (g_{\text{max}} + g_{\text{min}})/2$, em que g_{min} é o nível de cinza mais baixo e g_{max} o nível de cinza mais alto diferente de zero.
- Esse ponto intermediário é atualizado pela média dos dois picos localizados à esquerda e à direita, ou seja, $\hat{g}^* = (g_{\text{pico1}} + g_{\text{pico2}})/2$.
- O limiar é então

$$T = (g_{\text{max}} - g_{\text{min}}) \sum_{g=\hat{g}^*}^{\hat{g}^*} p_g$$

- Um outro enfoque similar ao de Otsu (1979) para selecionar valores de limiar emprega a entropia.
- A maximização da entropia da imagem limiarizada é interpretada como um indicativo de transferência máxima de informação.
- No método de Pun (1980), procura-se a maximização da entropia da imagem, dividindo-se o seu histograma em duas distribuições de probabilidade representando as duas classes presentes na imagem, ou seja, fundo e objeto.

- As entropias $H_b(T)$ e $H_w(T)$ associadas aos pixels do fundo e dos objetos, respectivamente, são dadas por

$$H_b(T) = - \sum_{i=0}^T p_i \log p_i \qquad H_w(T) = - \sum_{i=T+1}^{L-1} p_i \log p_i$$

sendo $\sum_{i=0}^{L-1} p_i = 1$, $p_i = \frac{n_i}{n}$, tal que n_i é o número de pixels com nível de cinza i , p_i representa a probabilidade do nível de cinza i ser encontrado na imagem, n é o número de pixels da imagem e L é o número de níveis de cinza da imagem.

- A escala de níveis de cinza é analisada de forma a encontrar um valor de limiar T , tal que $T = H_b(T) + H_w(T)$ seja máximo, procurando-se uma boa separação entre o objeto e o fundo da imagem, ou seja

$$T = \arg \max [H_b(T) + H_w(T)] \qquad (37)$$

- Kapur et al. (1985) propõem que as entropias $H_b(T)$ e $H_w(T)$ sejam calculadas como

$$H_b(T) = - \sum_{i=0}^T \frac{p_i}{p_1(T)} \log \frac{p_i}{p_1(T)}$$
$$H_w(T) = - \sum_{i=T+1}^{L-1} \frac{p_i}{p_2(T)} \log \frac{p_i}{p_2(T)}$$

sendo

$$p_1(T) = \sum_{i=0}^T p_i \qquad p_2(T) = \sum_{i=T+1}^{L-1} p_i \qquad p_1(T) + p_2(T) = 1$$

- A partir das duas entropias $H_b(T)$ e $H_w(T)$, realiza-se um processo de seleção exaustiva do limiar na escala de cinza, com o objetivo de encontrar o valor que maximize a soma das entropias, como no método definido por Pun, dado pela equação 37.

- Tsai (1985) utilizou a preservação de momentos para obter um valor de limiar sem a necessidade de iteração.
- O momento de ordem k , denotado m_k , de uma imagem f com L níveis de cinza é definido como

$$m_k = \sum_{i=0}^{L-1} p_i i^k \quad k = 0, 1, 2, 3 \quad (38)$$

sendo $p_i = n_i/n$, tal que n_i é o número de pixels com nível de cinza i , n é o número de pixels da imagem, L é o número de níveis de cinza da imagem e $m_0 = 1$.

- A imagem f pode ser vista como uma versão suavizada de uma imagem ideal g que apresenta histograma bimodal, cujos picos estão em h_1 e h_2 , $h_1 < h_2$.
- O método seleciona um limiar T tal que, se todas as intensidades abaixo do limiar na imagem f forem trocadas por h_1 e todas as intensidades acima do limiar forem trocadas por h_2 , então os três primeiros momentos de f serão preservados na imagem g .

- Considerando p_1 e p_2 as frações dos pixels que estão abaixo e acima do limiar na imagem f , então os três primeiros momentos da imagem binária g são dados por

$$m_k = p_1 h_1^k + p_2 h_2^k \quad k = 0, 1, 2, 3 \quad (39)$$

- Preservando os momentos e utilizando o fato de que $p_1 + p_2 = 1 = m_0$, um conjunto de quatro equações contendo os parâmetros p_1 , p_2 , h_1 e h_2 é definido como

$$p_1 h_1 + p_2 h_2 = m_1$$

$$p_1 h_1^2 + p_2 h_2^2 = m_2$$

$$p_1 h_1^3 + p_2 h_2^3 = m_3$$

$$p_1 + p_2 = 1$$

- No caso de um histograma bimodal, as equações são resolvidas por meio das expressões

$$h_1 = \frac{-c_1 - \sqrt{c_1^2 - 4c_0}}{2} \quad h_2 = \frac{-c_1 + \sqrt{c_1^2 - 4c_0}}{2} \quad p_1 = \frac{\begin{vmatrix} 1 & 1 \\ m_1 & h_2 \end{vmatrix}}{\begin{vmatrix} 1 & 1 \\ h_1 & h_2 \end{vmatrix}}$$

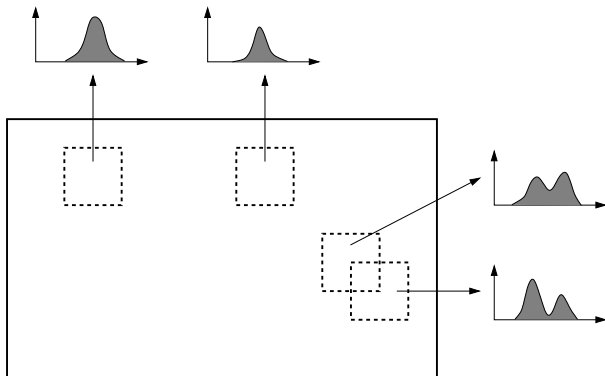
em que

$$c_0 = \frac{\begin{vmatrix} -m_2 & m_1 \\ -m_3 & m_2 \end{vmatrix}}{\begin{vmatrix} m_0 & m_1 \\ m_1 & m_2 \end{vmatrix}} \quad c_1 = \frac{\begin{vmatrix} m_0 & -m_2 \\ m_1 & -m_3 \end{vmatrix}}{\begin{vmatrix} m_0 & m_1 \\ m_1 & m_2 \end{vmatrix}}$$

- O valor do limiar T é escolhido como o nível de cinza mais próximo da fração p_1 no histograma da imagem f .

Limiarização Local

- Em casos nos quais a utilização de um único limiar para segmentar toda a imagem não produza bons resultados, a *limiarização local* pode ser uma alternativa mais apropriada.
- Uma forma comum de realizar a limiarização local é analisar as intensidades dos pixels em uma região da imagem para determinar limiares locais. Esse processo é ilustrado na figura a seguir.



Limiarização Local

- Algumas medidas estatísticas simples para calcular um limiar local são a média dos valores de intensidade em uma vizinhança local da imagem

$$T = \text{média}_v(p) \quad (40)$$

a mediana dos valores

$$T = \text{mediana}_v(p) \quad (41)$$

ou a média dos valores mínimo e máximo

$$T = \frac{\min_v(p) + \max_v(p)}{2} \quad (42)$$

em que v é uma vizinhança local ao ponto p na imagem.

- Um problema crítico encontrado nas técnicas de limiarização local é a escolha do tamanho da janela.
- O tamanho selecionado deve ser suficientemente grande para garantir uma boa estimativa do valor da média, entretanto, sem sofrer distorções por causa da não-uniformidade do fundo.

- Yasuda et al. (1980) propuseram um método que expande a escala de níveis de cinza da imagem, seguido de uma suavização não-linear que preserve bordas.
- A suavização consiste em substituir cada pixel pela média de seus oito vizinhos, dado que o intervalo do pixel (definido como a diferença entre os valores máximo e mínimo locais) está abaixo de um limiar T_1 .
- Um limiar adaptativo é aplicado tal que o valor de cada pixel seja atribuído como fundo se o intervalo local estiver abaixo de um limiar T_2 ou se o valor do pixel estiver acima de uma média local, ambos calculados sobre uma vizinhança de $n \times n$ pixels. Caso contrário, a escala de nível de cinza deve ser expandida convenientemente.
- Finalmente, a imagem é binarizada atribuindo-se um pixel como objeto se seu valor mínimo em uma vizinhança de 3×3 pixels estiver abaixo de T_3 ou se sua variância local estiver acima de T_4 .
- Kamel e Zhao (1989) sugerem os valores $T_1 = 50$, $n = 16$, $T_2 = 16$, $T_3 = 128$ e $T_4 = 35$.

- No método de Bernsen (1986), o limiar para cada pixel (x, y) é calculado como

$$T(x, y) = \frac{z_{\min} + z_{\max}}{2} \quad (43)$$

em que z_{\min} e z_{\max} são os valores de níveis de cinza mínimo e máximo em uma vizinhança de $n \times n$ pixels centrada em (x, y) .

Limiarização Local

- No método de Niblack (1986), o valor do limiar em um pixel (x, y) é baseado na média local e no desvio padrão calculados sobre uma vizinhança de tamanho $n \times n$ pixels, expresso como

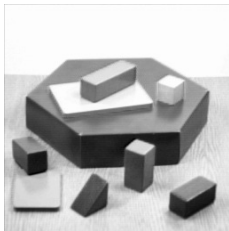
$$T(x, y) = \mu(x, y) + k \sigma(x, y) \quad (44)$$

em que $\mu(x, y)$ e $\sigma(x, y)$ são a média e o desvio padrão dos valores sobre a vizinhança, respectivamente.

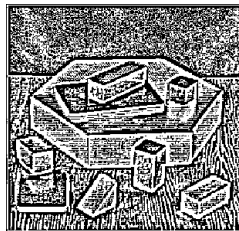
- O valor para a constante k é ajustado conforme o tipo de aplicação ou característica da imagem sendo processada.
- A escolha desse valor proporciona um equilíbrio entre supressão de ruído e preservação de detalhes, além de atribuir menor ou maior presença de contorno do objeto na imagem limiarizada.
- O tamanho da vizinhança deveria ser suficientemente pequeno para preservar detalhes locais, mas, ao mesmo tempo, suficientemente grande para suprimir ruído.
- Trier e Jain (1995) sugerem que resultados satisfatórios podem ser obtidos com uma janela de tamanho $n = 15$ e $k = -0.2$.

Limiarização Local

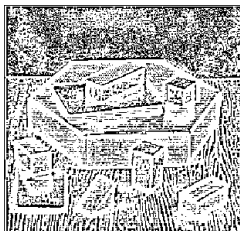
- Exemplo de imagem limiarizada pelo método de Niblack com diferentes valores de ajuste, para um mesmo tamanho de janela.



(a) imagem original



(b) $n = 25, k = 0.1$



(c) $n = 25, k = 0.5$



(d) $n = 25, k = 0.8$

- O método de Sauvola e Pietaksinen (2000) procura melhorar os resultados do método de Niblack, particularmente para imagens de documentos apresentando má iluminação.
- O limiar se adapta de acordo com a média local e o desvio padrão sobre uma janela de tamanho $n \times n$ pixels.
- O limiar em um pixel (x, y) é calculado como

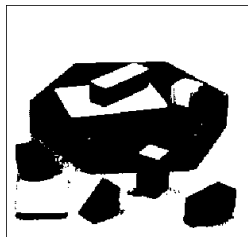
$$T(x, y) = \mu(x, y) \left[1 + k \left(\frac{\sigma(x, y)}{R} - 1 \right) \right] \quad (45)$$

em que $\mu(x, y)$ e $\sigma(x, y)$ são definidos como em Niblack.

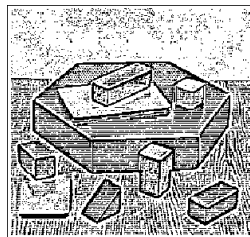
- Sauvola e Pietaksinen sugerem valores de $k = 0.5$ e $R = 128$ para as imagens utilizadas em seus experimentos.
- Dessa forma, a contribuição do desvio padrão torna-se adaptativa.

Limiarização Global × Local

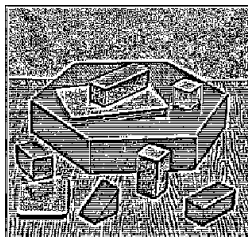
- Comparação entre diferentes técnicas de limiarização.



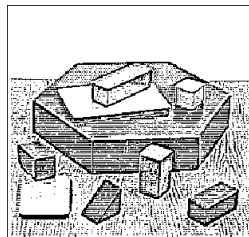
(a) global ($T = 179$)



(b) Bernsen ($n = 9$)



(c) Niblack ($n = 9$, $k = 0.01$)



(d) Sauvola e Pietakinen ($n = 9$,
 $k = 0.5$, $R = 128$)