

TESE  
DE MESTRADO

TÉCNICAS DE ENTROPIA EM PROCESSAMENTO DE IMAGENS

**Israel Andrade Esquef**

MESTRADO EM INSTRUMENTAÇÃO CIENTÍFICA  
CENTRO BRASILEIRO DE PESQUISAS FÍSICAS  
RIO DE JANEIRO, DEZEMBRO DE 2002



# Agradecimentos

Em primeiro lugar, gostaria de expressar o meu agradecimento aos meus orientadores Márcio Portes de Albuquerque e Marcelo Portes de Albuquerque por todos os ensinamentos e críticas, pela confiança que depositaram em mim e pela amizade que demonstraram durante estes anos de convivência.

Agradeço a minha família, mãe, pai, irmãos e avós pelo incentivo, suporte e apoio durante toda a minha vida.

A todo o pessoal da CAT, demonstro a minha gratidão por todo este tempo de agradável convívio e por todo apoio, direto ou indireto, que me foi dado. Aprendi muito com todos que compõem esta super equipe.

Um agradecimento especial aos meus amigos da primeira turma do Mestrado em Instrumentação Científica, Alex Mello, Eduardo Shigueo e Salvador Belmonte. A união e a competência desta turma foram muito importantes que este novo mestrado se consolide no CBPF.

Aos professores e idealizadores deste novo mestrado, em especial ao Prof. Anibal Caride, agradeço pelo empenho, dedicação e ensinamentos transmitidos durante as disciplinas e discussões extra classe. Agradeço ainda ao Prof. Constantino Tsallis e ao Fúlvio Baldovin pelas críticas construtivas e sugestões ao trabalho.

Agradeço aos muitos amigos que fiz dentro do CBPF durante estes anos: Roberto Sarthour, Jorge L. Gonzalez, Armando Takeuchi, Alex Rosa, Angelo, Ariel, Clayton, Deyse, João, Flávio, Aline, . . . enfim, todos com quem convivi desde que comecei como aluno de IC neste centro. Agradeço também ao amigo Luiz Wagner Biscainho.

Devo agradecer ainda a Juliana, André, Munir, Kátia, Flávio, Fábio, Rufino, Gustavo e aos muitos amigos em Campos que souberam entender os momentos em que precisei estar ausente.

Agradeço a todos que compõem o LCFIS na UENF, em especial ao Prof. Helion Vargas, pelo crédito e apoio incondicional que me foi dado para a conclusão desta tese.

Por fim, agradeço a Centro Brasileiro de Pesquisas Físicas (CBPF) pelo suporte e apoio para o desenvolvimento deste trabalho.



# Resumo

Este estudo concerne a utilização de métodos entrópicos aplicados como técnicas de processamento digital de imagens. O conceito de entropia, e mais recentemente o de entropia generalizada, foi aplicado como base para a segmentação de imagens e reconhecimento de padrões. A importância da análise quantitativa de imagens em diversas áreas da ciência tem motivado o desenvolvimento contínuo de novas técnicas de processamento de sinais. Este trabalho propõe novos métodos de segmentação de imagens revendo e ampliando as técnicas clássicas e também uma nova técnica de reconhecimento de padrões baseada no cálculo da entropia das formas presentes na imagem. Uma análise detalhada dos métodos entrópicos, com resultados em casos simulados e sistemas reais é apresentada e mostra que as técnicas de segmentação propostas são mais robustas e eficientes. Em termos do reconhecimento de padrões, testes mostram que a introdução da entropia relativa generalizada permite uma melhora na eficiência do classificador, através de um ajuste fino do parâmetro de não extensividade  $q$ .



# Abstract

This thesis concerns the use of entropy methods applied to digital image processing. The concept of entropy, and more recently the generalized entropy, was applied for image segmentation and pattern recognition. The importance of quantitative image analysis in many areas of material science has motivated the development of these techniques. In this work we present new methods for image segmentation based on the entropy of the image regions and a new technique of shape recognition based on entropy of each pattern in the image. A detailed analysis is presented for optimum image thresholding for simulated cases and real images to support the validity of the concepts used. Essentially this methods consist of viewing the image as two sets corresponding to two independent classes. Finally, for pattern recognition, the total efficiency was improved using the generalized cross-entropy foundations, where the parameter  $q$  is used as an adjustable parameter of the classifier.





# Sumário

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Imagem e Entropia . . . . .	3
1.2	O Documento . . . . .	5
<b>2</b>	<b>Fundamentos de Teoria da Informação</b>	<b>7</b>
2.1	Informação, Incerteza e Entropia . . . . .	8
2.1.1	Fonte Discreta de Informação . . . . .	8
2.1.2	Medida de Informação . . . . .	9
2.1.3	Entropia de uma Fonte de Informação . . . . .	10
2.2	Entropia Relativa . . . . .	12
2.3	Entropia Generalizada em Sistemas de Informação . . . . .	13
2.3.1	Conceitos Fundamentais da Entropia Não Extensiva . . . . .	13
2.3.2	Entropia de Shannon Generalizada . . . . .	15
2.3.3	Entropia Relativa Generalizada . . . . .	16
<b>3</b>	<b>Processamento Digital de Imagens</b>	<b>19</b>
3.1	Representação da Imagem Digital . . . . .	20
3.2	Etapas do Processamento de Imagens . . . . .	21
3.2.1	Aquisição de Imagens Digitais . . . . .	22
3.2.2	Técnicas de Pré-processamento . . . . .	23
3.2.3	Segmentação . . . . .	24
3.2.4	Pós-processamento . . . . .	27

3.2.5	Extração de Atributos . . . . .	29
3.2.6	Classificação e Reconhecimento . . . . .	31
<b>4</b>	<b>Segmentação Entrópica de Imagens</b>	<b>34</b>
4.1	Principais Abordagens . . . . .	34
4.2	Métodos Estatísticos . . . . .	36
4.2.1	Método de Binarização por Soma de Entropias . . . . .	37
4.2.2	Método de Binarização por Entropia Relativa . . . . .	41
4.3	Método Detecção de Bordas utilizando	
	Operador Entropia . . . . .	45
4.3.1	Definições Teóricas do Método . . . . .	46
<b>5</b>	<b>Segmentação de Imagens - Aplicações utilizando Entropia</b>	
	<b>Generalizada</b>	<b>49</b>
5.1	Imagens e Não Extensividade . . . . .	49
5.2	Método de Binarização Utilizando Entropia Não Extensiva . . . . .	52
5.2.1	Aplicações do Método e Discussão de Resultados . . . . .	54
5.2.2	Conclusões . . . . .	69
5.3	Método de Binarização Utilizando Entropia	
	Relativa Generalizada . . . . .	71
5.3.1	Aplicações do Método e Discussão de Resultados . . . . .	76
5.3.2	Resultados na Imagem . . . . .	77
5.3.3	Conclusões . . . . .	79
<b>6</b>	<b>Reconhecimento de Padrões Utilizando Entropia</b>	<b>82</b>
6.1	Representação do Objeto . . . . .	83
6.2	Método de Classificação . . . . .	84
6.3	Descrição de Similaridade . . . . .	85
6.4	Aplicação do Método . . . . .	87
6.5	Considerações sobre os resultados . . . . .	89
6.5.1	Considerações na representação pelo Histograma . . . . .	93

6.6	Classificador por entropia não extensiva . . . . .	95
6.6.1	A região ótima de $q$ – dois jogos de probabilidades. . . . .	100
6.7	Conclusão . . . . .	101
<b>7</b>	<b>Conclusão</b>	<b>103</b>
<b>A</b>	<b>Demonstrações Matemáticas</b>	<b>106</b>
<b>B</b>	<b>Métodos de Segmentação de Imagens</b>	<b>115</b>
<b>C</b>	<b>Trabalho Publicado</b>	<b>118</b>
	<b>Bibliografia</b>	<b>121</b>

# Lista de Figuras

2.1	<i>Informação Própria e Entropia da fonte binária como função da probabilidade.</i>	11
3.1	<i>Imagem monocromática “Goldhill” com destaque para uma região de <math>17 \times 17</math> pixels</i>	20
3.2	<i>Ilustração de tipos de conectividade dos pixels vizinhos ao pixel central <math>i_0</math>. Conectividade <math>B_8</math> apresenta 8 vizinhos, sendo 4 de bordas e 4 diagonais. Conectividade <math>B_4</math> apresenta apenas os pixels de borda.</i>	21
3.3	<i>Etapas de um sistema de PDI.</i>	22
3.4	<i>Exemplo de um pre-processamento simples: (A) Imagem original corrompida com ruído gaussiano, (B) Imagem após a aplicação de um filtro mediana para redução do ruído, e (C) Imagem final, após a aplicação de um filtro passa-altas para realce dos contornos.</i>	23
3.5	<i>Exemplo de imagens com histogramas diferenciados. (A) Imagem de baixo contraste e seu histograma de luminância. (B) Imagem de alto contraste e seu histograma de luminância.</i>	25
3.6	<i>Exemplo de imagens com histogramas bimodal e multimodal. A imagem (A) possui um histograma bimodal típico, com a presença de dois picos bem definidos. A imagem (B) possui um histograma do tipo multimodal, apresentando mais de duas regiões de concentração dos pixels da imagem.</i>	26
3.7	<i>Duas abordagens para segmentação. (A) Imagem original em níveis de cinza. (B) Imagem segmentada através de uma binarização. (C) Imagem segmentada por detecção de bordas.</i>	27

3.8	<i>Operação de erosão (a) e dilatação (b) aplicada a um elemento estruturante B. Obtemos a forma Y. <math>\mathbf{x}_1</math> é um ponto de X que também pertence a Y. Em (a) <math>\mathbf{x}_2</math> pertence a X, mas não a Y devido a operação de erosão. Por outro lado, em (b) <math>\mathbf{x}_2</math> passa a pertencer a Y devido a dilatação. <math>\mathbf{x}_3</math> não pertence a ambos X e Y. . . . .</i>	28
3.9	<i>Imagem ‘Labelizada’: (a) Imagem original composta por regiões contíguas de pixels. (b) Imagem final após o processo de rotulação. As cores são utilizadas para auxiliar na visualização das regiões. . . . .</i>	30
3.10	<i>Principais atributos de região, ou seja, dos objetos independentes presentes na imagem.</i>	31
4.1	<i>(a)Imagem Original - “cameraman” (b)PDF da Imagem Original (c)Imagem Binária (t = 160) (d)Função Critério para Binarização . . . . .</i>	39
4.2	<i>(a)Imagem Original (b)PDF da Imagem Original (c)Imagem Binária (t = 91) (d)Função critério para Binarização. A posição de menor distância é escolhida como ponto de ‘threshold’ dos níveis de cinza. . . . .</i>	43
4.3	<i>Perfil de uma linha da imagem (y = 120) ilustrando as discontinuidades de luminância entre pixels. . . . .</i>	46
4.4	<i>Exemplo de máscaras para o cálculo do operador entropia: A máscara de 3 × 3 pixels e B máscara de 7 × 7 pixels. . . . .</i>	47
4.5	<i>(a)Imagem Original (b)Imagem segmentada por detecção de bordas utilizando o operador entropia. . . . .</i>	48
5.1	<i>Imagens e não extensividade. Imagem de um ruído gaussiano 2D (Imagem A) e imagem complexa de uma retina humana obtida por tomografia (Imagem B) . . . .</i>	51
5.2	<i>Segmentação ótima: o processo de binarização entrópica deve ser capaz de encontrar o melhor ponto de separação entre o “objeto” e o “fundo”. . . . .</i>	54
5.3	<i>Evolução do histograma representativo das regiões de luminância da imagem (“objeto” e “fundo”). O histograma destacado corresponde a uma linha do gráfico de superfície que ilustra a evolução. . . . .</i>	55
5.4	<b>Caso 1.</b> <i>A gaussiana 1 é mantida fixa enquanto a gaussiana 2 é deslocada. Duas situações são observadas neste caso. No início os dois picos são facilmente destacados até o extremo onde os picos se encontram misturados entre si. . . . .</i>	56

5.5	<i>Caso 2.</i> A gaussiana 1 é mantida fixa, na mesma altura, enquanto modificamos a largura da gaussiana 2. Este processo equivale a imagens onde um dos objetos tem uma variação em sua reflexibilidade ou é muito ruidoso. . . . .	57
5.6	<i>Caso 3.</i> A gaussiana 1 é mantida fixa enquanto modificamos progressivamente a altura da gaussiana 2. Este processo representam imagens em que objetos e fundo tem probabilidades muito diferentes. . . . .	57
5.7	<i>Caso 4.</i> A gaussiana 1 está centrada em 64 e uma segunda gaussiana de largura e altura relativamente pequena é adicionada à primeira. Ao final, a segunda gaussiana praticamente não existe comparativamente à primeira. É um caso de difícil segmentação pois equivale àquelas imagens com uma região pouco provável, normalmente com muito ruído na sua região de luminância. . . . .	58
5.8	(a) Evolução do histograma para simulação do processo de segmentação para o caso 1. A linha escura no gráfico corresponde ao ponto de corte $t$ escolhido. (b) Superfícies relativa a soma das entropias ( $A \cup B$ ) utilizadas para determinação do ponto máximo para cada uma das situações do histograma. . . . .	59
5.9	(a) Evolução das regiões “objeto” e “fundo” no processo de segmentação para o caso 2. A linha escura corresponde ao ponto de corte $t$ . O ponto de segmentação evolui conforme aumentamos $q$ . (b) Superfícies utilizadas para determinação do ponto de segmentação para cada uma das situações do histograma. . . . .	60
5.10	(a) Evolução das regiões “objeto” e “fundo” no processo de segmentação para o caso 3. A linha escura corresponde ao ponto de corte no histograma. (b) Superfícies utilizadas para a determinação do ponto de segmentação para cada uma das situações do histograma. . . . .	61
5.11	(a) Evolução das funções gaussianas no processo de segmentação para o caso 4. O pico 2 se encontra praticamente submerso no pico 1 ou pela presença de ruídos na imagem. Valores de $q$ muito baixos criam duas regiões de máximos e a escolha do ponto de corte se torna instável. (b) Superfícies utilizadas para a determinação do ponto máximo para cada uma das situações do histograma. . . . .	62

5.12	<i>Caso 1 : A imagem apresenta duas regiões com dois picos idênticos. O processo de segmentação neste caso não depende de <math>q</math>. O valor de corte é encontrado para <math>t = 140</math>. Os métodos Two-Peaks e Iterative Selection apresentam também bons resultados. . . . .</i>	63
5.13	<i>Caso 1 : Dois picos sobrepostos - O processo de binarização entrópica determina o ponto de corte em <math>t = 147</math>. . . . .</i>	64
5.14	<i>Caso 2 : A presença de um pico mais largo que o outro faz com que o ponto de corte escolhido pelo processo de binarização entrópica tenda a subir na gaussiana mais larga. Este é o único caso em que o binarizador entrópico apresenta resultados inferiores aos dois outros processos de binarização. . . . .</i>	65
5.15	<i>Caso 3 : Imagem com dois picos independentes e com alturas muito diferentes. A segmentação entrópica, para eliminar a presença dos ruídos na imagem original e amplificados no método "Two Peaks", precisa de um <math>q</math> pequeno (<math>q=0.1</math>). . . . .</i>	66
5.16	<i>Caso 2 + Caso 3: Imagem com dois picos independentes e com alturas e larguras muito diferentes. A segmentação entrópica não extensiva apresenta resultados diferentes em função de <math>q</math>. Em (a) <math>q = 0.5</math> e o ponto de corte é <math>t = 127</math>. Em (b) <math>q = 3</math> e ponto de corte (<math>t = 31</math>) é melhor ajustado entre as duas regiões. . . . .</i>	67
5.17	<i>Caso 4 - objeto imerso em ruído : O histograma desta imagem é composta por um grande e largo pico que praticamente mascara a presença do outro, completamente dentro de uma região ruidosa de luminância. O método de segmentação entrópica funciona bem para estes casos, com um valor de <math>q</math> elevado. A curva para determinação do valor máximo de threshold tem um máximo em 183. . . . .</i>	68
5.18	<i>Ponto de segmentação ótimo é definido no ponto de encontro entre as duas gaussianas, <math>a</math> e <math>b</math>. . . . .</i>	75
5.19	<i>Resultado da segmentação por entropia relativa para os casos 1, 2, 3 e 4. . . . .</i>	77
5.20	<i>Imagem de uma paisagem. O histograma não segue a priori nenhuma distribuição específica. Neste caso, o parâmetro de não extensividade <math>q</math> irá alterar a distância estatística de Kullback-Leibler e modificará o ponto de corte <math>\hat{t}</math>. . . . .</i>	78
5.21	<i>Resultado da segmentação para a imagem com uma distribuição desconhecida. O parâmetro <math>q</math> permite um ajuste fino da posição de corte (<math>q = 1.1</math>). . . . .</i>	79

5.22	<i>Gráfico do comportamento da distância Kullback-Leibler generalizada entre gaussianas, para vários valores de <math>q</math>.</i> . . . . .	80
6.1	<i>Seqüência de tratamento da imagem para o processo de classificação</i> . . . . .	83
6.2	<i>Etapa de descrição do objeto. Apresentamos uma (ou mais) amostras ao sistema de descrição que irá construir os jogos de probabilidades utilizados posteriormente no processo de classificação das formas desconhecidas.</i> . . . . .	84
6.3	<i>Representação da distribuição de probabilidades de três formas básicas.</i> . . . . .	86
6.4	<i>Etapa de classificação dos objetos desconhecidos. Apresentamos uma amostra ao sistema de classificação. O sistema calcula a semelhança entre o jogo de probabilidades deste objeto com aqueles obtidos no processo de descrição da forma. A menor distância indica o padrão referente a forma.</i> . . . . .	87
6.5	<i>Jogos de probabilidades correspondentes as 10 classes (dígitos) utilizados no processo de classificação. Os histogramas representam a distribuição de distância entre os pixels de borda e o centro de massa de cada objeto. Estes histogramas servem como modelo de padronização das formas.</i> . . . . .	89
6.6	<i>O gráfico apresenta a evolução da distância entre os jogos de probabilidades, de um círculo perfeito àqueles corrompidos por adição de ruído gaussiano. Quanto mais ruído é adicionado mais “longe” o objeto fica do círculo perfeito. O ponto inicial de distância zero, se encontra fora dos limites do gráfico (i.e., acima de 35 dB).</i> . . . . .	90
6.7	<i>Varição da forma do “objeto” círculo com o acréscimo de ruído em sua borda, e os respectivos jogos de probabilidades calculados pela distância dos pixels de borda e o centro de massa da forma. Existe um limiar de decisão no qual não aceitamos (S.V.H.) mais que a forma equivale a um círculo.</i> . . . . .	91
6.8	<i>Medida de eficiência do classificador entrópico. Aos objetos, desconhecidos pelo classificador, adicionamos ruído gaussiano de média zero e desvio padrão (<math>\sigma</math>). O resultado da classificação é comparado com àquele esperado. A eficiência é medida pelo comportamento da taxa de acerto do classificador em função da relação sinal/ruído.</i> . . . . .	92



6.9	<i>Taxa de Acerto do classificador em função da relação sinal/ruído (SNR) para diferentes resoluções do histograma (10, 20, 40, 60, 80 e 100). Os melhores casos estão situados para histogramas com poucos bins. . . . .</i>	93
6.10	<i>(a) Forma “8” padrão utilizada no processo de comparação e outras formas “8” apresentada ao classificador com diferentes intensidade de ruído, e resolução do histograma igual a 20 intervalos. O classificador reconhece todos os objetos como pertence a classe correta. . . . .</i>	93
6.11	<i>Avaliação dos melhores pontos de funcionamento do classificador para três Taxas de Acerto (100, 90 e 80%). Podemos definir o limiar de funcionamento para uma taxa de acerto de 100% com 20 bins de resolução do histograma em 28.8 dB. O gráfico mostra que o sistema é eficiente com uma taxa acima de 80% das amostras com um ruído adicionado equivalente a 11.5 pixels em uma forma de altura igual 48 pixels de amplitude. . . . .</i>	94
6.12	<i>Os gráficos representam a excursão da medida de distância Kullback-Leibler, <math>D_q(p : p')</math>, entre uma forma padrão perfeita e uma na qual foi adicionado ruído, de diferentes amplitudes, em sua borda. Quatro valores para o parâmetro <math>q</math> são apresentados. A excursão observada para a distância <math>D_q(p : p')</math> é ampliada para valores de <math>q &gt; 1</math> (sistemas subextensivos) e atenuada para valores de <math>0 &lt; q \leq 1</math> (sistemas superextensivos). . . . .</i>	97
6.13	<i>Variação da entropia generalizada para as formas ruidosas da fig.6.12. Os valores são normalizados pela entropia máxima para cada <math>q</math>. Quando <math>q &gt; 1</math> forma tende mais rapidamente a ser considerada como uma forma aleatória. . . . .</i>	98
6.14	<i>Taxa de Acerto em função da relação SNR, para diferentes valores de <math>q</math>. (a) Comportamento para todos os casos (<math>q &gt; 1</math>, <math>q = 1</math> e <math>0 &lt; q \leq 1</math>). Duas regiões são destacadas: <math>0 &lt; q \leq 1</math> e <math>q &gt; 1</math>. (b) Comportamento para os casos <math>0 &lt; q \leq 1</math> e <math>q = 1</math> ampliados. Taxa de acerto é ligeiramente melhor para <math>q = 0.67</math>. . . . .</i>	99

6.15	<i>Evolução da medida de distância <math>D_q(p_a, p_b)</math> entre dois jogos de probabilidades para uma distribuição normal em função do parâmetro de não extensividade <math>q</math>, para diferentes valores de <math>\sigma_b</math>. Quando utilizamos <math>0 &lt; q \leq 1</math> a curva tem um comportamento crescente mais controlado, tanto para pequenos valores de desvio padrão quanto para grandes valores.</i>	100
A.1	<i>Decomposição da opção em três possibilidades</i>	108
A.2	<i>Escolhas com resultados equiprováveis (<math>m = 2</math>).</i>	109
A.3	<i>Escolhas com resultados equiprováveis (<math>m = 3</math>).</i>	109
A.4	<i>Diagrama de escolhas para o caso (II).</i>	111
B.1	<i>Exemplo de segmentação com o método Iterative Selection. A figura apresenta o histograma da imagem original e o <math>t_{otimo}</math> definido pelo algoritmo.</i>	116
B.2	<i>Exemplo de segmentação com o método Two Peaks. A figura apresenta o histograma da imagem original e o histograma modificado para detecção do segundo pico. O <math>t_{otimo}</math> é definido pelo algoritmo no valor mínimo (vale) entre os picos detectados.</i>	117

# Lista de Tabelas

6.1	<i>Distância entre os objetos e padrões calculados pela equação 6.3 (Distância Kullback-Leibler) entre jogos de probabilidades definidos pela equação 6.2. . . . .</i>	88
-----	--	----



# Capítulo 1

## Introdução

Atualmente a análise quantitativa de imagens é uma ferramenta importante para diferentes áreas, dentre elas podemos citar: ciência dos materiais (metalurgia, microscopia, nanoscopia, magnetismo), ciências biológicas (biologia, genética, botânica), ciências da terra (geologia), geografia (meteorologia, fotografias aéreas e de satélites), astronomia, robótica, etc. O que faz a análise de imagens uma disciplina comum a estas diferentes áreas é que imagens são na realidade um suporte físico para troca e transporte de informações. Esta INFORMAÇÃO pode estar associada a uma medida (neste caso falamos de um sinal em associação a um fenômeno físico), ou pode estar associada a um nível cognitivo (neste caso falamos de conhecimento). Uma imagem contém uma quantidade imensa de informações que um observador humano interpreta normalmente de um modo global e quantitativo. Processar uma imagem consiste em transformá-la sucessivamente com o objetivo de extrair a INFORMAÇÃO nela presente. Estas transformações vão desde o sinal numérico até tratamentos de mais alto nível, que correspondem ao sentido cognitivo da imagem. Podemos dividir este domínio em duas áreas: *Processamento de Imagem* (P.I.) e *Visão Computacional* (V.C.).

As técnicas de P.I. são aquelas que realizam transformações na imagem porém não se preocupando com o conteúdo de informação nela presente. Um dos primeiros interesses na aplicação dessas técnicas data do início do século passado, através da melhoria da qualidade de fotografias transmitidas entre os EUA e a Europa. Contrariamente, as técnicas de V.C., são aquelas que vão “tentar” extrair da imagem as

principais informações nela presente. Estas técnicas ganharam um verdadeiro impulso com a evolução dos computadores entre as décadas de 1960 e 1970. A crescente performance no hardware dos computadores popularizou esses métodos e diversas ferramentas podem ser encontradas atualmente nos computadores pessoais em laboratórios de pesquisa. Quando aplicamos estas técnicas em ambientes de física experimental, isto é, como um sistema complementar de medida, é importante que a análise seja quantitativa e relacionada diretamente ao fenômeno físico estudado. A maioria das técnicas desenvolvidas para estas duas áreas, P.I e V.C, são complementares e podem atuar de modo cooperativo. Neste trabalho o termo *Processamento Digital de Imagens* (P.D.I.) será utilizado para caracterizar estas duas áreas em conjunto.

Inúmeras áreas de pesquisas em P.D.I. tentam compreender ou imitar os métodos de processamento da informação visual do Sistema Visual Humano (S.V.H.). O S.V.H. é certamente um sistema complexo e eficiente para realizar reconhecimento [1]. Em alguns casos específicos, sistemas de análise de imagem por computador podem ser mesmo mais eficientes que o S.V.H. Isto nos faz supor que o P.D.I. é uma ferramenta de extrema dependência da aplicação e de suas características particulares.

O Processamento Digital de Imagens pode ser subdividido em algumas etapas, iniciando-se na formação e digitalização da imagem, e posteriormente em etapas como pré processamento, segmentação, pós processamento e finalmente a extração de atributos, medidas, classificação e reconhecimento.

Dentre as diversas etapas do P.D.I., a *segmentação* é considerada uma etapa crítica em todo o processo. Esta é a primeira etapa em que leva-se em consideração as informações presentes na imagem, e é também chamada de transformação Sinal-Símbolos. Segmentar uma imagem consiste em dividi-la em diferentes regiões, que serão posteriormente utilizadas para a identificação e extração de informações relevantes. Esta característica torna a segmentação um ponto crítico, porque quaisquer erros ou distorções produzidos durante o processo são transmitidos às demais etapas, intensificando a possibilidade de obtenção de resultados indesejados. Uma etapa posterior à segmentação é o *Reconhecimento*. Nesta parte do processamento os objetos segmentados são classi-

ficados a partir de informações obtidas na própria imagem.

Em laboratórios de física experimental, o PDI é uma valiosa ferramenta para análise quantitativa. Nessas aplicações o PDI age como uma ferramenta auxiliar de medida, trazendo um conjunto de técnicas específicas, normalmente dependente do tipo de fenômeno físico estudado na imagem.

## 1.1 Imagem e Entropia

Por volta de 1850, o cientista alemão Rudolph Clausius introduziu pela primeira vez o conceito de entropia com o objetivo de demonstrar a direcionalidade de processos físicos. Quando dois corpos trocam calor, sem que ocorram outros efeitos além da variação de suas temperaturas, eles acabam por atingir o equilíbrio mútuo a uma temperatura intermediária entre as suas temperaturas iniciais. Deste fato e da conservação de energia resulta que, quando os corpos atingem o equilíbrio, perde-se definitivamente parte da capacidade de conversão em trabalho. Como esta perda não pode ser recuperada, dizemos que há irreversibilidade externa na troca de calor. Clausius desenvolveu uma formulação matemática que permite exprimir o efeito da irreversibilidade em termos da quantidade de calor trocada e das temperaturas dos dois corpos, através de uma função denominada *entropia*. A irreversibilidade é descrita pelo crescimento monotônico dessa função que tende para um valor máximo, atingido quando se dá o equilíbrio térmico entre os dois corpos. De acordo com Clausius, a energia do universo seria constante enquanto a entropia do universo tenderia a um valor máximo.

Anos depois, em torno de 1870, o físico austríaco Ludwig von Boltzmann utilizou os métodos estatísticos propostos por Maxwell para redefinir o conceito de entropia, relacionando com ordem e desordem em escala atômica e estabelecendo uma relação entre a entropia e a análise estatística do movimento molecular. Boltzmann publicou uma série de artigos que descreviam a segunda lei da termodinâmica, as leis da mecânica e teoria das probabilidades aplicada aos movimentos dos átomos. Deste modo, Boltzmann mostrou que a entropia, que tinha sido compreendida previamente apenas de forma fenomenológica (através das equações de trabalho-energia da física clássica), era uma

medida estatística. Isto conduziu ao conceito de estado de equilíbrio de um sistema termodinâmico como o estado mais provável do sistema.

Um novo conceito de entropia foi definido pelo matemático americano Claude Shannon para descrição de sistemas de informação. Em 1948, com a publicação do artigo “The Mathematical Theory of Communication”, Shannon deu início ao que atualmente conhecemos como Teoria da Informação. Shannon definiu uma medida de entropia, que quando aplicada a uma fonte de informação, poderia determinar a capacidade do canal para a transmissão da fonte em bits codificados. A entropia definida por Shannon é bastante similar à definida pelos físicos. O consistente trabalho realizado por Boltzmann e Gibbs na física estatística foi uma fonte de inspiração para a adoção do termo entropia em sistemas de informação. Neste novo contexto, a entropia está relacionada a incerteza de se obter uma informação e a capacidade informativa da fonte.

O conceito de entropia para sistemas de informação também foi aplicado em sistemas de processamento de imagem. Durante as últimas duas décadas, foram desenvolvidas diversas técnicas fundamentadas no cálculo da entropia de fontes de informação, representadas neste sistemas por imagens digitais. Estes estudos foram direcionados basicamente para as áreas de compressão e segmentação. O interesse pela compressão de dados vem do fato que imagens digitais necessitam de uma quantidade relativamente elevada de *bits* para a sua representação, armazenamento e transmissão. Na parte de segmentação, o conceito de entropia foi utilizado em métodos de binarização de imagens, que na prática corresponde a classificação dos pixels da imagem em pertencentes a objetos ou fundo. Essa mesma classificação pode também ser vista como uma seleção de pixels que são relevantes no tratamento daqueles que podem ser descartados. Nos métodos de binarização, uma função critério é definida com base no cálculo da entropia da imagem e um nível ou limiar ótimo é definido através da maximização ou minimização da função, dependendo do algoritmo adotado.

A proposta deste trabalho é aplicar os conceito de entropia em técnicas de segmentação e em reconhecimento de padrões em imagens. As principais técnicas de segmentação que utilizam a entropia são abordadas e descritas, buscando ainda avaliar



uma extensão destas técnicas através da utilização da entropia generalizada de Tsallis. O conceito de entropia não extensiva, apesar de ter sido desenvolvido como uma generalização da tradicional expressão de Boltzmann para sistemas físicos não extensivos, tem despertado o interesse de investigação de sua aplicação em sistemas de informação. Alguns trabalhos têm sido publicados recentemente propondo a extensão da utilização do formalismo não extensivo em sistemas de informação. Este foi o ponto de partida para o desenvolvimento deste trabalho que propõe uma primeira abordagem da estatística não extensiva em sistemas de processamento de imagens. Desde o início, alguns questionamentos se fizeram presentes. Um sistema de informação pode apresentar características não-extensivas? Uma imagem digital pode ser considerada um sistema não-extensivo? Métodos entrópicos deveriam apresentar resultados diferenciados para certas classes de imagens? Quais? Os diversos trabalhos na área da teoria da informação e da teoria de sinais que vêm sendo publicados atualmente, nos motivaram a contribuir com o esclarecimento destas questões, propondo, quando for o caso, a utilização do formalismo de entropia generalizada no processamento de imagens.

## 1.2 O Documento

O segundo capítulo da tese introduz os conceitos fundamentais da teoria da informação, com ênfase para o conceito de entropia de uma fonte de informação. Ainda neste capítulo são apresentados os conceitos básicos da estatística não extensiva ou estatística Tsallis, a generalização da entropia de Shannon e Kullback-Leibler. O Capítulo 3 é dedicado as técnicas de processamento digital de imagens, descrevendo as etapas de processamento, alguns exemplos e o formalismo adotado. O capítulo 4 apresenta as principais técnicas de segmentação entrópica, enquanto que no Capítulo 5 são apresentados métodos que utilizam a entropia não extensiva. Neste capítulo são apresentados os resultados da segmentação com métodos entrópicos em diferentes tipos de histogramas simulados e em imagens reais. Os resultados da adoção do formalismo não extensivo é discutido e avaliado. O Capítulo 6 é dedicado ao reconhecimento de formas básicas em imagens através do cálculo da entropia. Por fim o Capítulo 7 apresenta as

conclusões e discussões sobre os resultados obtidos neste trabalho.

# Capítulo 2

## Fundamentos de Teoria da Informação

A Teoria da Informação é uma ramificação da teoria de probabilidades introduzida por Claude Shannon com a publicação do artigo “The Mathematical Theory of Communication” [2] em 1948, apresentando um novo modelo matemático para o estudo de sistemas de comunicação. Os principais objetivos de Shannon eram descobrir as leis que regulam os sistemas usados para comunicar e manipular a informação e definir medidas quantitativas para a informação e para a capacidade de determinados sistemas transmitirem, armazenarem e processarem a informação.

Uma das inovações mais importantes do modelo introduzido por Shannon, foi considerar os componentes de um sistema de comunicação (fontes de informação, canais de comunicação) como elementos probabilísticos.

Alguns dos problemas abordados por Shannon estão relacionados com a descoberta de melhores métodos para utilizar os sistemas de comunicação existentes e as melhores formas de separar a informação desejada (sinal) da informação desprezível (ruído). Um outro tema abordado é a definição de limites superiores para as possibilidades do meio de transporte de informação, também chamado um canal de comunicação.

Shannon propôs uma forma de medição quantitativa da informação fornecida por um evento probabilístico, baseada na tradicional expressão de entropia de Boltzmann

(1896) presente na termodinâmica e física estatística. Foi Shannon quem primeiro relacionou entropia e informação. Em seu modelo de comunicação (fonte-canal-receptor)[2], a quantidade de informação transmitida em uma mensagem é função de previsibilidade da mensagem. A noção de entropia está ligada ao grau de desorganização existente na fonte de informação. Quanto maior a desordem, maior o potencial de informação desta fonte. Uma fonte que responda com uma única e mesma mensagem a toda e qualquer pergunta não transmite informação, já que não há redução de incerteza.

## 2.1 Informação, Incerteza e Entropia

Uma *fonte de informação* é um modelo matemático para um sistema físico que produz uma sucessão de símbolos de maneira aleatória chamados *eventos*. Os símbolos produzidos podem ser números reais como valores de voltagens provenientes de um transdutor, números binários de dados computacionais<sup>1</sup>, etc. O espaço contendo todos os eventos possíveis é usualmente chamado de *alfabeto* da fonte de informação, ao qual é atribuído um conjunto de probabilidades de ocorrência.

### 2.1.1 Fonte Discreta de Informação

Uma fonte discreta de informação gera símbolos de um alfabeto

$$\mathcal{A} = \{x_i, i = 1, 2, \dots, k\} \quad (2.1)$$

com probabilidade de ocorrência  $p_i$  tal que  $\sum_{i=1}^k p_i = 1$ .

Os símbolos gerados são estatisticamente independentes, de modo que a probabilidade de ocorrência de qualquer seqüência gerada pela fonte é dada pelo produto das probabilidades de ocorrência dos símbolos que a constituem. Se uma fonte emite

---

<sup>1</sup>No caso do sistema físico ser uma imagem digital, pode-se considerar como fonte de informação os valores das voltagens emitidas por um sensor CCD (charge-coupled device), relativas à ocorrência dos fótons incidentes em cada célula.

uma seqüência de dois símbolos  $\alpha$  e  $\beta$ , com probabilidades  $p_\alpha$  e  $p_\beta$ , respectivamente, a probabilidade da seqüência gerada é definida como sendo

$$p^{\alpha\cup\beta} = p^\alpha \cdot p^\beta \quad (2.2)$$

caso os símbolos  $\alpha$  e  $\beta$  sejam estatisticamente independentes.

### 2.1.2 Medida de Informação

Quando se descreve um processo de seleção de um objeto entre vários existentes, aparece naturalmente as noções de informação e incerteza. Se um sistema é capaz de emitir 3 símbolos distintos A, B e C e esperamos a ocorrência do primeiro evento, mantemos uma *incerteza* sobre qual símbolo aparecerá. Quando o primeiro símbolo é emitido, a incerteza desaparece e podemos considerar que houve um *ganho de informação*.

Se considerarmos a fonte de informação definida em 2.1, o ganho de informação associado a ocorrência de um evento é chamada de *informação própria* de cada evento  $x_i$  e é definida como

$$I(x_i) = \log \left( \frac{1}{p_i} \right), \quad (2.3)$$

representando uma forma intuitiva de medição quantitativa de informação, mesmo sendo *informação* um conceito relativamente subjetivo. Esta forma de medir informação apresenta as seguintes características:

- $I(x_i) = 0$  se  $p_i = 1$

O ganho de informação resultante da ocorrência do evento único é nulo.

- $I(x_i) \geq 0$

A ocorrência de qualquer evento produz um ganho de informação, exceto no caso de uma fonte que emite um único símbolo.

- $I(x_i) > I(x_j)$  se  $p_i < p_j$

Quanto menor a probabilidade de ocorrência de um símbolo, maior é o ganho de informação.

E se considerarmos a ocorrência simultânea de dois eventos estatisticamente independentes  $x_i$  e  $x_j$ , o ganho de informação total é definido pela soma das informações próprias de cada um dos eventos

$$\begin{aligned} I(x_i, x_j) &= -\log(p_i \cdot p_j) \\ &= -\log p_i - \log p_j \\ &= I(x_i) + I(x_j) \end{aligned} \tag{2.4}$$

A base da função logaritmo, presente em 2.3, determina a unidade de medida de informação. A rigor, a base pode ser qualquer número maior que 1 [2], sendo usual a utilização da base 2 para sistemas digitais de informação. O logaritmo de base 2 define a unidade binária (bit) como a informação própria associada a cada um dos símbolos de uma fonte binária com eventos equiprováveis:

$$I(0) = I(1) = -\log_2\left(\frac{1}{2}\right) = 1 \text{ bit} \tag{2.5}$$

### 2.1.3 Entropia de uma Fonte de Informação

A entropia de uma fonte discreta de informação é dada pela esperança matemática da informação própria dos símbolos da fonte, ou seja, o produto do ganho de informação de cada símbolo pela sua probabilidade de ocorrência. A entropia da fonte é sensível a quantidade de símbolos que a fonte é capaz de emitir. Para uma fonte discreta de informação com  $k$  símbolos e probabilidade  $p_i = \{p_1, p_2, \dots, p_k\}$ , a entropia é definida como sendo

$$S = E\{I(x_i)\} = \sum_{i=1}^k p_i \cdot \log\left(\frac{1}{p_i}\right)$$

ou na forma mais utilizada na literatura

$$S = - \sum_{i=1}^k p_i \cdot \log p_i \quad (2.6)$$

A entropia de um sistema binário, que apresenta apenas dois estados possíveis com probabilidades  $p$  e  $q = 1 - p$ , pode ser representada como uma função de  $p$ ,

$$S_2(p) = S_2(p, 1 - p) = -p \cdot \log_2(p) - (1 - p) \cdot \log_2(1 - p) \quad (2.7)$$

e está ilustrada na figura 2.1, que também apresenta a informação própria em bits como função de  $p$ . Quanto menos provável é a ocorrência de um evento, maior é sua informação própria.

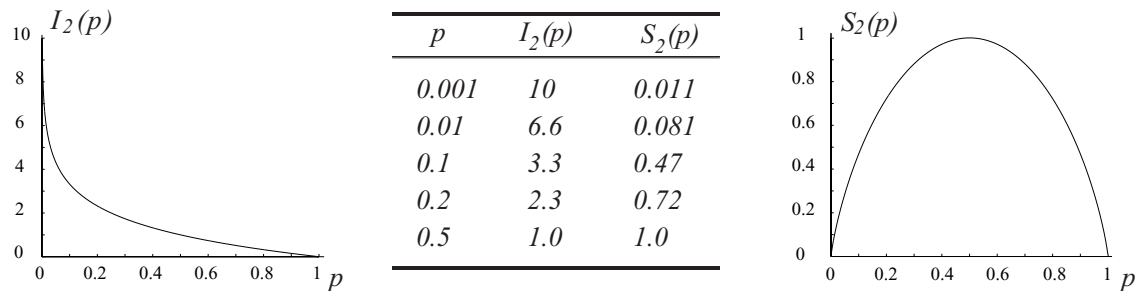


Figura 2.1: Informação Própria e Entropia da fonte binária como função da probabilidade.

Observando a figura 2.1, é interessante notar que:

1. Quando  $p = 0$  a entropia é nula ( $S = 0$ ), pois  $x \log x \rightarrow 0$  quando  $x \rightarrow 0$
2.  $S = 0$  quando  $p = 1$  para um único evento, definido como valor mínimo da entropia.
3. A entropia atinge o valor máximo  $S_{max} = 1 \text{ bit/simbolo}$ , quando  $p = q = \frac{1}{2}$  ou seja, quando os símbolos são equiprováveis.

Portanto, os valores para a entropia de uma fonte de informação com  $k$  símbolos é limitada segundo a desigualdade a seguir

$$0 \leq S \leq \log_2 k \quad (2.8)$$

em que o  $\log_2$  representa a função logaritmo na base 2.

## 2.2 Entropia Relativa

Nesta seção será introduzido o conceito de *entropia relativa*. A entropia relativa é a medida de uma distância estatística entre duas distribuições definidas sobre um mesmo alfabeto. Em estatística, isto significa o valor esperado do logaritmo da relação entre as probabilidades. A entropia relativa é definida como sendo

$$D_{KL}(p : p') = \sum_{i=1}^k p_i \cdot \log \frac{p_i}{p'_i} \quad (2.9)$$

e é também conhecida como *Distância Kullback-Leibler*, *Entropia Kullback-Leibler* ou *Divergência I*. Na definição acima, assumimos por convenção (baseada em argumentos de continuidade) que  $0 \cdot \log \frac{0}{p'} = 0$  e  $p \cdot \log \frac{p}{0} = \infty$ .

A entropia relativa é sempre não negativa, satisfazendo a desigualdade de Gibbs:

$$\sum_{i=1}^k p_i \cdot \log \frac{1}{p_i} \leq \sum_{i=1}^k p_i \cdot \log \frac{1}{p'_i} \quad (2.10)$$

considerando  $p_i$  uma distribuição de probabilidades qualquer e  $p'_i$  uma outra distribuição que satisfaz a condição a seguir

$$\sum_{i=1}^k p'_i \leq 1 \quad (2.11)$$

Apesar da entropia relativa ser chamada de distância Kullback-Leibler, esta não pode ser considerada uma distância verdadeira entre duas distribuições. A entropia relativa não apresenta simetria,



$$\sum_{i=1}^k p_i \cdot \log \frac{p_i}{p'_i} \neq \sum_{i=1}^k p'_i \cdot \log \frac{p'_i}{p_i} \quad (2.12)$$

e portanto não pode ser considerada uma distância métrica. Mesmo assim, é usual pensar na entropia relativa como uma medida de “distância” entre duas distribuições estatísticas.

Uma versão simétrica de entropia relativa, conhecida como divergência  $J$ , foi introduzida por H. Jeffreys [3] e é definida como a soma das duas divergências diretas (divergências  $I$ ):

$$\begin{aligned} D(p : p') &= D_{KL}(p : p') + D_{KL}(p' : p) \\ &= \sum_{i=1}^k p_i \cdot \log \frac{p_i}{p'_i} + \sum_{i=1}^k p'_i \cdot \log \frac{p'_i}{p_i} \end{aligned} \quad (2.13)$$

é válido notar que a divergência  $J$  é simétrica em relação aos argumentos[4], de forma que  $D(p : p') = D(p' : p)$ .

## 2.3 Entropia Generalizada em Sistemas de Informação

### 2.3.1 Conceitos Fundamentais da Entropia Não Extensiva

Durante aproximadamente 120 anos, o conceito de entropia tem sido descrito através de uma expressão particular, chamada *Entropia Boltzmann-Gibbs*

$$S = k \log W \quad (2.14)$$

onde a entropia ( $S$ ) é o produto da constante de Boltzmann ( $k$ ) pelo logaritmo de microestados ( $W$ ) do sistema.

O conceito de entropia é de fundamental importância na termodinâmica, mecânica estatística e teoria da informação. Recentemente, estudos de sistemas físicos que en-

volvem a presença de efeitos não extensivos têm despertado um grande interesse, principalmente porque tais sistemas não são convenientemente descritos pela mecânica estatística de Boltzmann-Gibbs. A presença de características não extensivas é comum em sistemas astrofísicos, sistemas magnéticos e sistemas que apresentam evolução temporal da entropia. De um modo simplificado, podemos afirmar que a não extensividade pode ocorrer quando:

- (i) as interações são de curto alcance sendo o sistema de tamanho finito e menor do que o alcance das interações ou
- (ii) com interações de longo alcance sendo o tamanho do sistema qualquer.

Com base neste contexto, acredita-se que a mecânica estatística atual possui limitações, existindo a necessidade de uma reformulação dos conceitos de entropia e extensividade.

Durante os últimos anos, uma nova expressão para entropia, proposta pelo físico Constantino Tsallis [5], tem sido considerada uma possível generalização da entropia de Boltzmann/Gibbs. Este novo formalismo, chamado de Entropia Tsallis ou Estatística Tsallis, tem sido aplicado a inúmeros sistemas, em diversas áreas da ciência, que vão desde a física do estado sólido até a teoria da informação [6] [7] [8] [9] [10] [11] [12]. A entropia Tsallis se adapta as características físicas de muitos sistemas físicos e ainda preserva as propriedades fundamentais da entropia na segunda lei da termodinâmica, ou seja, que a entropia do universo aumenta com o tempo em todos os processos físicos. Este novo formalismo para a termodinâmica e física estatística está fundamentado na expressão para a entropia sugerida por Tsallis

$$S_q = k \cdot \sum_{i=1}^W p_i^q \cdot \frac{(1 - p_i^{1-q})}{q - 1} \quad (2.15)$$

ou na forma mais utilizada na literatura,

$$S_q = k \frac{1 - \sum_{i=1}^W p_i^q}{q - 1} \quad (2.16)$$

em que  $k$  é uma constante positiva (a qual é atribuído o valor unitário),  $q$  é um número

real,  $W$  é o número total de microestados e  $p_i$  é o conjunto de probabilidades associado aos estados. Pode-se facilmente demonstrar que no limite em que  $q \rightarrow 1$  a equação 2.16 retorna para a expressão 2.6, a entropia de Boltzmann/Gibbs/Shannon<sup>2</sup>.

$$S = - \sum_{i=1}^W p_i \cdot \log p_i \quad (2.17)$$

### 2.3.2 Entropia de Shannon Generalizada

Nas seções anteriores deste capítulo, foram definidos os conceitos fundamentais da teoria da informação. Foi dada uma ênfase ao conceito de medida de informação proposta por Shannon a partir de um auto questionamento: “Podemos definir uma quantidade que possa medir, de alguma maneira, quanta informação é ‘produzida’ em um processo, ou melhor, a que taxa a informação é produzida?” A partir daí, Shannon define que para uma medida deste tipo ser possível, uma função  $S$  dependente de probabilidades  $p_i$  para  $W$  eventos deveria satisfazer as seguintes propriedades:

1.  $S$  deveria ser contínua em  $p_i$ .
2. Se a probabilidade de todos os eventos forem iguais, ou seja  $p_i = \frac{1}{W}$ , então  $S$  deveria ser uma função monotônica crescente. Com eventos equiprováveis, há maior poder de escolha, ou maior incerteza, quando existir um número grande de eventos possíveis.
3. Para dois subsistemas estatisticamente independentes  $A$  e  $B$  a função  $S$  para o sistema composto  $A + B$  deveria apresentar a propriedade aditiva, de forma que:
$$S(A + B) = S_A + S_B.$$

Tendo em conta estas propriedades, Shannon enuncia o seguinte teorema:

**Teorema 1** – A única função que satisfaz as três condições anteriores é da forma:

$$S = -K \sum_i p_i \cdot \log p_i \quad (2.18)$$

---

<sup>2</sup>Esta demonstração matemática está presente no Apendice A.

em que  $K$  é uma constante positiva que, por conveniência em sistemas de informação, é considerada de valor unitário.

Recentemente, Santos [13] generalizou o conhecido teorema de Shannon, que define a equação 2.18 como uma medida quantitativa da informação de uma fonte. É conhecido da estatística Tsallis que  $S_q$  definido anteriormente na equação 2.16 satisfaz as seguintes condições:

1.  $S_q$  é contínua em  $p_i$ , para  $0 < p_i < 1$ .
2. Para um conjunto de  $W$  de eventos equiprováveis, ou seja,  $p_i = \frac{1}{W}$ , então  $S_q$  é uma função monotônica crescente.
3. Para dois subsistemas estatisticamente independentes  $A$  e  $B$  a entropia generalizada  $S_q$  do sistema composto  $A + B$  satisfaz a relação de pseudo-aditividade

$$S_q(A + B) = S_q(A) + S_q(B) + (1 - q) \cdot S_q(A) \cdot S_q(B) \quad (2.19)$$

Tendo em conta as três condições anteriores, Santos define em [13] o seguinte teorema:

**Teorema 2** – A única expressão que satisfaz de forma simultânea todas as condições acima é a entropia generalizada Tsallis:

$$S_q = k \frac{1 - \sum_{i=1}^W p_i^q}{q - 1} \quad (2.20)$$

Deste forma, fica definido que a entropia Tsallis pode ser utilizada como uma medida de informação adequada para a utilização em sistemas de informação que apresentam características não extensivas<sup>3</sup>.

### 2.3.3 Entropia Relativa Generalizada

Inicialmente, torna-se necessário apresentar a derivação da expressão da entropia relativa pelo caso convencional. Assumindo que um conjunto de eventos  $W$  com pro-

---

<sup>3</sup>O Apêndice A apresenta a demonstração do Teorema 1, como definido por Shannon. A demonstração do Teorema 2 pode ser encontrada em [13]

babilidades  $p_i$ , considerando o índice  $i$  para os eventos  $W$ , a entropia de Shannon é definida como na equação 2.18. Considerando a quantidade  $I_i = -\log p_i$  como a informação própria de cada evento, conforme definido na seção 2.1.2, e tomando  $p_i$  e  $p'_i$  como probabilidades para dois conjuntos de eventos, podemos afirmar que a diferença de informação obtida através destas duas medidas é

$$\Delta I_i = -(\log p'_i - \log p_i) \quad (2.21)$$

A taxa média da informação modificada pode ser obtida através da expressão a seguir

$$D(p : p') = \sum_i p_i \cdot \Delta I_i = \sum_i p_i \cdot \log \frac{p_i}{p'_i} \quad (2.22)$$

Esta é a definição convencional para o ganho de informação Kullback-Leibler, também conhecido como entropia relativa.

Em recente trabalho publicado, L. Borland *et al* [14] generalizou o ganho de informação Kullback-Leibler para a estatística não extensiva. Uma medida Kullback-Leibler generalizada se deriva naturalmente da aplicação do formalismo da entropia Tsallis no lugar do convencional de Shannon. A partir da equação 2.15 podemos definir portanto, a informação própria não extensiva de cada evento, ou seja,  $I_i^q = -\frac{(p_i^{1-q}-1)}{1-q}$ . Considerando novamente  $p_i$  e  $p'_i$  como as probabilidades para dois conjuntos de eventos medidos, a diferença de informação entre as medidas é

$$\Delta I_i^q = \left[ \frac{1}{(1-q)} \right] \cdot [(1 - p_i^{1-q}) - (1 - p_i'^{1-q})] \quad (2.23)$$

A taxa média da informação modificada pode ser obtida através da expressão a seguir

$$D_{KL_q}(p : p') = \sum_i \frac{p_i^q}{1-q} \cdot (p_i^{1-q} - p_i'^{1-q}) \quad (2.24)$$

que representa a entropia relativa generalizada. Esta expressão pode ser escrita de

outra forma, considerando a função q-logarítmica,  $\log_q(p) = \frac{p^{1-q}-1}{1-q}$  definida em [15], resultando em

$$D_q(p : p') = \sum_i p_i \cdot \log_q \frac{p_i}{p'_i} \quad (2.25)$$

e é válido ressaltar que  $\log_1(p)$  retorna à expressão convencional  $\log(p)$ .

## Capítulo 3

# Processamento Digital de Imagens

As imagens são produzidas por uma variedade de dispositivos físicos, tais como câmeras e vídeo câmeras, equipamentos de radiografia, microscópios eletrônicos, magnéticos e de força atômica, radares, equipamento de ultra-som, entre vários outros. A produção e utilização de imagens podem ter diversos objetivos, que vão do puro entretenimento até aplicações militares, médicas ou tecnológicas. O objetivo da análise de imagens, seja por um observador humano ou por uma máquina, é extrair informações úteis e relevantes para cada aplicação desejada.

Em geral, a imagem pura, recém adquirida pelo dispositivo de captura, necessita de transformações e realces que a torne mais adequada para que se possa extrair o conteúdo de informação desejada com maior eficiência. O Processamento Digital de Imagens (PDI) é uma área da eletrônica/teoria de sinais em que imagens são convertidas em matrizes de números inteiros, sendo que cada elemento desta matriz é composta por um elemento fundamental: o *pixel* (uma abreviação de *picture element*). A partir desta matriz de pixels que representa a imagem, diversos tipos de processamento digital podem ser implementados por algoritmos computacionais. A aplicação destes algoritmos realizam as transformações necessárias para que se possa, por exemplo, obter uma imagem com os realces pretendidos ou extrair atributos ou informações pertinentes. Assim como definido no capítulo 1, o PDI é considerado neste trabalho como a união das áreas de processamento de imagem e visão computacional.

Este capítulo tem como objetivo apresentar alguns conceitos fundamentais do PDI,

procurando introduzir o formalismo e as notações que serão utilizados nos capítulos posteriores. Inicialmente será apresentada a forma de representação da imagem digital e em seguida serão descritas as etapas fundamentais de um sistema de PDI.

### 3.1 Representação da Imagem Digital

Uma imagem monocromática é uma função bidimensional  $f(x, y)$  da intensidade luminosa, onde  $x$  e  $y$  denotam coordenadas espaciais, que por convenção:  $x = [1, 2, \dots, M]$  e  $y = [1, 2, \dots, N]$ . O valor de  $f$  no ponto  $(x, y)$  é proporcional ao brilho (ou nível de cinza) da imagem neste ponto, como ilustrado na figura 3.1. Esta figura <sup>1</sup> apresenta uma região em destaque em que se pode observar os pixels e os níveis de cinza ou níveis de luminância de cada um deles.



Figura 3.1: Imagem monocromática “Goldhill” com destaque para uma região de  $17 \times 17$  pixels

#### Pixel e Conectividade

Um pixel é o elemento básico em uma imagem. A forma mais comum para o pixel é a forma retangular ou quadrada. O pixel é também um elemento de dimensões finitas na representação de uma imagem digital. Frequentemente, a organização de uma imagem

---

<sup>1</sup>A imagem “Goldhill” é frequentemente utilizada para testes e demonstrações em PDI



sob a forma de uma matriz de pixels é feita em uma simetria quadrada. Isto se deve a facilidade de implementação eletrônica, seja dos sistemas de aquisição seja dos sistemas de visualização de imagens. Este tipo de organização provoca o aparecimento de dois problemas importantes nas técnicas de processamento. Em primeiro lugar um pixel não apresenta as mesmas propriedades em todas as direções, isto é, ele é anisotrópico. Esta propriedade faz com que um pixel tenha quatro vizinhos de borda e quatro vizinhos de diagonal, ilustrado na figura 3.2. Esta propriedade obriga que seja definido o tipo de conectividade que será utilizada, ou B4 (considerando apenas os vizinhos de borda) ou B8 (considerando os vizinhos de borda e os de diagonal). O segundo problema é consequência direta do primeiro, ou seja, as distâncias entre um ponto e seus vizinhos não é a mesma para qualquer tipo de vizinho. Será igual a 1 para vizinhos de borda e  $\sqrt{2}$  para aqueles na diagonal.

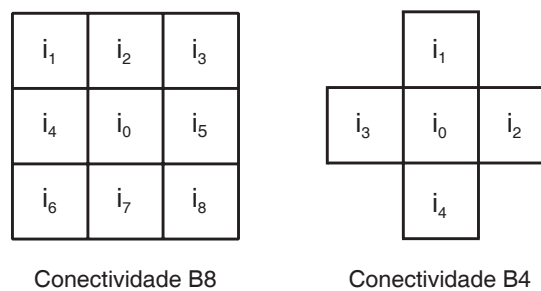


Figura 3.2: Ilustração de tipos de conectividade dos pixels vizinhos ao pixel central  $i_0$ . Conectividade B8 apresenta 8 vizinhos, sendo 4 de bordas e 4 diagonais. Conectividade B4 apresenta apenas os pixels de borda.

## 3.2 Etapas do Processamento de Imagens

Um sistema de processamento de imagens é constituído de diversas etapas, tais como: formação e aquisição da imagem, digitalização, pré-processamento, segmentação, pós-processamento, extração de atributos, classificação e reconhecimento, como ilustra a figura 3.3. A seguir, serão apresentadas breves descrições de cada uma destas etapas.

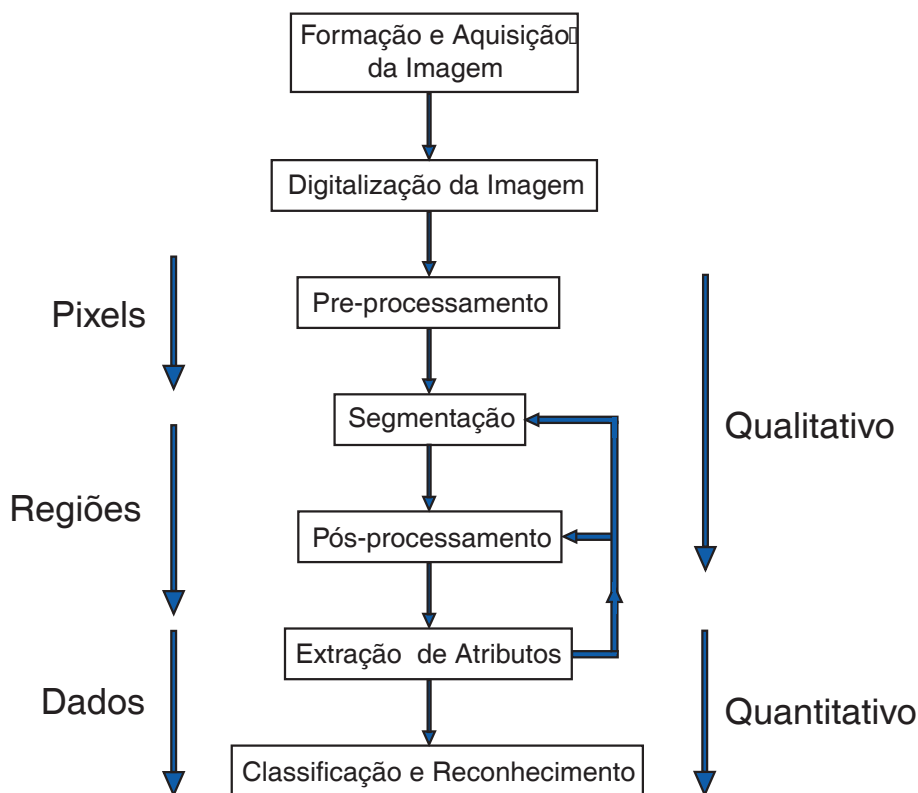


Figura 3.3: *Etapas de um sistema de PDI.*

### 3.2.1 Aquisição de Imagens Digitais

Dois elementos são necessários para a aquisição digital de imagens. O primeiro é um dispositivo físico que deve ser sensível ao espectro de energia eletromagnético, como por exemplo ao espectro de raio-x, luz ultravioleta, visível, ou infravermelha. Este dispositivo transdutor deve produzir em sua saída um sinal elétrico proporcional ao nível de energia percebido. O segundo, chamado digitalizador, é um dispositivo que converte o sinal elétrico analógico produzido na saída do sensor em um sinal digital.

Em laboratórios de física experimental diversos processos levam a uma organização bi-dimensional de dados, buscando uma representação dos mesmos sob a forma de imagens. A grande vantagem é que esta representação permite que ao observarmos todo o conjunto de dados, nós possamos bucar correlações espaciais entre eles.

### 3.2.2 Técnicas de Pré-processamento

As técnicas de pré-processamento têm a função de melhorar a qualidade da imagem. Estas técnicas envolvem duas categorias principais: métodos que operam no domínio espacial e métodos que operam no domínio da frequência. Técnicas de processamento no domínio espacial baseiam-se em filtros que manipulam o plano da imagem, enquanto que as técnicas de processamento no domínio da frequência se baseiam em filtros que agem sobre o espectro da imagem. É comum para realçar determinadas características de uma imagem, combinar vários métodos que estejam baseados nestas duas categorias. A figura 3.4 ilustra um pré-processamento simples: a aplicação de um filtro mediana, para redução de ruído e em seguida um filtro passa-altas, para realce nos contornos ou bordas dos objetos na imagem.

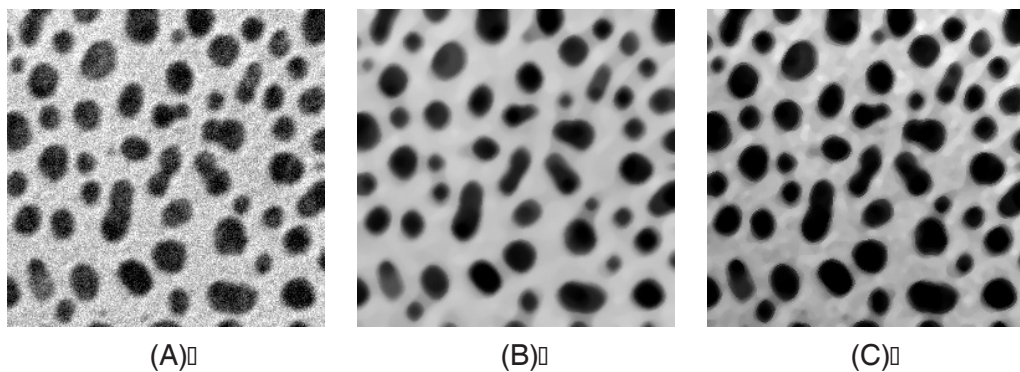


Figura 3.4: *Exemplo de um pre-processamento simples: (A) Imagem original corrompida com ruído gaussiano, (B) Imagem após a aplicação de um filtro mediana para redução do ruído, e (C) Imagem final, após a aplicação de um filtro passa-altas para realce dos contornos.*

### O Histograma de Luminância

O histograma de uma imagem digital com  $k$  níveis de cinza é definido por uma função discreta

$$p(k) = \frac{n_k}{n} \quad (3.1)$$

em que o argumento  $k$  representa os níveis de luminância discretos,  $n_k$  representa o número de pixels na imagem com intensidade  $k$  e  $n$  é o número total de pixels da imagem, ou seja,  $n = M \times N$ . De forma simplificada, podemos afirmar que o histograma de luminância de uma imagem representa a contagem dos níveis de cinza da imagem, podendo informar a distribuição dos pixels dentro dos  $k$  níveis possíveis. O histograma pode ser considerado como uma função distribuição de probabilidades, obedecendo aos axiomas e teoremas da teoria de probabilidades, *i.e.* que  $\sum_k p(k) = 1$ .

O histograma da imagem digital é uma ferramenta bastante útil na etapa de pré-processamento, fornece uma visão estatística sobre a distribuição dos pixels, sobre o contraste da imagem e os níveis de iluminação. Além disso, o histograma é bastante utilizado na etapa de segmentação, principalmente em técnicas que se utilizam da similaridade entre os pixels. O histograma é utilizado com frequência como sendo uma distribuição estatística dos pixels (“luminancia”) na imagem, como por exemplo no caso das técnicas que o utilizam para calcular a entropia da imagem.

Na figura 3.5 são apresentadas duas imagens e seus histogramas, sendo que a imagem (A) é uma imagem de baixo contraste, enquanto a imagem (B) possui um maior contraste.

A figura 3.6 apresenta outras duas imagens com histogramas de tipos bimodal e multimodal. O histograma bimodal é classificado desta forma devido a presença de dois picos, ou duas regiões de luminância com maior incidência de pixels da imagem. Este é o caso típico de imagens que apresentam objetos e fundo de maneira bem definida, como a imagem (A) da figura 3.6. O histograma multimodal apresenta os pixels distribuídos em mais de duas regiões de níveis de cinza, o que se pode perceber através dos picos no histograma em torno destes valores de luminância. Este é o caso da imagem (B) da figura 3.6, que apresenta três picos bem definidos.

### 3.2.3 Segmentação

Segmentar uma imagem significa, de modo simplificado, separar a imagem como um todo nas partes que a constituem e que se diferenciam entre si. É usual denominar

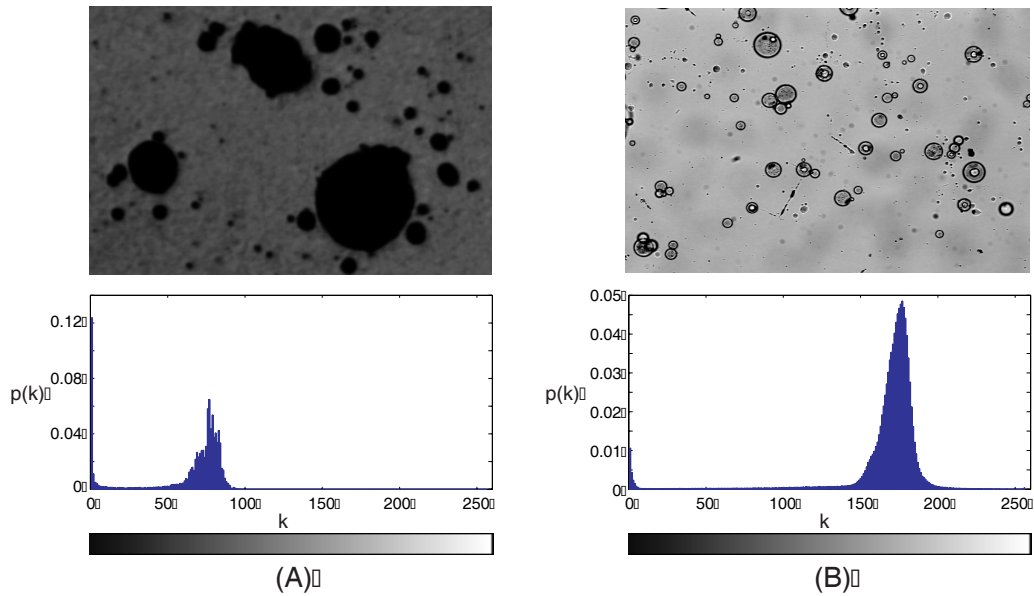


Figura 3.5: *Exemplo de imagens com histogramas diferenciados. (A) Imagem de baixo contraste e seu histograma de luminância. (B) Imagem de alto contraste e seu histograma de luminância.*

“objetos” da imagem os grupos de pixels de interesse, ou que fornecem alguma informação para o PDI. Da mesma forma, a denominação “fundo” da imagem é utilizada para o grupo de pixels que podem ser desprezados ou que não têm utilidade no PDI. Essas denominações “objeto” e “fundo” possuem uma conotação bastante subjetiva, podendo se referir a grupos de pixels que formam determinadas regiões na imagem sem que representem um objeto, de modo literal, presente na imagem processada.

A segmentação é considerada, dentre todas as etapas do processamento de imagens, a etapa mais crítica do tratamento da informação. É na etapa de segmentação que são definidas as regiões de interesse para processamento e análise posteriores. Como consequência deste fato, quaisquer erros ou distorções presentes nesta etapa se refletem nas demais etapas, de forma a produzir ao final do processo resultados não desejados que podem contribuir de forma negativa para a eficiência de todo o processamento.

Deve ser ressaltado que não existe um modelo formal para a segmentação de imagens. A segmentação é um processo empírico e adaptativo, procurando sempre se adequar às características particulares de cada tipo de imagem e aos objetivos que se pretende alcançar. Apesar de existir uma grande diversidade de técnicas de seg-

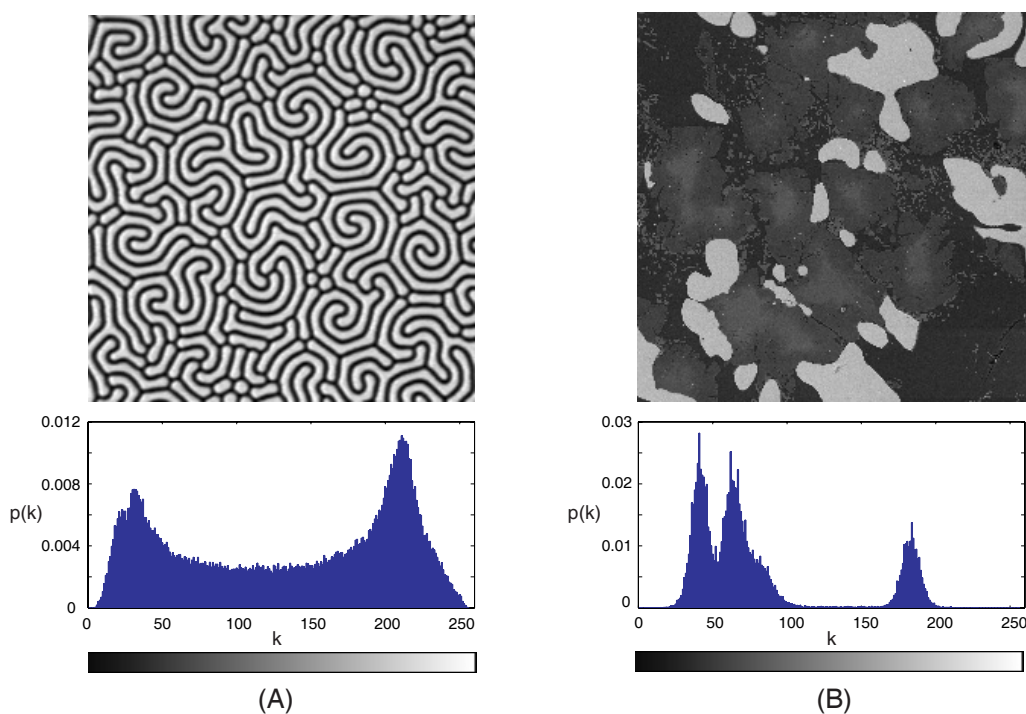


Figura 3.6: Exemplo de imagens com histogramas bimodal e multimodal. A imagem (A) possui um histograma bimodal típico, com a presença de dois picos bem definidos. A imagem (B) possui um histograma do tipo multimodal, apresentando mais de duas regiões de concentração dos pixels da imagem.

mentação de imagens, ainda assim existe atualmente, um grande interesse no estudo e desenvolvimento de novas técnicas.

De um modo geral, as técnicas de segmentação utilizam duas abordagens principais: a similaridade entre os pixels e a descontinuidade entre eles. Sem dúvida, a técnica baseada em similaridade mais utilizada é a chamada *binarização*. A binarização de imagens ou *image thresholding* é uma técnica eficiente e simples do ponto de vista computacional, sendo portanto largamente utilizada em sistemas de visão computacional. Este tipo de segmentação é utilizado quando as amplitudes dos níveis de cinza são suficientes para caracterizar os “objetos” presentes na imagem. Na binarização, um nível de cinza é considerado como um limiar de separação entre os pixels que compõem os objetos e o fundo. Nesta técnica, se obtém como saída do sistema uma imagem binária, i.e., uma imagem com apenas dois níveis de luminância: preto e branco. A de-

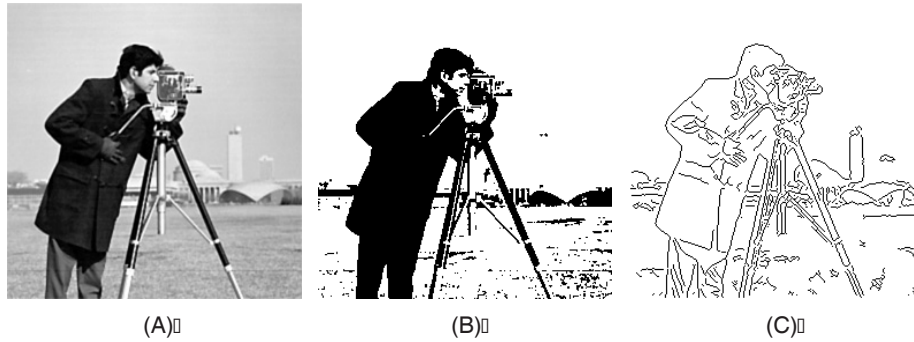


Figura 3.7: *Duas abordagens para segmentação. (A) Imagem original em níveis de cinza. (B) Imagem segmentada através de uma binarização. (C) Imagem segmentada por detecção de bordas.*

terminação deste limiar de modo otimizado para segmentação da imagem é o objetivo principal dos diversos métodos de binarização existentes.

As técnicas baseadas em descontinuidade entre os pixels procuram determinar variações abruptas do nível de luminância entre pixels vizinhos. Estas variações, em geral, permitem detectar o grupo de pixels que delimitam os contornos ou bordas dos objetos na imagem. A técnica de segmentação baseada em descontinuidade mais utilizada é a chamada detecção de bordas. A figura 3.7 apresenta dois exemplos de segmentação, uma binarização e uma detecção de bordas. A imagem (A) é a imagem original em níveis de cinza, a imagem (B) foi segmentada por binarização e a imagem (C) foi segmentada por detecção de bordas.

### 3.2.4 Pós-processamento

O pós-processamento geralmente é a etapa que sucede a segmentação. É nesta etapa que os principais defeitos ou imperfeições da segmentação são devidamente corrigidos. Normalmente, estes defeitos da segmentação são corrigidos através de técnicas de Morfologia Matemática, com a aplicação em seqüência de filtros morfológicos que realizam uma análise quantitativa dos pixels da imagem.

## Operações Morfológicas Básicas

A Morfologia Matemática (MM) é uma das grandes áreas do Processamento Digital de Imagens. Todos os métodos descritos pela MM são fundamentalmente baseados em duas linhas: os operadores booleanos de conjuntos (união, interseção, complemento etc.) e a noção de forma básica, chamado de “elemento estruturante”. As operações são realizadas sempre entre a imagem e o elemento estruturante. A forma do elemento estruturante é função do tratamento desejado e do tipo de conectividade adotada (B4 ou B8).

Dois operadores básicos são utilizados na maior parte das técnicas de MM: a erosão e a dilatação. Consideremos por exemplo, um objeto  $X$  como um grupo de pixels  $x$  delimitado por uma linha tracejada, como mostrado na figura 3.8. A operação de erosão consiste em eliminar do conjunto  $X$  os pixels  $\mathbf{x}$  em função do elemento estruturante  $B$ , tal que:

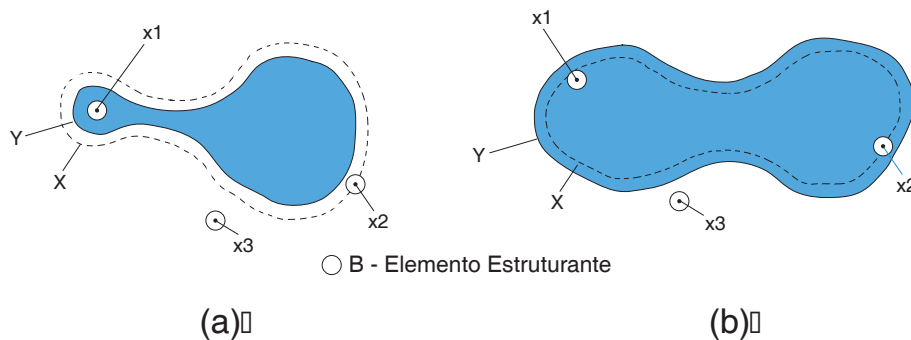


Figura 3.8: Operação de erosão (a) e dilatação (b) aplicada a um elemento estruturante  $B$ . Obtemos a forma  $Y$ .  $\mathbf{x}_1$  é um ponto de  $X$  que também pertence a  $Y$ . Em (a)  $\mathbf{x}_2$  pertence a  $X$ , mas não a  $Y$  devido a operação de erosão. Por outro lado, em (b)  $\mathbf{x}_2$  passa a pertencer a  $Y$  devido a dilatação.  $\mathbf{x}_3$  não pertence a ambos  $X$  e  $Y$ .

$$Y = E^B(X) \rightarrow Y = \{x/B(x) \subset X\} \quad (3.2)$$

onde  $B(x)$  é o elemento estruturante centrado no pixel  $\mathbf{x}$ . Na prática este procedimento corresponde a construir um novo conjunto de pontos  $Y$ , a partir do conjunto



$X$ , tal que o elemento estruturante esteja inserido totalmente em  $X$ . A figura 3.8(a) apresenta o resultado da operação de erosão para 3 pontos distintos da imagem,  $\mathbf{x}_1$ ,  $\mathbf{x}_2$  e  $\mathbf{x}_3$ . O elemento estruturante  $\mathbf{B}$  (representado por um círculo), centrado em  $\mathbf{x}_1$ , está totalmente inserido em  $X$ . O ponto  $\mathbf{x}_1$  pertencerá a forma resultante  $Y$ . O ponto  $\mathbf{x}_2$ , na borda de  $X$ , e  $\mathbf{x}_3$  fora de  $X$  não farão parte de  $Y$ .

A operação dual da Erosão é a Dilatação. Esta operação consiste em dilatar o objeto  $X$  com o elemento estruturante  $B$  tal que:

$$Y = D^B(X) \rightarrow Y = \left(E^B(X^c)\right)^c = \{x/B(x) \cap \subset X \neq 0\} \quad (3.3)$$

em que  $c$  representa o complemento da operação booleana. A figura 3.8(b) apresenta o resultado da operação de dilatação para 3 pontos da imagem.

A operação de erosão permite separar objetos que se tocam. Ao contrario, a operação de dilatação permite preencher furos no interior de um objeto ou mesmo ligá-los. Este resultado dependerá da forma do elemento estruturante. Como as duas operações são iterativas é possível realizar uma seqüência de  $N$  operações de erosão e dilatação sucessivas ou mesmo alternadas. A operação de abertura, uma erosão seguida de uma dilatação, permite eliminar pequenas partículas na imagem (partículas do tamanho do elemento estruturante) e suavizar o contorno dos objetos. Inversamente, a operação de fechamento, uma dilatação seguida de uma erosão, permite fechar canais estreitos que separam objetos, ou suprimir os pequenos furos no seu interior.

### 3.2.5 Extração de Atributos

A etapa final de um sistema de processamento de imagens é aquela em que se extrai as informações úteis da imagem processada. Quando o objetivo do processamento é obter informações numéricas, realiza-se a extração de atributos da imagem.

#### Rotulação ou *Labelização*

A etapa chamada Labelização ou Rotulação é uma etapa intermediária na extração de atributos. Após a etapa de segmentação obtemos uma imagem onde as regiões

correspondentes aos “objetos” estão separadas daquelas correspondentes ao “fundo” da imagem. Neste ponto do sistema de processamento, as regiões de interesse estão contiguamente agrupadas por pixels que se tocam. O próximo passo é dar um rótulo (ou *label*) para cada um desses grupos de pixels. Esta identificação permitirá posteriormente parametrizar os objetos segmentados calculando para cada região de pixels contíguos um parâmetro específico, como área ou perímetro por exemplo. A figura 3.9 apresenta um exemplo desta técnica para uma imagem constituída de células bem delimitadas entre si. O processo de segmentação separa as regiões pertencentes às células daquelas pertencentes às regiões entre células (fundo), criando um delimitador entre elas. A etapa de “labelização” cria um rótulo que identifica cada uma dessas regiões para que os processos seguintes de tratamento da informação sejam concentrados em cada uma das regiões que receberam um rótulo.

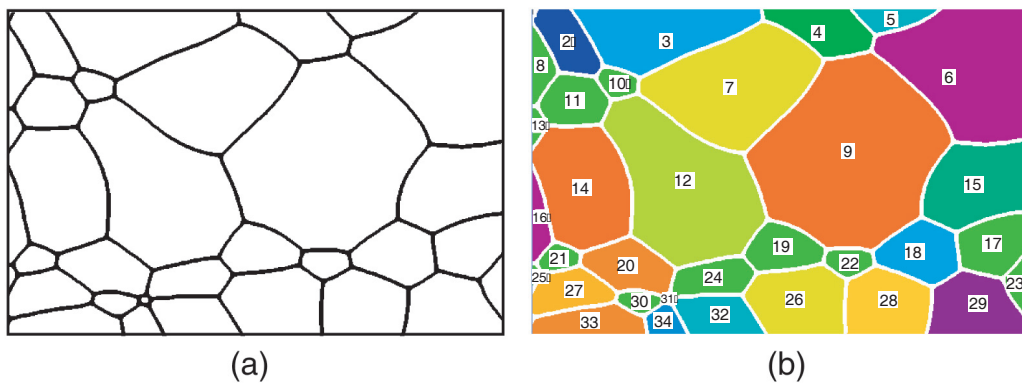


Figura 3.9: Imagem ‘Labelizada’: (a) Imagem original composta por regiões contíguas de pixels. (b) Imagem final após o processo de rotulação. As cores são utilizadas para auxiliar na visualização das regiões.

### Atributos da Imagem

Existem basicamente duas classes de medidas: *i*) atributos da imagem como um todo (*field features*), por exemplo número de objetos, área total de objetos, etc. e *ii*) atributos de região (*region features*) que se referem aos objetos independentemente, por exemplo área, perímetro, forma, etc. Os atributos de região podem ser muito sofisticados, permitindo uma nova separação dos objetos em classes de similaridades,

em função dos parâmetros medidos. A figura 3.10 apresenta os principais atributos de região que podem ser extraídos de uma imagem, após as etapas de segmentação e pós-processamento.

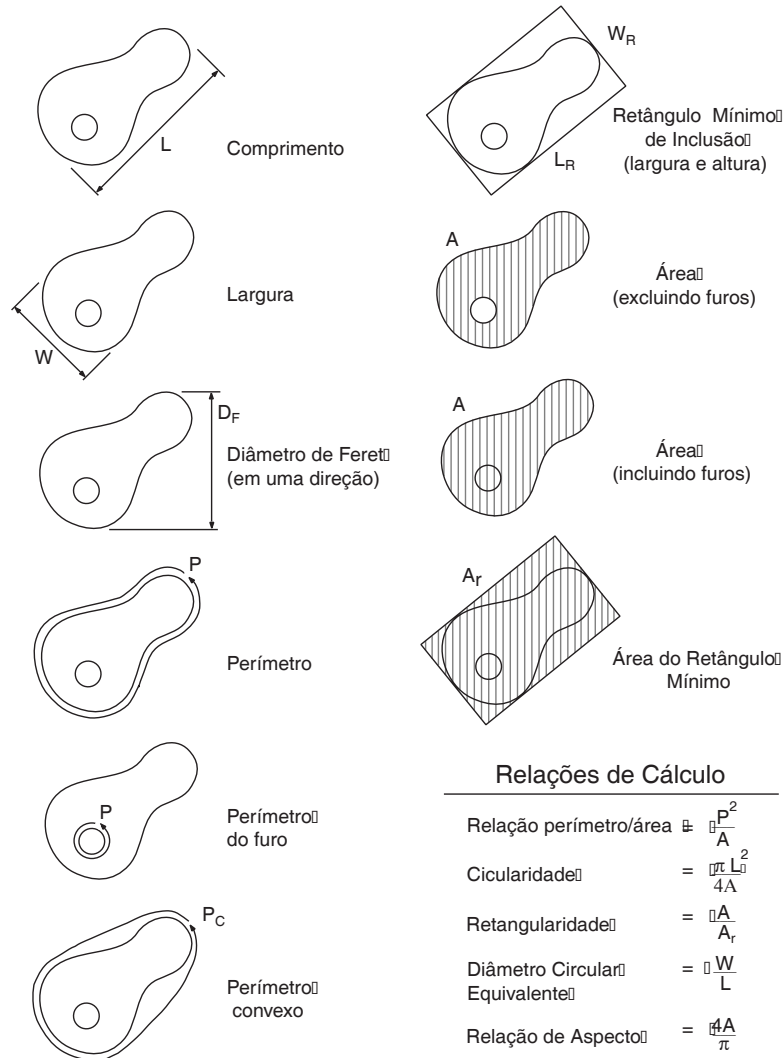


Figura 3.10: Principais atributos de região, ou seja, dos objetos independentes presentes na imagem.

### 3.2.6 Classificação e Reconhecimento

O objetivo do reconhecimento é realizar, de forma automática, a “identificação” dos objetos segmentados na imagem. Existem duas etapas no processo de classificação de formas: o aprendizado e o reconhecimento propriamente dito. Na maior parte dos sistemas de reconhecimento de formas, os parâmetros provenientes da etapa de extração

de atributos são utilizados para construir um espaço de medida à  $N$  dimensões. Os sistemas de aprendizado irão definir uma função discriminante que separe eficientemente todas as formas representadas neste espaço de medida.

O espaço de medidas é um espaço de dimensão  $N$  onde cada dimensão corresponde a um parâmetro (atributo). A escolha e a qualidade dos parâmetros para construir este espaço é essencial para a realização de um bom processo de reconhecimento de forma. A escolha de um grande número de parâmetros leva a um espaço de grande dimensão e a uma difícil etapa de aprendizado. Um pequeno espaço de medidas pode acarretar em uma baixa caracterização da forma e a muitos erros no processo de reconhecimento. Em alguns casos pode ser interessante normalizar cada eixo para construir um espaço que seja bem distribuído e facilite o processo de classificação.

Podemos dividir o processo de aprendizado em dois tipos: os métodos supervisionados e não supervisionados. No método supervisionado, o classificador, em sua fase de aprendizado, recebe informações de como as classes devem ser identificadas. Por exemplo, em um sistema de reconhecimento de caracteres, existirá classes independentes para todas as letras do alfabeto. Por exemplo a classe das letras “**A**” será representada pela ocorrência da letra A em suas diferentes variações. O aprendizado consiste então em apresentar ao sistema uma série de objetos “teste” no qual suas classes são previamente conhecidas e definir uma função que separe todas as classes entre si.

Podemos dizer que o sistema de aprendizado supervisionado age sob a supervisão de um outro sistema de reconhecimento (externo por exemplo) que identificou anteriormente os objetos testes é permitirá a construção correta de seu espaço de medida e sua função discriminante. Durante este processo devemos modificar os parâmetros que compõem o espaço de medida e permitir um melhor ajuste da função discriminante, objetivando sempre que o sistema possa realizar com mais eficiência o processo de classificação. Ao final, é possível determinar a função discriminante responsável pela separação das diversas classes. Este processo pode ser lento e de elevado custo computacional. Somente então os objetos desconhecidos serão fornecidos a este classi-

ficador, na fase de reconhecimento.

No caso em que a classificação não é supervisionada, o classificador receberá os objetos desconhecidos e, a partir da medida dos diferentes parâmetros (atributos dos objetos presentes na imagem), ele tentará alocá-los em diferentes classes. A identificação de classes é usualmente realizada a partir da identificação de agrupamentos em “clusters” de objetos no espaço de medidas.

Existem atualmente vários métodos de reconhecimento de formas. Entre eles podemos citar os métodos baseados em propriedades estatísticas dos objetos (classificador de Bayes), os métodos baseados em técnicas de distância entre os objetos na imagem e suas formas padrões (como as redes neurais artificiais) ou ainda a descrição da forma através de um dicionário ou uma linguagem básica. Neste último caso é definido uma seqüência de elementos básicos que representem as formas dos objetos. Em seguida é construída uma linguagem e formada uma gramática. O reconhecimento é neste caso um reconhecimento sintático baseado nesta gramática [16]. O reconhecimento por meio de redes neurais artificiais é um método bastante atraente, pois consegue dar um maior grau de liberdade a função discriminante. Nesta tese apresentamos de forma introdutória, no capítulo 6, um método de classificação de formas baseado na técnica de entropia cruzada entre objetos da imagem e padrões previamente definidos.



# Capítulo 4

## Segmentação Entrópica de Imagens

Neste capítulo são apresentadas técnicas de segmentação utilizando o cálculo da entropia da imagem. A entropia da imagem pode ser calculada a partir de um jogo de probabilidades que, em geral, é obtido através do histograma de luminância. A rigor, qualquer distribuição de probabilidades obtida através da imagem pode ser considerada para o cálculo da entropia. Isto quer dizer que a entropia de uma imagem não está limitada apenas a utilização dos valores de luminância, podendo ser consideradas outras quantidades extraídas da imagem, como por exemplo atributos espaciais.

Nos métodos de segmentação apresentados neste capítulo, a entropia da imagem será calculada considerando o histograma de luminância como uma distribuição de probabilidades. Os métodos de segmentação que utilizam o cálculo de entropia da imagem apresentam vantagens em relação a outros métodos de segmentação globais devido a menor sensibilidade a presença de ruídos na imagem, e conseqüentemente exigem pouco ou nenhum processamento anterior a etapa de segmentação.

### 4.1 Principais Abordagens

A técnicas de segmentação de imagens em níveis de cinza seguem basicamente duas abordagens: a descontinuidade entre pixels da imagem e a similaridade entre eles.

Os métodos que utilizam a descontinuidade entre pixels procuram detectar variações abruptas do nível de cinza entre pixels vizinhos como um indicativo da presença de um

objeto a ser destacado ou segmentado. A principal técnica de segmentação baseada em descontinuidade é a detecção de bordas. O método mais simples para detecção de bordas é o cálculo do gradiente de uma imagem através de uma máscara de convolução tipo passa-altas que destaca contornos de objetos presentes na imagem.

Existem muitas formas de realizar a detecção de bordas em imagens. Entretanto, a maioria destes métodos podem ser agrupados em duas categorias: Gradiente e Laplaciano. Os métodos de Gradiente detectam as bordas procurando o máximo e o mínimo na primeira derivada da imagem. Como exemplos de métodos de Gradiente, podemos citar: Roberts, Prewitt e Sobel. Os métodos de Laplaciano procuram por passagens por zero na derivada segunda da imagem para detectar as bordas. Um exemplo desta abordagem é o método Marrs-Hildreth [17]. O Apêndice B apresenta uma breve descrição dos métodos citados anteriormente, com as máscaras de convolução de cada um deles e alguns resultados ilustrativos.

Os métodos de segmentação que utilizam a similaridade entre os pixels procuram destacar grupos de pixels que possuem valores similares de iluminação ou níveis de cinza. O principal método de segmentação por similaridade é também o mais utilizado em sistemas de visão computacional, conhecido como *image thresholding* ou binarização. A segmentação por este método é definida por um limiar global de iluminação que deve distinguir dois grupos de pixels, “objeto” e “fundo”, e formar uma imagem que apresenta apenas dois níveis de cinza, ou seja, uma imagem binária. A palavra objeto destacada anteriormente possui um significado bastante subjetivo neste contexto e representa a região ou o conjunto de regiões de pixels que se pretende destacar ou segmentar em uma imagem.

A binarização se destaca entre os diversos métodos de segmentação de imagens devido ser relativamente simples e de fácil implementação computacional, através de algoritmos rápidos. Por este motivo é o método mais utilizado em sistemas de visão automatizados que atuam em aplicações médicas e industriais.

Existem diversas técnicas de segmentação baseadas em similaridade e descontinuidade e ainda outras aproximações que utilizam morfologia matemática, casamento de padrões,



etc. Nenhuma destas técnicas pode ser considerada capaz de segmentar todos os tipos de imagem, não existindo portanto um modelo formal de segmentação, mas diversas abordagens para os diversos casos e objetivos específicos.

## 4.2 Métodos Estatísticos

Vários métodos de binarização tem sido propostos com o objetivo de realizar a segmentação de imagens. Muitos destes métodos utilizam a informação sobre a iluminação da imagem presente no histograma de níveis de cinza. A forma do histograma, ou seja, as regiões formadas devido a maior ocorrência de determinados níveis de cinza é largamente utilizada em métodos de binarização automática. Um exemplo é o método que procura mínimos locais entre dois ou mais picos de histogramas bimodais, trimodais, etc. Estes métodos baseados na forma do histograma são extremamente sensíveis ao ruído e freqüentemente exigem etapas de pré-processamento a fim de ‘suavizar’ o histograma da imagem.

Os métodos entrópicos que utilizam o histograma da imagem não determinam o limiar ótimo para a binarização através da forma do histograma. Nestes métodos, a entropia da imagem é calculada utilizando-se o histograma normalizado pelo número total de pixels da imagem, de modo a equivaler a uma distribuição de probabilidades. O primeiro método a utilizar o histograma normalizado como um conjunto probabilidades para o cálculo da entropia da imagem foi proposto por T. Pun [18]. Posteriormente, J. N. Kapur [19] evoluiu as considerações teóricas e práticas do método inicial proposto por Pun, desenvolvendo um novo e eficiente método de binarização entrópica utilizando o histograma da imagem.

Uma outra classe de algoritmos de binarização entrópica, também baseados no histograma, aplicam o conceito de entropia relativa como uma abordagem para a segmentação de imagens. O conceito de entropia relativa definida por Kullback [20], também conhecida como divergência direta, é utilizado em métodos que tentam minimizar a distância teórica entre duas distribuições estatísticas. Na pratica procura-se minimizar a distância estatística entre uma distribuição relativa à imagem original em

níveis de cinza e uma distribuição relacionada com a imagem binária ou com um modelo teórico adotado [21] [22] [23] [24].

Os métodos de binarização entrópica baseados no histograma da imagem, envolvem a otimização de uma função critério definida a partir do cálculo da entropia da imagem, seja a entropia de Shannon ou a entropia relativa de Kullback-Leibler.

#### 4.2.1 Método de Binarização por Soma de Entropias

O método da Soma de Entropias, proposto por Kapur [19], utiliza a propriedade de aditividade da entropia de Shannon e é baseado na maximização da medida de informação entre as duas classes. Se consideramos a imagem digital um sistema que pode ser decomposto em dois subsistemas  $A$  e  $B$  estatisticamente independentes, a entropia do sistema apresenta a propriedade da aditividade, de forma que

$$S = S_A + S_B \quad (4.1)$$

Considerando uma imagem digital com  $M \times N$  pixels, define-se um conjunto de probabilidades  $p$  relativo aos níveis de luminância da imagem, ou seja, relativo à frequência de ocorrência dos níveis de cinza para o conjunto total de pixels da imagem. O histograma de luminância da imagem  $h$ , normalizado pelo número total de pixels da imagem, representa uma boa aproximação para uma distribuição de probabilidades de uma imagem digital. Utilizando esta abordagem, Kapur propôs um método de binarização através do cálculo da entropia da imagem considerando que os pixels da imagem poderiam ser classificados como pertencentes ao “objeto” e ao “fundo”. Ao definir duas classes estatisticamente independentes (objeto e fundo) como componentes de um sistema único (a imagem digital), o método considera que as características de luminância dos pixels que compõem as classes objeto e fundo são independentes entre si.

Considerando  $p = \{p_i\} = \frac{h_i}{M \times N}$  como um conjunto de probabilidades fechado, de modo que  $\sum_i p_i = 1$ , a entropia do sistema com  $k$  eventos pode ser calculada através da expressão

$$S_k = - \sum_{i=1}^k p_i \log p_i \quad (4.2)$$

Para um sistema composto por dois subsistemas A e B, define-se dois conjuntos de probabilidades normalizados:

$$p^A : \frac{p_1}{P_t}, \frac{p_2}{P_t}, \dots, \frac{p_t}{P_t} \quad (4.3)$$

$$p^B : \frac{p_{t+1}}{1 - P_t}, \frac{p_{t+2}}{1 - P_t}, \dots, \frac{p_k}{1 - P_t} \quad (4.4)$$

onde  $P_t = \sum_{i=1}^t p_i$  e o limite superior  $t$  representa o nível de luminância que “separa” a imagem em classes objeto e fundo.

A normalização aplicada aos conjuntos de probabilidades garante que os dois novos conjuntos de probabilidades sejam fechados, de modo que  $\sum p^A = 1$  e  $\sum p^B = 1$ . Como exemplo, se considerarmos um conjunto de probabilidades  $p = \{0.25 \ 0.25 \ 0.25 \ 0.25\}$  e definirmos dois subconjuntos  $p_a = \{0.25\}$  e  $p_b = \{0.25 \ 0.25 \ 0.25\}$  teremos portanto

$$p^A = \left\{ \frac{0.25}{0.25} \right\} \quad e \quad p^B = \left\{ \frac{0.25}{0.75}, \frac{0.25}{0.75}, \frac{0.25}{0.75} \right\} \quad (4.5)$$

que satisfazem a condição de  $\sum_i p_i = 1$ .

As entropias associadas a cada uma das distribuições definidas anteriormente são deduzidas a seguir

$$\begin{aligned} S_A &= \sum_{i=1}^t \frac{p_i}{P_t} \log \frac{p_i}{P_t} \\ &= -\frac{1}{P_t} \left[ \sum_{i=1}^t p_i \log p_i - \sum_{i=1}^t p_i \log P_t \right] \\ &= \frac{1}{P_t} [S_t + P_t \log P_t] = \log P_t + \frac{S_t}{P_t} \end{aligned} \quad (4.6)$$

$$S_B = \sum_{i=t+1}^k \frac{p_i}{1 - P_t} \log \frac{p_i}{1 - P_t}$$

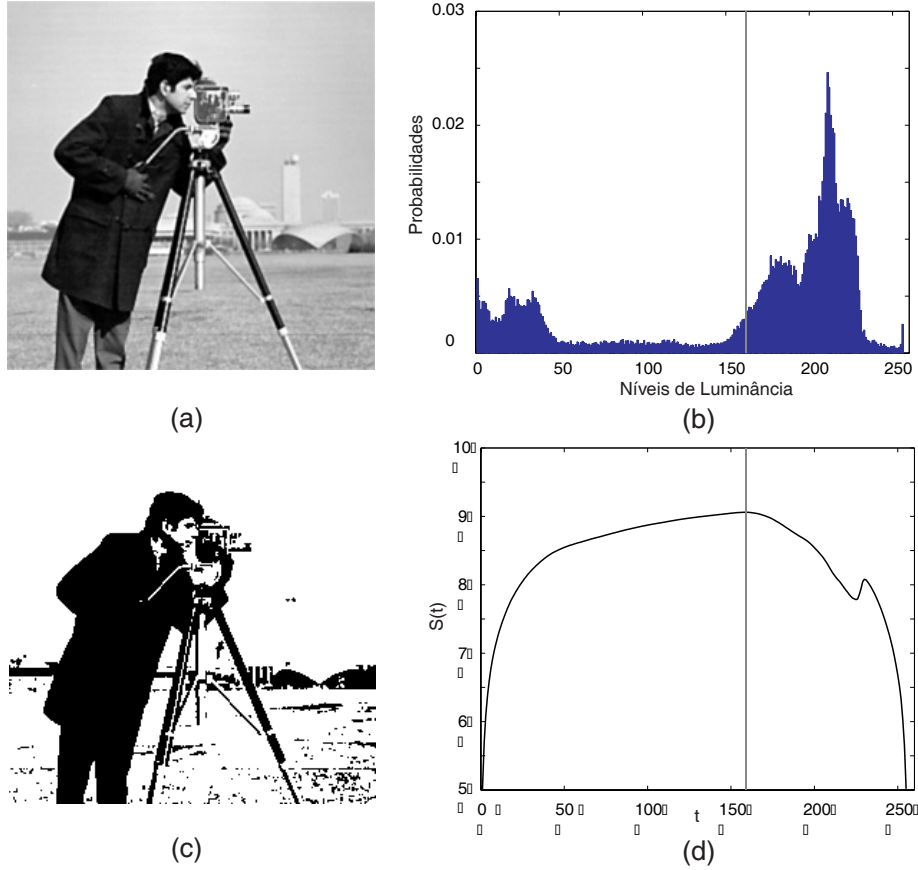


Figura 4.1: (a) Imagem Original - "cameraman" (b) PDF da Imagem Original (c) Imagem Binária ( $t = 160$ ) (d) Função Critério para Binarização

$$\begin{aligned}
 &= -\frac{1}{1 - P_t} \left[ \sum_{i=t+1}^k p_i \log p_i - \sum_{i=t+1}^k p_i \log (1 - P_t) \right] \\
 &= \log (1 - P_t) + \frac{S_k - S_t}{1 - P_t}
 \end{aligned} \tag{4.7}$$

considerando que  $S_t = -\sum_{i=1}^t p_i \log p_i$  de modo análogo à equação 4.2.

Aplicando a propriedade aditiva da entropia definida em 4.1 e tornando esta equação dependente do parâmetro  $t$  ( $t = \text{thresholding level}$ ), definimos  $S(t)$  como uma função critério para seleção do nível ótimo para a binarização das imagens

$$\begin{aligned}
 S(t) &= S_A + S_B = \log P_t + \log (1 - P_t) + \frac{S_t}{P_t} + \frac{S_k - S_t}{1 - P_t} \\
 &= \log P_t (1 - P_t) + \frac{S_t}{P_t} + \frac{S_k - S_t}{1 - P_t}
 \end{aligned} \tag{4.8}$$

O valor discreto de  $t$  que maximiza a função  $S(t)$  é considerando o nível ótimo de binarização.

Este método de binarização entrópica é amplamente utilizado em programas de tratamento de imagens, citando como exemplo o software Visilog da Noesis [25], que implementa exatamente esta técnica de segmentação assim como uma extensão para o *multithresholding*<sup>1</sup>. Algumas extensões deste método podem ser encontradas em [26] [27] [23].

O algoritmo apresentado a seguir descreve como obter o limiar ótimo de segmentação pelo método da soma máxima da entropia das 2 regiões (objeto e fundo).

Begin

Max\_value = LowValue;

For t=2 to t=k-1 do

    Begin

        Calcule  $p^A$ ,  $i = 1, 2, \dots, t$  pela equação 4.3;

        Calcule  $p^B$ ,  $i = s + 1, s + 2, \dots, k$  pela equação 4.4;

        Calcule  $S_A$  pela equação 4.6;

        Calcule  $S_B$  pela equação 4.7;

        Calcule  $S(t)$  pela equação 4.8;

        Se Max\_value <  $S(t)$

            Begin

                Max\_value =  $S(t)$ ;

                Threshold =  $t$ ;

            End

    End

End

---

<sup>1</sup>Técnica que define mais de um limiar ( $t$ ) para a segmentação, para casos em que o histograma apresenta mais de duas regiões (3 picos, por exemplo).

## 4.2.2 Método de Binarização por Entropia Relativa

A entropia Kullback-Leibler pode ser utilizada para determinar o ponto ótimo de segmentação de uma imagem. Brink [28], Li e Lee [24] e N. Pal [21] utilizaram a entropia relativa (também chamada de divergência direta de Kullback) para a seleção do ponto ótimo de binarização da imagem. Os dois primeiros trabalharam diretamente sobre os níveis de cinza da imagem. N. Pal propôs uma modificação do algoritmo trabalhando com o histograma de luminância da imagem e um histograma teórico, como dois jogos de probabilidade a serem aproximados.

Em sua definição inicial Kullback propôs uma distância teórica,  $D$ , entre dois jogos de probabilidade  $p$  e  $p'$ . Sendo  $p = p_1, p_2, \dots, p_k$  e  $p' = p'_1, p'_2, \dots, p'_k$  duas distribuições de probabilidade, então a entropia relativa,  $D_{KL}(p : p')$ , entre os dois jogos é definida como sendo

$$D_{KL}(p : p') = \sum_{i=1}^k p_i \cdot \log \frac{p_i}{p'_i} \quad (4.9)$$

Se  $\sum_i p_i = \sum_i p'_i = 1$  teremos, baseado na desigualdade de Gibbs ( $-\sum p_i \log p_i \leq -\sum p_i \log p'_i$ ) que a distância  $D_{KL}(p : p') \geq 0$ . A distância  $D_{KL}(p : p')$  será zero quando  $p_i = p'_i, \forall i = 1, 2, 3, \dots, k$ . A função distância  $D_{KL}(p : p')$  não é simétrica, de modo que  $D_{KL}(p : p') \neq D_{KL}(p' : p)$ . Uma versão simétrica pode ser escrita da seguinte forma:

$$D(p : p') = D_{KL}(p : p') + D_{KL}(p' : p) \quad (4.10)$$

Tomando como partida que o objetivo é segmentar a imagem em duas regiões (objeto e fundo), podemos considerar a região de níveis de cinza relativa ao objeto, sendo aquela composta pelos pixels no intervalo  $[0, t]$  e a região relativa ao fundo da imagem aquela composta pelos pixels de  $[(t + 1), k]$ .

O método proposto por N. Pal usa a versão simétrica da entropia relativa de Kullback-Leibler:

$$D(p : p') = \sum_{i=1}^k p_i \cdot \log \frac{p_i}{p'_i} + \sum_{i=1}^k p'_i \cdot \log \frac{p'_i}{p_i} \quad (4.11)$$

Tomando o parâmetro  $t$  como o ponto de segmentação ótimo, os dois jogos de probabilidades podem ser definidos como:

$P_O = \{p_1^O, p_2^O, p_3^O, \dots, p_t^O\}$  o jogo de probabilidades dos pixels pertencentes ao objeto e

$P_B = \{p_{t+1}^B, p_{t+2}^B, p_{t+3}^B, \dots, p_k^B\}$  o jogo de probabilidades dos pixels pertencentes ao fundo da imagem, considerando:

$$p_i^O = \frac{h_i}{P_t}, \quad i = 1, 2, 3, \dots, t; \quad P_t = \sum_i^t h_i \quad (4.12)$$

$$p_i^B = \frac{h_i}{1 - P_t}, \quad i = t + 1, t + 2, t + 3, \dots, k; \quad (4.13)$$

onde  $M \times N$  representa o número total de pixels da imagem.

É importante notar que os dois jogos de probabilidades são independentes, i.e.:

$$\sum_i^t p_i^O = \sum_{i=t+1}^k p_i^B = 1 \quad (4.14)$$

Um segundo ponto importante nesta técnica é a definição do jogo de probabilidade  $p'$ . Este jogo de probabilidades é normalmente baseado em algum modelo teórico (ou alguma informação a priori que podemos ter da imagem). Este jogo de probabilidades será utilizado efetivamente na medida de distância entre o modelo ( $p'$ ) e a imagem real ( $p$ ). Vamos assumir o jogo de probabilidade  $P'_O = \{p'_1{}^O, p'_2{}^O, p'_3{}^O, \dots, p'_t{}^O\}$  como a distribuição de probabilidades modelo para o objeto e  $P'_B = \{p'_{t+1}{}^B, p'_{t+2}{}^B, p'_{t+3}{}^B, \dots, p'_k{}^B\}$  para o fundo da imagem.

A entropia relativa entre objeto e fundo da imagem e objeto e fundo do modelo seria então definida por  $D_O(t) = D(P_O : P'_O)$  e  $D_B(t) = D(P_B : P'_B)$  respectivamente. A entropia total seria então escrita pela seguinte expressão:

$$\begin{aligned} D(t) &= D_O(t) + D_B(t) \\ &= \sum_{i=1}^t p_i^O \cdot \log \frac{p_i^O}{p'_i{}^O} + \sum_{i=1}^t p'_i{}^O \cdot \log \frac{p'_i{}^O}{p_i^O} \end{aligned}$$

$$+ \sum_{i=t+1}^k p_i^B \cdot \log \frac{p_i^B}{p_i'^B} + \sum_{i=t+1}^k p_i'^B \cdot \log \frac{p_i'^B}{p_i^B} \quad (4.15)$$

para achar o valor ótimo de binarização da imagem devemos minimizar a distância  $D$  em função de  $t$ .

Porem, é importante ainda especificar a função modelo para o jogo de probabilidades do objeto e do fundo,  $P'_O$  e  $P'_B$ . N. Pal propõe que o jogo de probabilidades que melhor aproxima o histograma de uma imagem está baseado em uma distribuição de Poisson,  $p_i'^O$  e  $p_i'^B$ . Ele assume que os níveis de cinza de uma região da imagem segue a distribuição de Poisson, como parâmetros  $\lambda_O$  e  $\lambda_B$  para objeto e fundo:

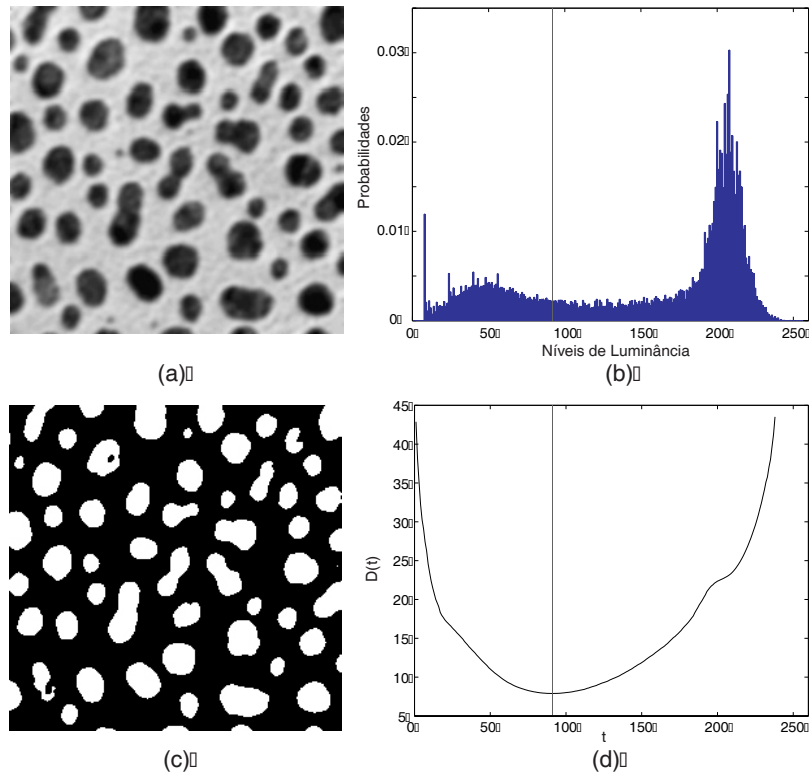


Figura 4.2: (a)Imagem Original (b)PDF da Imagem Original (c)Imagem Binária ( $t = 91$ ) (d)Função critério para Binarização. A posição de menor distância é escolhida como ponto de 'threshold' dos níveis de cinza.

$$p_i'^O = \frac{e^{-\lambda_O} \cdot \lambda_O^i}{i!} \quad i = 1, 2, 3, \dots, t \quad (4.16)$$

e

$$p_i'^B = \frac{e^{-\lambda_B} \cdot \lambda_B^i}{i!} \quad i = t + 1, t + 2, t + 3, \dots, k; \quad (4.17)$$



onde  $\lambda_O$  e  $\lambda_B$  são estimados a partir da imagem (ou de seu histograma) pelas seguintes fórmulas:

$$\lambda_O = \frac{\sum_{i=1}^t i \cdot h_i}{\sum_{i=1}^t h_i} \quad (4.18)$$

e

$$\lambda_B = \frac{\sum_{i=t+1}^k i \cdot h_i}{\sum_{i=t+1}^k h_i} \quad (4.19)$$

O limiar de segmentação ótima baseado na técnica da entropia relativa de Kullback-Leibler pode ser definido pelo seguinte algoritmo:

```

Begin
Min_value = HighValue;
For t=2 to t=k-1 do
    Begin
        Calcule  $p_i^O$ ,  $i = 1, 2, \dots, t$  pela equação 4.12;
        Calcule  $p_i^B$ ,  $i = s + 1, s + 2, \dots, k$  pela equação 4.13;
        Calcule  $\lambda_O$  pela equação 4.18;
        Calcule  $\lambda_B$  pela equação 4.19;
        Calcule  $p_i^O$ ,  $i = 1, 2, \dots, t$  pela equação 4.16;
        Calcule  $p_i^B$ ,  $i = t + 1, t + 2, \dots, k$  pela equação 4.17;
        Calcule  $D(t)$  pela equação 4.15;
        Se Min_value >  $D(t)$ 
        Begin
            Min_value =  $D(t)$ ;
            Threshold =  $t$ ;
        End
    End
End
End

```

## 4.3 Método Detecção de Bordas utilizando Operador Entropia

A detecção de bordas é o método mais utilizado quando a descontinuidade entre os pixels de uma imagem é a abordagem mais adequada para a segmentação. Uma borda em uma imagem digital pode ser definida como sendo o limite entre duas regiões que possuem níveis de cinza relativamente distintos. Este tipo de abordagem para segmentação pressupõe que as regiões, ou “objetos” da imagem, sejam suficientemente homogêneos em luminância. Isto garante que a transição entre estas regiões possam ser determinadas com base apenas nas descontinuidades entre os níveis de cinza. A figura 4.3 ilustra as descontinuidades presente apenas em uma linha selecionada na imagem ( $y = 120$ ).

O método que é apresentado nesta seção utiliza o *operador entropia*, introduzido por A. Shiosaki em [29]. Neste método, a entropia de uma região da imagem é calculada utilizando os valores de intensidade dos pixels desta região, de modo equivalente a um histograma de luminância. O cálculo da entropia permite detectar descontinuidades na luminância da imagem. Uma região de pixels com valores de luminância semelhantes, vão ficar concentrados em um único ‘bin’ do histograma, o que tende a fornecer um valor de entropia nulo. Uma outra região que apresente valores de luminância bastante distintos, denotando uma descontinuidade na imagem, vai apresentar um histograma com estas ocorrências distribuídas entre os ‘bins’. Esta região que apresenta este tipo de histograma possui um valor de entropia maior. Os valores de entropia calculados para cada região representam novos valores de luminância para a imagem segmentada, depois de devidamente ajustados aos limites dos níveis de cinza.

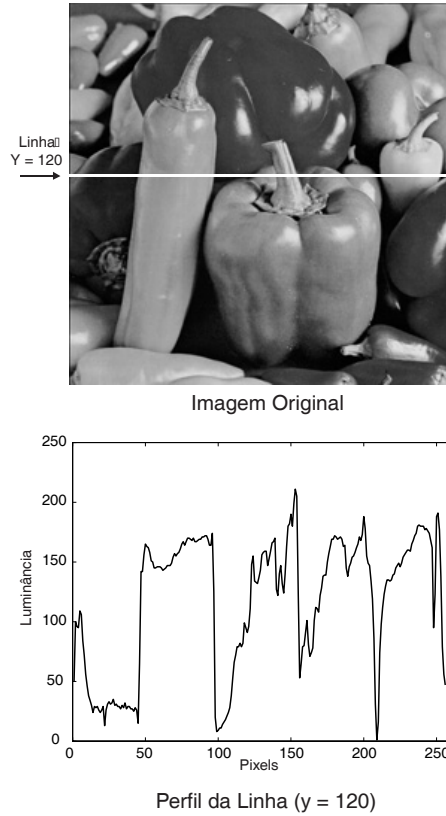


Figura 4.3: *Perfil de uma linha da imagem ( $y = 120$ ) ilustrando as discontinuidades de luminância entre pixels.*

### 4.3.1 Definições Teóricas do Método

A região da imagem é selecionada por uma máscara de convolução, como as apresentadas na figura 4.4. A entropia da região é calculada tendo como referência o pixel central  $a_0$ . Para cada região da imagem, é atribuído ao pixel central da máscara o valor da entropia calculado para a região. Utilizando a técnica de “sliding window” em duas dimensões, o operador entropia é aplicado em toda a imagem, criando uma nova imagem  $S$  que possui os valores de entropia de cada região. A seguir serão apresentadas algumas definições do método:

- Sinal amostrado puro: imagem em níveis de cinza

$$I(i, j) : i = [1, \dots, N] , j = [1, \dots, M] \quad (4.20)$$

em que  $N$  e  $M$  representam as dimensões da imagem.

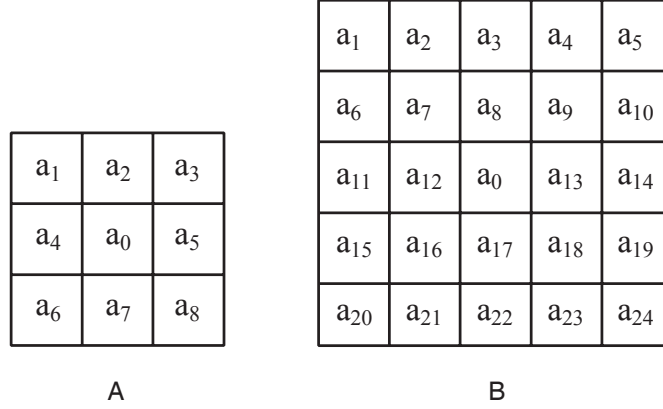


Figura 4.4: Exemplo de máscaras para o cálculo do operador entropia: A máscara de  $3 \times 3$  pixels e B máscara de  $7 \times 7$  pixels.

- Dimensão da janela:

$$w[w_i, w_j] : w_i \leq N, w_j \leq M \quad (4.21)$$

- Passo de deslocamento:

$$\Delta \leq w \quad (4.22)$$

Deste modo, as definições formais para a janela deslocada são:

$$\begin{aligned} W(\{n, m\} : w : \Delta) &= \{I(i, j); i = 1 + n\Delta, \dots, w_i + \Delta n, j = 1 + m\Delta, \dots, w_j + \Delta m\}, \\ n &= 0, 1, 2, \dots, [N/\Delta] - w_i + 1 \\ m &= 0, 1, 2, \dots, [M/\Delta] - w_j + 1 \end{aligned} \quad (4.23)$$

Para cada janela  $W$ , o cálculo da entropia é realizado através da expressão

$$S = \sum_{i=1}^k p_i \cdot \log p_i \quad (4.24)$$

em que  $k$  é o valor da luminância, ou seja,  $k = [0, \dots, 255]$ , no caso de uma imagem de 8 bits. A probabilidade  $p$  é obtida através histograma  $p_L$ , calculado somente para os pixels da máscara, e normalizado pela soma total da luminância destes pixels,  $P_c$ . A probabilidade para o cálculo da entropia pode ser definida como

$$p_i = \frac{p_L}{P_c} = \left\{ \frac{p_1}{P_c}, \frac{p_2}{P_c}, \dots, \frac{p_k}{P_c} \right\} \quad (4.25)$$

onde  $p_L = p_1, p_2, \dots, p_k$  e  $P_c = \sum_{i=1}^k p_L$

A imagem segmentada é obtida calculando a entropia, parametricamente dependente da janela  $W$  definida anteriormente em 4.23.

$$S(W) = \sum_{i=1}^k p(W) \cdot \log p(W) \quad (4.26)$$

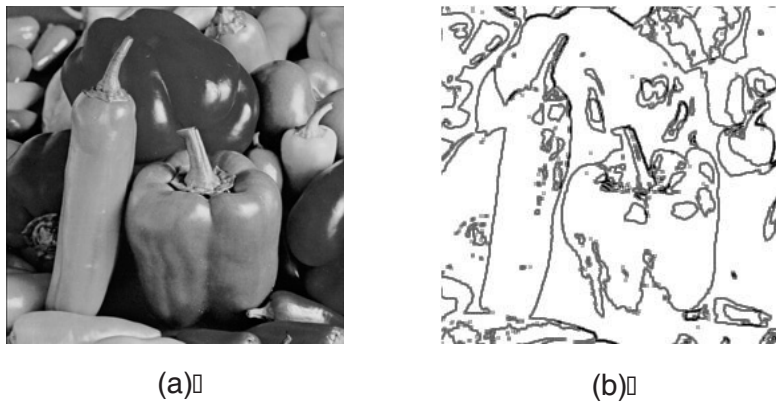


Figura 4.5: (a) Imagem Original (b) Imagem segmentada por detecção de bordas utilizando o operador entropia.

Esta expressão fornece um sinal de duas dimensões com os valores da entropia calculada. A normalização deste valores para os limites dos níveis de cinza, ou seja, 0 e 255 para uma imagem de 8 bits, fornece como saída do sistema a imagem segmentada por detecção de bordas. A figura 4.5 apresenta o resultado da aplicação deste método em uma imagem clássica (*peppers*) para avaliação de métodos de detecção de bordas.



# Capítulo 5

## Segmentação de Imagens - Aplicações utilizando Entropia Generalizada

Neste capítulo serão apresentadas técnicas de segmentação de imagens utilizando a entropia generalizada. As principais abordagens adotadas para se segmentar uma imagens nas suas partes constituintes são as que utilizam a descontinuidade entre os pixels e a similaridade entre eles, como descrito anteriormente em 3.2.3.

O objetivo deste capítulo é apresentar técnicas de segmentação que utilizam a abordagem da similaridade, aplicando o conceito de entropia não extensiva. Pela primeira vez se propõe a utilização da entropia não extensiva em processamento de imagens. Foram desenvolvidas duas técnicas de binarização que utilizam o cálculo da entropia não extensiva, considerando a hipótese de que algumas classes de imagens podem ser classificadas como sistemas de informação com características não extensivas.

### 5.1 Imagens e Não Extensividade

O principal questionamento deste trabalho é: porque utilizar este novo formalismo em sistemas de informação, e de modo mais específico, em segmentação de imagens? Nos diversos casos em que o formalismo proposto por Tsallis tem sido aplicado, os

sistemas físicos apresentavam características de não extensividade, como interações de longo alcance, memória de longo tempo e estrutura fractal. O formalismo Tsallis se mostrava mais adequado para medir a entropia destes sistemas, demonstrando se ajustar de forma mais coerente com os modelos teóricos. De um modo geral, os resultados obtidos com  $q \neq 1$  apresentavam um diferencial qualitativo, se comparados com os resultados obtidos pelo formalismo tradicional ( $q = 1$ ).

A utilização deste novo conceito de entropia generalizada, que surgiu dentro da física estatística, tem se estendido ao longo dos anos para outras áreas da ciência, dentre elas a teoria da informação [30] [31] [14]. No contexto da teoria da informação, apresenta-se a hipótese de que imagens complexas podem apresentar características não extensivas. A formação da imagem digital é um processo físico em que a incidência de fótons em uma matriz CCD, ou em um outro dispositivo de aquisição, determina níveis de luminância para cada célula do sensor ou do pixel da imagem, depois que estas quantidades analógicas são devidamente digitalizadas. A formação da imagem pode ainda se dar por processos não óticos, como em microscópios eletrônicos ou de força atômica. Nestes casos, a interação de forças entre o material analisado e uma ponta de prova produz sinais elétricos que vão equivaler a valores de luminância para os pixels da imagem. A formação de objetos em uma imagem é portanto um processo complexo, que sugere a existência de correlações espaciais e correlações de luminância entre os pixels que os formam.

A forma do objeto é definida por um agrupamento de pixels que possuem valores de luminância correlacionados, tanto aqueles que o compõem quanto aos que delimitam suas fronteiras. É razoável considerar portanto, que em algumas classes de imagens, estas correlações espaciais e de luminância possam caracterizar estes sistemas de formação como sistemas não extensivos.

As duas imagens na figura 5.1 representam um exemplo de como os níveis de luminância dos pixels em diferentes classes de imagens estão correlacionados. A imagem A foi gerada a partir de um ruído com distribuição gaussiana em duas dimensões. Nesta imagem, os níveis de luminância de cada pixel não possuem nenhuma correlação com os



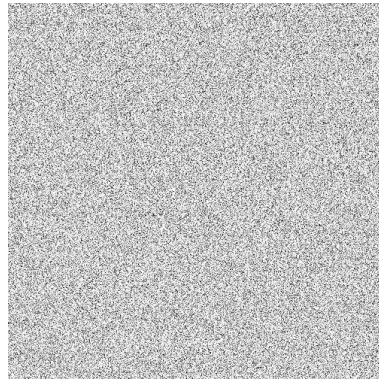


Imagem A - Ruído Gaussiano 2D



Imagem B - Retina Humana(Tomografia)

Figura 5.1: *Imagens e não extensividade. Imagem de um ruído gaussiano 2D (Imagem A) e imagem complexa de uma retina humana obtida por tomografia (Imagem B)*

demais, já que estes valores foram definidos de forma aleatória. Este é um caso extremo em que a referida imagem não fornece nenhuma informação. Na imagem B, que mostra uma tomografia de uma retina, podemos perceber que os pixels que formam uma rede de capilares e terminações nervosas possuem níveis de luminância correlacionados.

É exatamente a similaridade na iluminação destes pixels e a descontinuidade com os pixels de fronteira que permitem ao sistema de visão humano classificar esta imagem como uma tomografia de uma retina e identificar as informações que ela carrega. Além das correlações de luminância, correlações espaciais também definem os pixels como pertencentes aos objetos ou ao fundo da imagem.

Se considerarmos que estas correlações entre pixels representam uma característica particular de cada sistema, ou de cada classe de imagem digital, sugerimos a hipótese de que em certas classes de imagens, tais correlações possam caracterizar o sistema de informação como um sistema não extensivo. Sendo esta hipótese verdadeira, a utilização da entropia generalizada pode apresentar resultados significativos, assim como os já obtidos em diversos sistemas físicos.

## 5.2 Método de Binarização Utilizando Entropia Não Extensiva

Este método propõe a aplicação do conceito de entropia não extensiva na técnica de binarização proposta em [19]. O objetivo da técnica é determinar o ponto ótimo de segmentação através da maximização da soma das entropias dos dois subsistemas que compõem a imagem: objeto e fundo.

Considerando que uma imagem digital possui  $k$  níveis de cinza, definimos uma distribuição de probabilidade para estes níveis de luminância  $p_i = p_1, p_2, \dots, p_k$ . Desta distribuição, derivamos duas distribuições de probabilidades, uma para o objeto e outra para o fundo. Um nível de luminância  $t$  é considerado como limiar teórico entre as duas classes de pixels. As duas distribuições ficam definidas como

$$p_A : \frac{p_1}{P^A}, \frac{p_2}{P^A}, \dots, \frac{p_t}{P^A} \quad (5.1)$$

$$p_B : \frac{p_1}{P^B}, \frac{p_2}{P^B}, \dots, \frac{p_k}{P^B} \quad (5.2)$$

$$\text{onde } P^A = \sum_{i=1}^t p_i \quad \text{e} \quad P^B = \sum_{i=t+1}^k p_i$$

A entropia Tsallis  $S_q$ , parametricamente dependente do nível de decisão  $t$ , é definida como o somatório das entropias de cada distribuição, respeitando a propriedade pseudo aditiva da entropia não extensiva para sistemas estatisticamente independentes, definida na Seção 2.3.2. Deste modo, temos:

$$S_q^A(t) = \frac{1 - \sum_{i=1}^t (p_A)^q}{q - 1} \quad (5.3)$$

$$S_q^B(t) = \frac{1 - \sum_{i=t+1}^k (p_B)^q}{q - 1} \quad (5.4)$$

$$\begin{aligned}
S_q(t) &= \frac{1 - \sum_{i=1}^t (p_A)^q}{q-1} + \frac{1 - \sum_{i=t+1}^k (p_B)^q}{q-1} + \\
&(1-q) \cdot \frac{1 - \sum_{i=1}^t (p_A)^q}{q-1} \cdot \frac{1 - \sum_{i=t+1}^k (p_B)^q}{q-1}
\end{aligned} \tag{5.5}$$

O objetivo é maximizar a soma da medida de informação (entropia) das duas classes consideradas (objeto e fundo). Quando  $S_q(t)$  é máxima, o nível de luminância  $t$  é considerado o nível de binarização ótimo. Isto pode ser interpretado como um indicativo de máxima transferência de informação, de forma que a imagem binária obtida no  $t$  ótimo, preserve a ‘informação’ presente na imagem original.

$$t_{opt} = argmax[S_q^A(t) + S_q^B(t) + (1-q) \cdot S_q^A(t) \cdot S_q^B(t)] \tag{5.6}$$

O algoritmo apresentado a seguir descreve como obter o limiar de segmentação ótima pelo método da entropia máxima não extensiva.

Begin

Max\_value = LowValue;

For t=2 to t=k-1 do

Begin

Calcule  $p_A$ ,  $i = 1, 2, \dots, t$  pela equação 5.1;

Calcule  $p_B$ ,  $i = s+1, s+2, \dots, k$  pela equação 5.2;

Calcule  $S_A$  pela equação 5.3;

Calcule  $S_B$  pela equação 5.4;

Calcule  $S(t)$  pela equação 5.5;

Se Max\_value <  $S(t)$

Begin

Max\_value =  $S(t)$ ;

Threshold =  $t$ ;

End

End

End

### 5.2.1 Aplicações do Método e Discussão de Resultados

Para essa investigação definimos uma metodologia de estudo por meio da evolução de um histograma que simule as informações relativa ao “objeto” e “fundo” através de dois picos com distribuição gaussiana. O objetivo é compreender o comportamento da segmentação entrópica não extensiva, suas possíveis áreas de utilização e seus limites.

Conforme apresentado anteriormente, em um processo de segmentação procuramos um valor de luminância  $t$  que separe duas regiões. Normalmente, chamamos de segmentação otimizada um procedimento que escolhe algum valor de  $t$  entre os picos do objeto e do fundo obtido por meio de uma minimização (ou maximização) de uma função critério definida para esse fim.

Neste estudo iremos simular o histograma de uma imagem contendo dois picos gaussianos. Uma distribuição gaussiana é determinada pela sua posição  $p$  e largura (desvio padrão)  $\sigma$ . Esse procedimento nos permite avaliar o resultado da segmentação através do controle de parâmetros como amplitude, posição e largura dos picos no histograma. Todos esses parâmetros têm um papel significativo na caracterização da imagem, como por exemplo: homogeneidade da iluminação da cena, tamanho da imagem e do objeto, reflexibilidade (rugosidade) do objeto e do fundo, ruídos na imagem, resolução etc., figura 5.2.

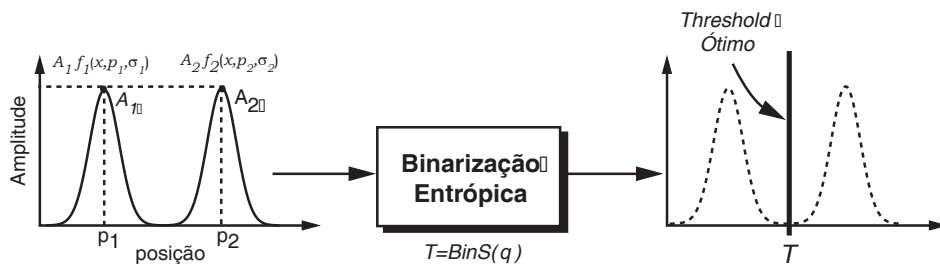


Figura 5.2: Segmentação ótima: o processo de binarização entrópica deve ser capaz de encontrar o melhor ponto de separação entre o “objeto” e o “fundo”.

Consideramos inicialmente algumas hipóteses de histogramas, estando interessados na prática em determinar o comportamento da função de segmentação para várias configurações dos parâmetros descritos acima. Definimos 4 casos particulares e fizemos evoluir uma função representando um histograma, figura 5.3, contendo dois picos com distribuição  $h(x) = f_f(p_1, \sigma_1) + f_o(p_2, \sigma_2)$  para  $h'(x) = f_f(p'_1, \sigma'_1) + f_o(p'_2, \sigma'_2)$  onde :

$$f(x, p, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left(-\frac{(x-p)^2}{2\sigma^2}\right) \quad (5.7)$$

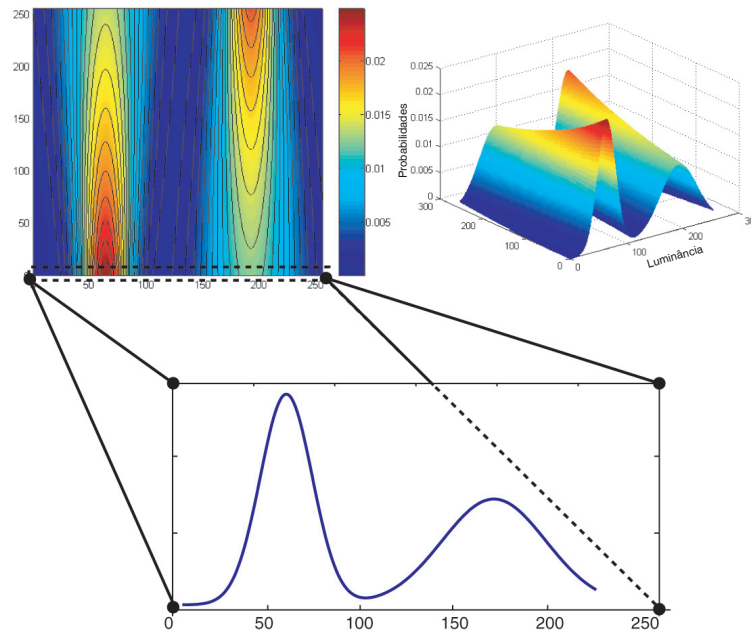


Figura 5.3: *Evolução do histograma representativo das regiões de luminância da imagem (“objeto” e “fundo”). O histograma destacado corresponde a uma linha do gráfico de superfície que ilustra a evolução.*

## Simulações

A seguir apresentamos os quatro casos utilizados para caracterizar processo de segmentação entrópica entre “objeto” e “fundo”:

- **Caso 1:** Posição do pico

Esse caso simula a variação de iluminação na cena e para estudá-lo fizemos evoluir

$h(x) = f_f(x, 64, 16) + f_o(x, 192, 16)$  para  $h'(x) = f_f(x, 64, 16) + f_o(x, 100, 16)$ , conforme apresentado na figura 5.4. É importante ressaltar que na configuração final, onde os picos do objeto e do fundo estão superpostos, algumas funções de segmentação clássicas, i.e. não entrópicas, têm dificuldade de encontrar o valor de  $t$  satisfatório.

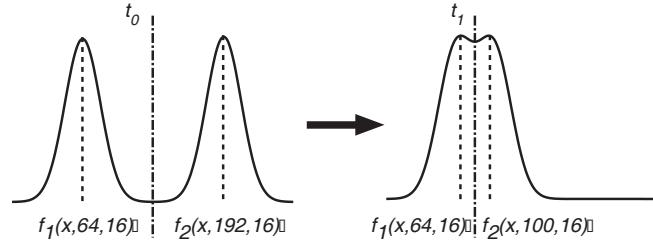


Figura 5.4: **Caso 1.** A gaussiana 1 é mantida fixa enquanto a gaussiana 2 é deslocada. Duas situações são observadas neste caso. No início os dois picos são facilmente destacados até o extremo onde os picos se encontram misturados entre si.

- **Caso 2:** Largura do pico

Significa uma variação na reflexibilidade no objeto, ou ainda uma variação de ruído neste. Nesse caso fizemos evoluir  $h(x) = f_f(x, 64, 16) + f_o(x, 192, 64)$  para  $h'(x) = f_f(x, 64, 16) + f_o(x, 192, 64)$ , conforme apresentado na figura 5.5. Este caso apresenta uma condição inicial ótima entretanto evolui para uma condição de difícil segmentação onde os picos encontram-se sobrepostos, porém com a presença de um mínimo local entre eles.

- **Caso 3:** Altura do pico

Esse é um caso típico onde uma das regiões é mais provável que a outra. Normalmente na imagem isto pode significar um objeto maior que o outro, ou uma grande região de intensidade luminosa homogênea. Para esse estudo fizemos evoluir  $h(x) = A \cdot f_f(x, 64, 16) + 1.5 \cdot A \cdot f_o(x, 192, 16)$  para  $h'(x) = A \cdot f_f(x, 64, 16) + 0.25 \cdot A \cdot f_o(x, 192, 64)$ , conforme apresentado na figura 5.6.

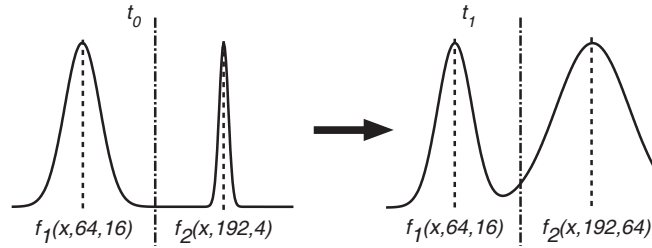


Figura 5.5: **Caso 2.** A gaussiana 1 é mantida fixa, na mesma altura, enquanto modificamos a largura da gaussiana 2. Este processo equivale a imagens onde um dos objetos tem uma variação em sua reflexibilidade ou é muito ruidoso.

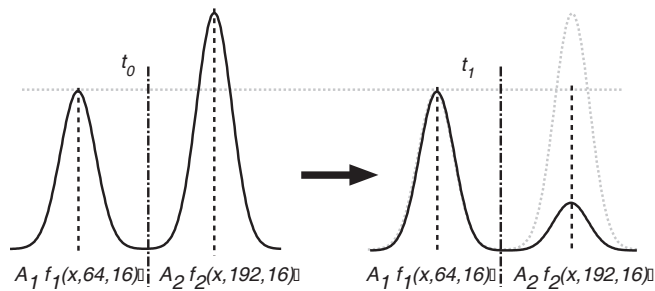


Figura 5.6: **Caso 3.** A gaussiana 1 é mantida fixa enquanto modificamos progressivamente a altura da gaussiana 2. Este processo representam imagens em que objetos e fundo tem probabilidades muito diferentes.

- **Caso 4:** Picos muito próximos - objeto imerso no ruído

Esse é um caso de difícil segmentação. Todos os métodos clássicos de segmentação funcionam com muita dificuldade para este caso. Um dos picos do histograma está praticamente submerso pelo outro ou pelo ruído. Para esse estudo fizemos evoluir  $h(x) = A_1 \cdot f_f(x, 64, 16) + A_2 \cdot f_o(x, 192, 16)$  para  $h'(x) = A_1 \cdot f_f(x, 64, 16) + 0.025 \cdot A_2 \cdot f_o(x, 192, 64)$ , conforme apresentado na figura 5.7

## Resultados das Simulações

Nesta seção, apresentamos os resultados das simulações realizadas a fim de compreender melhor o processo de binarização da imagem pelo histograma, por meio das técnicas de segmentação entrópica convencional e não extensiva.

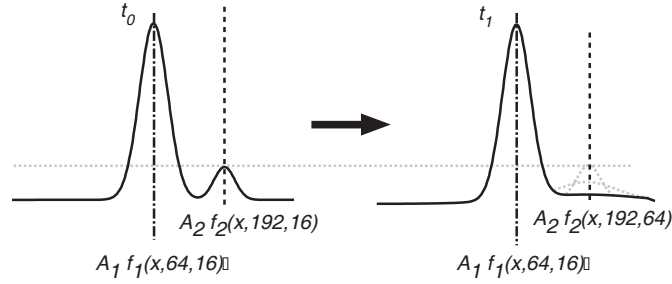


Figura 5.7: **Caso 4.** A gaussiana 1 está centrada em 64 e uma segunda gaussiana de largura e altura relativamente pequena é adicionada à primeira. Ao final, a segunda gaussiana praticamente não existe comparativamente à primeira. É um caso de difícil segmentação pois equivale àquelas imagens com uma região pouco provável, normalmente com muito ruído na sua região de luminância.

Para todos os casos apresentamos a seguir, 5 gráficos de superfície, lado esquerdo das figuras subsequentes, destacando o comportamento da trajetória do ponto de corte  $t$  (*threshold*) em função dos parâmetros apresentados acima e do parâmetro de não extensividade  $q$ . Outros 5 gráficos também são apresentados, a direita das figuras, mostrando as superfícies da soma das entropia através da relação de pseudo-aditividade, equação 5.5, na qual é possível decidir o valor de  $t$  que será utilizado para corte.

- **Caso 1:** posição do *pico*  $2 \times q$

Nesta figura podemos observar que a trajetória percorrida por  $t$  independe do valor do parâmetro  $q$ . Nessa situação, em que a amplitude e largura do pico são constantes,  $t$  sempre é escolhido no valor intermediário às duas gaussianas. Essa característica da segmentação entrópica é importante, pois os métodos de segmentação existentes, normalmente não fornecem bons resultados quando os picos estão muito próximos. É importante observar também que a entropia tende a uma superfície plana, com pequenas variações, quando aumentamos o valor de  $q$ .

- **Caso 2:** Largura do *pico*  $2 \times q$

Neste caso a trajetória ideal do ponto de corte  $t$  seria o vale entre os dois picos. Podemos observar que para valores de  $q$  pequenos a trajetória se move grada-



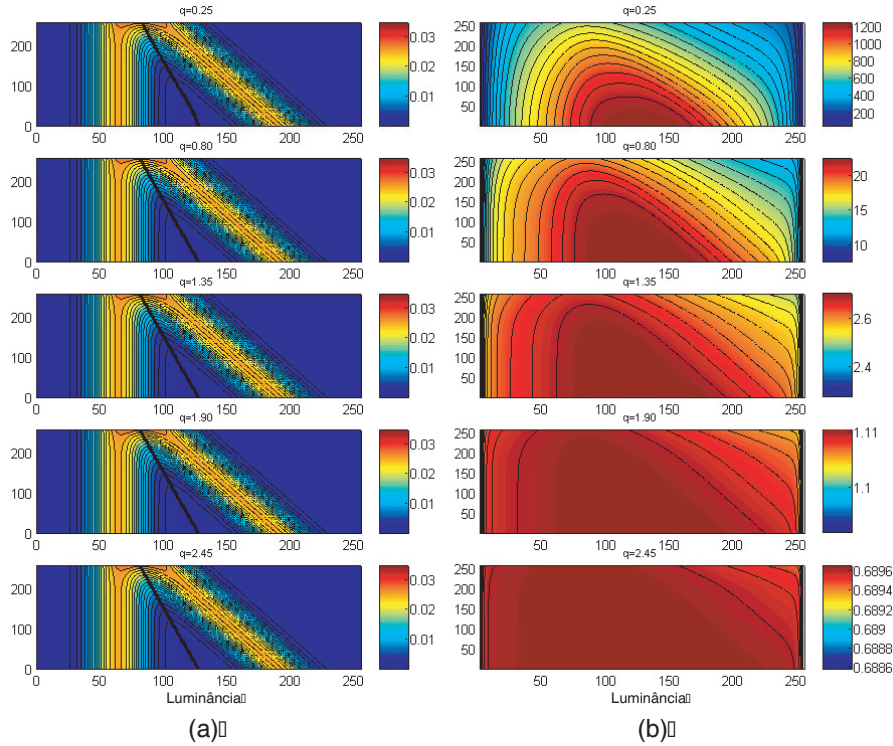


Figura 5.8: (a) Evolução do histograma para simulação do processo de segmentação para o caso 1. A linha escura no gráfico corresponde ao ponto de corte  $t$  escolhido. (b) Superfícies relativa a soma das entropias  $(A \cup B)$  utilizadas para determinação do ponto máximo para cada uma das situações do histograma.

tivamente para cima do pico 2, conforme sua largura aumenta. Por outro lado a trajetória calculada apresenta um movimento mais abrupto para valores de  $q$  próximo ou maiores que 1.

Dessa forma, podemos deduzir que quando existe uma diferença na largura das gaussianas,  $t$  depende de  $q$ . Esse aumento de  $q$  quando  $\sigma_2$  cresce pode ser compreendido devido ao aumento da entropia da região a direita de  $t$ , de forma que o sistema compensa o ponto de corte reposicionando  $t$ . O parâmetro  $q$  tem então a função de levar  $t$  para o ponto ótimo de separação das duas gaussianas.

- **Caso 3:** Altura do  $pico2 \times q$

A trajetória desejada, como nos casos anteriores, deveria ter algum valor de  $t$  no

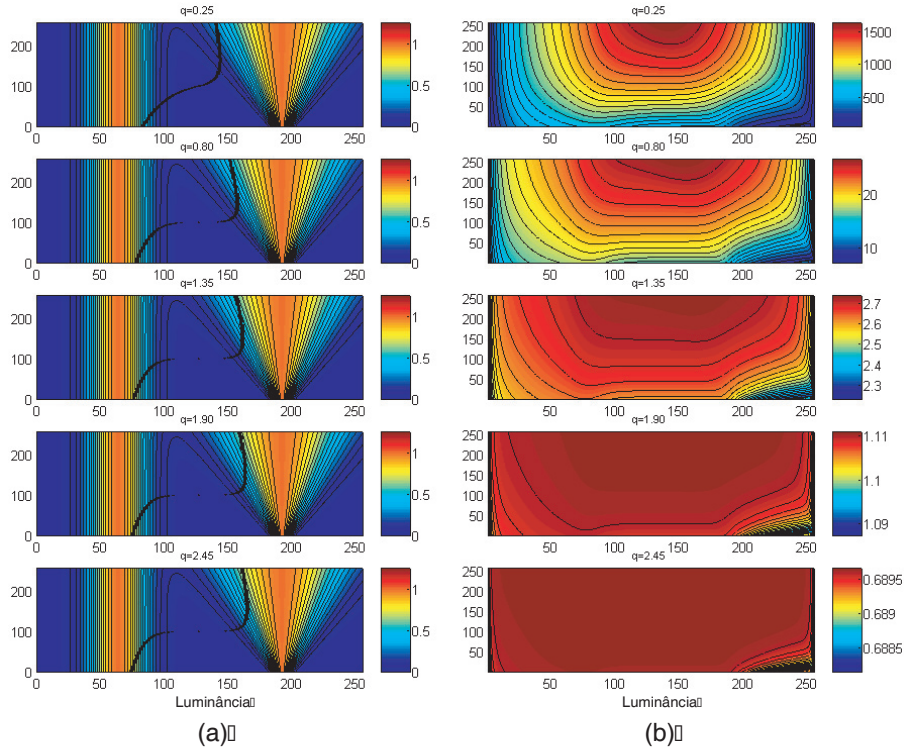


Figura 5.9: (a) Evolução das regiões “objeto” e “fundo” no processo de segmentação para o caso 2. A linha escura corresponde ao ponto de corte  $t$ . O ponto de segmentação evolui conforme aumentamos  $q$ . (b) Superfícies utilizadas para determinação do ponto de segmentação para cada uma das situações do histograma.

vale entre os dois picos (figura 5.10). Nessa figura, podemos observar uma dependência de  $t$  para diferentes valores de  $q$ , entretanto, apesar desta dependência, a segmentação é sempre bem sucedida. Porém, o corte realizado ligeiramente em cima das gaussianas irá certamente provocar uma classificação errada desses poucos pixels. É importante ressaltar que a altura do pico praticamente não altera o valor da entropia para um mesmo  $q$ .

- **Caso 4:** Picos muito próximos - Objeto imerso no ruído

Esse caso é normalmente um caso de difícil segmentação para a maior parte dos algoritmos automatizados. O pico que define um dos objetos é muito pequeno e pode estar misturado com ruídos presente na imagem. O resultado da segmentação para diversos modelos da segunda gaussiana é apresentado na

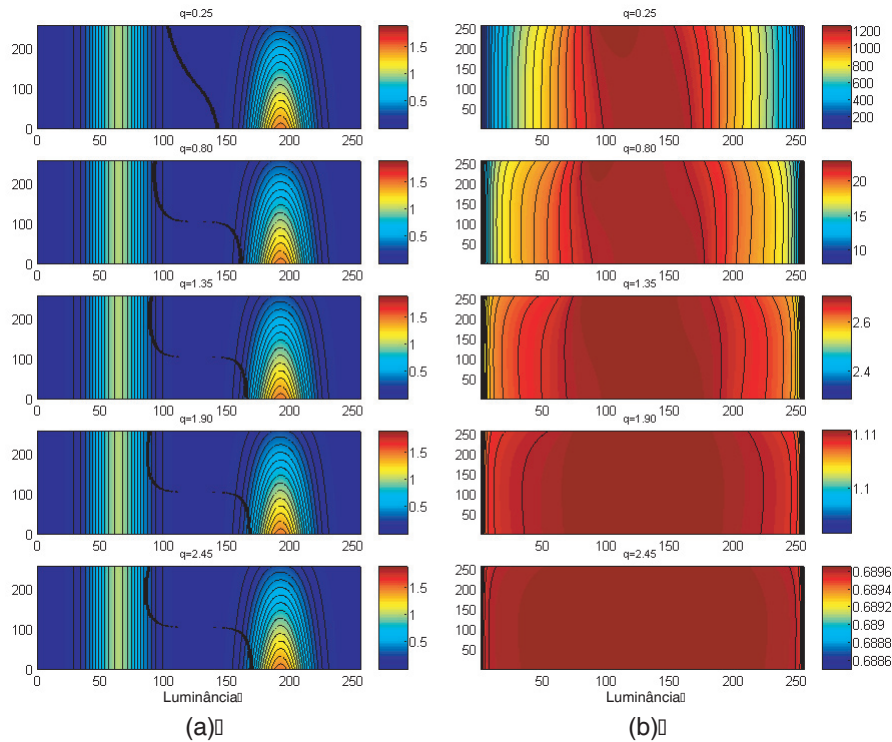


Figura 5.10: (a) *Evolução das regiões “objeto” e “fundo” no processo de segmentação para o caso 3. A linha escura corresponde ao ponto de corte no histograma.* (b) *Superfícies utilizadas para a determinação do ponto de segmentação para cada uma das situações do histograma.*

figura 5.11, até mesmo para a condição em que esta está totalmente submersa na outra ou em ruídos na imagem. Podemos observar na figura que valores de  $q$  muito próximos de zero fazem aparecer uma segunda região de máximo antes da primeira gaussiana, tornando o processo de escolha do ponto de corte instável. Valores de  $q$  próximos ou maiores que 1 são mais adaptados e funcionarão para todos esses tipos de imagens.

## Resultados em Imagens

Nesta seção apresentamos a técnica de segmentação entrópica não extensiva em imagens reais, constituídas normalmente por objeto e fundo. Para isso, escolhemos imagens contendo 256 níveis de cinza e dimensões de  $256 \times 256$  pixels (64K pixels). Tais imagens foram selecionadas porque apresentam propriedades relevantes em seus histogramas de

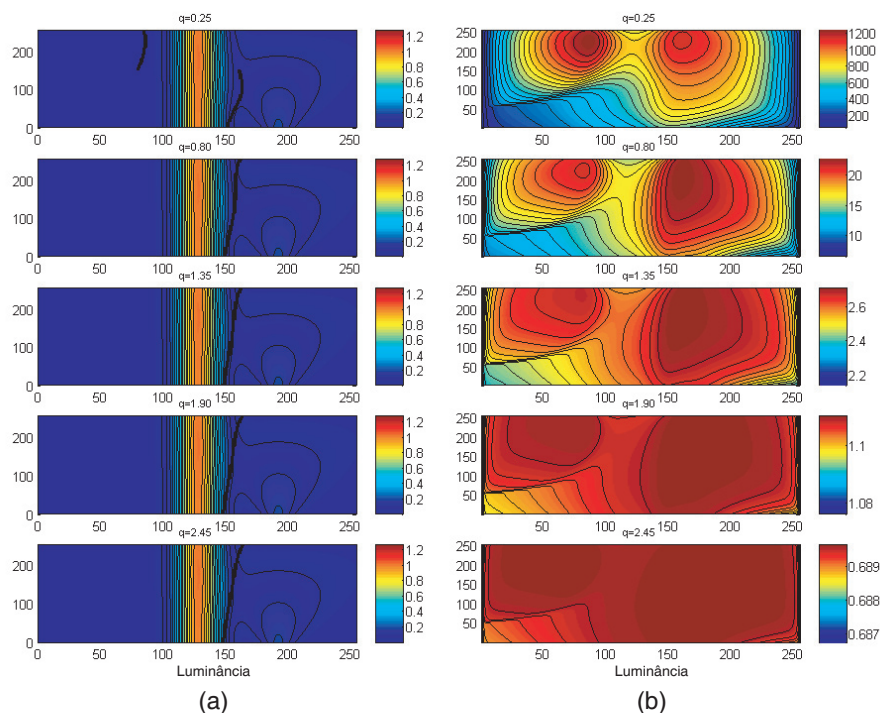


Figura 5.11: (a) Evolução das funções gaussianas no processo de segmentação para o caso 4. O pico 2 se encontra praticamente submerso no pico 1 ou pela presença de ruídos na imagem. Valores de  $q$  muito baixos criam duas regiões de máximos e a escolha do ponto de corte se torna instável. (b) Superfícies utilizadas para a determinação do ponto máximo para cada uma das situações do histograma.

luminância, i.e. estão de acordo com as características dos casos estudados anteriormente. Todos os histogramas foram filtrados por um filtro mediano de ordem 15 para diminuir a intensidade do ruído no mesmo.

Os resultados são comparados com dois outros métodos<sup>1</sup>: *Two Peaks* (dois picos) e *Iterative Selection* (seleção iterativa) [32].

- Caso 1: Posição do Pico

A figura 5.12 apresenta uma imagem de microscopia ótica de uma amostra magnética observada por efeito magneto-ótico, contendo em seu histograma dois picos referentes as regiões de domínios magnéticos. A segmentação desejável nesse caso é qualquer valor entre os dois picos, e o valor de  $t$  escolhido para a

<sup>1</sup>Uma breve descrição dos métodos *Two Peaks* e *Iterative Selection* é apresentada no Apêndice B

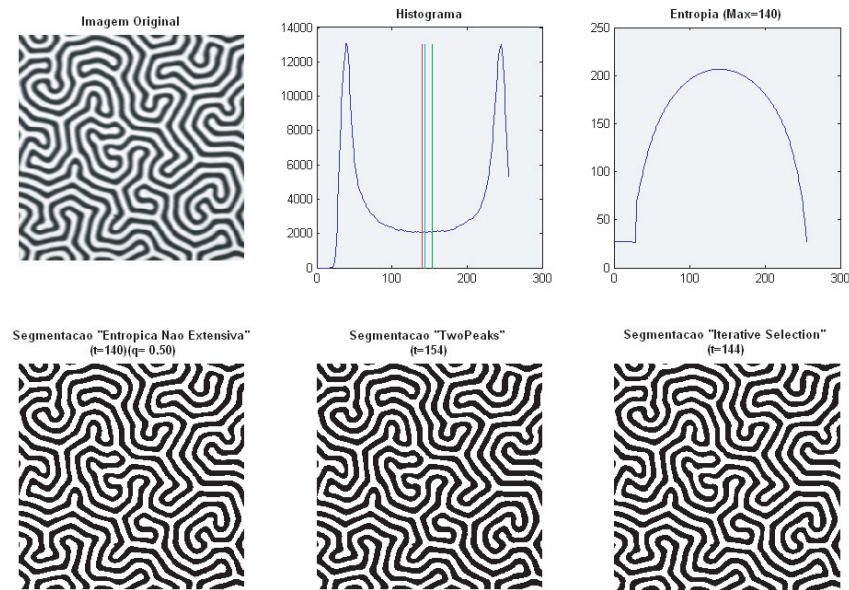


Figura 5.12: *Caso 1* : A imagem apresenta duas regiões com dois picos idênticos. O processo de segmentação neste caso não depende de  $q$ . O valor de corte é encontrado para  $t = 140$ . Os métodos *Two-Peaks* e *Iterative Selection* apresentam também bons resultados.

segmentação está de acordo com as simulações realizadas.

A Figura 5.13 apresenta uma imagem de uma amostra de ferrofluido observada por iluminação transmitida, e seu histograma de luminância, que apresenta duas regiões com distribuição gaussiana próximas. A imagem original é bastante ruidosa, fazendo com que os dois picos no histograma sejam largos. A segmentação desejada neste caso é separar as regiões celulares (“blobs”) do fundo. Como no caso anterior o processo de segmentação entrópica consegue encontrar o vale entre os dois picos com sucesso. Da mesma forma, o parâmetro de não extensividade  $q$  não interfere no processo. O algoritmo *Two Peaks* não consegue encontrar o vale e a imagem final, por ele gerada, fica relativamente comprometida.

- Caso 2: Largura dos Picos

A figura 5.14 exhibe uma imagem de satélite da região metropolitana da cidade do Rio de Janeiro. O objetivo da segmentação nesta imagem é separar a área equivalente a baía do continente. Neste caso, a utilização da segmentação entrópica

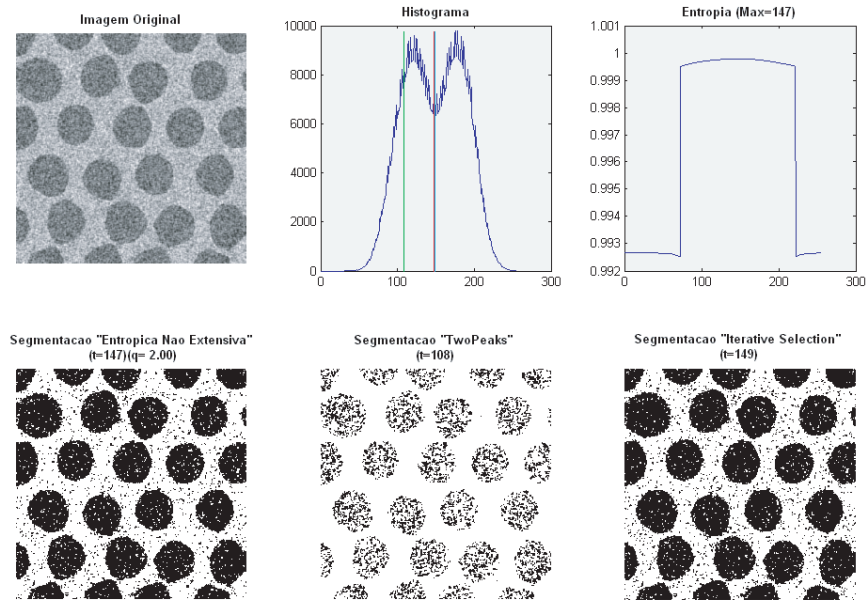


Figura 5.13: *Caso 1 : Dois picos sobrepostos - O processo de binarização entrópica determina o ponto de corte em  $t = 147$ .*

se refere ao segundo caso simulado, i.e., o histograma de luminância contém dois picos, sendo um deles mais largo que o outro.

O parâmetro  $q$  interfere na posição do ponto de corte. Para valores de  $q$  próximos de zero o ponto de corte se aproxima do vale entre as duas regiões. Para valores de  $q > 1$  o ponto de corte tende a subir sobre a segunda região comprometendo parcialmente o resultado da segmentação. Como apresentado na figura 5.14 os resultados obtidos diretamente pelos métodos *Two Peaks* e *Iterative Selection* podem ser interpretados como mais completos, pois eliminam regiões de ruídos na imagem final. É possível observar no gráfico do histograma que o ponto de corte escolhido pela técnica entrópica não extensiva ( $q = 0.10$ ) sobe parcialmente em cima do segundo pico (região onde o pico é mais largo). A região que fica à direita (mal classificada) será responsável pela presença de ruídos (pontos pretos) na imagem final. Este ruído será maior com o aumento do valor de  $q$ .

- Caso 3: Altura dos Picos

A figura 5.15 apresenta uma imagem do crescimento dentrítico de uma amostra



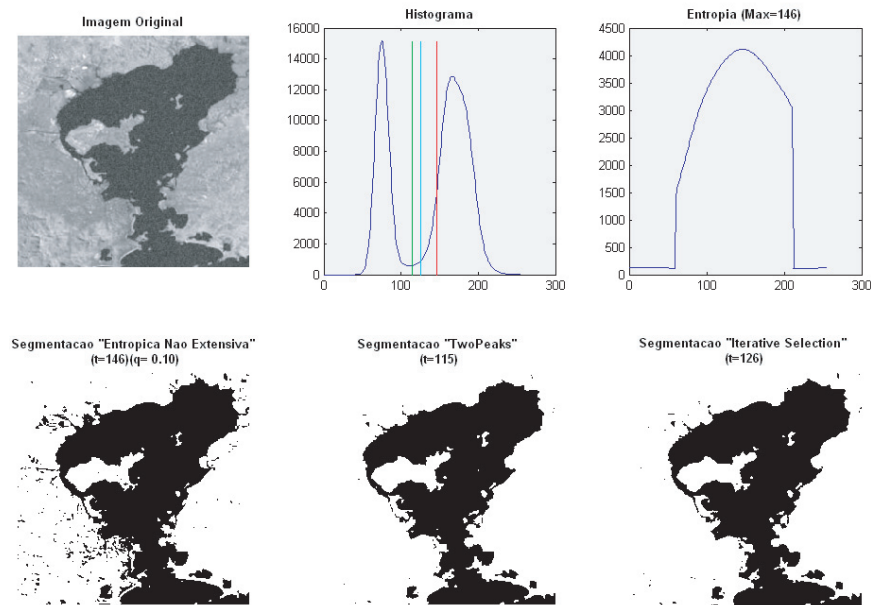


Figura 5.14: *Caso 2* : A presença de um pico mais largo que o outro faz com que o ponto de corte escolhido pelo processo de binarização entrópica tenda a subir na gaussiana mais larga. Este é o único caso em que o binarizador entrópico apresenta resultados inferiores aos dois outros processos de binarização.

de succinonitrila em um capilar. O objetivo é segmentar a parte relativa ao crescimento (folhas) do fundo da imagem. Essa imagem apresenta um histograma de luminância com dois picos independentes com probabilidades diferentes. É tipicamente equivalente ao caso 3 apresentado anteriormente. Nestas condições a aplicação da segmentação entrópica tem um desempenho essencial por meio da definição do parâmetro  $q$ . Nesse caso, podemos observar que a utilização de um valor de  $q = 3$  permite a seleção de uma pequena região nos limites do histograma. Isto é possível porque a técnica permite que o pequeno pico, presente na borda do histograma, tenha um papel importante no aumento da entropia do sistema. A modificação dada pelo fato de que  $p^q > p$ , para  $q < 1$  para a região de menor intensidade do histograma, faz com que o ponto de corte se movimente para próximo do pico mais baixo (região de menor probabilidade no histograma). Ainda neste sentido, apresentamos a imagem da figura 5.16,(a) e (b), uma flor,

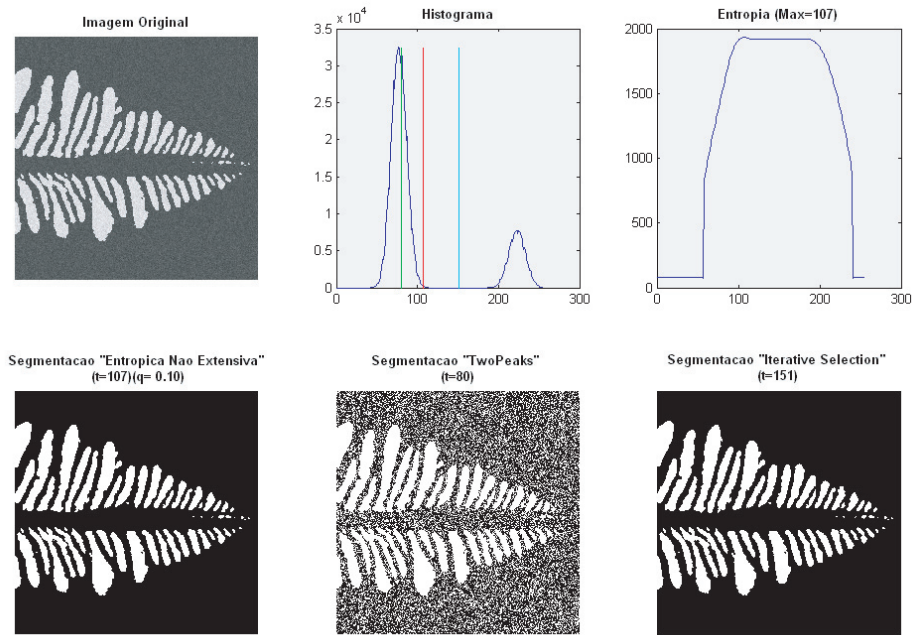
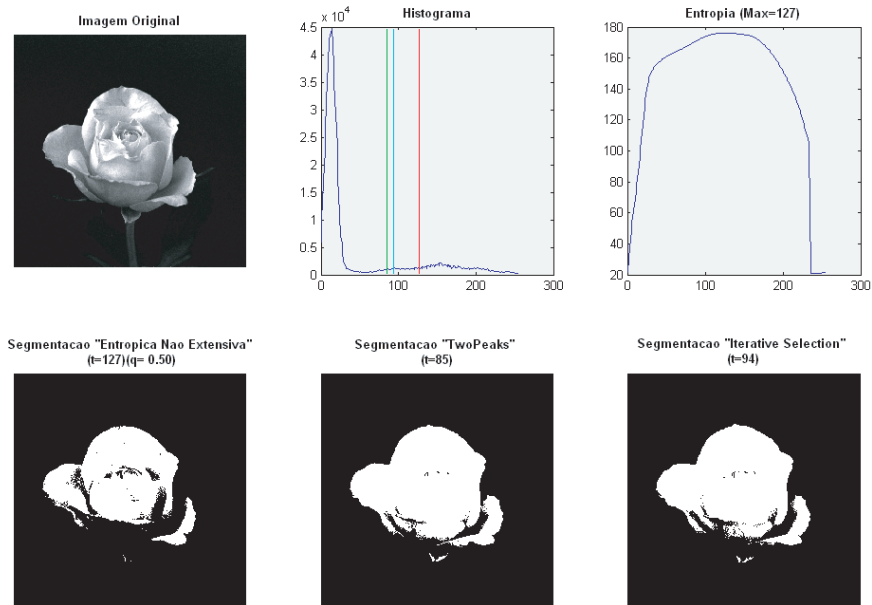


Figura 5.15: *Caso 3 : Imagem com dois picos independentes e com alturas muito diferentes. A segmentação entrópica, para eliminar a presença dos ruídos na imagem original e amplificados no método “Two Peaks”, precisa de um  $q$  pequeno ( $q=0.1$ ).*

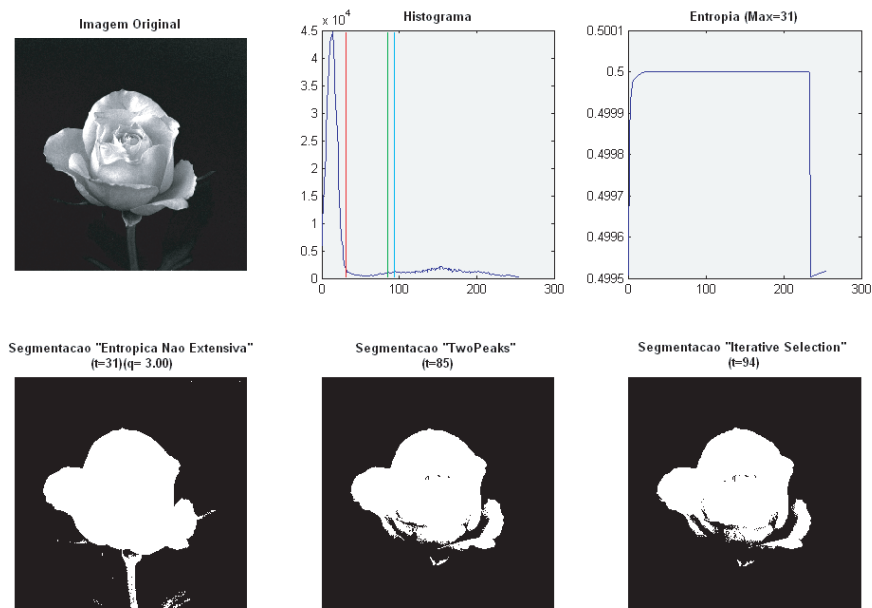
com uma distribuição irregular de luminosidade ao seu redor, devido a iluminação da cena em um fundo quase homogêneo e escuro. Este caso equivale na prática a uma mistura dos casos 2 e 3 apresentados anteriormente. O objetivo é separar a flor do fundo escuro. No entanto alguns algoritmos irão se preocupar em segmentar somente a parte superior do objeto. Na figura 5.16(a) utilizamos um  $q = 0.50$  e o ponto de corte é bem próximo de centro da segunda região. Toda região do caule é deixado de fora por todos os métodos. Na figura 5.16(b) utilizamos um valor de  $q = 3$ , reposicionando o ponto de corte entre as duas regiões, fundo e objeto. A técnica entrópica não extensiva possibilitou neste caso o reposicionamento do valor de *threshold* para o vale entre as duas regiões. Os outros dois métodos permitem apenas a segmentação da parte superior da flor. Esta possibilidade de ajustar o ponto de corte é um dos pontos fortes da técnica entrópica não extensiva.

- Caso 4: Objeto imerso em ruído





(a)



(b)

Figura 5.16: *Caso 2 + Caso 3: Imagem com dois picos independentes e com alturas e larguras muito diferentes. A segmentação entrópica não extensiva apresenta resultados diferentes em função de  $q$ . Em (a)  $q = 0.5$  e ponto de corte é  $t = 127$ . Em (b)  $q = 3$  e ponto de corte ( $t = 31$ ) é melhor ajustado entre as duas regiões.*

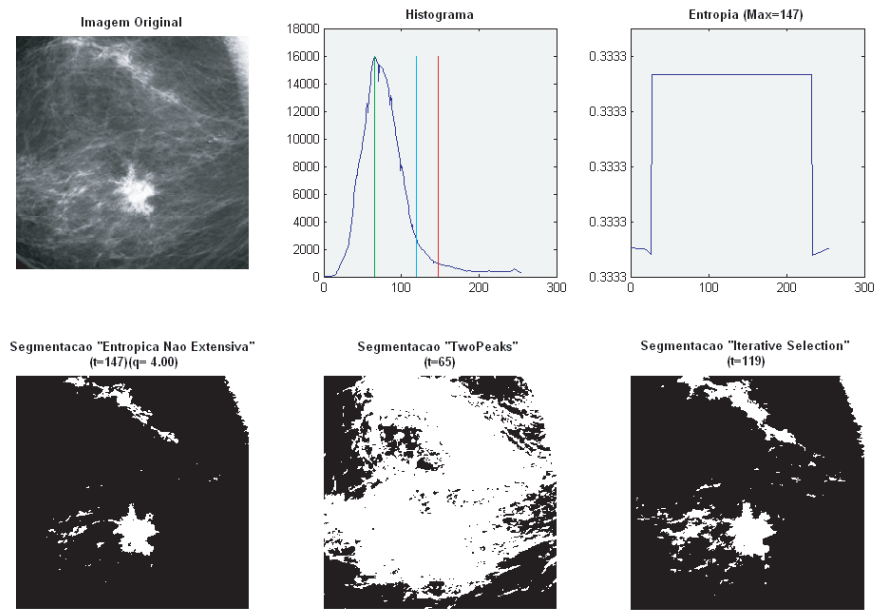


Figura 5.17: *Caso 4 - objeto imerso em ruído* : O histograma desta imagem é composta por um grande e largo pico que praticamente mascara a presença do outro, completamente dentro de uma região ruidosa de luminância. O método de segmentação entrópica funciona bem para estes casos, com um valor de  $q$  elevado. A curva para determinação do valor máximo de threshold tem um máximo em 183.

A figura 5.17 apresenta uma imagem médica de um tumor observado em uma mamografia. Essa imagem apresenta um histograma de luminância, em que os pixels que compõem o tumor formam um pequeno pico à direita, praticamente submerso no ruído. O objetivo é segmentar a região do tumor em relação ao fundo. Nestas condições, a aplicação da segmentação entrópica tem um desempenho essencial por meio da definição do parâmetro  $q$ . Nesse caso, podemos observar que a utilização de um valor de  $q = 3$  permite a seleção de uma pequena região nos limites do histograma. Isto é possível porque a técnica permite que o pequeno pico, presente na borda do histograma, tenha um papel importante no aumento da entropia da região.

## 5.2.2 Conclusões

O método de segmentação entrópica não extensiva é uma generalização do método proposto inicialmente por Kapur e quando aplicado para alguns casos específicos pode apresentar resultados qualitativamente melhores. Neste caso a soma das entropias dos subsistemas A e B (objeto e fundo) é levada em conta através da expressão pseudo aditiva proposta na estatística não extensiva. Em segmentação de imagens não podemos dizer qual técnica é mais apropriada para funcionar para todos os casos. Em algumas situações, como naquelas apresentadas para os casos 3 e 4, podemos certamente afirmar que o processo de segmentação entrópica pode apresentar resultados bastante satisfatórios, e o fato de permitir uma modificação do parâmetro  $q$  no cálculo da entropia nos permite ajustar precisamente onde se quer segmentar a imagem (mais para o vale ou mais para dentro da maior região do histograma). Esta característica pode ser certamente um diferencial na montagem de todo o processo de tratamento da informação presente na imagem.

Em PDI o processo de binarização é um ponto chave no sucesso de extração da informação presente na imagem. Ademais, o objetivo é sempre desenvolver uma técnica adaptada e que possam ser realizadas de forma automática para várias imagens de uma mesma classe. As técnicas entrópicas (tanto convencionais como não extensivas) funcionam bem para a maior parte dos casos, sendo que são sensivelmente melhores quando as imagens são compostas por objetos pouco prováveis, cujos os picos no histograma se misturam ao pico do fundo da imagem. Os métodos não extensivos consideram o aparecimento de um novo parâmetro,  $q$ , que permite um ajuste fino do histograma de acordo com a equação da entropia não extensiva. A trajetória da binarização por entropia para os vários casos apresentados nos mostra que o valor de  $t$  sempre tende a ficar sobre a gaussiana de maior largura. O parâmetro de não extensividade pode ser considerado um ajuste para levar o ponto de corte  $t$  a valores que separam regiões independentes.

Cabe lembrar que considerar correto o processo de binarização dependerá muito das etapas posteriores no processo de tratamento da informação presente na imagem

analisada. Nos casos em que é possível, uma análise específica de eficiência deveria ser realizada completamente, avaliando toda a cadeia de tratamento da imagem. Cabe ressaltar também que o sucesso desta técnica dependerá do fato de que o histograma de luminância possa descrever eficientemente as regiões de “objeto” e “fundo” na imagem.

Um segundo ponto chave é o questionamento de como devemos interpretar imagens diferentes que apresentem o mesmo histograma. Devemos esperar um resultado igual para o ponto de corte? Na prática, o cálculo do histograma (regiões da imagem com mesma luminância) é aqui o processo que define as correlações espaciais da imagem. Podemos observar assim que imagens de natureza diferente podem apresentar histogramas iguais, e o ponto de corte deverá ser o mesmo. Porém devemos ressaltar que o resultado da segmentação da informação presente na imagem será relativamente diferente, pois os pixels nos dois casos estarão certamente em regiões diferentes da imagem.

## 5.3 Método de Binarização Utilizando Entropia Relativa Generalizada

O método apresentado na seção anterior não resolve satisfatoriamente o processo de binarização para todos os casos. A principal dificuldade vem do fato de que, ao buscarmos maximizar a soma das entropias para as duas regiões da imagem, estamos procurando por histogramas planos e que na prática representam regiões ruidosas ou dispersas na imagem, com entropia máxima.

O método apresentado nesta seção irá trabalhar com a entropia relativa tentando aproximar um jogo de probabilidade teórico àquele encontrado na própria imagem para objeto e fundo. Nesta seção utilizaremos uma distribuição gaussiana para aproximar os dois jogos de probabilidades (objeto e fundo), onde os parâmetros relativos são obtidos no próprio histograma da imagem, utilizando como função de aproximação a entropia relativa generalizada.

A Entropia Relativa ou Divergência Direta é utilizada em métodos de segmentação de imagem que envolvem a medida de distância entre duas distribuições estatísticas. Em geral, uma função critério é definida para a obtenção do ponto ótimo de segmentação, através da comparação entre a imagem original e a imagem segmentada. As distribuições estatísticas são frequentemente obtidas através do histograma de luminância da imagem original e do histograma da imagem segmentada ou de modelos teóricos para histogramas sintetizados.

O método proposto é baseado na binarização de imagens utilizando a Entropia Relativa Generalizada (ERG) através do formalismo não extensivo. Este formalismo foi apresentado na seção 2.3.3, e define a Distância Kullback-Leibler Generalizada entre dois jogos de probabilidade  $p$  e  $p'$ . Sendo  $p = p_1, p_2, \dots, p_k$  e  $p' = p'_1, p'_2, \dots, p'_k$  duas distribuições de probabilidade, então a ERG,  $D_{KL_q}(p : p')$ , entre os dois jogos é definida como sendo:

$$D_{KL_q}(p : p') = \sum_i \frac{p_i^q}{1-q} \cdot (p_i^{1-q} - p_i'^{1-q}) \quad (5.8)$$

que pode ser re-escrita se considerarmos a função q-logarítmica,  $\log_q(p) = \frac{p^{1-q}-1}{1-q}$  definida em [15], ficando

$$D_{KL_q}(p : p') = \sum_i p_i \cdot \log_q \frac{p_i}{p'_i} \quad (5.9)$$

A versão simétrica da divergência Kullback-Leibler pode ser escrita da seguinte forma:

$$D_q(p : p') = \sum_i p_i \cdot \log_q \frac{p_i}{p'_i} + \sum_i p'_i \cdot \log_q \frac{p'_i}{p_i} \quad (5.10)$$

Esta versão simétrica será utilizada para o cálculo da distância estatística entre os dois jogos de probabilidades.

Neste método, o jogo de probabilidades  $p$  é obtido através do histograma de luminância da imagem original, como definido anteriormente na seção 3.2.2. O segundo jogo de probabilidades  $p'$  será obtido através de um modelo gaussiano obtido no histograma, que será definido levando se em consideração alguma informação a priori sobre a imagem.

Como o objetivo deste método é segmentar a imagem em duas regiões (objeto e fundo), consideramos um valor  $t$  de luminância que teoricamente separa os pixels da imagem nestas duas regiões. Deste modo, podemos considerar que os pixels do intervalo  $[0, t]$  pertencem ao objeto e os pixels do intervalo  $[(t + 1), k]$  pertencem ao fundo da imagem. O jogo de probabilidades  $p$  fica dividido em dois jogos de probabilidades, sendo  $p_A$  para o objeto e  $p_B$  para o fundo, definidos como  $p_A = \{p_1^A, p_2^A, p_3^A, \dots, p_t^A\}$  o jogo de probabilidades dos pixels pertencentes ao objeto e  $p_B = \{p_{t+1}^B, p_{t+2}^B, p_{t+3}^B, \dots, p_k^B\}$  o jogo de probabilidades dos pixels pertencentes ao fundo da imagem, e considerando que:

$$p_i^A = \frac{h_i}{P_t}, \quad i = 1, 2, 3, \dots, t; \quad P_t = \sum_i^t h_i \quad (5.11)$$

$$p_i^B = \frac{h_i}{1 - P_t}, \quad i = t + 1, t + 2, t + 3, \dots, k; \quad (5.12)$$

em que  $k$  é o número máximo de níveis de cinza da imagem e  $h$  é o histograma e a normalização por  $P_t$  é importante porque torna  $p_i^B$  e  $p_i^A$  jogos de probabilidades fechados, i.e.:

$$\sum_{i=1}^t p_i^A = \sum_{i=t+1}^k p_i^B = 1 \quad (5.13)$$

Um segundo ponto importante nesta técnica, é a definição do jogo de probabilidades  $p'$ . Este jogo de probabilidades é normalmente baseado em algum modelo teórico (ou alguma informação a priori que podemos ter da imagem). Neste método, assumimos o jogo de probabilidade  $p'_A = \{p'_1, p'_2, p'_3, \dots, p'_t\}$  como a distribuição de probabilidades modelo para o objeto e  $p'_B = \{p'_{t+1}, p'_{t+2}, p'_{t+3}, \dots, p'_k\}$  para o fundo da imagem.

A especificação da função modelo para o jogo de probabilidades do objeto e do fundo,  $p'_A$  e  $p'_B$  será baseada em uma distribuição gaussiana <sup>2</sup>:

$$p_i^A = \frac{1}{\sqrt{2\pi}\sigma_a} \exp^{-\frac{(x-x_a)^2}{2\sigma_a^2}} \quad i = 1, 2, 3, \dots, t \quad (5.14)$$

e

$$p_i^B = \frac{1}{\sqrt{2\pi}\sigma_b} \exp^{-\frac{(x-x_b)^2}{2\sigma_b^2}} \quad i = t + 1, t + 2, t + 3, \dots, k; \quad (5.15)$$

onde  $x_a$  e  $x_b$  são obtidos pela posição onde se encontram os valores máximos ( $A_a$  e  $A_b$ ) de cada lado do histograma e  $\sigma_a$  e  $\sigma_b$  correspondem ao valor a meia largura a meia

---

<sup>2</sup>N. Pal [21] propõe que o jogo de probabilidades que melhor aproxima o histograma de uma imagem está baseado em uma distribuição de Poisson. Ele assume que os níveis de cinza de uma região da imagem seguem uma distribuição de Poisson como parâmetros  $\lambda_A$  e  $\lambda_B$ , equivalente aos valores médios das regiões segmentadas para objeto e fundo. No capítulo 4, apresentamos em detalhes o trabalho proposto por N.Pal. Para maiores detalhes sobre as distribuições estatísticas veja [33][34].

altura (HWHH),<sup>3</sup> obtidos varrendo o próprio histograma da região (esquerda e direita), definido pelas seguintes fórmulas:

$$x_a = \operatorname{argmax}(h_i) \quad i = 1, 2, 3, \dots, t; \quad (5.16)$$

$$x_b = \operatorname{argmax}(h_i) \quad i = t + 1, t + 2, t + 3, \dots, k; \quad (5.17)$$

$$\sigma_a = \operatorname{arg}(p'_i{}^A = \hat{\sigma}) \quad (5.18)$$

e

$$\sigma_b = \operatorname{arg}(p'_i{}^B = \hat{\sigma}) \quad (5.19)$$

onde  $\hat{\sigma} \approx 0.606$ .

O limiar de definição dos dois jogos será definido no ponto  $t$  onde é encontrada a menor distância estatística entre os dois jogos de probabilidades, definida pela seguinte soma :

$$D_{KL_q}(t) = D_{KL_q}(p_A : p'_A) + D_{KL_q}(p_B : p'_B) \quad (5.20)$$

Ao encontrar os jogos de probabilidade ( $p'_A$  e  $p'_B$ ), que estão mais próximos dos jogos de probabilidades do objeto e do fundo ( $p_A$  e  $p_B$ ), temos na realidade uma representação estatística do conjunto de pixels que compõem as duas regiões. Podemos assumir como ponto de corte ótimo  $t$ , o ponto onde as gaussianas  $p'_A$  e  $p'_B$  se encontram, figura 5.18. Isto pode ser feito igualando-se as equações 5.14 e 5.15. Duas soluções são encontradas e dadas pela seguinte equação<sup>4</sup>:

$$\hat{t}_{1,2} = \frac{x_b \sigma_a^2 - x_a \sigma_b^2 \mp \sigma_a \sigma_b \sqrt{(x_a - x_b)^2 - 2 \log\left(\frac{A_a}{A_b}\right) (\sigma_a^2 + \sigma_b^2)}}{\sigma_a^2 - \sigma_b^2} \quad (5.21)$$

---

<sup>3</sup> “Half Width at Half Height”: A função de densidade de probabilidade normal tem seu valor máximo em  $0.3989/\sigma$  e sua meia largura a meia altura a  $0.2420/\sigma$ . O valor estimado  $\hat{\sigma}$  é obtido pela razão entre esses dois valores:  $\hat{\sigma} = 0.606604$ .

<sup>4</sup>Outras soluções também existem quando as gaussianas tendem a zero, porém são desprezadas neste caso.



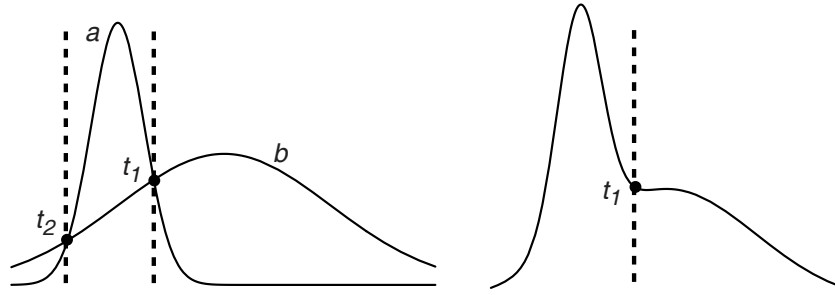


Figura 5.18: *Ponto de segmentação ótimo é definido no ponto de encontro entre as duas gaussianas, a e b.*

Como duas gaussianas são sempre estimadas pelo algoritmo, aquela mais à direita será sempre a gaussiana  $b$  e a mais a esquerda a gaussiana  $a$ . Dessa forma, a solução que nos interessa é sempre  $\hat{t}_1$ , pois esta raiz terá sempre seu valor entre as duas gaussianas.

$$\hat{t} = \hat{t}_1 \quad (5.22)$$

Desta forma, a aproximação dos jogos de probabilidade é feita através da técnica da entropia relativa generalizada e o ponto de corte é definido na junção das duas gaussianas. O algoritmo fica então:

```

Begin
Min_value = HighValue;
For t=2 to t=k-1 do
  Begin
    Calcule  $p_i^A$ ,  $i = 1, 2, \dots, t$  pela equação 5.11;
    Calcule  $p_i^B$ ,  $i = s + 1, s + 2, \dots, k$  pela equação 5.12;
    Calcule  $x_a$  pela equação 5.16;
    Calcule  $x_b$  pela equação 5.17;
    Calcule  $\sigma_a$  pela equação 5.18;
    Calcule  $\sigma_b$  pela equação 5.19;
    Calcule  $p_i^A$ ,  $i = 1, 2, \dots, t$  pela equação 5.14;
    Calcule  $p_i^B$ ,  $i = t + 1, t + 2, \dots, k$  pela equação 5.15;
  
```

```

    Calcule  $D_q(t)$  pela equação 5.20;
    Se  $\text{Min\_value} > D_q(t)$ 
    Begin
         $\text{Min\_value} = D_q(t)$ ;
    End
    Calcule  $\hat{t}$  pela equação 5.21;
     $\text{Threshold} = \hat{t}$ ;
End
End

```

### 5.3.1 Aplicações do Método e Discussão de Resultados

Realizamos os mesmos estudos apresentados na seção anterior (seção 5.2.1) para os casos 1, 2, 3 e 4. Todos os casos não apresentam variações do ponto de corte em função do parâmetro de não extensividade  $q$ . Os resultados da trajetória de  $\hat{t}$  são apresentados na figura 5.19. Para todos os casos o processo de encontro do ponto de corte é achado diretamente através das gaussianas resultantes da menor distância estatística entre os dois jogos de probabilidade. O processo entrópico neste caso foi utilizado para fins de comparação entre os dois jogos. Durante todo o processo, o parâmetro  $q$  não irá influenciar diretamente na escolha do ponto de corte  $\hat{t}$ , pois o objetivo é escolher dois jogos de probabilidades que melhor se aproximem dos jogos de probabilidade que representem o objeto e o fundo. No processo de busca da melhor distribuição gaussiana, modifica-se a todo instante sua posição, largura e altura para que esta se aproxime ao máximo da região do histograma que esta sendo segmentada. No entanto, o parâmetro  $q$  irá influenciar diretamente no processo de comparação das distribuições. No limite, a eficiência do método será dependente desta comparação.

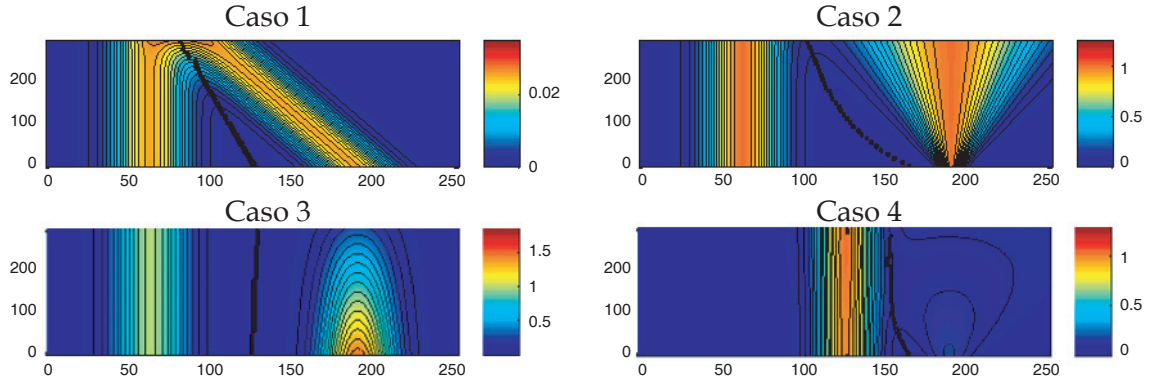


Figura 5.19: Resultado da segmentação por entropia relativa para os casos 1, 2, 3 e 4.

### 5.3.2 Resultados na Imagem

Desta forma, o processo de cálculo da distância estatística será fundamental para a determinação do ponto de corte  $\hat{t}$ . Em situações em que a formação da imagem pode ser modelizada por distribuições gaussianas (ou de poisson) o processo encontrará sempre uma jogo de probabilidade que melhor satisfaça o critério de aproximação e encontrará as duas gaussianas em  $D_q(t)_{\text{minimo}}$  e, conseqüentemente, o ponto de *threshold* na interseção entre elas. Porém, quando a imagem tiver uma distribuição que não possa ser aproximada por uma distribuição gaussiana o ponto de corte poderá oscilar em função de uma melhor sintonia no parâmetro de não extensividade.

Na prática é muito difícil encontrar imagens com distribuições estatísticas específicas. Porém, podemos encontrar algumas imagens em que o histograma não segue, a priori, nenhuma distribuição específica e avaliar o comportamento deste método nesta situação. A imagem apresentada na figura 5.20 corresponde a uma paisagem onde a presença de uma região de brilho intenso altera sensivelmente seu histograma. Neste caso existirá uma região onde o parâmetro de não extensividade irá alterar a distância mínima e irá permitir uma segmentação melhor da imagem.

Nesta imagem, a presença de vários picos no histograma irá certamente complicar o encontro do ponto ótimo de corte pelo método da aproximação estatística. O fato de  $q \rightarrow 1$  (ou mesmo  $q > 1$ ) faz com que a distância Kullback-Leibler generalizada,

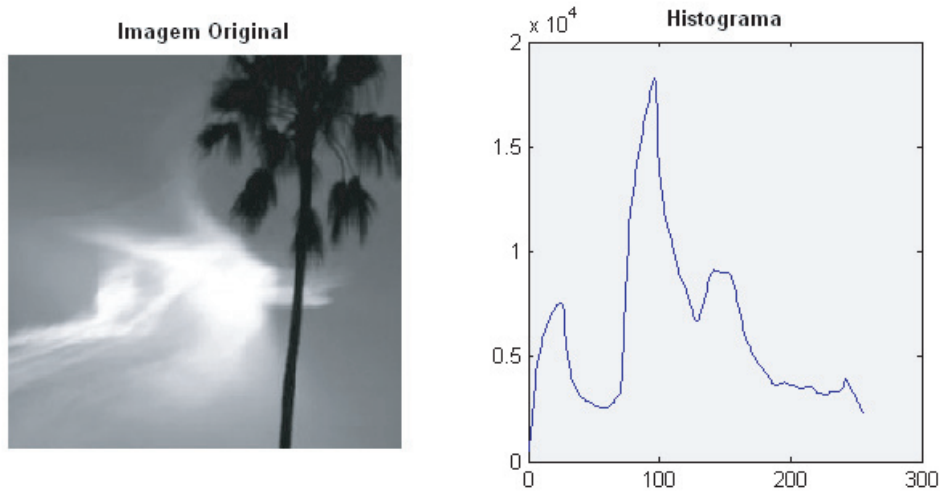


Figura 5.20: Imagem de uma paisagem. O histograma não segue a priori nenhuma distribuição específica. Neste caso, o parâmetro de não extensividade  $q$  irá alterar a distância estatística de Kullback-Leibler e modificará o ponto de corte  $\hat{t}$ .

$D_q(t)$  seja amplificada<sup>5</sup>. Esta característica é um ponto importante e positivo em todo o processo, porém estamos caminhando em uma região de grande sensibilidade no cálculo. Na prática, durante a implementação dos algoritmos, precisamos ter um controle detalhado do cálculo numérico quando  $q$  assume valores maiores que 1, pois a equação se torna instável devido a regiões de zeros no histograma. Na imagem da figura 5.21 o nível de segmentação definido pelo método foi  $\hat{t}=132$  para  $q < 1$ . Ao levarmos  $q$  para a região acima de 1 ( $q = 1.1$ ), conseguimos ajustar o ponto de escolha das gaussianas  $p'_A$  e  $p'_B$  para uma outra região de mínimo na curva  $D_q(t)$ , permitindo a escolha de outro ponto de corte  $\hat{t}=34$ , figura 5.21. Esta amplificação diferenciada para os dois vales se deve às diferenças encontradas nas suas redondezas, pois estes irão definir as gaussianas representativas de suas regiões.

A figura 5.22 apresenta em detalhes este comportamento, através de uma simulação do cálculo de  $D_q(t)$  para duas gaussianas conhecidas. O gráfico apresenta a evolução da distância estatística para vários tipos de gaussianas, para diferentes valores de  $q$ . A

---

<sup>5</sup>Pode ser observado pela equação de entropia de Kullback-Leibler generalizada, equação 5.8, que para  $q > 1$ , esta se torna instável na presença de zeros no histograma, tanto para a imagem quanto para o modelo teórico.

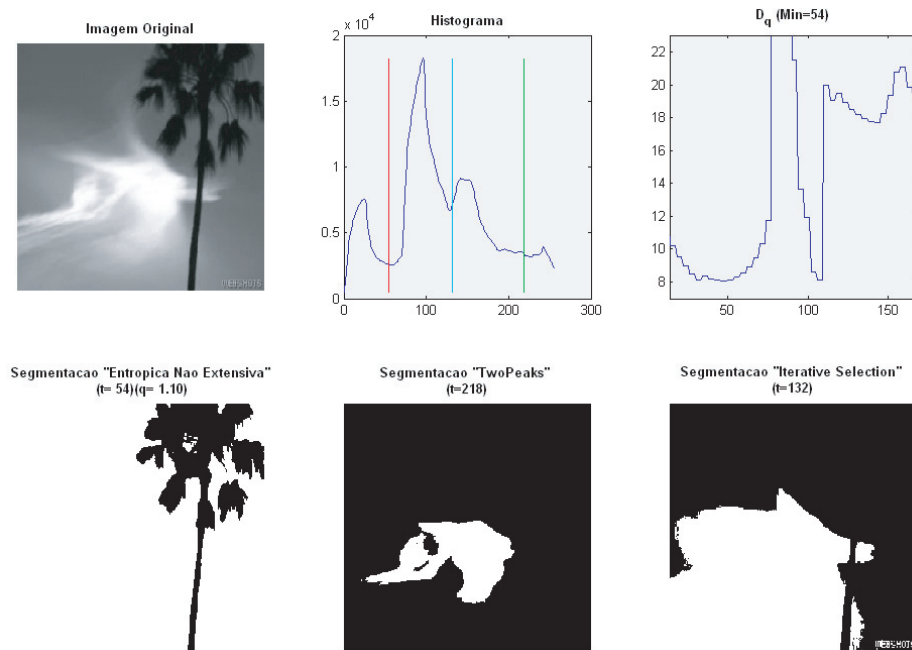


Figura 5.21: Resultado da segmentação para a imagem com uma distribuição desconhecida. O parâmetro  $q$  permite um ajuste fino da posição de corte ( $q = 1.1$ ).

medida que  $q$  aumenta, o ponto de mínimo da curva tende a ser uma curva pontual, com os jogos de probabilidades tendo distância zero para o único caso em que as distribuições são efetivamente iguais. Cabe ressaltar então que a escolha do parâmetro de não extensividade  $q$  significa definir um valor de  $q$  tal que as variações de distância sejam interessantes para fazer variar o ponto de corte. No entanto, podemos supor que adotar valores de  $q$  próximos a 0.9 ou a 1 deve a escolha ser mais apropriada, pois seria esta a região com o melhor compromisso entre proximidade e distância entre os jogos de probabilidade, ao comparamos jogos de probabilidade  $p$  e  $p'$ . Na imagem da figura 5.21, existe uma competição clara para decidir entre dois vales. A ampliação da distância para os dois casos será diferente fazendo com que o primeiro caso seja menor que o segundo, a partir de um determinado valor de  $q$ .

### 5.3.3 Conclusões

O funcionamento ideal de um sistema de segmentação é agir precisamente para qualquer tipo de imagem e em perfeita harmonia com os processos de tratamento da in-

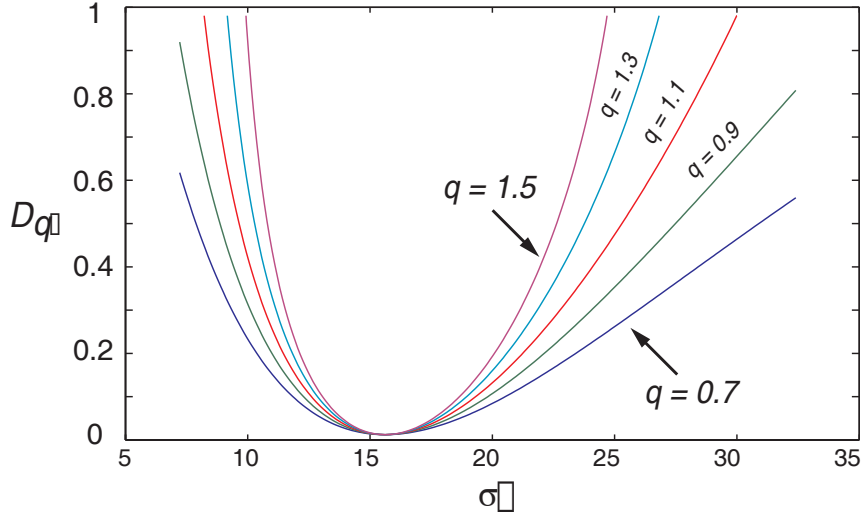


Figura 5.22: Gráfico do comportamento da distância Kullback-Leibler generalizada entre gaussianas, para vários valores de  $q$ .

formação posteriores. Na prática, a idealização de tal sistema é ainda uma tarefa muito distante para a maior parte dos sistemas de segmentação disponíveis hoje na literatura. No entanto, podemos dizer que o processo de segmentação entrópica, baseado no modelo apresentado nesta seção, é robusto e pode funcionar para uma variedade muito grande de imagens com ótimo resultado final.

A robustez do método é devido principalmente ao fato de podermos aproximar a distribuição gaussiana, para os histogramas obtidos em diversos tipos de imagens. Porém, não existe para estes casos uma dependência direta do parâmetro de não extensividade. O parâmetro  $q$  irá influenciar principalmente, no processo de escolha da proximidade entre os jogos envolvidos no cálculo da distância estatística. Nas situações onde o histograma das imagens não possam ser modelizados por uma distribuição gaussiana, o parâmetro  $q$  poderá ser um importante ponto de sintonia do processo, permitindo ajustes finos de onde se quer realizar o corte nas imagens. No entanto, uma precaução adicional deve ser tomada, pois nestes casos estaremos trabalhando em regiões instáveis da equação de distância Kullback-Leibler generalizada.

Um proposta de ampliação deste método é de utilizarmos o conjunto  $p' = p'_A + p'_B$  ao invés de utilizarmos as distâncias separadamente, para assim compará-lo diretamente com  $p$  (o histograma original da imagem). Este método poderia trazer mais

novidades no cálculo de  $D_q$  mínimo. Um segundo ponto interessante é que este método nos fornece os dois jogos de probabilidade que melhor representam a região correspondente ao objeto e ao fundo da imagem. Isto nos permite estimar o erro de classificação das duas regiões. Nas situações em que as gaussianas estão superpostas, é praticamente inevitável classificar erroneamente as regiões da imagem (i.e. pixels correspondentes aos objetos são classificados como de fundo e pixels correspondentes ao fundo são classificados como de objetos). Este erro poderia ser um interessante ponto de apoio para as etapas posteriores ao processamento da informação.





# Capítulo 6

## Reconhecimento de Padrões

### Utilizando Entropia

As técnicas de entropia podem ser utilizadas como base para efetuarmos reconhecimento e classificação de formas. O resultado obtido após a fase de segmentação consiste em um conjunto de pixels prontos para serem processados com o objetivo de realização de medidas específicas. É comum em PDI nos interessarmos em classificar diferentes objetos em diferentes classes. A classificação pode permitir um nível a mais na realização de medidas específicas, i.e., medidas em função da categoria do objeto. Diversas técnicas de reconhecimento e análise de formas têm sido desenvolvidas e são hoje utilizadas nas mais variadas aplicações, e.g., redes neurais artificiais, métodos estatísticos, métodos baseados nos atributos extraídos dos objetos, etc. Neste capítulo estaremos interessados na aplicação da técnica de entropia relativa como ferramenta discriminante no processo de classificação de formas entre diferentes classes. Normalmente em processos de representação e descrição de formas podemos nos basear em técnicas que utilizam a informação da forma do objeto ou da intensidade luminosa dos pixels que compõem o objeto. O método aqui descrito está baseado na representação externa do objeto, i.e., em seu contorno. Cabe ressaltar no entanto que esta técnica, dependente da função densidade de probabilidade escolhida para representar o objeto, e pode ser aplicada nas mais variadas situações.

## 6.1 Representação do Objeto

Como dito no capítulo 3, em PDI um objeto pode ser definido como um conjunto de pixels que representam algum interesse específico na cadeia de tratamento da informação visual. Em um processo de reconhecimento, devemos planejar o sistema através da escolha de parâmetros que possam de alguma forma caracterizar corretamente o objeto. No classificador aqui apresentado, a representação do objeto é feita por seus pixels de borda. Utilizamos para tal, a detecção de contornos em uma imagem binária. O procedimento realizado para a obtenção dos pixels que servirão para a classificação consiste em: binarização, identificação do conjunto de pixels contíguos (técnica conhecida como “labelização”), detecção de borda para cada objeto identificado e classificação, figura 6.1.

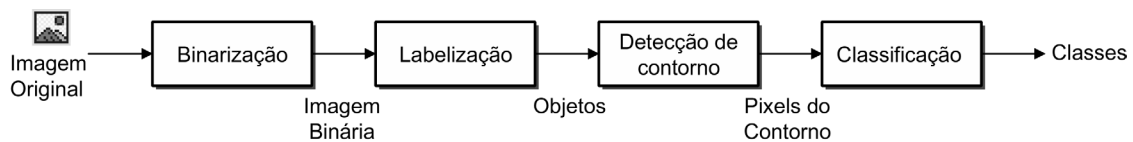


Figura 6.1: *Seqüência de tratamento da imagem para o processo de classificação*

A técnica de detecção dos pontos do contorno ( $CO$ ) de cada objeto ( $O$ ) utilizada foi baseado na seguinte operação morfológica:

$$CO = X - (E^B(O)) \quad (6.1)$$

onde  $B$  é o elemento estruturante correspondente a uma matriz de dimensão  $3 \times 3$ . Esta técnica permite a detecção de uma borda com um único pixel de largura em conectividade  $B8$ . Esta operação equivale na prática à subtração do objeto segmentado pela erosão do seu conjunto de pixels uma única vez. Este procedimento é hoje implementado por diversos softwares de processamento digital de imagens. Um exemplo típico é a função “outline” no programa Scion Image [35].

## 6.2 Método de Classificação

A construção das classes dos objetos exige o processamento do algoritmo descrito anteriormente, além da determinação do histograma que caracterize a forma. Esta técnica implementa uma “descrição do objeto” do tipo supervisionada pois temos que ter um conjunto de objetos que representem suas classes e possam determinar as funções densidade de probabilidade. Nesta etapa também é importante definir a resolução do histograma utilizada para seu armazenamento, figura 6.2.

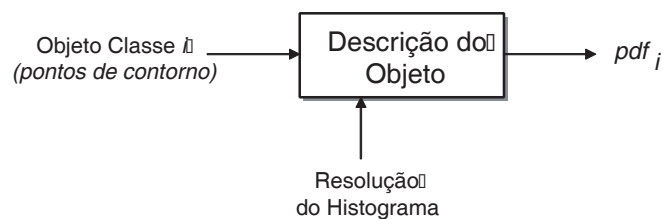


Figura 6.2: *Etapa de descrição do objeto. Apresentamos uma (ou mais) amostras ao sistema de descrição que irá construir os jogos de probabilidades utilizados posteriormente no processo de classificação das formas desconhecidas.*

A etapa de descrição consiste na elaboração dos histogramas representativos dos objetos. O tratamento da forma inclui todo o processo de segmentação da borda e cálculo da função densidade de probabilidade relativo à distância dos pixels de borda até o centro de massa da forma. A normalização realizada nos pixels de borda do objeto  $(x_N, y_N)$  proveniente da imagem digital é:

$$\begin{aligned}x' &= x - \bar{x} \\y' &= y - \bar{y} \\x_N &= \frac{(x' - x_{min})}{x_{mean}} - 1 \\y_N &= \frac{(y' - y_{min})}{y_{mean}} - 1\end{aligned}$$

em que  $\bar{x}$  e  $\bar{y}$  correspondem à média dos pixels de borda do objeto nos respectivos eixos na imagem digital. A normalização pelos valores médios ( $x_{mean}$  e  $y_{mean}$ ) é importante

para evitar as flutuações que possam aparecer na borda do objeto e que alterariam de forma significativa o histograma de distâncias. O cálculo da distância é realizado pela distância euclidiana ( $D_E$ ) entre cada ponto da borda do objetos e o centro de massa (origem). A distância máxima esperada, baseada no processo de normalização será  $2\sqrt{2}$ . O ajuste da resolução e do limite superior do histograma será uma etapa fundamental no processo de ajuste das funções de probabilidade, e deve ser caracterizado para cada caso. Desta forma podemos determinar a distribuição de probabilidade das distâncias calculadas, por:

$$p_i = \frac{d_E}{N}, \quad \text{sendo } \sum_i p_i = 1 \quad (6.2)$$

em que  $d_E$ , corresponde ao histograma das distâncias euclidianas calculadas, em função de um determinado numero de *bins*, e  $N$  corresponde ao total de pixels na borda do objeto. A escolha de um correto número de *bins* para o histograma poderá influenciar o processo de classificação. Este tópico será discutido mais adiante na seção dos resultados obtidos. A figura 6.3 apresenta a densidade de probabilidade para formas geométricas simples (triângulo, quadrado e círculo). A figura do círculo corresponde a um pico na região de seu raio, pois todas as distâncias contribuem igualmente para o mesmo intervalo do histograma. Este processo é bastante próximo da representação da forma por sua assinatura, definida em [16], sendo que aqui a representação em função do ângulo não é levada em conta.

### 6.3 Descrição de Similaridade

Medidas de similaridade podem ser estabelecidas através de diversos níveis de complexidade em uma imagem, desde a comparação direta entre pixels de uma mesma região até a comparação de toda a imagem. Nosso objetivo aqui é realizar o teste de semelhança através da entropia dos objeto após a fase de segmentação.

Como dito anteriormente, a entropia relativa é utilizada em métodos que envolvam a medida de distâncias entre duas distribuições estatísticas. Este formalismo foi apresentado na seção 2.2 através da medida da Distância Kullback-Leibler entre dois jogos

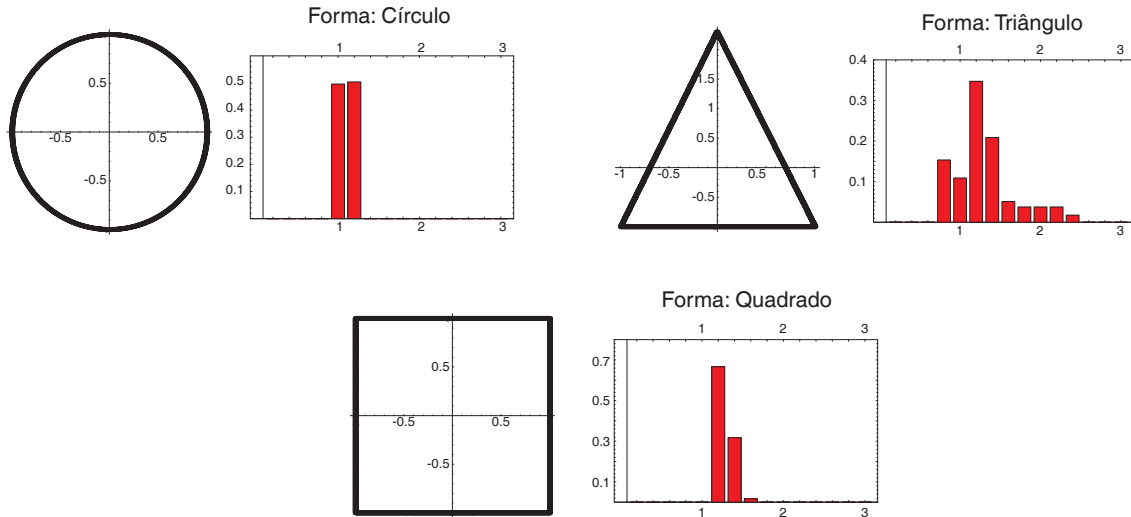


Figura 6.3: Representação da distribuição de probabilidades de três formas básicas.

de probabilidade. A proposta descrita era de ter uma medida do ganho de informação através da diferença entre os dois jogos  $p$  e  $p'$ . A modificação da informação é dada por  $\Delta I = -(\log p' - \log p)$ .

No nosso caso, a construção do classificador está baseada no limiar máximo para o ganho da informação, i.e., na prática iremos procurar por um jogo de probabilidade cujo o ganho de informação seja o menor possível, fazendo com que os jogos  $p$  e  $p'$  estejam o mais próximo possível, em termos de distância estatística. Consideramos para o cálculo as formas padrões utilizadas para definição do jogo de probabilidades representativas de cada objeto. A similaridade será descrita então pela aproximação do objeto desconhecido, que tenha sofrido o mesmo método de cálculo de sua densidade de probabilidade ao jogo dos objetos padronizados anteriormente. Os jogos de probabilidade são calculados aqui através do histograma de distâncias euclidianas entre os pixels de borda e o centro de massa do objeto, conforme descrito na seção anterior. A menor distância estatística entre os dois jogos será aquela do objeto mais próximo em termos de sua forma, ver figura 6.4.

Sendo  $p_i$  o jogo de probabilidade que represente todos os padrões previamente conhecidos, temos para o cálculo da distância estatística:

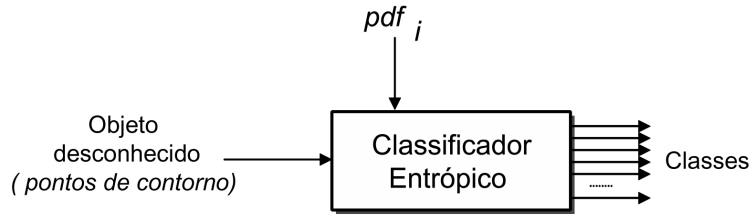


Figura 6.4: *Etapa de classificação dos objetos desconhecidos. Apresentamos uma amostra ao sistema de classificação. O sistema calcula a semelhança entre o jogo de probabilidades deste objeto com aqueles obtidos no processo de descrição da forma. A menor distância indica o padrão referente a forma.*

$$D_i(p : p') = \sum_j p \cdot \log\left(\frac{p}{p'_i}\right) + \sum_j p'_i \cdot \log\left(\frac{p'_i}{p}\right) \quad i = 1, 2, 3, \dots, F \quad (6.3)$$

para  $\sum p_i = \sum p'_i = 1$

em que  $F$  representa o número total de classes padronizadas a serem comparadas com o objeto a ser classificado e é normalmente, um número previamente conhecido. O objeto será então escolhido pela menor distância estatística entre os dois jogos de probabilidade, pela seguinte expressão:

$$\text{Objeto} \in \text{Classe}_i / i = \text{argmin} \{D_i\} \quad (6.4)$$

Nós realizamos alguns testes experimentais em um sistema de reconhecimento de caracteres (padrões numéricos) afim de demonstrar a eficiência e detalhar a robustez desse método.

## 6.4 Aplicação do Método

Considere dez classes representadas pelos dígitos “0” a “9”. Cada dígito corresponde a um objeto de 48 pixels de altura em uma imagem de 72dpi's<sup>1</sup>. Estes dez dígitos

---

<sup>1</sup>Este tamanho adotado na representação dos objetos é bastante útil, pois objetos com pouca resolução prejudicam a eficiência do classificador. Os sistemas de classificação disponíveis no Mercado, para conseguir uma alta taxa de acerto, trabalham normalmente com uma imagem de 300dpi, tendo para um caractere de 3mm de altura um pouco mais de 30 pixels na vertical.

podem estar presentes em qualquer posição e em qualquer orientação na imagem. Os objetos são localizados através do processo de “labelização”. O objetivo do sistema de classificação é associar cada dígito a uma das 10 classes previamente conhecidas utilizando a distância Kullback-Leibler entre dois jogos de probabilidade. Para tal, utilizaremos os procedimentos descritos nas seções anteriores. As dez classes são descritas pelos jogos de probabilidade apresentados na figura 6.5. Cabe ressaltar a proximidade entre o jogo de probabilidade do dígito “6” e “9”. Esta equivalência é devido ao próprio processo de cálculo do jogo de probabilidades, que faz com que tenhamos independência da representação em função da rotação do objeto na imagem. O resultado foi obtido através de um histograma com 40 intervalos (cada um com 0.15 unidades normalizadas de distância). Os dez histogramas serão utilizados para representar cada forma e serão a base de comparação com aqueles obtidos na imagem.

A tabela 6.1 representa o resultado da apresentação dos jogos de probabilidade correspondente aos objetos em relação àqueles das formas numéricas padrão. Cada célula da matriz representa a distância estatística entre os objetos entre si. A distância igual a zero na diagonal da matriz indica que o classificador consegue identificar objetos semelhantes a ele.

		Padrões/Classes									
		0	1	2	3	4	5	6	7	8	9
Objetos	0	0	8.31	8.27	5.78	6.40	7.49	3.50	7.94	5.80	3.70
	1	8.31	0	1.25	2.92	4.63	2.07	4.48	2.31	5.29	4.32
	2	8.27	1.25	0	0.98	2.46	0.51	2.74	0.57	3.34	2.64
	3	5.78	2.92	0.98	0	1.12	0.39	1.14	0.86	1.22	1.02
	4	6.40	4.63	2.46	1.12	0	1.47	0.69	1.36	1.01	0.61
	5	7.49	2.07	0.51	0.39	1.47	0	1.59	0.57	1.71	1.48
	6	3.50	4.48	2.74	1.14	0.69	1.59	0	1.89	0.77	0.04
	7	7.94	2.31	0.57	0.86	1.36	0.57	1.89	0	2.01	1.88
	8	5.80	5.29	3.34	1.22	1.01	1.71	0.77	2.01	0	0.66
	9	3.70	4.32	2.64	1.02	0.61	1.48	0.04	1.88	0.66	0

Tabela 6.1: *Distância entre os objetos e padrões calculados pela equação 6.3 (Distância Kullback-Leibler) entre jogos de probabilidades definidos pela equação 6.2.*

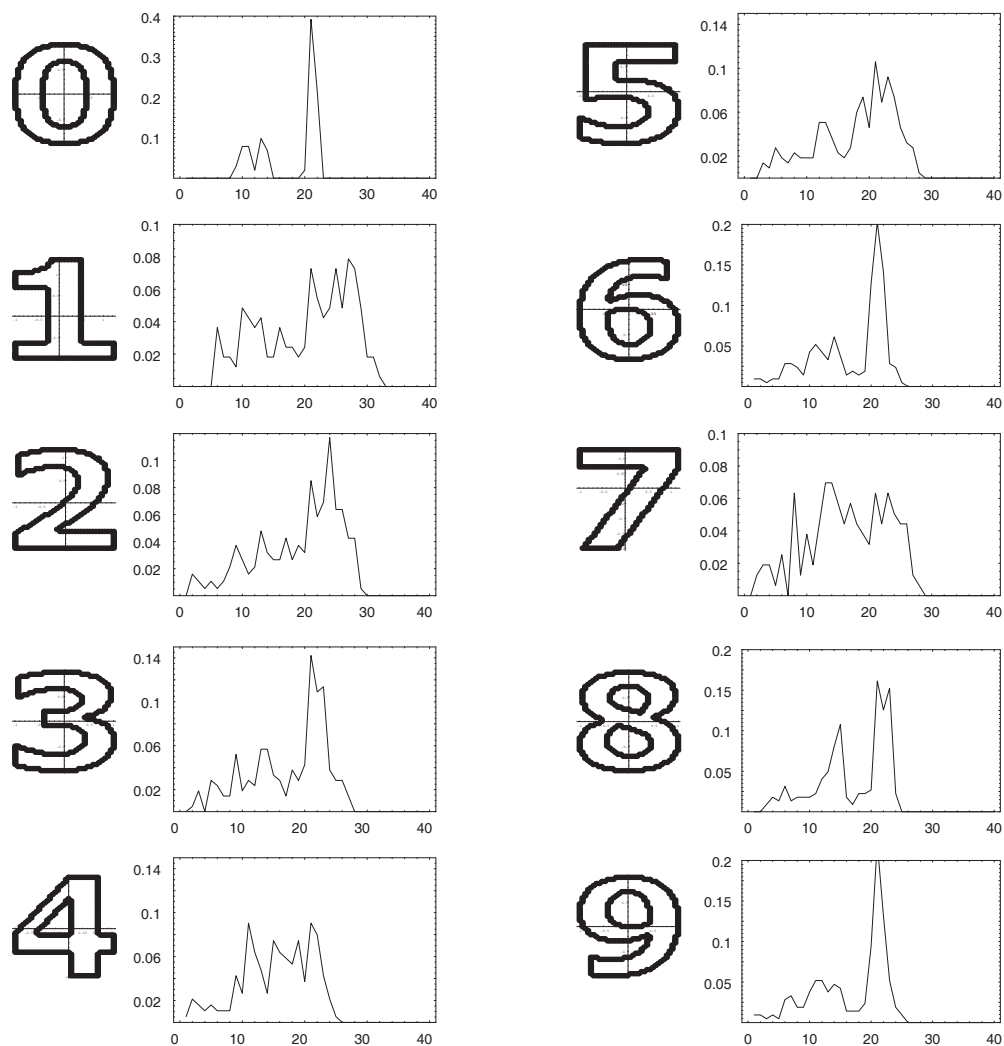


Figura 6.5: Jogos de probabilidades correspondentes as 10 classes (dígitos) utilizados no processo de classificação. Os histogramas representam a distribuição de distância entre os pixels de borda e o centro de massa de cada objeto. Estes histogramas servem como modelo de padronização das formas.

## 6.5 Considerações sobre os resultados

Uma análise interessante é adicionar ruído à borda dos objetos afim de observar o comportamento do classificador em função da amplitude do mesmo. O resultado irá caracterizar o classificador em função da relação *senal/ruído* (SNR). A inserção de ruído na borda dos objetos é muito comum no processo de formação da imagem digital e na segmentação da informação. Esta análise poderá nos informar a eficiência



do classificador em função do ruído e estimar seu limite de funcionamento. Neste processo adicionamos variáveis aleatórias aos pixels de borda do objeto segundo uma distribuição normal, com média zero e desvio padrão ( $\sigma$ ). O cálculo da relação SNR é dado por:

$$SNR = 20 \cdot \log\left(\frac{1}{\sigma}\right) \quad (6.5)$$

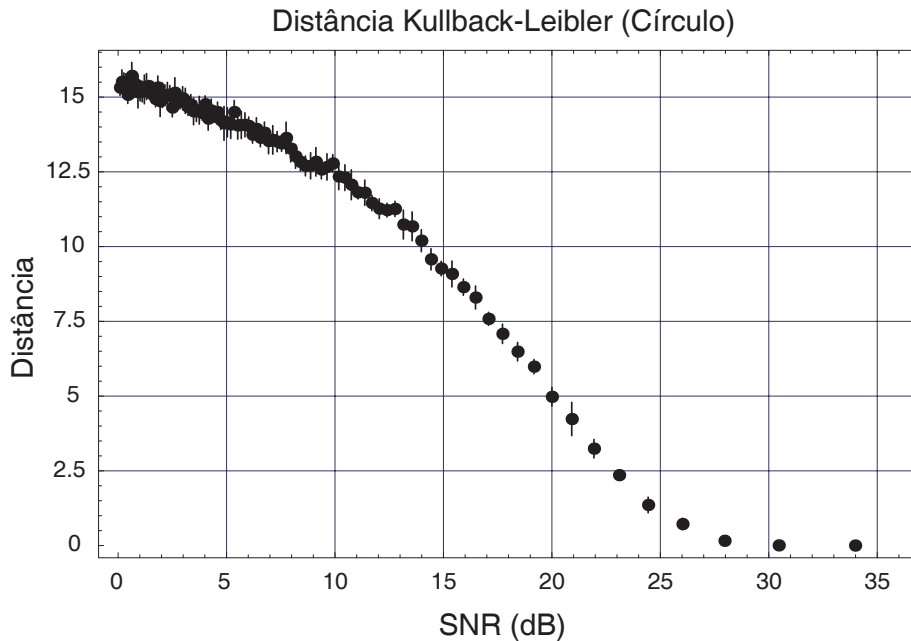


Figura 6.6: O gráfico apresenta a evolução da distância entre os jogos de probabilidades, de um círculo perfeito àqueles corrompidos por adição de ruído gaussiano. Quanto mais ruído é adicionado mais “longe” o objeto fica do círculo perfeito. O ponto inicial de distância zero, se encontra fora dos limites do gráfico (*i.e.*, acima de 35 dB).

Uma primeira análise avalia a deformação da forma, através da adição do ruído em seus pixels de borda, e seu distanciamento da forma básica original. O gráfico da figura 6.6 apresenta a evolução da distância Kullback-Leibler para o círculo, e a figura 6.7 apresenta a evolução de sua forma na imagem.

Com o aumento da amplitude do ruído (redução da relação SNR) temos um decréscimo da distância entre os dois jogos de probabilidades. Um limiar de decisão entre um forma básica poderia ser adotado entre os pontos 19.17 dB e 14.89 dB.

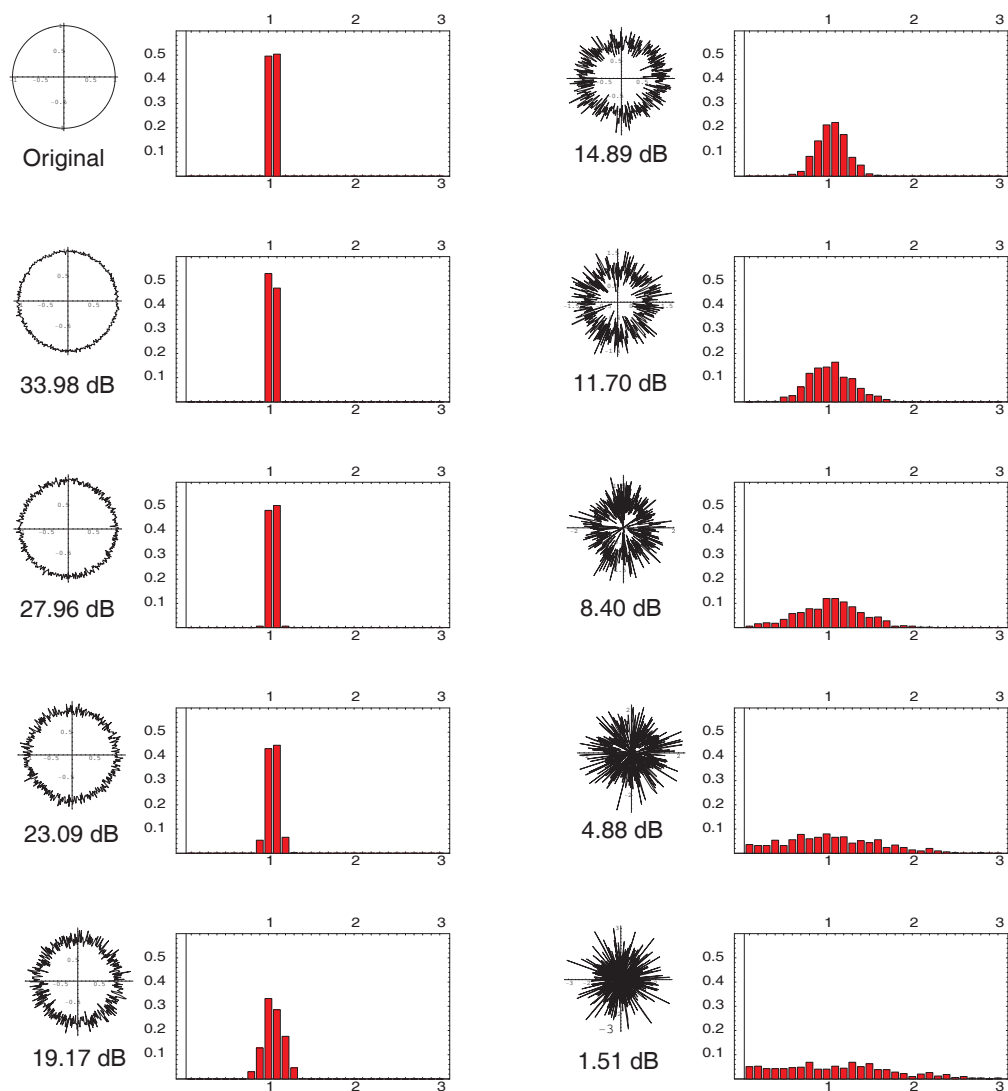


Figura 6.7: Variação da forma do “objeto” círculo com o acréscimo de ruído em sua borda, e os respectivos jogos de probabilidades calculados pela distância dos pixels de borda e o centro de massa da forma. Existe um limiar de decisão no qual não aceitamos (S.V.H.) mais que a forma equivale a um círculo.

Este comportamento, apresentado para formas simples, pode ser aplicado ao classificador numérico, a fim de testar sua performance. Este estudo permite acompanhar a taxa de acerto do classificador em função da amplitude do ruído adicionado aos objetos apresentados e estimar o erro do classificador, figura 6.8.

O gráfico SNR x Taxa de Acerto apresentado na figura 6.6 mostra que a partir de

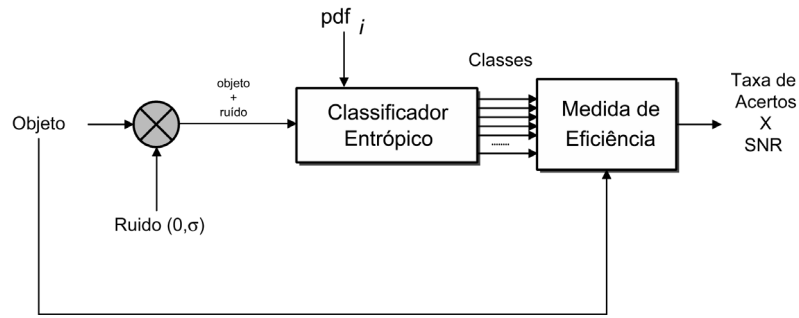


Figura 6.8: Medida de eficiência do classificador entrópico. Aos objetos, desconhecidos pelo classificador, adicionamos ruído gaussiano de média zero e desvio padrão ( $\sigma$ ). O resultado da classificação é comparado com àquele esperado. A eficiência é medida pelo comportamento da taxa de acerto do classificador em função da relação sinal/ruído.

uma relação SNR de  $18.41 \text{ dB}$ <sup>2</sup> o sistema começa a trabalhar abaixo de 80% de acertos para uma resolução do histograma igual a 10 intervalos. Este valor (80%) foi por nós adotado como um limiar ótimo de decisão no processo de classificação, acreditando que classificadores que trabalhem em um limiar inferior a 80% sejam efetivamente sistemas com pouca eficiência no reconhecimento dos objetos. Como adicionamos ruído em toda a borda do objeto, este resultado nos leva a concluir que os acréscimos na variação da forma apresentada ao classificador, podem ter um limite de até 11.5 pixels (23% da altura do objeto) a mais que o tamanho máximo da forma (altura do objeto é igual aqui a 48 pixels). Qualquer variação superior a este limite fará com que o classificador aponte resultados com uma taxa de acertos abaixo dos 80% esperados.

A figura 6.10 apresenta um exemplo de forma enviada ao classificador, com ruídos em sua borda, utilizada neste processo. O classificador consegue corretamente classificar o dígito “8” ruidoso, com  $18 \text{ dB}$  (figura 6.10d), na classe dos dígitos “8”.

<sup>2</sup>O sinal é 8.33 vezes superior ao ruído.

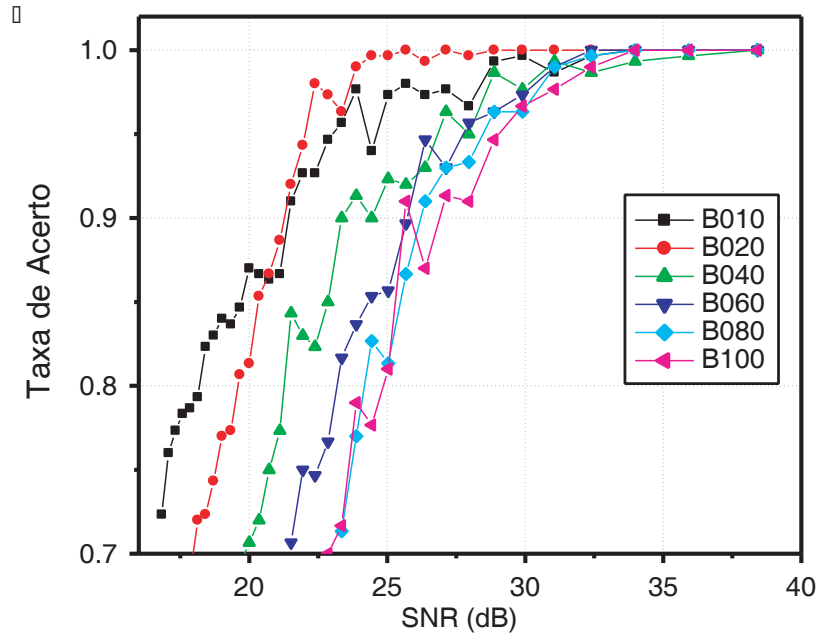


Figura 6.9: Taxa de Acerto do classificador em função da relação sinal/ruído (SNR) para diferentes resoluções do histograma (10, 20, 40, 60, 80 e 100). Os melhores casos estão situados para histogramas com poucos bins.

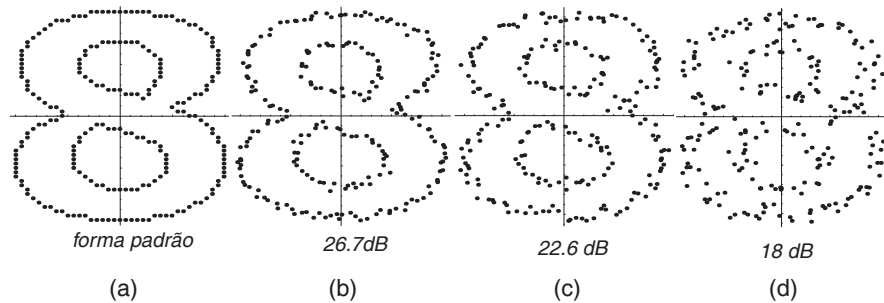


Figura 6.10: (a) Forma “8” padrão utilizada no processo de comparação e outras formas “8” apresentadas ao classificador com diferentes intensidade de ruído, e resolução do histograma igual a 20 intervalos. O classificador reconhece todos os objetos como pertence a classe correta.

### 6.5.1 Considerações na representação pelo Histograma

Em todos os processos de representação de densidades de probabilidades por histogramas, uma questão chave é a escolha correta dos intervalos (*bins*) para a acumulação correta da ocorrência dos eventos. No processo de classificação aqui apresentado um

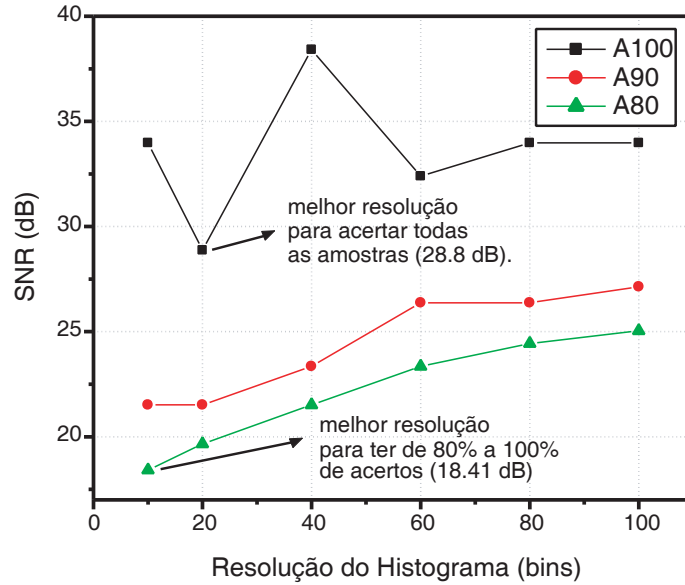


Figura 6.11: Avaliação dos melhores pontos de funcionamento do classificador para três Taxas de Acerto (100, 90 e 80%). Podemos definir o limiar de funcionamento para uma taxa de acerto de 100% com 20 bins de resolução do histograma em 28.8 dB. O gráfico mostra que o sistema é eficiente com uma taxa acima de 80% das amostras com um ruído adicionado equivalente a 11.5 pixels em uma forma de altura igual 48 pixels de amplitude.

ponto importante para a eficiência desse método é a determinação correta dos intervalos do histograma, que representará a função densidade de probabilidade que descreve a forma, através das distâncias euclidianas entre borda e centro do objeto. A utilização da resolução do histograma correta é então um ponto chave para aumentar a eficiência do método. Um histograma com muita resolução (muitos *bins*) será sensível a pequenas variações (detalhes) da forma e poderá ser decisivo no processo de classificação de formas muito parecidas. Um histograma com pouca resolução (poucos *bins*) poderá ser muito eficiente se estas variações (detalhes) forem apenas ruídos em todo o processo.

Realizamos um estudo para diferentes resoluções do histograma em função da relação SNR para três situações distintas:

- i) quando o sistema acertou completamente todos os padrões apresentados a ele - 100%;

*ii)* quando o sistema classificou corretamente acima de 90% dos padrões apresentados a ele e,

*iii)* quando o sistema acertou acima de 80% dos padrões apresentados.

Foram apresentados ao classificador para este estudo 300 variações dos objetos dentro de uma mesma amplitude de ruído. O resultado da eficiência da classificação é apresentado no gráfico da figura 6.11 para diferentes taxas de acerto. Este gráfico mostra que para trabalhar com uma eficiência de 100% o ideal é utilizarmos uma resolução do histograma de 20 *bins*, sabendo que o classificador começará a errar a partir de um ruído de 28.8 *dB* (sinal é 26.3 vezes superior ao ruído). Se aceitarmos trabalhar com uma taxa de 90% de acertos poderemos ter tanto 10 quanto 20 intervalos para o histograma e o classificador funcionará até 21.51 *dB* (sinal é 11.6 vezes superior ao ruído). Para o limite de 80% o ideal é trabalhar com 10 *bins* e o classificador irá acertar as amostras apresentadas até 18.41 *dB* (sinal é 8.3 vezes superior ao ruído). Dessa forma, para esse sistema, seria correto trabalhar com uma resolução do histograma igual a 20 *bins*, pois esta é a curva que tem a melhor relação entre a taxa de acerto e amplitude do ruído, além de ser a curva que mais acerta para todos os casos (100%), ver gráfico da figura 6.11 <sup>3</sup>. Este valor é também atraente pois torna todo o processo de cálculo das equações de Kullback-Leibler mais rápido.

## 6.6 Classificador por entropia não extensiva

As técnicas não extensivas permitem um ajuste da entropia através do parâmetro  $q$ . Esta modificação nos permite avaliar o processo de classificação em função de  $q$ .

---

<sup>3</sup>É comum nos métodos automáticos de construção do histogramas utilizarmos para o número de *bins* a raiz quadrada do número máximo de eventos. No nosso caso, o número de evento equivale ao número de pontos na borda dos objetos, igual em média para todas as formas a 201 pontos. O número de intervalos encontrados para obter a melhor performance do classificador foi de 20 *bins*, próximo do sugerido. Porém cabe ressaltar que não existe qualquer garantia que a adoção desse número seja ideal para todos os casos, sendo o procedimento correto verificar a melhor performance do classificador em função das variações em sua entrada.

Esta etapa tem como objetivo buscar entre os valores de  $q$ , aquele em que o sistema de classificação apresenta a melhor eficiência no reconhecimento das formas. No caso não extensivo, o ganho (modificação) de informação é dado pela seguinte relação entre os jogos de probabilidade,  $\Delta\sigma_q = [\frac{1}{(1-q)}] \cdot (1 - p'(1 - q)) - (1 - p(1 - q))$ . Dessa forma, o objetivo continua sendo buscar padrões cujo o acréscimo de informações seja o menor possível, garantido o reconhecimento dos objetos em suas respectivas classes. Como apresentado anteriormente na seção 2.3.3, a medida de distância através da relação Kullback-Leibler generalizada pode ser obtida a partir da seguinte equação:

$$D_{KL_q}(p : p') = \sum_i \frac{p_i^q}{1 - q} \cdot (p_i^{1-q} - p_i'^{1-q}) \quad (6.6)$$

A soma entre as duas distâncias  $D_{KL_q}(p : p') + D_{KL_q}(p' : p)$  garantem que a medida de distância entre o objeto padrão e aquele apresentado ao classificador sejam equivalentes. A função discriminante utilizada será então a versão simétrica da Distância Kullback-Leibler (ou Divergência  $J$ ) generalizada:

$$D_q(p : p') = D_{KL_q}(p : p') + D_{KL_q}(p' : p) \quad (6.7)$$

Uma primeira análise consiste em avaliar a evolução da medida de distância Kullback-Leibler para o objeto “círculo” onde é adicionado ruído, como apresentado na figura 6.7 para o caso  $q = 1$  (extensivo). Os gráficos da figura 6.12 apresentam o comportamento da função  $D_q(p : p')$  para diferentes valores do parâmetro de não extensividade, especialmente para casos superextensivos ( $0 < q < 1$ ) e subextensivos ( $q > 1$ ).

Para uma relação SNR alta (sinal  $\gg$  ruído), a medida de distância, para todos os casos, é próxima de zero. No entanto ao adicionarmos ruído na borda do objeto temos um distanciamento da forma original. Este comportamento é observado para todos os casos. No entanto, para  $0 < q \leq 1$  a medida de distância se afasta continuamente com o crescimento do ruído. Para os casos em que  $q > 1$  existe um ponto de máximo situado em torno de  $15 \text{ dB}$ <sup>4</sup>. A partir desse ponto, ao adicionarmos ruído, a distância tende a

---

<sup>4</sup>Esta região equivale àquela cuja a expectativa da maioria das pessoas as quais perguntamos ao apresentar a figura 6.12, qual relação SNR o objeto “círculo” inicial deixava de ter uma forma circular

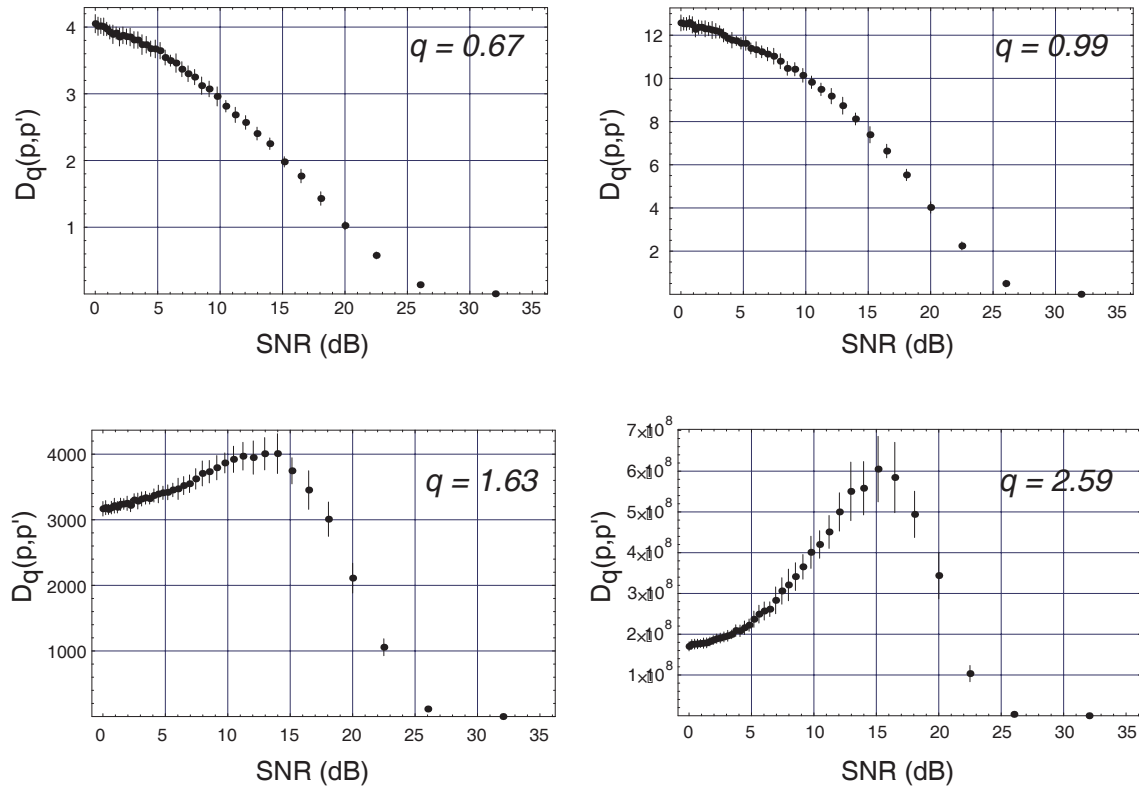


Figura 6.12: Os gráficos representam a excursão da medida de distância Kullback-Leibler,  $D_q(p : p')$ , entre uma forma padrão perfeita e uma na qual foi adicionado ruído, de diferentes amplitudes, em sua borda. Quatro valores para o parâmetro  $q$  são apresentados. A excursão observada para a distância  $D_q(p : p')$  é ampliada para valores de  $q > 1$  (sistemas subextensivos) e atenuada para valores de  $0 < q \leq 1$  (sistemas superextensivos).

se aproximar de novo da forma inicial. Na prática, nesta região estamos comparando o objeto original com um outro de forma aleatória que não se assemelha mais a uma forma circular. Uma forma aleatória, i.e., onde todos os intervalos do histograma de distâncias tem a mesma probabilidade, pode ser medida através de sua entropia. O gráfico da figura 6.13 apresenta a evolução da entropia da forma ruidosa em função da relação SNR. Os valores são normalizados pelo valor máximo da entropia para cada valor de  $q$ . A entropia máxima pode ser observada nos gráficos para valores de  $q > 1$ . O ajuste nos valores de  $p_i$ , ( $p_i^q$ ) na fórmula da entropia generalizada modifica a e passava para um “objeto” de forma aleatória.



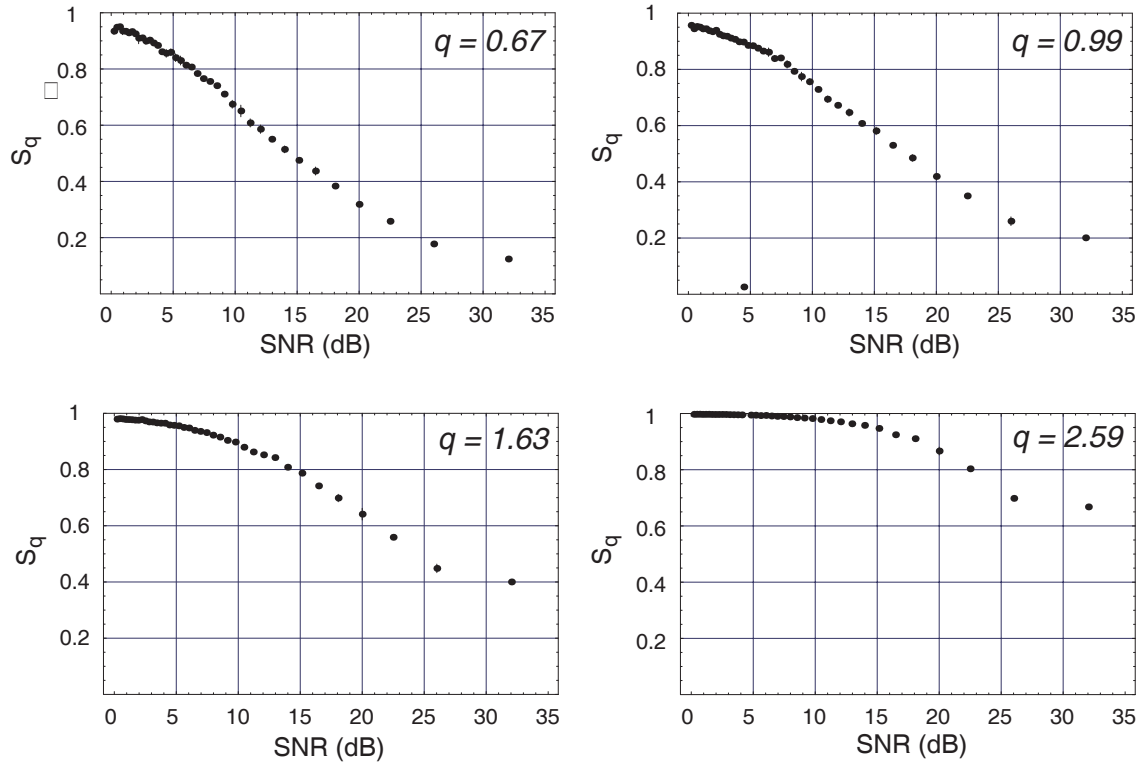


Figura 6.13: *Variação da entropia generalizada para as formas ruidosas da fig.6.12. Os valores são normalizados pela entropia máxima para cada  $q$ . Quando  $q > 1$  forma tende mais rapidamente a ser considerada como uma forma aleatória.*

distribuição original para uma equivalente a de uma distribuição equiprovável, quando  $q > 1$ . O mesmo não acontece para  $0 < q < 1$ , onde se privilegia eventos raros no histograma ( $p_i^q > p_i$ , se  $0 < q < 1$ ) e a entropia continua crescente, sem saturação, em todo o intervalo. A amplitude da barra de erro, para  $q > 1$ , é também superior quando comparada com  $0 < q \leq 1$ . Este efeito irá certamente prejudicar a eficiência do classificador para um conjunto muito grande de classes a serem separadas.

O próximo passo é avaliar a eficiência do classificador em função de  $q$ . Este resultado é apresentado na figura 6.14. Na figura 6.14a é apresentado a taxa de eficiência do classificador para diversos valores de  $q$ . A curva equivalente a medida de distância Kullback-Leibler, para  $q = 1$  (caso extensivo, entropia de Shannon) é adicionada ao gráfico para fins de comparação. Existem dois regimes distintos neste gráfico, para

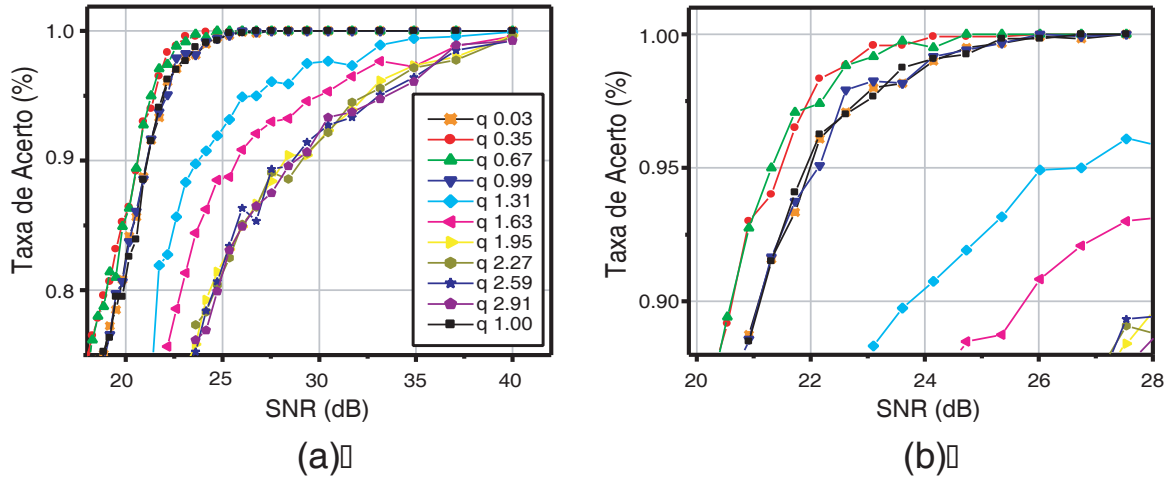


Figura 6.14: Taxa de Acerto em função da relação SNR, para diferentes valores de  $q$ . (a) Comportamento para todos os casos ( $q > 1$ ,  $q = 1$  e  $0 < q \leq 1$ ). Duas regiões são destacadas:  $0 < q \leq 1$  e  $q > 1$ . (b) Comportamento para os casos  $0 < q \leq 1$  e  $q = 1$  ampliados. Taxa de acerto é ligeiramente melhor para  $q = 0.67$ .

$q > 1$  e para  $0 < q \leq 1$ . Todos os casos superextensivos ( $0 < q < 1$ ) apresentaram eficiência superior aos casos subextensivos ( $q > 1$ ). As curvas com  $q > 1$  tem uma perda de eficiência mesmo para uma relação SNR muito elevada ( $> 35$  dB, i.e., sinal é praticamente 60 vezes superior ao ruído).

Na figura 6.14b o mesmo gráfico é apresentado, explorando a região dos casos superextensivos e extensivo ( $0 < q \leq 1$ ). Neste gráfico podemos observar que duas curvas se destacam,  $q = 0.35$  e  $q = 0.67$ , tendo esta última uma eficiência ligeiramente superior. Isto comprova que o sistema de classificação consegue aumentar sua eficiência se estiver utilizando o método de decisão baseado na entropia não extensiva.

Quando  $q > 1$  pequenas flutuações no jogo de probabilidade, devido a ruídos de baixa amplitude na borda das formas, pode acarretar grandes variações na medida de distância  $D_q(p : p')$ , levando inclusive a distâncias maiores que àquelas esperadas para as distâncias com outras classes. O processo decisório sendo baseado na escolha da menor distância ficará então prejudicado e o classificador cometerá um número muito grande de erros. Quando  $0 < q \leq 1$  as variações são menores fazendo com que o sistema

tenha um comportamento mais estável. Este comportamento nos levou a analisar a equação de medida de distância Kullback-Leibler para a comparação de dois jogos de probabilidades conhecidos e controlados.

### 6.6.1 A região ótima de $q$ – dois jogos de probabilidades.

Construímos dois jogos de probabilidade  $p_a$  e  $p_b$ , baseados em duas distribuições normais com o mesmo valor médio,  $\mu$ , e desvio padrão igual a  $\sigma_a$  e  $\sigma_b$ , figura 6.15. Iremos manter a primeira distribuição com  $\sigma_a$  fixo e modificar progressivamente o desvio padrão da segunda e avaliar o comportamento da função  $D_q(p_a, p_b)$  para diferentes valores de  $q$ .

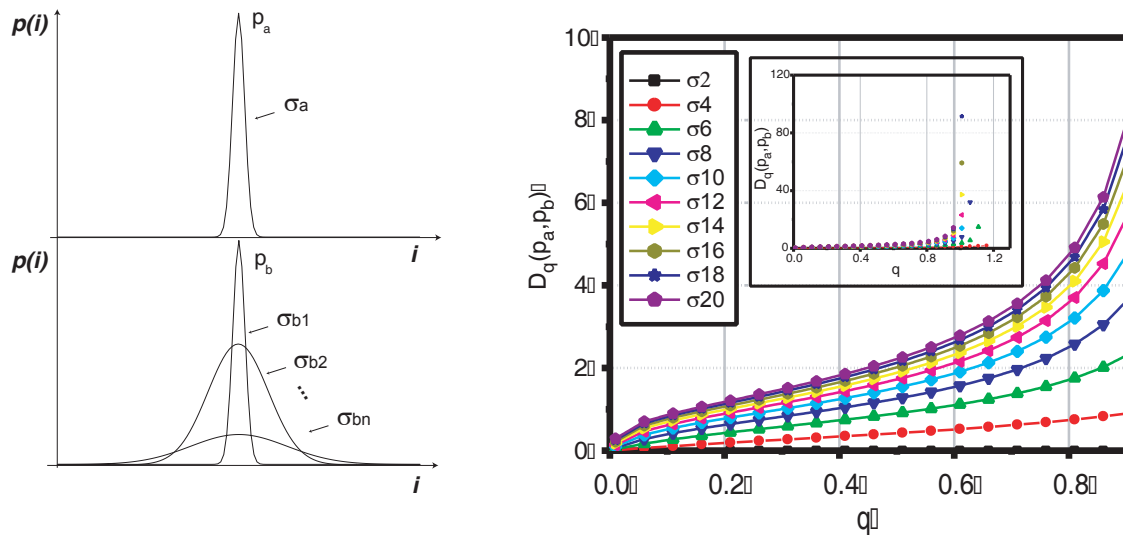


Figura 6.15: Evolução da medida de distância  $D_q(p_a, p_b)$  entre dois jogos de probabilidades para uma distribuição normal em função do parâmetro de não extensividade  $q$ , para diferentes valores de  $\sigma_b$ . Quando utilizamos  $0 < q \leq 1$  a curva tem um comportamento crescente mais controlado, tanto para pequenos valores de desvio padrão quanto para grandes valores.

O resultado é apresentado no gráfico da figura 6.15. Para o intervalo  $0 < q < 1$ , uma pequena modificação no desvio padrão do jogo  $p_b$ , acarreta uma pequena variação na medida de distância Kullback-Leibler entre os dois jogos. A medida que  $q$  se aproxima (e passa) de 1 existe uma grande excursão da medida de distância para uma pequena

modificação no desvio padrão. Este comportamento leva o classificador de formas entrópico a apresentar um número maior de erros no processo de reconhecimento, pois esta medida pode ser maior que aquela obtida na comparação com outros padrões. O comportamento da distância  $D_q(p_a, p_b)$  é menos sensível na região entre 0.2 e 0.7 em que o comportamento das curvas é praticamente linear em função de  $q$ .

## 6.7 Conclusão

Neste capítulo utilizamos técnicas entrópicas como ponto chave no processo de reconhecimento de padrões, comparando as formas através da medida de distância Kullback-Leibler entre dois jogos de probabilidades. A descrição do objeto ou regiões segmentadas é uma etapa fundamental em qualquer processo de reconhecimento de padrões. As etapas que nos levaram a segmentação da imagem e após na representação do “objeto” por seu contorno podem e irão certamente influenciar diretamente no resultado final. A robustez de qualquer sistema de classificação dependerá de quanto este será influenciado pelas possíveis falhas no processo de segmentação. Para caracterizar esta robustez analisamos o classificador entrópico para varias formas ruidosas. O classificador apresentou uma eficiência de 100% de acertos para uma relação SNR máxima de 28.8  $dB$ . O estudo da eficiência do classificador por entropia não extensiva apresentou uma resultado melhor, para  $q = 0.67$ , com uma taxa de acerto de 100% para uma relação SNR de 24.1  $dB$ . Para uma taxa de acerto acima de 80% a relação SNR pode ser de até 18.41  $dB$ .

De forma complementar, podemos afirmar que classificadores entrópicos tem melhor eficiência com métodos não extensivos e devem utilizar como parâmetro de não extensividade  $q$  no intervalo de 0.2 a 0.7. Em um sistema de classificação podemos escolher entres vários jogos de probabilidades que caracterizem melhor o sistema em termos da informação relevante no processo de reconhecimento. No limite, esperamos uma definição específica do jogo de probabilidade que possa indicar melhor o processo de decisão do classificador, certamente existem várias. Esta definição pode ser um dos desafios abertos por este trabalho.

O melhor resultado para o caso não extensivo, pode ser uma comprovação das correlações de longo alcance da forma (borda) dos objetos. Certamente os pontos pertencentes a borda, por fazerem parte do mesmo objeto, estão correlacionados entre si. Os resultados encontrados para a eficiência deste método são bastante encorajadores, pois o alto índice de acerto para grandes variações na forma apresentada é um diferencial quando comparados a outros métodos.

Uma outra vantagem deste método é ter independência de rotação e do tamanho da forma na imagem, porem estudos específicos de orientação da forma devem ainda ser realizados com o intuito de caracterizar a eficiência do método em função de sua orientação. Cabe ressaltar que esta característica não vem especificamente do método entrópico, mas da representação do jogo de probabilidades que caracteriza o objeto pela distância dos pontos de borda a seu centro de massa. Porém, a discretização imposta pela representação da imagem na forma digital insere pontos na borda dos objeto podendo modificar de forma sensível seu tamanho e sua borda. Este procedimento poderá alterar de forma significativa o histograma e desta forma a eficiência da classificação.

Uma limitação no processo de classificação esta relacionado ao fato do classificador sempre indicar uma saída entres todas as possíveis, i.e., ele terá sempre um valor mínimo de distância estatística entre os dois jogos de probabilidades, mesmo que esta não indique a classe correta. Um estudo específico que possa indicar este limiar de decisão é um interessante ponto de continuidade deste trabalho. Além disso, uma técnica complementar que auxilie o processo decisório poderia estar baseada numa coletânea de funções de vários atributos extraídos dos objetos, todas baseadas em medidas entrópicas dos objetos na imagem.



# Capítulo 7

## Conclusão

Neste trabalho utilizamos os conceitos de entropia em várias aplicações onde os dados estão organizados sob a forma de imagens. Mais especificamente, a proposta da tese foi aplicar esse conceito em técnicas de segmentação de imagens e reconhecimento de formas. Na cadeia de processamento da informação visual estas são duas etapas importantes na caracterização da informação presente na imagem. Imagens são na realidade um suporte físico para troca de informações. Esta informação está associada aos objetos presentes na imagem e sua interpretação correta dependerá, na prática, do sistema visual humano como um todo. Através da entropia é possível realizar medidas que possam exprimir a quantidade de informação em um processo de seleção de um objeto entre vários existentes. O tratamento da informação, em toda a cadeia do tratamento da imagem, se ajusta quase sempre a este processo, seja na separação do sinal pertinente daquele não interessante, seja no reconhecimento ou na classificação de objetos em classes. Desta forma, a análise da imagem através de técnicas entrópicas é uma abordagem consistente e como descrito neste trabalho apresenta resultados interessantes e motivadores.

Imagens podem ser caracterizadas por seu conteúdo não extensivo. As correlações de longo alcance na imagem correspondem as regiões que compõem o próprio objeto (ou fundo da imagem). Pontos de um mesmo objeto estão correlacionados por sua forma ou/e por sua intensidade luminosa (cor ou textura). A técnica não extensiva irá utilizar esta característica através do cálculo da entropia generalizada, dependente do

parâmetro  $q$ . A associação do parâmetro  $q$  a uma característica física do processo de formação da imagem (ou dos objetos) é ainda um desafio. Nesse trabalho o  $q$  funcionou como um importante parâmetro de ajuste, melhorando (adaptando) os resultados de segmentação e classificação de formas. Porém, estudos específicos poderiam ser realizados no sentido de se obter funções de probabilidades que pudessem caracterizar diferentemente os objetos, levando talvez a uma relação direta entre a forma do objeto e o parâmetro  $q$ .

No caso de segmentação de imagens, podemos afirmar que as técnicas entrópicas (convencionais e não extensivas) podem ser utilizadas como métodos de separação das regiões para a maior parte dos casos (classes de imagens). Porém, naqueles casos em que as regiões correspondentes aos objetos são pouco prováveis (em termos de sua representação no histograma de luminância), pois os picos estão submersos por uma outra região muito provável ou em seu próprio ruído, podemos dizer que os métodos entrópicos funcionam melhor (ver casos 3 e 4 apresentados no capítulo 5). O processo de segmentação entrópica pode apresentar resultados bastante satisfatórios, e o fato de permitir uma modificação do parâmetro de não extensividade  $q$  no cálculo da entropia nos permite ajustar precisamente onde se quer segmentar a imagem. Esta característica pode ser certamente um diferencial na montagem de todo o processo de tratamento da informação, objetivo principal de todo processo de tratamento da imagem.

No caso da segmentação entrópica por entropia relativa podemos afirmar é que esta é uma técnica robusta e irá funcionar para uma variedade muito grande de imagens com ótimo resultado final para a classificação das regiões. Nos casos em que se puder aproximar o histograma da imagem por uma distribuição gaussiana não existirá uma dependência do método em função do parâmetro de não extensividade. Em casos onde a imagem não segue uma distribuição específica,  $q$  poderá ser mais uma vez um parâmetro de ajuste do ponto de segmentação.

Podemos encarar o processo de segmentação da imagem com um processo de classificação dos pixels em 2 ou mais regiões. Na prática, o cálculo do histograma (regiões da imagem que tem a mesma intensidade luminosa) é o processo que define as cor-



relações espaciais da imagem. Assim, imagens de natureza diferente podem apresentar o mesmo histograma, e o ponto de corte deverá neste caso ser o mesmo. No entanto, o resultado da segmentação da imagem será relativamente diferente, pois a conclusão da etapa de segmentação implica em classificar os pixels da imagem em regiões distintas. Logo, o resultado final poderá ser relativamente diferente, pois dependerá da distribuição espacial dos mesmos na imagem. O processo de binarização entrópica deve ser visto como um processo de separação “objeto” e “fundo” pela distribuição de sua luminância, maximizando a soma de informações de processos independentes. Este ponto nos leva certamente a caracterizar o “objeto” dentro do contexto onde ele se encontra, colocando em relação o objeto e o fundo da imagem, através do histograma e através da entropia.

A utilização de técnicas entrópicas como base para o reconhecimento de padrões foi uma proposta inovadora deste trabalho. Utilizamos a medida de distância Kullback-Leibler entre dois jogos de probabilidades, uma representando o objeto obtido na imagem analisada e outro representando um padrão previamente conhecido (classe) como função critério para o processo de reconhecimento. O objetivo era encontrar dentro de um conjunto de jogos de probabilidade, aquele com menor acréscimo de informação (o mais próximo). Neste trabalho utilizamos para representar o objeto a distribuição de distância euclidiana entre os pontos de borda e o centro de massa de cada forma. Procuramos caracterizar o sistema de classificação em função do ruído adicionado às formas a ele apresentadas. Os testes realizados com a entropia convencional e não extensiva são bastante encorajadores. Podemos afirmar que classificadores entrópicos tem uma eficiência melhor com métodos não extensivos e devem utilizar  $q$  no intervalo de 0.2 a 0.7, pois o sistema de classificação fica mais estável e conseqüentemente consegue aumentar a taxa de acertos para formas mais ruidosas. Uma possível extensão deste trabalho seria alterar o cálculo da função que determina o jogo de probabilidades que define a forma. Um estudo específico com diferentes funções critérios, maximizando a taxa de acerto, pode ser uma complemento importante para esta técnica.



# Apêndice A

## Demonstrações Matemáticas

### Entropia Generalizada para $q = 1$

Nesta seção, apresentamos a demonstração do retorno da expressão da entropia generalizada no limite em que  $q \rightarrow 1$ . A entropia não extensiva é definida como sendo

$$S_q = \frac{1 - \sum_i p_i^q}{q - 1} \quad (\text{A.1})$$

Para  $q = 1$  temos uma indeterminação do tipo  $\frac{0}{0}$ . Aplicando a regra de *L'Hospital*:

$$S_1 = \lim_{q \rightarrow 1} S_q = \frac{d\left(1 - \sum_i p_i^q\right)}{\frac{d(q-1)}{dq}} = \frac{d(-p^q)}{dq} \quad (\text{A.2})$$

Resolvendo a derivada anterior:

$$S_1 = \frac{d(-p^q)}{dq} = -p^q \cdot \log p \quad (\text{A.3})$$

assim, temos que:

$$S = \lim_{q \rightarrow 1} S_q = - \sum_i p_i \cdot \log p_i \quad (\text{A.4})$$

que é a expressão convencional da entropia Boltzmann/Gibbs/Shannon.

## Entropia Relativa Generalizada para $q = 1$

Nesta seção, apresentamos a demonstração do retorno da expressão da entropia relativa generalizada no limite em que  $q \rightarrow 1$ . A entropia relativa generalizada é definida como sendo

$$D_q(p : p') = \sum_i \frac{p_i^q}{1 - q} \cdot (p_i^{1-q} - p_i'^{1-q}) \quad (\text{A.5})$$

em que  $p$  e  $p'$  são jogos de probabilidades no intervalo  $[0, 1]$ .

Para  $q = 1$  temos uma indeterminação do tipo  $\frac{0}{0}$ . Aplicando a regra de *L'Hospital*:

$$D_1(p : p') = \lim_{q \rightarrow 1} D_q(p : p') = \frac{\frac{d(p^q \cdot (p^{1-q} - p_i'^{1-q}))}{dq}}{\frac{d(1-q)}{dq}} = \frac{d(\frac{p' \cdot p^q}{p'^q})}{dq} \quad (\text{A.6})$$

Resolvendo a derivada anterior:

$$D_1(p : p') = \frac{d(\frac{p' \cdot p^q}{p'^q})}{dq} = p \cdot \log\left(\frac{p}{p'}\right) \quad (\text{A.7})$$

temos que:

$$D(p : p') = \lim_{q \rightarrow 1} D_q(p : p') = \sum_i p \cdot \log\left(\frac{p}{p'}\right) \quad (\text{A.8})$$

que é a expressão convencional da entropia relativa Kullback-Leibler.

## Demonstração do *Teorema 1* - Entropia Shannon

Nesta seção é apresentada a demonstração do **Teorema 1** proposto por Shannon, seção 2.3.2. A dedução de expressão para a entropia de Shannon será feita a partir das propriedades a seguir:

- **P1.** A função  $S$  deve ser contínua em  $p_i$
- **P2.** Se  $p_i = \frac{1}{n}$  então  $S(n)$  deve ser uma função monotônica crescente. Para acontecimentos equiprováveis há uma maior incerteza sobre o resultado, quando o número de resultados possíveis é maior.

- **P3.** Se uma escolha for subdividida em escolhas sucessivas, então a entropia original deve ser a soma ponderada das duas entropias individuais. O sentido disto é ilustrado na figura A.1. À esquerda temos três possibilidades  $p_1 = \frac{1}{2}$ ,  $p_2 = \frac{1}{3}$  e  $p_3 = \frac{1}{6}$ . À direita, começamos por escolhas de probabilidade  $\frac{1}{2}$  e a segunda fase faz-se uma nova escolha com probabilidades  $\frac{2}{3}$  e  $\frac{1}{3}$ . Os resultados finais têm as mesmas probabilidades que anteriormente. Exige-se, neste caso especial que

$$S\left(\frac{1}{2}, \frac{1}{3}, \frac{1}{6}\right) = S\left(\frac{1}{2}, \frac{1}{2}\right) + \frac{1}{2}S\left(\frac{2}{3}, \frac{1}{3}\right) \quad (\text{A.9})$$

O coeficiente  $\frac{1}{2}$  é o peso introduzido porque a segunda escolha só ocorre metade das vezes.

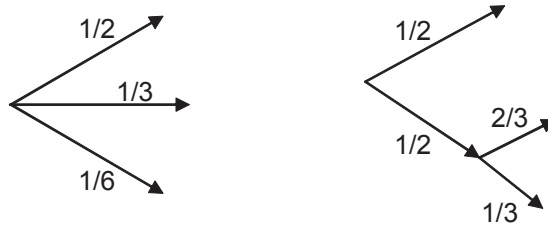


Figura A.1: *Decomposição da opção em três possibilidades*

Para a demonstração de que a única função que satisfaz estas três condições anteriores é

$$S = -k \sum_i p_i \log(p_i) \quad (\text{A.10})$$

tomamos dois resultados que derivam da **P3**:

(I) - Supondo que se faz uma escolha entre  $s^m$  resultados equiprováveis e que  $m$  escolhas são feitas entre  $s$  resultados equiprováveis. Pela propriedade **P3** de Shannon, tem-se que

$$S\left(\frac{1}{s^m}, \dots, \frac{1}{s^m}\right) = mS\left(\frac{1}{s}, \dots, \frac{1}{s}\right) \quad (\text{A.11})$$

Vejamos o caso em que  $s = 2$ :

Se  $m = 2$  então, têm-se de relacionar uma escolha entre  $2^2 = 4$  resultados equiprováveis e duas escolhas com 2 resultados equiprováveis, o que pode ser observado na figura A.2.

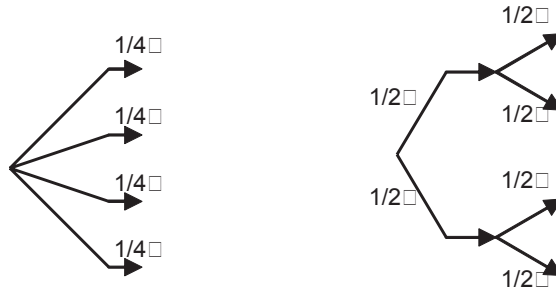


Figura A.2: Escolhas com resultados equiprováveis ( $m = 2$ ).

Pela propriedade **P3**, temos:

$$S\left(\frac{1}{2^2}, \frac{1}{2^2}, \frac{1}{2^2}, \frac{1}{2^2}\right) = S\left(\frac{1}{2}, \frac{1}{2}\right) + \frac{1}{2}S\left(\frac{1}{2}, \frac{1}{2}\right) + \frac{1}{2}S\left(\frac{1}{2}, \frac{1}{2}\right) = 2S\left(\frac{1}{2}, \frac{1}{2}\right) \quad (\text{A.12})$$

Se  $m = 3$  temos que relacionar uma escolha dentre  $2^3 = 8$  resultados equiprováveis e três escolhas com 2 resultados equiprováveis, o que pode ser observado de forma esquemática na figura A.3.

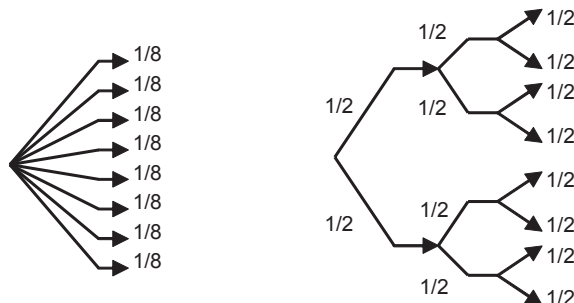


Figura A.3: Escolhas com resultados equiprováveis ( $m = 3$ ).

Logo,

$$\begin{aligned}
& S\left(\frac{1}{2^3}, \frac{1}{2^3}, \frac{1}{2^3}, \frac{1}{2^3}, \frac{1}{2^3}, \frac{1}{2^3}, \frac{1}{2^3}, \frac{1}{2^3}\right) = \\
& S\left(\frac{1}{2}, \frac{1}{2}\right) + \frac{1}{2}S\left(\frac{1}{2}, \frac{1}{2}\right) + \frac{1}{2}S\left(\frac{1}{2}, \frac{1}{2}\right) + \frac{1}{4}S\left(\frac{1}{2}, \frac{1}{2}\right) + \frac{1}{4}S\left(\frac{1}{2}, \frac{1}{2}\right) + \frac{1}{4}S\left(\frac{1}{2}, \frac{1}{2}\right) + \frac{1}{4}S\left(\frac{1}{2}, \frac{1}{2}\right) \\
& = 3 \cdot S\left(\frac{1}{2}, \frac{1}{2}\right) \tag{A.13}
\end{aligned}$$

Qualquer que seja o inteiro  $m$ , temos que

$$S\left(\frac{1}{2^m}, \dots, \frac{1}{2^m}\right) = m \cdot S\left(\frac{1}{2}, \frac{1}{2}\right) \tag{A.14}$$

Para o caso de  $s = 3$ , qualquer que seja o inteiro que tomarmos (em cada uma das  $m$  escolhas existem 3 resultados equiprováveis), teremos que

$$S\left(\frac{1}{3^m}, \frac{1}{3^m}\right) = m \cdot S\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) \tag{A.15}$$

Assim, para qualquer inteiros  $s$  e  $m$  que tomemos, teremos que

$$S\left(\frac{1}{s^m}, \dots, \frac{1}{s^m}\right) = m \cdot S\left(\frac{1}{s}, \dots, \frac{1}{s}\right) \tag{A.16}$$

(II) - Considere-se o caso particular em que se faz uma escolha entre 10 resultados equiprováveis e o caso em que se faz uma escolha entre 3 resultados possíveis com probabilidades  $\frac{2}{10}, \frac{3}{10}, \frac{5}{10}$  e, se o primeiro resultado for o escolhido realiza-se uma outra escolha entre 2 resultados equiprováveis, ou se o terceiro resultado for escolhido, realiza-se uma outra escolha entre 5 resultados equiprováveis. Este processo é ilustrado na esquema da figura A.4.

Cada um dos resultados finais tem probabilidade  $\frac{1}{10}$  de ocorrer, logo

$$\begin{aligned}
S\left(\frac{1}{10}, \frac{1}{10}, \dots, \frac{1}{10}\right) &= S\left(\frac{2}{10}, \frac{3}{10}, \frac{5}{10}\right) + \frac{2}{10}S\left(\frac{1}{2}, \frac{1}{2}\right) + \\
& \frac{3}{10}S\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) + \frac{5}{10}S\left(\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}\right) \tag{A.17}
\end{aligned}$$

Generalizando, se considerarmos o caso em que se faz uma escolha entre  $n$  resultados possíveis com probabilidades  $p_1, p_2, \dots, p_n$  sendo

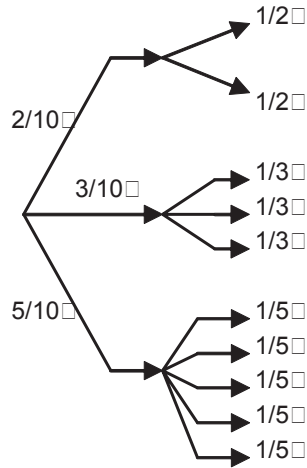


Figura A.4: Diagrama de escolhas para o caso (II).

$$p_i = \frac{n_i}{\sum_i n_i} \quad (\text{A.18})$$

e no caso de se ter escolhido o resultado  $i$ , faz-se uma nova escolha entre  $n_i$  resultados equiprováveis. Estas escolhas podem ser interpretadas como uma divisão de uma escolha com resultados equiprováveis, em duas escolhas sucessivas. Assim,

$$S\left(\frac{1}{\sum_i n_i}, \dots, \frac{1}{\sum_i n_i}\right) = S(p_1, \dots, p_n) + \sum_{i=1}^n p_i \cdot S\left(\frac{1}{n_i}, \dots, \frac{1}{n_i}\right) \quad (\text{A.19})$$

Tendo como base estes casos particulares, podemos passar a dedução formal da expressão da entropia de Shannon:

Considere-se  $S\left(\frac{1}{n}, \dots, \frac{1}{n}\right) = A(n)$  e dois números inteiros  $s$  e  $t$ . Do resultado em (I) sabemos que podemos decompor a escolha de  $s^m$  resultados equiprováveis numa série de  $m$  escolhas, todas equiprováveis, obtendo

$$A(s^m) = m \cdot A(s) \quad (\text{A.20})$$

da mesma forma,

$$A(t^n) = n \cdot A(t) \quad (\text{A.21})$$



Podemos escolher um  $n$  arbitrariamente grande e um  $m$  que satisfaça

$$s^m \leq t^n \leq s^{(m+1)} \quad (\text{A.22})$$

$$\log s^m \leq \log t^n \leq \log s^{(m+1)} \quad (\text{A.23})$$

$$\frac{\log s^m}{n \log s} \leq \frac{\log t^n}{n \log s} \leq \frac{\log s^{(m+1)}}{n \log s} \quad (\text{A.24})$$

$$\frac{m}{n} \leq \frac{\log t}{\log s} \leq \frac{m+1}{n} \quad (\text{A.25})$$

logo

$$\frac{m}{n} - \frac{1}{n} \leq \frac{\log t}{\log s} \leq \frac{m}{n} + \frac{1}{n} \quad (\text{A.26})$$

Fazendo  $\epsilon = \frac{1}{n}$  arbitrariamente pequeno,

$$-\epsilon \leq \frac{\log t}{\log s} - \frac{m}{n} \leq \epsilon \quad (\text{A.27})$$

Por outro lado, atendendo à propriedade P2 a à desigualdade  $s^m \leq t^n \leq s^{(m+1)}$ , temos que

$$A(s^m) \leq A(t^n) \leq A(s^{(m+1)}) \quad (\text{A.28})$$

$$m \cdot A(s) \leq n \cdot A(t) \leq (m+1) \cdot A(s) \quad (\text{A.29})$$

dividindo por  $n \cdot A(s)$

$$\frac{m}{n} \leq \frac{A(t)}{A(s)} \leq \frac{m}{n} + \frac{1}{n} \quad (\text{A.30})$$

$$\frac{m}{n} - \frac{1}{n} \leq \frac{A(t)}{A(s)} \leq \frac{m}{n} + \frac{1}{n} \quad (\text{A.31})$$

$$-\epsilon \leq \frac{A(t)}{A(s)} - \frac{m}{n} \leq \epsilon \quad (\text{A.32})$$

De A.27 e A.32 podemos escrever

$$-2\epsilon \leq \frac{A(t)}{A(s)} - \frac{\log t}{\log s} \leq 2\epsilon \quad (\text{A.33})$$

ou seja,

$$\left| \frac{A(t)}{A(s)} - \frac{\log t}{\log s} \right| \leq 2\epsilon, \forall \epsilon > 0 \quad (\text{A.34})$$

Pelo que  $\frac{A(t)}{A(s)} = \frac{\log t}{\log s}$  ou ainda,  $A(t) \propto \log t$ . Portanto, a função  $A(t)$  será da forma  $A(t) = k \log t$ .

Logo,

$$S\left(\frac{1}{t}, \dots, \frac{1}{t}\right) = k \log t \quad (\text{A.35})$$

Para que a propriedade P2 se verifique, a constante  $k$  tem que ser positiva. Considerando agora o resultado (II)

$$S\left(\frac{1}{\sum_i n_i}, \dots, \frac{1}{\sum_i n_i}\right) = S(p_1, \dots, p_n) + \sum_{i=1}^n p_i \cdot S\left(\frac{1}{n_i}, \dots, \frac{1}{n_i}\right) \quad (\text{A.36})$$

e aplicando a equação A.35, obtém-se

$$k \log \sum_{i=1}^n n_i = S(p_1, \dots, p_n) + \sum_{i=1}^n p_i \cdot S\left(\frac{1}{n_i}, \dots, \frac{1}{n_i}\right) \quad (\text{A.37})$$

$$S(p_1, \dots, p_n) = k \log \sum_{i=1}^n n_i - \sum_{i=1}^n p_i k \log n_i \quad (\text{A.38})$$

$$S(p_1, \dots, p_n) = k \sum_{i=1}^n p_i \log \sum_{i=1}^n n_i - \sum_{i=1}^n p_i k \log n_i \quad (\text{A.39})$$

$$S(p_1, \dots, p_n) = -k \sum_{i=1}^n p_i \left[ -\log \sum_{i=1}^n n_i + \log n_i \right] \quad (\text{A.40})$$

$$S(p_1, \dots, p_n) = -k \sum_{i=1}^n p_i \log \frac{n_i}{\sum_{i=1}^n n_i} \quad (\text{A.41})$$

e pela equação A.18

$$S(p_1, \dots, p_n) = -k \sum_{i=1}^n p_i \log p_i \quad (\text{A.42})$$

Aplicando apenas as propriedades enunciadas por Shannon, chegamos a sua equação. Segundo Shannon, as quantidades da forma

$$S = - \sum_i p_i \log p_i \quad (\text{A.43})$$

desempenham um papel fundamental na teoria da informação. A esta quantidade, chamou entropia do conjunto de probabilidades  $p_1, \dots, p_n$ . A constante  $k$  quantifica meramente a escolha de uma unidade de medida.



# Apêndice B

## Métodos de Segmentação de Imagens

Este Apêndice apresenta uma breve descrição dos métodos de segmentação *Two Peaks* e *Iterative Selection*. Estes foram os métodos não entrópicos utilizados para fins de comparação de resultados das técnicas de segmentação, apresentadas no Capítulo 5. Os métodos de binarização procuram determinar um nível ótimo de luminância para segmentar a imagem, fornecendo como saída uma imagem binária.

### Método *Iterative Selection*

O algoritmo *iterative selection* foi proposto por Ridler e Calvard (1978), como um método de binarização automática. Neste método, define-se inicialmente um valor inicial para o *threshold* ( $t$ ), que em geral é a média da luminância de toda a imagem, com a aproximação para o inteiro mais próximo.

A partir da escolha inicial de  $t$ , calcula-se a média dos níveis de cinza dos pixels acima de  $t$  ( $t_b$ ) e abaixo de  $t$  ( $t_o$ ). A seguir, re-estima-se a posição de  $t$  utilizando o nível de cinza médio entre as duas médias calculadas,  $(t_b + t_o)/2$ . Repete-se este procedimento até que o valor de  $t$  computado não se modifique, convergindo para o valor  $t_{otimo}$ .

Este procedimento é facilmente implementado computacionalmente utilizando o

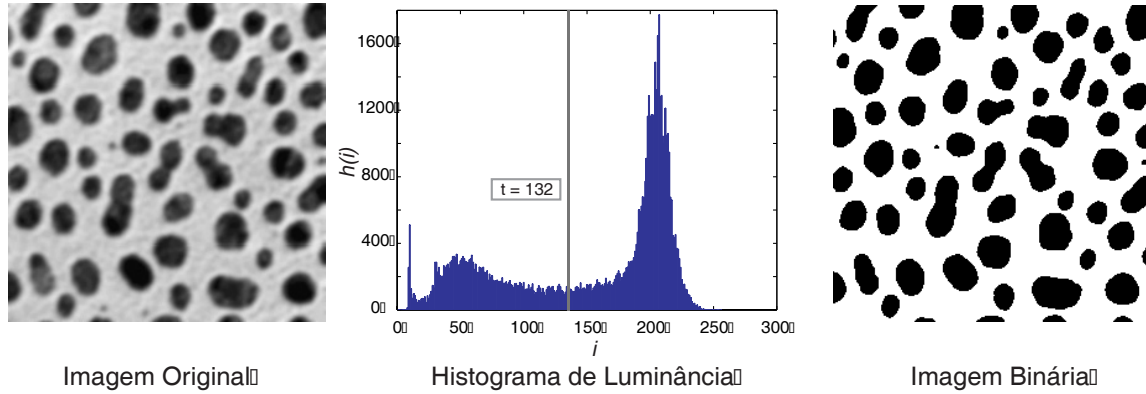


Figura B.1: *Exemplo de segmentação com o método Iterative Selection. A figura apresenta o histograma da imagem original e o  $t_{otimo}$  definido pelo algoritmo.*

histograma de luminância de imagem ( $h$ ), de modo que:

$$t_{otimo} = \frac{\sum_{i=1}^t i \cdot h_i}{2 \sum_{i=1}^t h_i} + \frac{\sum_{j=t+1}^k j \cdot h_j}{2 \sum_{j=t+1}^k h_j} \quad (\text{B.1})$$

em que  $h$  é o histograma de luminância da imagem que possui  $k$  níveis de cinza.

## Método *Two Peaks*

O algoritmo *two peaks* [32] implementa um método de binarização automática localizando dois picos no histograma da imagem e definindo um ponto ótimo entre os dois picos para a binarização.

A localização do segundo pico é um ponto crítico em métodos que tentam encontrar vales entre picos. O problema principal é que o segundo pico em amplitude no histograma tende a ser na região (ou no *bin*) vizinha a região (ou *bin*) que representa o primeiro pico. Este método minimiza este efeito multiplicando os valores do histograma  $h$  pela raiz quadrada da distância em relação ao primeiro pico, conforme a expressão a seguir:

$$P_2 = \max\{(i - x_1)^2 \cdot h_i\} \quad (\text{B.2})$$

em que  $x_1$  é a posição ou argumento do primeiro pico no histograma  $h_i$  e  $P_2$  é o valor do segundo pico encontrado. A posição do segundo pico é definida por  $x_2 = \operatorname{argmax}\{(i - x)^2 \cdot h_i\}$ .

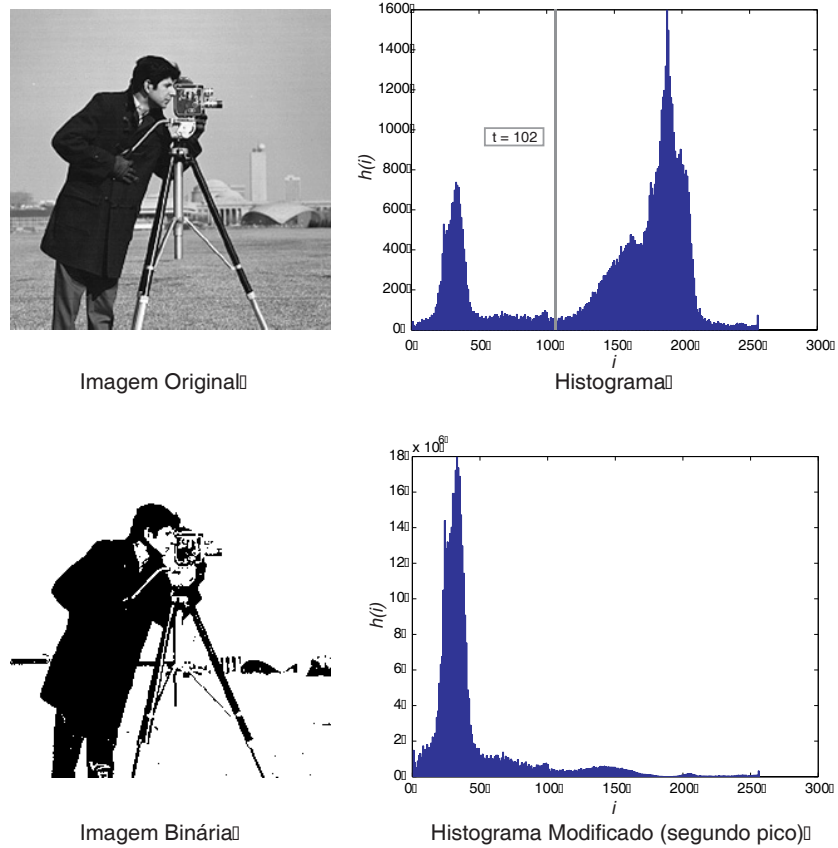


Figura B.2: Exemplo de segmentação com o método *Two Peaks*. A figura apresenta o histograma da imagem original e o histograma modificado para detecção do segundo pico. O  $t_{otimo}$  é definido pelo algoritmo no valor mínimo (vale) entre os picos detectados.

Este procedimento modifica o histograma original da imagem atenuando, em amplitude, a região próxima ao primeiro pico encontrado, dando preferência a picos secundários distantes do pico máximo.





# Apêndice C

## Trabalho Publicado

*Proceedings*

*XV Brazilian Symposium on  
Computer Graphics and Image Processing*

*7-10 October 2002*

*Fortaleza-CE, Brazil*

*Edited by*

*Luiz Marcos Garcia Gonçalves*

*Soraia Raupp Musse*

*Sponsored by*

*Brazilian Computer Society (SBC)*

*Brazilian National Council for Research (CNPq)*

*Brazilian Commission for Higher Education (CAPES)*



Los Alamitos, California

Washington • Brussels • Tokyo

---

# Nonextensive Entropic Image Thresholding

ISRAEL ESQUEF, MARCELO ALBUQUERQUE, MÁRCIO ABUQUERQUE,

CBPF - Centro Brasileiro de Pesquisas Físicas - Rua Xavier Sigaud 150, Urca, 22290-180 Rio de Janeiro, RJ, Brasil  
 {israel,marcelo,mpa}@cbpf.br

**Abstract.** In image processing, one of the most efficient techniques for image segmentation is entropy-based thresholding. In this work it was applied a generalized entropy formalism that represents a recent development in statistical mechanics. We propose, for the first time, an image thresholding method using a nonextensive entropy regarding the presence of nonadditive information content in some image classes. Preliminary results are shown.

## 1 Nonextensive Entropy

The entropy of a discrete source is often obtained from the probability distribution  $p = \{p_i\}$ , and the Shannon entropy may be described as  $S = -\sum_{i=1}^k p_i \ln(p_i)$ , being  $k$  the total number of states. If we consider that a physical system can be decomposed in two statistical independent subsystems A and B, the Shannon entropy has the extensive property (additivity)  $S(A+B) = S(A) + S(B)$ . This formalism has been shown to be restricted to the Boltzmann-Gibbs-Shannon (BGS) statistics.

However, for nonextensive physical systems, some kind of extension appears to become necessary. Tsallis [1] has proposed a generalization of the BGS statistics which is useful for describing the thermostistical properties of nonextensive systems. It is based on a generalized entropic form,

$$S_q = \frac{1 - \sum_{i=1}^k (p_i)^q}{q-1} \quad (1)$$

where the real number  $q$  is a entropic index that characterizes the degree of nonextensivity. This expression recovers to BGS entropy in the limit  $q \rightarrow 1$ . Tsallis entropy has a nonextensive property for statistical independent systems, defined by the following pseudo additivity entropic rule

$$S_q(A+B) = S_q(A) + S_q(B) + (1-q) \cdot S_q(A) \cdot S_q(B) \quad (2)$$

## 2 The Thresholding Technique

Let  $p_i = p_1, p_2, \dots, p_k$  be the probability distribution for an image with  $k$  gray-levels. From this distribution, we derive two probability distributions, one for the object (class A) and the other for the background (class B), given by

$$p_A : \frac{p_1^t}{P^t}, \frac{p_2^t}{P^t}, \dots, \frac{p_t^t}{P^t} \quad \text{and} \quad p_B : \frac{p_{t+1}^t}{P^t}, \frac{p_{t+2}^t}{P^t}, \dots, \frac{p_k^t}{P^t}$$

$$\text{where } P^A = \sum_{i=1}^t p_i \quad \text{and} \quad P^B = \sum_{i=t+1}^k p_i$$

The Tsallis entropy of order  $q$  for each distribution is defined as

$$S_q^A(t) = \frac{1 - \sum_{i=1}^t (p_i)^q}{q-1} \quad S_q^B(t) = \frac{1 - \sum_{i=t+1}^k (p_i)^q}{q-1} \quad (3)$$

The Tsallis entropy  $S_q(t)$  is parametrically dependent upon the threshold value  $t$  for the foreground and background. It is formulated as the sum each entropy, allowing the pseudo-additive property, defined in equation (2). We try to maximize the information measure between the two classes (object and background). When  $S_q(t)$  is maximized, the luminance level  $t$  that maximizes the function is considered to be the optimum threshold value [2].

$$t_{opt} = \operatorname{argmax}[S_q^A(t) + S_q^B(t) + (1-q) \cdot S_q^A(t) \cdot S_q^B(t)] \quad (4)$$

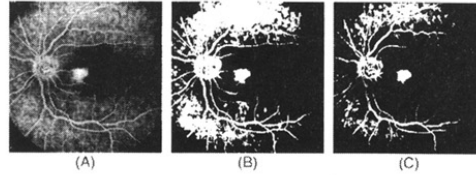


Figure 1: (A) Grayscale 8 bit image, (B) Binary image using Shannon entropy ( $t = 121$ ), (C) Binary image using Tsallis entropy,  $q = 5$  ( $t = 171$ ).

## 3 Conclusions

The preliminary results obtained confirm the viability of using the nonextensive entropy formalism in image thresholding and other segmentation techniques. We believe that Tsallis entropy may trigger some practical future applications in such an area of image processing and recognition.

## References

- [1] C. Tsallis, J. Statistical Phys., 52, 480-487, (1988).
- [2] J. N. Kapur *et al*, C.V.G.I.P. 29, 273-285, (1985).
- [3] T. Yamano, Entropy 2001, 3, 280-292, (2001).



# Referências Bibliográficas

- [1] D. Marr. *Vision*. W. H. Freeman and Company, 1996.
- [2] C. E. Shannon. A mathematical theory of communication. *Bell Sys. Tech. J.*, vol.**27**:p.379–423 e 623–656, Julho e Outubro 1948.
- [3] H. Jeffreys. *Theory of Probability*. Clarendon Press, Oxford, 1983.
- [4] J. E. Cavanaugh. A large-sample model selection criterion based on Kullback's symmetric divergence. *Statistics and Probability Letters*, vol.**42**:p.333–343, 1999.
- [5] C. Tsallis. Possible generalization of Boltzmann-Gibbs statistics. *J. Stat. Phys.*, vol.**52**:p.479–487, 1988.
- [6] L. C. Sampaio, M. P. Albuquerque, and F. S. Menezes. Nonextensivity and Tsallis statistics in magnetic systems. *Physical Review B*, vol.**55**(9):p.5611–5614, Março 1997.
- [7] I. S. Oliveira. Some metallic properties in the framework of Tsallis generalized statistics. *Eur. Phys. J. B*, vol.**14**(1):p.43–46, 2000.
- [8] M. T. Martin, A. R. Plastino, and A. Plastino. Tsallis-like information measure and the analysis of complex signals. *Physica A*, vol.**275**:p.262–271, 1999.
- [9] S. Tong, A. Bezerianos, J. Paul, Y. Zhu, and N. Thakor. Nonextensive entropy measure of eeg following brain injury from cardiac arrest. *Physica A*, vol.**305**(3-4):p.619–628, 2002.

- [10] P. T. Landsberg and V. Vedral. Distributions and channel capacities in generalized statistical mechanics. *Phys. Lett. A*, vol.**247**:p.211–217, 1998.
- [11] A. Capurro, L. Diambra, D. Lorenzo, O. Macadar, M. T. Martin, C. Mostaccio, A. Plastino, E. Rofman, M. E. Torres, and J. Velluti. Tsallis entropy and cortical dynamics: The analysis of eeg signals. *Physica A*, vol.**257**:p.149–155, 1998.
- [12] F. Michael and M. D. Johnson. Financial market dynamics. *Physica A*, In Press:p., 2002.
- [13] R. J. V. Santos. Generalization of Shannon’s theorem for Tsallis entropy. *J. Math. Phys.*, vol.**38**(8):p.4104–4107, 1997.
- [14] L. Borland, A. R. Plastino, and C. Tsallis. Information gain within nonextensive thermostatistics. *J. Math. Phys.*, vol.**39**(12):p., Dezembro 1998.
- [15] E. P. Borges. On a q-generalization of circular and hyperbolic functions. *J. Phys. A*, vol.**31**(23):p.5281–5288, 1998.
- [16] R. C. Gonzalez and R. E. Woods. *Digital image processing*. Addison Wesley, 1993.
- [17] D. Marr and E. Hildreth. *Theory of edge detection*. Proc.R.Soc. (London), 1980.
- [18] T. Pun. Entropic thresholding: a new approach. *Comp. Graph. Image Process.*, vol.**16**:p.210–239, 1981.
- [19] J. N. Kapur, P. K Sahoo, and A. K. C. Wong. A new method for gray-level picture thresholding using the entropy of the histogram. *Comp. Vision Graph. Image Process.*, vol.**29**:p.273–285, 1985.
- [20] S. Kullback. *Information Theory and Statistics*. Dover Publication, 1968.
- [21] N. R. Pal. On minimum cross-entropy thresholding. *Pattern Recognition.*, vol.**29**(4):p.575–580, 1996.
- [22] A. D. Brink. Minimum spatial entropy threshold selection. *IEE Proc.- Vis. Image Signal Process.*, vol.**142**(3):p.128–132, Junho 1995.

- [23] A. S. Abutaleb. Automatic thresholding of gray-level pictures using two-dimensional entropy. *Comp. Vision Graph. Image Process.*, vol.**47**:p.22–32, 1989.
- [24] C. H. Li and C. K. Lee. Minimum cross entropy thresholding. *Pattern Recognition*, vol.**26**(4):p.617–625, 1993.
- [25] Noesis. Visilog - software de processamento de imagens. <http://www.noesisvision.com>, Abril 2001.
- [26] P. Sahoo, C. Wilkins, and J. Yeager. Threshold selection using renyi's entropy. *Pattern Recognition*, vol.**30**(1):p.71–84, Maio 1997.
- [27] A. D. Brink. Thresholding of digital images using two dimensional entropies. *Pattern Recognition*, vol.**25**(8):p.803–808, 1991.
- [28] A. D. Brink and N. E. Pendock. Minimum cross-entropy threshold selection. *Pattern Recognition*, vol.**19**(1):p.179–188, 1996.
- [29] A. Shiosaki. Edge extration using entropy operator. *Comp. Vision Graph. Image Process.*, vol.**36**:p.1–9, Março 1986.
- [30] T. Yamano. A possible extension of Shannon's information theory. *Entropy 2001 - Versão eletrônica disponível em (www.mdpi.org/entropy/)*., vol.**3**:p.280–292, 2001.
- [31] T. Yamano. Information theory based in nonadditive information content. *Physical Review E*, vol.**63**:p.46105/1 – 46105/7, Março 2001.
- [32] J. R. Parker. *Algorithms for Image Procesing and Computer Vision*. John Wiley and Sons, Inc., 1996.
- [33] Wolfram Research. Resource library. <http://mathworld.wolfram.com/>, Novembro 2002.
- [34] P. Z. Peebles. *Probability, Random Variables, and Random Signal Principles*. McGraw-Hill, Inc - 3 edição, 1993.

- [35] Scion Corporation. Scion image 4.02 - software de processamento de imagens com distribuição gratuita. <http://www.scioncorp.com>, Agosto 2002.
- [36] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. John Wiley Sons, Inc., 1991.
- [37] M. Coster and J. L. Chermant. *Precis D'analyse D'Images*. Presses du CNRS, 1989.
- [38] M. Goossens, F. Mittelback, and A. Samarin. *The Latex Companion*. Addison-Wesley, 1994.
- [39] Márcio Portes de Albuquerque. *Analyse par Traitement d'Images de Domaines Magnetiques*. PhD thesis, Institut National Polytechnique de Grenoble - INPG, 1995.
- [40] Marcelo Portes de Albuquerque. *Mesure Optimisée de Densités d'Aimantation*. PhD thesis, Institut National Polytechnique de Grenoble - INPG, 1999.
- [41] R. M. Haralick and L. G. Shapiro. *Image Segmentation Techniques*. Addison-Wesley Pub Co, 1985.
- [42] I. Shinya. *Video Microscopy*. Plenum Press, 1986.
- [43] B. P. Lathi. *Sistemas de comunicação*. Editora Guanabara S. A., 1987.
- [44] A. K. Jain. *Fundamentals of digital image processing*. Prentice Hall, 1989.
- [45] J. Gomes and L. Velho. *Computação Gráfica: Imagem*. Série de Computação e Matemática, IMPA, 1994.
- [46] C. A. Lindley. *Practical Image Processing In C: Acquisition, Manipulation, Storage*. John Wiley and Sons, 1991.
- [47] K. R. Castleman. *Digital Image Processing*. Prentice Hall, 1995.
- [48] H. R. Myler and A. R. Weeks. *The pocket handbook of image processing algorithms in C*. Prentice Hall, 1993.



- [49] O. Marques Filho and H. Vieira Neto. *Procesamento Digital de Imagens*. BRAS-PORT Livros e Multimidia Ltda, 1999.
- [50] S. Goldman. *Information Theory*. Dover Publication, 1968.
- [51] S. Guiasu. *Information Theory with Applications*. McGraw-Hill Inc., 1977.
- [52] A. I. Khinchin. *Mathematical Foundation of Information Theory*. Dover Publication, 1957.
- [53] C. H. Li and P. K. S. Tam. An iterative algorithm for minimum cross entropy thresholding. *Pattern Recognition Letters*, vol.**19**:p.771–776, Março 1998.
- [54] J. Lin. Divergence measure based on the Shannon entropy. *IEEE Transactions on Information Theory*, vol.**37**(1):p.145–150, Janeiro 1991.
- [55] W. Tatsuaki and S. Takeshi. When non-extensive entropy becomes extensive. *Physica A*, vol.**301**:p.284–290, 2001.
- [56] Y. J. Zhang. A survey on evaluation methods for image segmentation. *Pattern Recognition*, vol.**29**(8):p.1335–1346, Dezembro 1996.
- [57] O. D. Trier and A. K. Jain. Goal-directed evaluation of binarization methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.**17**(12):p.1191–1201, Dezembro 1995.
- [58] J. S. Weska, R. N. Nagel, and A. Rosenfeld. A threshold selection technique. *IEEE Trans. Comput.*, vol.**23**:p.1322–1326, 1974.
- [59] J. S. Weska. A survey of threshold selection techniques. *Comput. Graph. Image Process.*, vol.**7**(259-265):p., 1979.
- [60] J. F. C. Lopera, J. M. Aroza, A. M. R. Pérez, and R. R. Roldán. An analysis of edge detection by using the Jensen-Shannon divergence. *J. Math. Imaging and Vision*, vol.**13**(1):p., 2000.

- [61] A. G. Shanbhag. Utilization of information measure as a means of image thresholding. *CVGIP: Graphical Models and Image Processing*, vol.**56**(5):p.414–419, Setembro 1994.
- [62] C. K. Leung and F. K. Lam. Maximum a posteriori spatial probability segmentation. *IEE Proc.- Vis. Image Signal Process.*, vol.**144**(3):p.161–167, Junho 1997.
- [63] S. Abe. Axioms and uniqueness theorem for Tsallis entropy. *Physics Letters A*, vol.**271**:p.74–79, 2000.
- [64] K. S. Fa. Note on generalization of Shannon theorem and inequality. *J. Phys. A: Math. Gen.*, vol.**31**:p.8159–8164, Maio 1998.
- [65] C. Tsallis. Entropic nonextensivity: a possible measure of complexity. *Chaos, Solitons and Fractals*, vol.**13**:p.371–391, 1988.
- [66] C. Tsallis. Some comments on boltzmann-gibbs statistical mechanics. *Chaos, Soliton and Fractals*, vol.**6**:p.539–559, 1995.
- [67] B. R. La Cour and W. C. Schieve. Tsallis maximum entropy principle and the law of large numbers. *Physical Review E*, vol.**62**(5):p.7494–7496, Novembro 2000.
- [68] O. S. Costa, A. H. Rodríguez, and G. J. Rodgers. Tsallis entropy and the transition to scaling in fragmentation. *Entropy 2001 - Versão eletrônica disponível em (www.mdpi.org/entropy/)*, vol.**2**:p.172–177, Dezembro 2000.
- [69] C. Tsallis and M. P. Albuquerque. Are citations of scientific papers a case of nonextensivity? *The European Physical Journal B*, vol.**13**:p.777–780, 2000.
- [70] J. C. D. MacKay. Information theory, inference and learning algorithms. <http://www.inference.phy.cam.ac.uk/mackay/itprnn/ps/>, Março 2002.
- [71] T. D. Schneider. Information theory primer. <http://www.lecb.ncifcrf.gov/toms/paper/primer>, Maio 2002.

- [72] M. Schmidt. The java imaging utilities - sítio privado e não comercial, disponibiliza imagens standard para teste de algoritmos de processamento de imagem. <http://jiu.sourceforge.net/testimages/index.html>, Maio 2002.
- [73] H. Touchette. When is a quantity additive, and when is it extensive? *Physica A*, vol.**350**:p.84–88, 2002.
- [74] P. L. Rosin. Unimodal thresholding. *Pattern Recognition*, vol.**34**:p.2083–2096, 2001.