

UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA ELÉTRICA E DE
COMPUTAÇÃO

Representação de Objetos Através de
Descritores do Contorno e Classificação
por Funções Discriminantes Lineares por
Partes

José Ricardo Porto Silva

Orientador: **Prof. Dr. Clésio Luis Tozzi**

Dissertação de Mestrado apresentada à
Faculdade de Engenharia Elétrica como
parte dos requisitos exigidos para
obtenção do título de Mestre em
Engenharia Elétrica, Área de
concentração: Automação Industrial

Campinas - SP Brasil
1996

Este exemplar corresponde a redação final da tese
defendida por José Ricardo Porto
Silva e aprovada pela Comissão
Julgada em 24 / 09 / 1996 21

UNIDADE	BC
N.º CHAMADA:	T/UNICAMP
	2138M
V.	Ex.
TOMBO BC/	32224
PROC.	281/27
C	☐
D	☒
PREÇO	R\$ 11,00
DATA	25/11/97
N.º CPD	

CM.00102958-2

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DA ÁREA DE ENGENHARIA - BAE - UNICAMP

Si38r

Silva, José Ricardo Porto

Representação de objetos através de descritores do contorno e classificação por funções discriminantes lineares por partes / José Ricardo Porto Silva.--Campinas, SP: [s.n.], 1996.

Orientador: Clésio Luis Tozzi.

Dissertação (mestrado) - Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação.

1. Análise de Fourier. 2. Fourier, Transformações de. 3. Reconhecimento de padrões. I. Tozzi, Clésio Luis. II. Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação. III. Título.

UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA ELÉTRICA E DE
COMPUTAÇÃO

Representação de Objetos Através de
Descritores do Contorno e Classificação
por Funções Discriminantes Lineares por
Partes

José Ricardo Porto Silva

Orientador: Prof. Dr. Clésio Luis Tozzi

Dissertação de Mestrado

Campinas - SP Brasil
1996

Sumário

LISTA DE FIGURAS	vii
LISTA DE TABELAS	xii
RESUMO	xiii
ABSTRACT	xiv
1 Introdução	1
2 Descritores	4
2.1 Descritores comuns	5
2.2 Parametrização do contorno	7
2.3 Descritores de Fourier	9
2.3.1 Transformações de coordenadas	12
2.3.2 Cálculo dos descritores de Fourier	13
2.3.2.1 Determinação do número de amostras	14
2.3.3 Normalizações	17
2.3.3.1 Normalização quanto à translação	17
2.3.3.2 Normalização quanto à rotação	18
2.3.3.3 Normalização quanto ao ponto inicial	19
2.3.3.4 Normalização quanto à reflexão	20
2.3.3.5 Normalização quanto à escala	22
2.3.3.6 Normalização completa	24
2.4 Descritores elípticos de Fourier	26
2.4.1 Determinação dos descritores elípticos de Fourier	27
2.4.1.1 Cálculo dos coeficientes	30
2.4.2 Normalizações	31

2.4.2.1 Normalização quanto à translação	32
2.4.2.2 Normalização quanto à rotação	32
2.4.2.3 Normalização quanto ao ponto inicial	33
2.4.2.4 Normalização quanto à reflexão	34
2.4.2.5 Normalização quanto à escala	34
2.4.2.6 Normalização completa	35
2.5. Descritores cossenoidais	35
2.5.1 Cálculo dos coeficientes	37
2.5.2 Normalizações	37
2.5.2.1 Normalização quanto à translação	37
2.5.2.2 Normalização quanto à rotação	38
2.5.2.3 Normalização quanto ao ponto inicial	39
2.5.2.4 Normalização quanto à reflexão	41
2.5.2.5 Normalização quanto à escala	42
2.5.2.6 Normalização completa	42
2.6 Momentos	43
3 Algoritmo para geração de funções discriminantes	46
3.1 Descrição do algoritmo proposto	48
3.1.1 Método rotacional	51
3.1.2 Pré-classificação	52
3.1.3 Geração dos pontos médios	53
3.1.4 Classificação	55
3.2 Resultados	58
3.3 Comparações	59
3.4 Conclusão	62
4 Aplicação dos descritores	64
4.1 Obtenção das imagens	65
4.1.1 Rastreamento do contorno	66

4.1.2 Pré-filtragem da imagem	66
4.2 Determinação dos descritores	69
4.2.1 Determinação dos descritores do contorno completo	71
4.2.2 Determinação dos descritores da base e ápice	72
 5 Conclusão	 91
 REFERÊNCIAS BIBLIOGRÁFICAS	 93
 APÊNDICE A - gráficos dos descritores	 97
 APÊNDICE B	 107
 APÊNDICE C - transformada de Hotelling	 110
 APÊNDICE D - amostras de folíolos terminais	 112

Lista de Figuras

2.1 - ilustração do efeito de "aliasing"	14
2.2 - diagrama em blocos do processo de reamostragem no domínio discreto	15
2.3 - diagrama em blocos simplificado do processo de reamostragem no domínio discreto	16
2.4 - ilustração do processo de reflexão em torno de um eixo vertical que passa pelo centróide do objeto	21
2.5 - exemplos da ambigüidade de representação dos descritores de Fourier	25
2.6 - ilustração dos parâmetros dos descritores elípticos de Fourier	29
2.7 - ilustração da reconstrução de um objeto a partir de suas elipses	30
2.8 - ilustração dos parâmetros dos descritores cossenoidais	36
2.9 - exemplo com $N=8$ do fluxograma do algoritmo de cálculo do comprimento do 1º segmento de reta dos descritores cossenoidais para os N deslocamentos circulares da sequência de entrada	40
3.1 - caso extremo de espaço de padrões com classes inter-penetrantes	47
3.2 - fluxograma do algoritmo para geração de funções discriminantes lineares por partes proposto	49
3.3 - resultado do algoritmo para geração de funções discriminantes proposto aplicado a dados sintéticos	55
3.4 - resultado do algoritmo para geração de funções discriminantes proposto aplicado a dados sintéticos	57

3.5 - resultado do algoritmo para geração de funções discriminantes proposto aplicado a dados reais.....	58
3.6 - resultado do algoritmo para geração de funções discriminantes proposto aplicado a dados reais.....	62
4.1 - ilustração dos conceitos de folha e folíolo	64
4.2 - histogramas dos descritores elíticos de Fourier da 2ª harmônica do contorno completo dos folíolos terminais	78
4.3 - histogramas dos descritores elíticos de Fourier da 3ª harmônica do contorno completo dos folíolos terminais	79
4.4 - histogramas dos descritores elíticos de Fourier da 4ª harmônica do contorno completo dos folíolos terminais	80
4.5 - histogramas dos descritores elíticos de Fourier da 5ª harmônica do contorno completo dos folíolos terminais	81
4.6 - gráfico das primeiras componentes das transformadas de Hotelling dos descritores elíticos de Fourier das harmônicas 2 a 5 do contorno completo dos folíolos terminais	82
4.7 - histogramas dos descritores de Fourier das harmônicas ímpares de 1 a 5 da base simetrizada dos folíolos terminais	83
4.8 - histogramas dos descritores de Fourier das harmônicas 7 e 9 da base simetrizada dos folíolos terminais	84
4.9 - histogramas dos descritores de Fourier das harmônicas ímpares de 1 a 5 do ápice simetrizado dos folíolos terminais	85
4.10 - histogramas dos descritores de Fourier das harmônicas 7 e 9 do ápice simetrizado dos folíolos terminais	86

4.11 - gráfico das primeiras componentes das transformadas de Hotelling dos descritores de Fourier das harmônicas ímpares de 1 a 9 da base e ápice simetrizados dos folíolos terminais	87
4.12 - gráfico das primeiras componentes das transformadas de Hotelling dos descritores de Fourier das harmônicas 1 e 3 da base e ápice simetrizados dos folíolos terminais	87
4.13 - gráfico das primeiras componentes das transformadas de Hotelling dos descritores elípticos de Fourier (com normalização à escala alterada) das harmônicas ímpares de 1 a 7 da base e ápice simetrizados dos folíolos terminais	88
4.14- gráfico do semi-eixo menor dos descritores elípticos de Fourier (com normalização quanto à escala alterada) da primeira harmônica da base e ápice simetrizados dos folíolos terminais	88
4.15- fluxograma do procedimento completo de identificação de um folíolo utilizando descritores de Fourier	89
4.16 - imagens dos passos intermediários do processo de identificação	90
A-1 - histogramas dos descritores de Fourier das harmônicas 1 a 3 do contorno completo dos folíolos terminais	97
A-2 - histogramas dos descritores de Fourier das harmônicas 4 e 5 do contorno completo dos folíolos terminais	98
A-3 - gráfico das duas primeiras componentes das transformadas de Hotelling dos descritores de Fourier das harmônicas 1 a 5 do contorno completo dos folíolos terminais	99
A-4 - gráfico das primeiras componentes das transformadas de Hotelling dos descritores de Fourier das harmônicas ímpares de 1 a 9 da base e ápice simetrizados dos folíolos terminais	99

A-5 - gráfico das primeiras componentes das transformadas de Hotelling dos descritores de Fourier das harmônicas ímpares de 1 a 7 da base e ápice simetrizados dos folíolos terminais	100
A-6 - gráfico das primeiras componentes das transformadas de Hotelling dos descritores de Fourier das harmônicas ímpares de 1 a 5 da base e ápice simetrizados dos folíolos terminais.....	100
A-7 - gráfico das primeiras componentes das transformadas de Hotelling dos descritores de Fourier das harmônicas ímpares 1 e 3 da base e ápice simetrizados dos folíolos terminais	101
A-8 - gráfico das primeiras componentes das transformadas de Hotelling dos descritores de Fourier da primeira harmônica da base e ápice simetrizados dos folíolos terminais	101
A-9 - gráfico das duas primeiras componentes das transformadas de Hotelling dos 7 momentos centrais invariantes do contorno completo dos folíolos terminais	102
A-10 - gráfico das primeiras componentes das transformadas de Hotelling dos 7 momentos centrais invariantes da base e ápice simetrizados dos folíolos terminais	102
A-11 - gráfico das duas primeiras componentes das transformadas de Hotelling dos 4 momentos invariantes da sequência unidimensional do contorno completo dos folíolos terminais	103
A-12 - gráfico das primeiras componentes das transformadas de Hotelling dos 4 momentos invariantes das sequências unidimensionais da base e ápice simetrizados dos folíolos terminais	103
A-13 - gráfico das primeiras componentes das transformadas de Hotelling dos descritores cossenoidais das harmônicas 2 a 5 da base e ápice simetrizados dos folíolos terminais	104

A-14 - gráfico do espaço de padrões com os dados da Fig. 3.3 e hiper-retângulos gerados pelo algoritmo de Ichino com $T_a=2$ e $T_c=0$	104
A-15 - gráfico do espaço de padrões com os dados da Fig. 3.4 e hiper-retângulos gerados pelo algoritmo de Ichino com $T_a=1$ e $T_c=0$	105
A-16 - gráfico do espaço de padrões com os dados da Fig. 3.5 e hiper-retângulos gerados pelo algoritmo de Ichino com $T_a=2$ e $T_c=1$	105
A-17 - gráfico do espaço de padrões com os dados da Fig. 3.6 e hiper-retângulos gerados pelo algoritmo de Ichino com $T_a=2$ e $T_c=1$	106
D-1 - dois folíolos terminais de plantas da espécie <i>Lonchocarpus Latifolius</i>	112
D-2 - três folíolos terminais de plantas da espécie <i>Lonchocarpus Campestris</i>	113
D-3 - dois folíolos terminais de plantas da espécie <i>Lonchocarpus Muehlbergianus</i>	114
D-4 - dois folíolos terminais de plantas da espécie <i>Lonchocarpus Guillemianus</i>	115

Lista de Tabelas

3.1 - ações corretivas a serem tomadas na etapa de pré-classificação do algoritmo para geração de funções discriminantes proposto	53
3.2 - índices de erros de quatro algoritmos de classificação aplicados aos dados mostrados na Fig. 3.3	59
3.3 - índices de erros de quatro algoritmos de classificação aplicados aos dados mostrados na Fig. 3.4	60
3.4 - índices de erros de quatro algoritmos de classificação aplicados aos dados mostrados na Fig. 3.5	60
3.5 - índices de erros de quatro algoritmos de classificação aplicados aos dados mostrados na Fig. 3.6	61
4.1 - erros de reconstruções parciais do contorno completo dos folíolos	71
4.2 - autovalores da transf. de Hotelling dos descritores de Fourier do contorno completo dos folíolos	72
4.3 - erros de reconstruções parciais da base simetrizada dos folíolos	73
4.4 - erros de reconstruções parciais do ápice simetrizado dos folíolos	74
4.5 - autovalores da transf. de Hotelling dos descritores de Fourier da base simetrizada dos folíolos	74
4.6 - autovalores da transf. de Hotelling dos descritores de Fourier do ápice simetrizado dos folíolos	75

Resumo

SILVA, J. R. P. **Representação de objetos através de descritores do contorno e classificação por funções discriminantes lineares por partes.** Campinas: DCA/FEEC/UNICAMP, 1996. (Dissertação de Mestrado)

Neste trabalho são estudados descritores de forma aplicados à classificação e identificação de objetos, incluindo particularmente os descritores de Fourier, descritores elípticos de Fourier e as normalizações necessárias para obtenção de suas invariâncias sob translação, rotação, escala, reflexão e ponto inicial de processamento. Também é proposto um algoritmo rápido e automático para geração de funções discriminantes lineares por partes que permitem a separação de classes no espaço de padrões para qualquer tipo de função de densidade de probabilidade.

A validade e eficiência dos métodos apresentados são verificadas para uma aplicação em identificação de espécies de plantas através de seus folíolos. Os desempenhos obtidos por cada método considerado são comparados, explicados e comentados, sendo mostrada a superioridade dos descritores de Fourier em relação aos outros nesta aplicação.

Palavras-chave: Análise de Fourier, Fourier-Transformações de, Reconhecimento de padrões

Abstract

This work presents a study of shape descriptors applied to objects classification and identification, including the Fourier descriptors, elliptic Fourier descriptors and the necessary normalizations to achieve translation, rotation, scale, reflection and starting point invariances. It is also proposed a fast and automatic algorithm to determine piecewise linear discriminant functions that allow the classes separation in the pattern space for any kind of probability density function.

The validity and efficiency of the presented methods are evaluated using an application that seeks plants identification by their folioles. The performances of each used method are compared, explained and commented, and the superiority of the Fourier descriptors among the others is shown for this application.

Capítulo 1

Introdução

Reconhecimento de objetos é um processo complexo onde a entrada geralmente é uma imagem e a saída é uma palavra ou conjunto de palavras (descrição do conteúdo da imagem na linguagem humana). Esse reconhecimento é feito utilizando algum conhecimento prévio dos possíveis objetos presentes na imagem (domínio do problema). O conhecimento possuído deve possibilitar o estabelecimento de métricas ou representações que permitam a separação dos objetos em classes ou grupos.

A automação do reconhecimento de objetos por meio de recursos computacionais vem sendo pesquisada e aplicada por vários anos tanto no âmbito acadêmico quanto industrial. Esta automação pode ser aplicada em áreas tão variadas e importantes quanto reconhecimento de peças industriais defeituosas, células cancerosas em imagens médicas, assinaturas em documentos, impressões digitais ou espécimes biológicos. É almejada pelos benefícios que pode trazer ao processo, que são: maior rapidez, menor índice de erros, independência do fator humano especializado e facilidade de utilização.

Porém, para que seja efetuada uma classificação automática de objetos, estes devem ser descritos de uma forma quantitativa, isto é, através de valores numéricos que possam ser ordenados e comparados com outros. Quando não for desejável ou possível atribuir essa representação numérica aos objetos devem ser estabelecidas regras e hierarquias específicas e rígidas que permitam que as máquinas manipulem outras formas de representação.

Como uma das características mais usuais e eficientes de reconhecimento usadas pelo homem é a forma, foram desenvolvidos descritores que representam a forma através de métodos matemáticos. Entre os descritores mais comuns podem ser citados: perímetro, área (estes dois primeiros não descrevem forma estritamente, mas

tamanho), circularidade ($\text{perímetro}^2/\text{área}$), assinatura (distância do contorno ao centróide em função do ângulo de revolução), momentos, esqueleto, regra da cadeia ("chain codes" e "shape numbers"), descritores sintáticos (primitivas ou blocos básicos que constituem o contorno), descritores de Fourier, descritores por wavelet e pontos de controle (ou representação por B-splines).

O processo de reconhecimento consiste em calcular um ou mais descritores para o objeto a ser identificado na imagem e atribuir-lhe a denominação da classe com a qual sua representação é mais semelhante segundo alguma medida de comparação.

Entretanto, no reconhecimento de objetos bidimensionais, várias digitalizações de um mesmo elemento podem gerar várias imagens diferentes, isto é, o objeto pode estar centrado em qualquer posição da imagem, pode ter qualquer orientação angular, devido à diferentes resoluções de digitalização (ou mesmo pode ser desejado comparar objetos de dimensões distintas), pode ter tamanhos variados e, ao ser digitalizado, pode estar com a face para cima ou para baixo, sendo então refletido ou espelhado. Além disso, devido à passagem do domínio espacial do objeto (contínuo) para o espaço imagem (discreto), ocorre a introdução de um erro de quantização, ou erro de discretização. Também podem existir outros problemas na imagem devidos ao processo de aquisição, tais como ruído, baixo contraste, diferentes ângulos de visão, oclusão, etc.

Um algoritmo de reconhecimento deve ser tolerante a todas essas contingências, isto é, deve fornecer o mesmo resultado em todos os casos citados, e ainda ser capaz de diferenciar objetos de classes diferentes mas próximas no espaço de padrões. Alguns métodos buscam obter essa robustez através de pré-processamentos na imagem para normalizá-la (Wunsch e Laine (1995), Wood (1996)), mas isto normalmente tem um alto custo computacional. Outros métodos utilizam descritores que possuem invariâncias ou baixa sensibilidade às diversas possibilidades citadas.

Um dos métodos que permitem a obtenção das invariância desejadas são os descritores de Fourier, que utilizam o conceito de frequência, isto é, a decomposição da função de entrada em funções harmônicas de diferentes velocidades de variação espacial. As harmônicas (ou frequências) inferiores carregam a informação das variação lentas, ou seja, a forma básica da função contorno utilizada para caracterização da forma, enquanto as harmônicas superiores representam as variações bruscas ou pequenos detalhes.

Como a forma básica existente nas possíveis imagens de um mesmo objeto deve ser a mesma, as harmônicas de ordem baixa devem permitir a sua correta classificação dado que as de ordem alta são devidas ao ruído de digitalização, limiarização e variação intra-classe no caso de objetos não idênticos.

Dentro do contexto apresentado, o objetivo deste trabalho é estudar descritores de forma aplicados ao reconhecimento de objetos, incluindo de forma especial algumas variações dos descritores de Fourier. Também é proposto como contribuição deste trabalho um algoritmo para geração de funções discriminantes lineares por partes que permitem a separação de classes no espaço de padrões mesmo quando as densidades de probabilidade não são corretamente modeladas por distribuições paramétricas conhecidas.

Para avaliar a eficiência dos métodos estudados foi escolhida uma aplicação que apresenta um desafio severo às invariâncias e capacidade de discriminação de qualquer descritor, que é a identificação de quatro espécies de plantas do mesmo gênero (*lonchocarpus campestris*, *lonchocarpus guilleminianus*, *lonchocarpus latifolius* e *lonchocarpus muehlbergianus*) através de seus folíolos. A identificação de espécies de seres vivos é um teste difícil para qualquer algoritmo de reconhecimento porque as amostras não são exatamente iguais, como no caso de objetos construídos em série pelo homem, e podem apresentar uma grande variação intra-classe devido às diferentes condições de crescimento, alimentação e competição. Nessa aplicação é verificada a eficácia e superioridade dos descritores de Fourier em relação a momentos, descritores elípticos de Fourier e descritores cossenoidais para essa categoria de objetos.

Este texto está dividido em cinco capítulos, sendo que no segundo são apresentados alguns descritores conhecidos e definidas as variações dos descritores de Fourier utilizadas. No terceiro capítulo é desenvolvido o algoritmo para geração de funções discriminantes lineares por partes que serão utilizadas para a delimitação das classes envolvidas no espaço de padrões. No capítulo 4 são descritos os detalhes relacionados à aplicação considerada e apresentados os resultados obtidos, e finalmente no quinto são expostas as conclusões.

CAPÍTULO 2

Descritores

Os descritores são formas matemáticas de representação de objetos ou seus contornos. São normalmente utilizados para fornecerem um valor (ou conjunto de valores) que melhor representa o objeto e que pode ser utilizado como medida de comparação (ou similaridade) com outros objetos.

Porém, um mesmo objeto ao ser digitalizado em situações diferentes pode não gerar a mesma imagem, ou seja, o objeto pode estar centrado em qualquer posição da imagem, pode apresentar qualquer orientação angular, devido à diferentes resoluções de digitalização (ou mesmo pode ser desejado comparar objetos de dimensões distintas) pode possuir tamanhos variados, e ao ser digitalizado, pode estar com a face para cima ou para baixo, sendo portanto refletido ou espelhado. Além disso, devido à passagem do domínio contínuo da realidade para o discreto da imagem ocorre a introdução de um erro, chamado erro de discretização ou erro de quantização. Também podem existir outros problemas, tais como ruído, baixo contraste, oclusão, etc.

Um algoritmo de reconhecimento deve ser tolerante a todas essas contingências, isto é, deve fornecer o mesmo resultado para todas as variações da imagem de um mesmo objeto, e ainda assim ser capaz de diferenciar um objeto de outro que pertença a outra classe. Portanto, devem ser realizadas filtragens e normalizações que propiciem essas invariâncias.

Na seção 2.1 deste capítulo é feita uma breve apresentação dos descritores mais comuns. Na seção 2.2 são apresentadas várias formas de parametrização de contornos (que são necessárias para o cálculo dos descritores), e nas seções seguintes são apresentados, pela ordem, os descritores de Fourier, descritores elípticos de Fourier, descritores cossenoidais e momentos, assim como a demonstração dos procedimentos necessários para obtenção de suas invariâncias à translação, rotação, escala, reflexão e ponto inicial de processamento.

2.1 Descritores comuns

O perímetro é um descritor escalar facilmente obtido porém, isoladamente, não é um bom descritor, pois objetos com um formato completamente diferente podem ter perímetros iguais ou de valor próximo. Ele não é invariante à escala, embora seja à rotação, reflexão, ponto inicial e deslocamento.

A área também é um descritor escalar de cálculo simples mas iguala objetos que podem ter formatos diferentes. Não é invariante à escala, embora seja à rotação, reflexão, ponto inicial e deslocamento.

A circularidade (Gonzalez e Woods (1993), Pratt (1991) e Rosenfeld e Kak (1976)) é definida como a razão entre o quadrado do perímetro e a área, este descritor é invariante à escala, à rotação, reflexão, ponto inicial e deslocamento. É calculado de forma simples mas pode falhar em diferenciar objetos semelhantes. É sensível à ruídos e discretização.

A assinatura (Gonzalez e Woods (1993) e Sonka et al. (1993)) é uma função unidimensional que representa o contorno bidimensional, sendo definida como a sequência das distâncias entre os pontos do contorno e o centróide do objeto ordenadas pelo ângulo de revolução. Pode ser tornada invariante à escala, rotação, deslocamento e ponto inicial. É sensível à ruídos, discretização e exige algum método de comparação de funções.

Os momentos estão entre os mais populares descritores de objetos, principalmente por serem conhecidos há décadas e por terem uma interpretação física. Podem ser calculados sete momentos invariantes (Bell (1965) e Hu (1962)) à escala, rotação e deslocamento. Porém essas invariâncias são válidas para a versão contínua dos momentos. Para uma imagem que é discreta por natureza, o ruído de discretização pode deturpar bastante os momentos, como discutido em Jain (1989) e Pratt (1991). Momentos invariantes da sequência de contorno unidimensional foram definidos em Gupta e Srinath (1987); são a extensão do método de assinatura exposto no parágrafo anterior e possuem um custo computacional menor que os momentos tradicionais.

O esqueleto pode ser calculado pela transformação de eixo médio (medial axis transformation, Blum (1967), Pavlidis (1977) e Schalkoff (1989)) ou por afinamento pelo algoritmo de Zhang e Suen (1984). É bastante sensível à ruído, mas um procedimento com multi-resolução pode diminuir a sensibilidade (Pizer et al. (1987) e Maragos (1989)). Porém, é usado basicamente quando não se deseja invariância à escala e rotação, além de ser de difícil comparação. Um exemplo de sua aplicação é a verificação automatizada das conexões de uma placa de circuito impresso.

A regra da cadeia mapeia a direção que se deve tomar a partir de um pixel para chegar ao seu vizinho mais próximo no contorno do objeto, percorrendo todo o perímetro até voltar ao ponto inicial. Pode ser definida numa vizinhança 4 ou 8-conectada. É dependente do ponto inicial, rotação e escala. Pode ser tornado invariante ao ponto inicial, e sua primeira diferença é invariante à rotação. Mas é sensível à ruído e resolução.

Os descritores sintáticos são um método de descrição qualitativa onde o contorno é dividido em algumas primitivas (formas geométricas simples definidas em função da família de objetos alvos, que geram uma gramática específica) de modo vagamente semelhante à regra da cadeia. Como ela, também é sensível ao ponto inicial e rotação, embora possa deixar de sê-lo. Mas ainda é dependente da escala e sensível à ruído.

Os descritores de Fourier são o método de comparação utilizado neste trabalho. Há muitas versões diferentes, mas o processo consiste basicamente na aplicação da DFT (transformada discreta de Fourier) nos pontos do contorno do objeto, obtendo no domínio frequência as desejadas invariâncias à escala, rotação, translação, ponto inicial e baixa sensibilidade à ruído, discretização e pequenas deformações nos objetos. Descritores de Fourier têm sido aplicados no reconhecimento de vários tipos de objetos, tais como caracteres escritos a mão (Granlund (1972)), caracteres chineses (Wang et al. (1994)) e árabes (Mahmoud (1994)), aeronaves (Lin e Chellapa (1987)), órgãos internos humanos (Staib e Duncan (1992)) e conchas marinhas (Ferson et al. (1985)). As incertezas e perdas de informação que Bookstein et al. (1982) imputa a este método de descrição são devidas à falta de correta normalização e escolha do tipo adequado de descritores de Fourier por parte deste autor, além da tentativa de discriminação mal orientada. Essas falhas foram individualmente apontadas e explicadas em Ehrlich et al. (1983), onde é

mostrado inclusive que os valores dos coeficientes das harmônicas podem ser interpretados sob o ponto de vista da morfometria biológica.

Um conjunto de descritores baseados na transformada por wavelets foi introduzido por Wunsch e Laine (1995), que mostram que os mesmos apresentam um desempenho ligeiramente superior ao dos descritores de Fourier quando aplicados ao reconhecimento de dígitos escritos à mão; apesar de (e talvez por) não serem invariantes à translação, ponto inicial e rotação, sendo as duas primeiras limitações atenuadas por normalização da imagem antes do cálculo dos descritores, e a última desprezada.

Curvas e contornos podem ser aproximados por B-splines, e essas são determinadas por pontos de controle. Estes pontos são os coeficientes da decomposição da curva em funções de base B-spline normalizadas. Esta representação pode ser tornada invariante à rotação, translação, escala e ponto inicial, mas as classificações de objetos assim descritos são mais rapidamente efetuadas através da DFT desses pontos de controle, numa variação do método dos descritores de Fourier, segundo Paglieroni e Jain (1987) e (1988).

2.2 Parametrização do contorno

Para que sejam aplicados os descritores de Fourier num contorno, este deve ser parametrizado. Existem várias parametrizações diferentes:

- pode-se parametrizar a curva que descreve o contorno segundo o comprimento do arco a partir de um ponto qualquer do contorno

1 - e utilizar a sequência das coordenadas dos pontos na forma vetorial [$x(t)$ $y(t)$] (Staib e Duncan (1992), Tseng e Shenck (1992) e Lin e Hwang (1987))

2 - ou na forma de número complexo $x(t) + jy(t)$ (Gonzalez e Woods (1993))

3 - pode ser calculado o ângulo da tangente ao contorno para cada ponto consecutivo deste, determinando uma sequência real $\phi(t)$ (Sonka et al. (1993))

- 4 - pode ser utilizada a diferença entre os ângulos $\varphi(t)$, definida como $d\varphi(t)$
- 5 - pode ser utilizada a distância de cada ponto (em seqüência) ao centróide do objeto, $r(t)$
- pode-se selecionar um número pré-definido de pontos do contorno igualmente espaçados angularmente
- 6 - e formar uma nova seqüência somente com as distâncias entre pontos e o centróide ordenados pelo ângulo de revolução, chamada de $r(\theta)$ (Wang et al. (1994))
- 7 - pode-se aproximar a curva do contorno por um polígono inscrito (regular ou não), e utilizar os comprimentos (nas direções x e y) dos lados do polígono em seqüência, $[\Delta x(t) \Delta y(t)]$ (Ferson et al. (1985) e Persoon e Fu (1986))

As formas (1) e (2) são bastante semelhantes entre si e as mais comuns, principalmente por serem mais diretas e simples que as outras. As representações (3) e (4) podem apresentar variações bruscas em cantos do contorno, o que induz ao conhecido fenômeno de Gibbs (Oppenheim (1989)), que é prejudicial à representação da curva no domínio freqüência. A forma (5) tem a vantagem da redução dos dois valores da forma (1) para um (ficando um dos valores implícito pela seqüência), mas possui a desvantagem de exigir o cálculo prévio do centróide e das distâncias entre este e cada ponto da curva. Na forma (6), outra versão muito difundida, além dos problemas da forma (5) ainda há a possibilidade de ocorrência de "aliasing" devido a uma amostragem abaixo da freqüência de Nyquist (que não é conhecida a priori), ou mesmo ambigüidade se a curva não possuir um valor único de $r(\theta)$ para cada ângulo θ . A forma (7) também é uma variação da forma (1) em que é realizada uma amostragem regular ou não dos pontos, e utilizada a diferença entre os valores $x(t)$ e $y(t)$ consecutivos. Dependendo da taxa de amostragem (regular ou não), também pode estar sujeita ao problema de "aliasing". Pelas razões apresentadas percebe-se que as formas (1) e (2) são as melhores candidatas à representação e, dentre estas, (2) foi escolhida por simplicidade de notação.

2.3 Descritores de Fourier

Os descritores de Fourier são calculados em duas etapas: a primeira consiste no cálculo da DFT (transformada discreta de Fourier) através do algoritmo de FFT (Oppenheim (1989)) do contorno parametrizado, e a segunda são as operações matemáticas sobre os coeficientes da transformada para obtenção das desejadas invariâncias à escala, rotação, translação, reflexão, ponto inicial e baixa sensibilidade a ruído, discretização, binarização e pequenas variações no objeto. De acordo com a teoria de descritores elípticos de Fourier (Kuhl e Giardian (1982) e Lin e Hwang (1987)), as duas funções periódicas $x(t)$ e $y(t)$, com $0 \leq t \leq 2\pi$, que formam o contorno podem ser expressas na forma matricial :

$$\text{Equação de síntese: } \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} a_0 \\ c_0 \end{bmatrix} + \sum_{k=1}^{\infty} \begin{bmatrix} a_k & b_k \\ c_k & d_k \end{bmatrix} \begin{bmatrix} \cos(kt) \\ \sin(kt) \end{bmatrix} \quad (2.1)$$

onde :

$$\text{Equações de análise: } a_0 = \frac{1}{2\pi} \int_0^{2\pi} x(t) dt \quad (2.2a)$$

$$c_0 = \frac{1}{2\pi} \int_0^{2\pi} y(t) dt \quad (2.2b)$$

$$a_k = \frac{1}{\pi} \int_0^{2\pi} x(t) \cos(kt) dt \quad (2.2c)$$

$$b_k = \frac{1}{\pi} \int_0^{2\pi} x(t) \sin(kt) dt \quad (2.2d)$$

$$c_k = \frac{1}{\pi} \int_0^{2\pi} y(t) \cos(kt) dt \quad (2.2e)$$

$$d_k = \frac{1}{\pi} \int_0^{2\pi} y(t) \sin(kt) dt \quad (2.2f)$$

Para o caso discreto:

$$\text{Equação de síntese: } \begin{bmatrix} x[n] \\ y[n] \end{bmatrix} = \sum_{k=0}^{N-1} \begin{bmatrix} a_k & b_k \\ c_k & d_k \end{bmatrix} \begin{bmatrix} \cos(2k\pi n / N) \\ \sin(2k\pi n / N) \end{bmatrix} \quad (2.3)$$

e

$$\text{Equações de análise: } a_k = \frac{1}{N} \sum_{n=0}^{N-1} x[n] \cos(2k\pi n / N) \quad (2.4a)$$

$$b_k = \frac{1}{N} \sum_{n=0}^{N-1} x[n] \sin(2k\pi n / N) \quad (2.4b)$$

$$c_k = \frac{1}{N} \sum_{n=0}^{N-1} y[n] \cos(2k\pi n / N) \quad (2.4c)$$

$$d_k = \frac{1}{N} \sum_{n=0}^{N-1} y[n] \sin(2k\pi n / N) \quad (2.4d)$$

onde N é o número de amostras das funções $x[n]$ e $y[n]$.

A utilização de colchetes na especificação do argumento das funções indica que estas são discretas, isto é, seus argumentos só podem assumir valores inteiros, enquanto que os parênteses identificam funções contínuas. Os coeficientes a_0 e c_0 foram inseridos dentro da matriz na Eq. (2.3) por simplicidade de notação. Pode-se perceber pelas equações de análise da versão discreta (Eqs. 2.4) que os coeficientes b_0 e d_0 são identicamente nulos. Os valores a_0 e c_0 são as coordenadas do centróide do contorno, isto é, a soma de todos os pontos do contorno dividido pelo número de elementos deste. As matrizes dos coeficientes das k harmônicas na Eq. (2.3) formam um conjunto redundante de $4N$ parâmetros segundo as relações da transformada de Hilbert (Oppenheim (1989)) devido ao fato das seqüências $x[n]$ e $y[n]$ serem reais.

A possibilidade da reconstrução das funções através dos coeficientes das harmônicas pode ser demonstrada inserindo as equações de análise (Eqs. 2.4) na de síntese (Eq. 2.3), isto é, chamando a reconstrução de $x[n]$ como $x'[n]$:

$$\begin{aligned} x'[n] &= \sum_{k=0}^{N-1} (a_k \cos(2k\pi n / N) + b_k \sin(2k\pi n / N)) \\ &= \sum_{k=0}^{N-1} \left(\left(\sum_{r=0}^{N-1} x[r] \cos(2k\pi r / N) \right) \cos(2k\pi n / N) + \left(\sum_{r=0}^{N-1} x[r] \sin(2k\pi r / N) \right) \sin(2k\pi n / N) \right) \end{aligned}$$

Invertendo a ordem dos somatórios, vem:

$$\begin{aligned} x'[n] &= \sum_{r=0}^{N-1} \left(x[r] \sum_{k=0}^{N-1} (\cos(2k\pi r / N) \cos(2k\pi n / N) + \sin(2k\pi r / N) \sin(2k\pi n / N)) \right) \\ &= \sum_{r=0}^{N-1} \left(x[r] \sum_{k=0}^{N-1} \cos(2k\pi(r-n) / N) \right) \end{aligned} \quad (2.5)$$

E como $\sum_{k=0}^{N-1} \cos(2k\pi(r-n) / N) = \delta[r-n]$ onde $\delta[i]$ é a função delta de Kroenecker

que vale $\delta[i] = \begin{cases} 1 & \text{se } i = 0 \\ 0 & \text{para } i \neq 0 \end{cases}$ a Eq. (2.5) pode ser reescrita como:

$$x'[n] = \sum_{r=0}^{N-1} x[r] \delta[r-n]$$

E aplicando a definição da função $\delta[i]$ vem:

$$x'[n] = x[n]$$

Ou seja, a função reconstruída é igual a função original.

De forma análoga à demonstração acima prova-se a validade da reconstrução para $y[n]$. Neste trabalho, por simplicidade de notação e cálculo (conforme mencionado anteriormente), foi utilizada a forma de número complexo para a seqüência de contorno, o que altera as Eq. (2.3) e (2.4) para:

$$\text{Equação de análise: } F[k] = \sum_{n=0}^{N-1} f[n] \exp(-j2k\pi n / N) \quad (2.6)$$

$$\text{Equação de síntese: } f[n] = \frac{1}{N} \sum_{k=0}^{N-1} F[k] \exp(j2k\pi n / N) \quad (2.7)$$

$$\text{onde } f[n] = x[n] + jy[n] \quad (2.8)$$

e $0 \leq k, n < N$, sendo N o número de pontos do contorno, o que produz um conjunto não redundante de $2N$ parâmetros (N harmônicas, sendo cada uma expressa por um número complexo). A demonstração dessa decomposição é semelhante à anterior, portanto não será apresentada.

2.3.1 Transformações de coordenadas

Nesta seção é mostrado como as variações na imagem de um objeto devidas ao processo de aquisição afetam o contorno parametrizado, e a partir disto, a mudança provocada nos coeficientes da transformada de Fourier. Um ponto (x,y) ao ser transladado, escalado e rotacionado em torno da origem tem suas novas coordenadas determinadas pela Eq. (2.9).

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = S \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \quad (2.9)$$

onde S é o fator de escala, θ o ângulo de rotação e Δx e Δy os deslocamentos em x e y respectivamente.

Como para o caso dos descritores de objetos deseja-se que sua rotação (e mudança de escala) seja realizada em torno do seu centróide (ou centro de massa) e não em torno da origem como determinado pela Eq. (2.9), esta equação deve ser reescrita como:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = S \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x - a_0 \\ y - c_0 \end{bmatrix} + \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} + \begin{bmatrix} a_0 \\ c_0 \end{bmatrix} \quad (2.10)$$

Ou, na forma de um número complexo :

$$x' + jy' = S \exp(j\theta) (x - a_0 + j(y - c_0)) + \Delta x + j\Delta y + a_0 + jc_0 \quad (2.11)$$

onde os valores a_0 e c_0 são as coordenadas do centróide do contorno.

Como a Eq. (2.10) é válida para cada ponto isoladamente, pode ser aplicada a toda a seqüência gerando um novo conjunto de pontos segundo a Eq. (2.12).

$$\begin{bmatrix} x'[n] \\ y'[n] \end{bmatrix} = S \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x[n] - a_0 \\ y[n] - c_0 \end{bmatrix} + \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} + \begin{bmatrix} a_0 \\ c_0 \end{bmatrix} \quad (2.12)$$

Pode-se agora usar a Eq. (2.12) para verificar a mudança na matriz dos coeficientes da transformada discreta de Fourier da equação de síntese (Eq. 2.3) que essas transformações provocarão. Ou seja, aplicando as operações determinadas pela Eq. (2.12) na Eq. (2.3) resulta:

$$\begin{bmatrix} x'[n] \\ y'[n] \end{bmatrix} = S \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \left(\sum_{k=1}^{N-1} \begin{bmatrix} a_k & b_k \\ c_k & d_k \end{bmatrix} \begin{bmatrix} \cos(2k\pi n / N) \\ \sin(2k\pi n / N) \end{bmatrix} \right) + \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} + \begin{bmatrix} a_0 \\ c_0 \end{bmatrix}$$

Inserindo o fator de escala S e a matriz de rotação no somatório vem:

$$\begin{bmatrix} x'[n] \\ y'[n] \end{bmatrix} = \sum_{k=1}^{N-1} S \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} a_k & b_k \\ c_k & d_k \end{bmatrix} \begin{bmatrix} \cos(2k\pi n / N) \\ \sin(2k\pi n / N) \end{bmatrix} + \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} + \begin{bmatrix} a_0 \\ c_0 \end{bmatrix}$$

Que pode ser reescrita na forma:

$$\begin{bmatrix} x'[n] \\ y'[n] \end{bmatrix} = \sum_{k=0}^{N-1} \begin{bmatrix} a'_k & b'_k \\ c'_k & d'_k \end{bmatrix} \begin{bmatrix} \cos(2k\pi n / N) \\ \sin(2k\pi n / N) \end{bmatrix}$$

onde

$$\begin{bmatrix} a'_k & b'_k \\ c'_k & d'_k \end{bmatrix} = S \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} a_k & b_k \\ c_k & d_k \end{bmatrix} \quad \text{para } k \neq 0$$

$$\text{e } a'_0 = a_0 + \Delta x \quad c'_0 = c_0 + \Delta y \quad b'_0 = d'_0 = 0$$

Isto é, as mudanças que estas transformações provocam na função contorno podem ser completamente absorvidas pelos coeficientes da transformada, sem afetar as funções de base desta.

2.3.2 Cálculo dos descritores de Fourier

Neste trabalho os descritores de Fourier foram calculados diretamente pela Eq. (2.6), isto é, as coordenadas (x,y) dos pontos do contorno foram reorganizadas na forma de números complexos, onde a abscissa $x[n]$ é a parte real, e a ordenada $y[n]$ é a parte imaginária, gerando a função complexa da Eq. (2.8). Isto permite que os

parâmetros $F[k]$ sejam calculados diretamente através de um algoritmo de FFT com a função $f[n]$ como entrada.

2.3.2.1 Determinação do número de amostras

Em todas as referências indicadas que tratam de descritores de Fourier o número de amostras (parâmetro N das Eqs. 2.3, 2.4, 2.6 e 2.7) sobre o contorno é uma constante pré-definida e igual para todos os objetos, sendo, portanto, utilizadas somente N amostras igualmente espaçadas do contorno adquirido. Isto permite que a escala dos objetos seja alterada simplesmente através de uma multiplicação por uma constante (ver Eqs. 2.9, 2.10 e 2.11). Mas, por ser pré-definido e invariável, esse número de amostras pode não satisfazer o critério de Nyquist, o que acarretará em "aliasing", deformação na reconstrução do objeto e possivelmente inviabilização do reconhecimento. Na Fig. 2.1 observa-se um exemplo de deformação devida a amostragem numa frequência inferior à exigida por Nyquist.

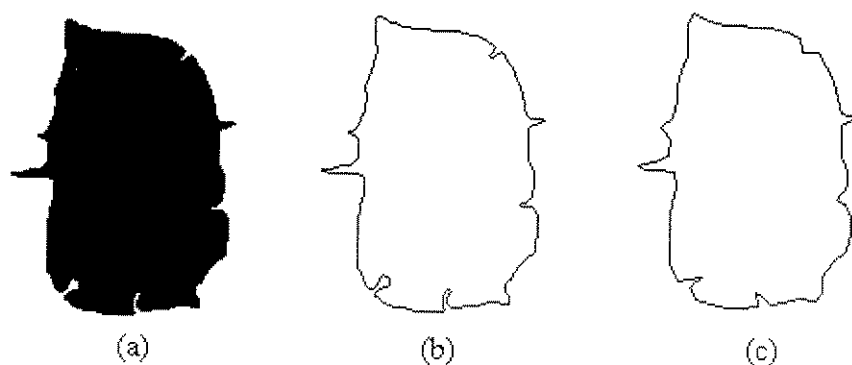


Figura 2.1: (a) imagem original, (b) contorno extraído, (c) reconstrução do contorno com "aliasing"

A Fig. 2.1b mostra o contorno extraído da Fig. 2.1a, e na Fig. 2.1c está o contorno reconstruído a partir de 64 amostras igualmente espaçadas. Percebe-se diferenças nas reentrâncias dos contornos na parte inferior e no canto superior direito. Esta distorção (que foi relativamente pequena nesse caso) aumenta de modo geral com a diferença entre o número de "pixels" amostrados e o número de "pixels" do contorno.

Para garantir que a frequência de amostragem seja maior que o dobro da maior frequência contida no contorno e evitar "aliasing", sem saber "a priori" essa frequência, o único método garantido é a super-amostragem, ou seja, amostrar todos os "pixels" do contorno do objeto, tendo portanto um número N diferente para cada objeto e também diferente para o mesmo objeto se ele for escalonado ou se a imagem for adquirida com resolução diferente. Se isso for feito, a diferença entre os descritores de Fourier não será somente o fator de escala S da Eq. (2.11), mas será dependente da relação entre o número de amostras de cada imagem do objeto. A demonstração da alteração dos descritores devida a essa amostragem com frequência variável e a normalização que produz invariância à escala para este caso está na seção 2.3.3.5.

Apesar de existirem algoritmos de FFT para números compostos e fatores primos (Oppenheim (1989)), foi utilizado o algoritmo tradicional, que exige que o número de amostras seja uma potência de dois. Para isso pode-se alterar a frequência de amostragem após a aquisição de todos os pontos do contorno (e sem distorcer o objeto) utilizando operações de reamostragem porque o critério de Nyquist foi respeitado, e portanto as amostras contém toda a informação do sinal e ele pode ser completamente recuperado a partir delas. Como a relação número de amostras adquiridas dividido pela potência de dois superior mais próxima não é necessariamente um número inteiro, deve-se combinar interpolação e dizimação para realizar a reamostragem, como demonstrado em Oppenheim (1989). Esse processo é todo realizado no domínio discreto e está representado no diagrama da Fig. 2.2.

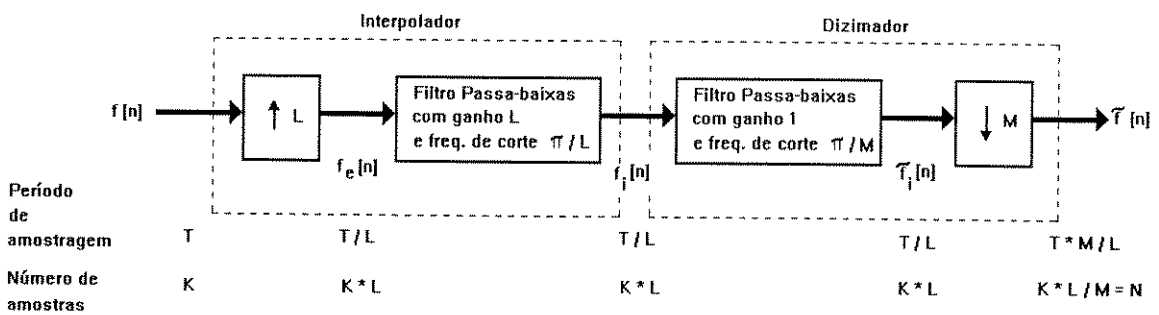


Figura 2.2: diagrama em blocos do processo de reamostragem no domínio discreto

Os dois filtros passa-baixas da Fig. 2.2 podem ser combinados num único filtro equivalente (multiplicados no domínio frequência) como apresentado no diagrama da Fig. 2.3.

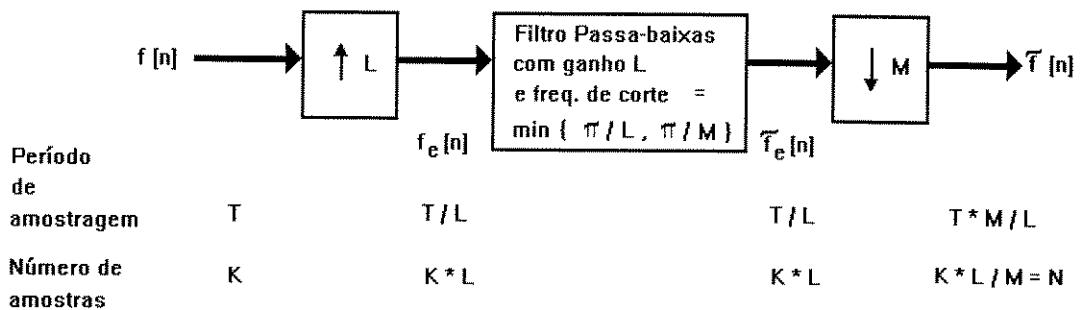


Figura 2.3: diagrama em blocos simplificado do processo de reamostragem no domínio discreto

O "Interpolador" da Fig. 2.2 realiza uma operação de aumento de frequência de amostragem ou "upsampling", onde o primeiro bloco é um expensor, cuja saída é:

$$f_e[n] = \begin{cases} f[n/L] & \text{para } n = 0, L, 2L, 3L, \dots \\ 0 & \text{para outros } n's \end{cases} \quad \text{onde } L \text{ é o fator de expansão.}$$

O filtro discreto do "Interpolador" realiza a interpolação que gera os valores para n não múltiplo de L .

O "Dizimador" da Fig. 2.2 realiza uma operação de diminuição de frequência de amostragem ou "downsampling", onde o filtro é necessário para garantir que o critério de Nyquist seja satisfeito para a nova frequência de amostragem, e o último bloco é um compressor, cuja saída é:

$$\tilde{f}[n] = \tilde{f}_i[nM] \quad \text{onde } M \text{ é o fator de compressão.}$$

Esses dois blocos que compõem o "Dizimador" fazem uma reamostragem do sinal a uma frequência menor (maior período).

Essa mudança de frequência de amostragem é exata, isto é, não introduz erro, mas exige filtros ideais. Como, na prática, aproximações razoáveis dos filtros ideais terão uma resposta impulsiva longa, implicando em muitos cálculos e elevado tempo de computação, neste trabalho é usada interpolação linear, isto é, uma amostra a ser criada entre duas já existentes é calculada através de uma média ponderada pela sua distância a estas últimas. Apesar de não ser preciso, esse método produz resultados bastante adequados e, pela proximidade das amostras, introduz pouco erro.

2.3.3 Normalizações

Para que o reconhecimento seja possível mesmo com todas as variações na imagem de um mesmo objeto (devidas ao processo de digitalização) mencionadas na introdução deve-se utilizar descritores invariantes (ou insensíveis) a todas essas mudanças. Isto é obtido realizando operações matemáticas nos descritores que produzem normalizações. Estas operações são deduzidas nas seções seguintes.

2.3.3.1 Normalização para invariância quanto à translação

Um ponto quando transladado no espaço bidimensional tem suas novas coordenadas obtidas pelo acréscimo do valor dos deslocamentos nas direções x e y, ou seja (pela convenção adotada e expressa na Eq. 2.8):

$$f'[n] = f[n] + \Delta x + j\Delta y \quad (2.13)$$

Utilizando a equação de análise da DFT (Eq. 2.6) para obtenção da transformada da seqüência transladada obtém-se:

$$\begin{aligned} F'[k] &= \sum_{n=0}^{N-1} f'[n] \exp(-j2k\pi n / N) = \sum_{n=0}^{N-1} (f[n] + \Delta x + j\Delta y) \exp(-j2k\pi n / N) \\ &= \sum_{n=0}^{N-1} f[n] \exp(-j2k\pi n / N) + \sum_{n=0}^{N-1} (\Delta x + j\Delta y) \exp(-j2k\pi n / N) \end{aligned}$$

Substituindo o primeiro somatório pela transformada da seqüência original (Eq. 2.6) e extraindo o fator constante do segundo somatório vem:

$$F'[k] = F[k] + (\Delta x + j\Delta y) \sum_{n=0}^{N-1} \exp(-j2k\pi n / N)$$

E como $\sum_{n=0}^{N-1} \exp(-j2k\pi n / N) = N\delta[k]$ onde $\delta[k]$ é a função delta de Kroenecker já

definida na seção 2.3, obtém-se:

$$F'[k] = F[k] + (\Delta x + j\Delta y)N\delta[k] \quad \text{para } 0 \leq k < N$$

onde N é o número de amostras da função $f[n]$ e também o número de harmônicas.

Portanto, a translação somente altera a componente de $k=0$ (DC) da DFT da seqüência das coordenadas dos "pixels" do contorno, permanecendo as demais constantes. Com isso, a invariância ao deslocamento é obtida desprezando-se $F[0]$.

2.3.3.2 Normalização quanto à rotação

A rotação de um ponto (x,y) por um ângulo θ em torno da origem e no sentido anti-horário num espaço bidimensional é obtida por:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x \cos \theta - y \sin \theta \\ x \sin \theta + y \cos \theta \end{bmatrix} \quad (2.14)$$

Utilizando a notação de número complexo a Eq. (2.14) pode ser reescrita como:

$$x' + jy' = x \cos \theta - y \sin \theta + j(x \sin \theta + y \cos \theta)$$

$$= (x+jy)(\cos \theta + j \sin \theta)$$

Ou seja, a rotação corresponde, nessa notação, a multiplicação da coordenada do ponto original pelo número complexo $(\cos \theta + j \sin \theta)$.

Definindo a seqüência de pontos rotacionados como $f'[n] = f[n](\cos \theta + j \sin \theta)$ e aplicando a Eq. (2.6) a esta nova seqüência obtém-se:

$$\begin{aligned} F'[k] &= \sum_{n=0}^{N-1} f'[n] \exp(-j2k\pi n / N) = \sum_{n=0}^{N-1} f[n](\cos \theta + j \sin \theta) \exp(-j2k\pi n / N) \\ &= (\cos \theta + j \sin \theta) \sum_{n=0}^{N-1} f[n] \exp(-j2k\pi n / N) \end{aligned}$$

Expriminado o número complexo $(\cos \theta + j \sin \theta)$ na sua forma polar e percebendo que o somatório é a DFT da seqüência original (Eq. 2.6), vem:

$$F'[k] = 1 \exp(j\theta) F[k] \quad \text{para } 0 \leq k < N$$

Portanto a rotação não altera o módulo da DFT, somente uma constante é adicionada a fase de todas as harmônicas. A invariância à rotação é obtida desprezando-se a fase da transformada da função contorno, ou seja, considerando somente o módulo da DFT.

2.3.3.3 Normalização quanto ao ponto inicial

Não se pode garantir que a seqüência das coordenadas dos "pixels" da borda dos objetos seja adquirida sempre a partir do mesmo ponto inicial no processo de amostragem, isto é, pode haver um deslocamento circular na função $f[n]$, que não altera os valores dessa função, mas apenas modifica o parâmetro n de cada um de seus pontos.

Supondo que $f[n] = f[n-p]$, isto é, a função $f[n]$ é igual a função $f[n]$ atrasada de p amostras (levando-se em conta a periodicidade de $f[n]$), então

$$F'[k] = \sum_{n=0}^{N-1} f[n-p] \exp(-j2k\pi n / N)$$

Sendo $h=n-p$:

$$\begin{aligned} F'[k] &= \sum_{h=-p}^{N-1-p} f[h] \exp(-j2k\pi(h+p) / N) \\ &= \exp(-j2k\pi p / N) \sum_{h=-p}^{N-1-p} f[h] \exp(-j2k\pi h / N) \\ &= \exp(-j2k\pi p / N) \left(\sum_{h=-p}^{-1} f[h] \exp(-j2k\pi h / N) + \sum_{h=0}^{N-1-p} f[h] \exp(-j2k\pi h / N) \right) \end{aligned}$$

Dada a periodicidade da exponencial complexa e de $f[n]$ com N vem:

$$F'[k] = \exp(-j2k\pi p / N) \left(\sum_{h=N-p}^{N-1} f[h] \exp(-j2k\pi h / N) + \sum_{h=0}^{N-1-p} f[h] \exp(-j2k\pi h / N) \right)$$

Reagrupando os somatórios resulta:

$$F'[k] = \exp(-j2k\pi p / N) \sum_{h=0}^{N-1} f[h] \exp(-j2k\pi h / N)$$

E utilizando a Eq. (2.6) vem:

$$F'[k] = \exp(-j2k\pi p / N) F[k] \quad \text{para } 0 \leq k < N$$

Novamente, o módulo da DFT não foi alterado, mas a fase das harmônicas não foi alterada de forma uniforme como no caso da rotação. Portanto, a invariância ao ponto inicial é obtida considerando-se somente o módulo da transformada discreta de Fourier da função contorno.

2.3.3.4 Normalização quanto à reflexão

Um objeto pode tanto ser digitalizado com a face para cima como para baixo, o que provoca uma transformação de reflexão na sua imagem. Portanto é necessário aplicar uma operação matemática sobre os descritores que produza invariância à reflexão. A reflexão em torno de um eixo vertical que passa pelo centróide do objeto (ou, mais exatamente, pelo centróide do contorno) pode ser modelada como uma reflexão em torno do eixo imaginário (jy), inversão da ordem da seqüência de pontos e translação na direção horizontal de um valor igual ao dobro da abscissa do centróide (ver Fig. 2.4), ou seja a nova função contorno é:

$$f'[n] = -f^*[N-n] + 2\text{Re}\{F[0]/N\} \quad (2.15)$$

onde $\text{Re}\{.\}$ significa a operação extração da parte real e o asterisco representa a operação conjugado.

A DFT da seqüência de pontos do objeto refletido é:

$$F'[k] = \sum_{n=0}^{N-1} f'[n] \exp(-j2k\pi n / N) = \sum_{n=0}^{N-1} (-f^*[N-n] + 2\text{Re}\{F[0]/N\}) \exp(-j2k\pi n / N)$$

$$= 2 \operatorname{Re}\{F[0] / N\} \sum_{n=0}^{N-1} \exp(-j2k\pi n / N) - \sum_{n=0}^{N-1} f^*[N-n] \exp(-j2k\pi n / N)$$

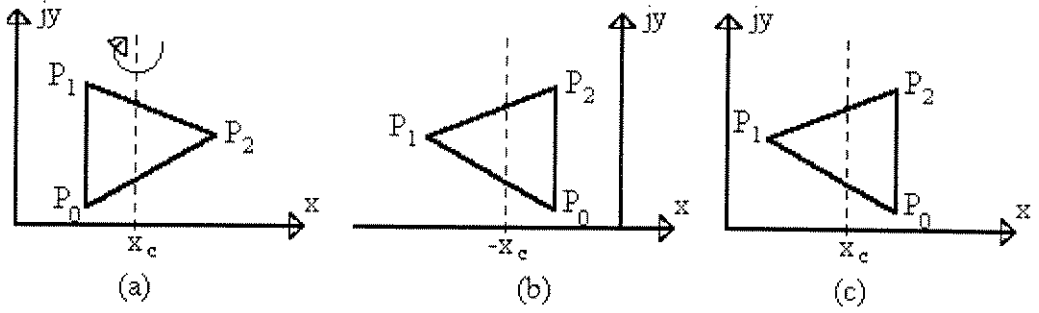


Figura 2.4: (a) objeto original e eixo de reflexão, (b) objeto refletido em torno do eixo jy com sequência de pontos invertida, (c) objeto da Fig. 2.4b transladado de $(2x_c, 0)$

Substituindo o primeiro somatório pelo seu resultado (indicado na seção 2.3.3.1) e sabendo que o produto dos conjugados é o conjugado do produto e que o conjugado da soma é igual à soma dos conjugados, vem:

$$F'[k] = 2 \operatorname{Re}\{F[0] / N\} N \delta[k] - \left(\sum_{n=0}^{N-1} f[N-n] \exp(j2k\pi n / N) \right)^*$$

Definindo $p = N - n$:

$$F'[k] = \frac{2}{N} \operatorname{Re}\{F[0]\} N \delta[k] - \left(\sum_{p=N}^1 f[p] \exp(j2k\pi(N-p) / N) \right)^* \quad (2.16)$$

Aproveitando a periodicidade da exponencial complexa com N e reordenando o somatório (pois a função $f[n]$ também tem período N) pode-se reescrever a Eq. (2.16) como:

$$\begin{aligned} F'[k] &= 2 \operatorname{Re}\{F[0]\} \delta[k] - \left(\sum_{p=0}^{N-1} f[p] \exp(j2k\pi(-p) / N) \right)^* \\ &= 2 \operatorname{Re}\{F[0]\} \delta[k] - \left(\sum_{p=0}^{N-1} f[p] \exp(-j2k\pi p / N) \right)^* \end{aligned}$$

Da definição de DFT (Eq. 2.6) vem:

$$F'[k] = 2 \operatorname{Re}\{F[0]\} \delta[k] - F^*[k] \quad \text{para } 0 \leq k < N$$

ou

$$F'[k] = \begin{cases} F[k] & \text{para } k = 0 \\ -F^*[k] & \text{para } 0 < k < N \end{cases}$$

Isto é, esta reflexão não altera a componente de $k=0$ da DFT (porque o centróide não foi alterado), mas aplica o conjugado e troca o sinal das outras harmônicas. A invariância a esta reflexão é obtida se considerarmos somente o módulo de $F'[k]$ para $0 < k < N$. Uma reflexão em torno de outro eixo qualquer pode ser decomposta numa combinação de rotação, translação e na reflexão calculada acima.

2.3.3.5 Normalização quanto à escala

O mesmo objeto pode ser digitalizado com resoluções diferentes, ou pode-se desejar que objetos com a mesma forma mas de tamanhos diferentes sejam agrupados na mesma classe, portanto deve-se definir uma operação que produza invariância à escala quando aplicada aos descritores.

Quando a imagem de um objeto é aumentada (ou reduzida) o número de pontos do seu contorno também é alterado pelo mesmo fator. O método utilizado em todas as referências que tratam de descritores de Fourier para manter o número de pontos constantes é a amostragem a intervalos regulares ao longo do contorno. Isso permite que a invariância dos descritores à escala seja efetuada por um simples fator de ajuste (valor S das Eqs. 2.9, 2.10 e 2.11), mas pode acarretar "aliasing" e conseqüente deformação na reconstrução do objeto, como explicado na seção 2.3.2.1. Neste trabalho o contorno dos objetos é super-amostrado, o que produz um número de amostras (N) possivelmente diferente para cada objeto. A dedução do procedimento para obtenção de invariância à escala nesse caso é mostrada a seguir.

Suponha um objeto com o centróide na origem. Uma redução do tamanho desse objeto faria no domínio discreto algo semelhante ao "downsampling" (ver Oppenheim (1989)). O valor de todas as amostras do contorno (suas coordenadas) seriam divididas pelo fator de redução, e como o perímetro também diminuiu na mesma proporção, o número de amostras da borda seria reduzido pelo mesmo fator.

Portanto há uma aproximação entre as amostras do contorno, e a subsequente eliminação daquelas que não mais possuem um índice n inteiro. Matematicamente, isso é traduzido como uma multiplicação entre $f[n]$ e um trem de impulsos de valor igual ao inverso do fator de redução, e reordenação desse produto de modo a eliminar as amostras nulas (que não existem na imagem reduzida). Essa operação é mostrada abaixo de forma equivalente para o domínio espacial e domínio frequência.

Domínio espacial

Domínio frequência

$$f'[n] = f[n] \frac{1}{a} \sum_{p=0}^{(N/a)-1} \delta[n - ap] \quad F'[k] = \frac{1}{N} F[k] \otimes \left(\frac{1}{a} \sum_{p=0}^{a-1} \frac{N}{a} \delta[k - \frac{N}{a} p] \right) \quad (2.17)$$

$$F'[k] = \frac{1}{a^2} F[k] \otimes \sum_{p=0}^{a-1} \delta[k - \frac{N}{a} p] \quad (2.18)$$

onde o símbolo $\otimes$ significa convolução circular de N pontos e a é o fator de redução.

Como os elementos de $f'[n]$ para $n \neq ap$ com $p=0,1,2,\dots, (N/a)-1$ são nulos, a DFT desta seqüência pode ser calculada sem eles (pois sua inclusão não altera o somatório na Eq. 2.6), isto é $F'[k]$ pode ser calculada através de:

$$F'[k] = \sum_{n=0}^{(N/a)-1} f'[an] \exp(-j2k\pi n / N) \quad \text{para } 0 \leq k < N/a \quad (2.19)$$

Definindo $N'=N/a$ pode-se reescrever a Eq. (2.19) como:

$$F'[k] = \sum_{n=0}^{N'-1} f'[an] \exp(-j2k\pi n / N')$$

Como a função contorno do objeto reduzido $g[n]=f'[an]$ para $0 \leq n < N/a$ resulta:

$$F'[k] = \sum_{n=0}^{N'-1} g[n] \exp(-j2k\pi n / N') = \text{DFT}_{N'}\{g[n]\} \quad \text{para } 0 \leq k < N/a = N' \quad (2.20)$$

E igualando a Eq. (2.18) à Eq. (2.20), obtém-se que a DFT de $g[n]$ (função contorno do objeto reduzido) é igual à convolução circular da DFT da função contorno do objeto original com um trem de impulsos dividida pelo quadrado do fator de redução.

Portanto, para que seja possível comparar as DFTs de contornos de tamanhos diferentes, como sabe-se o comprimento de cada DFT (que é igual ao número de amostras do contorno), deve-se fazer a convolução da maior sequência com um trem de impulsos (como o da versão frequência da Eq. 2.17), dividir todas as harmônicas pelo quadrado da relação entre os comprimentos e considerar somente as harmônicas de ordem inferior ao comprimento da maior DFT dividido pelo fator de redução. Também pode-se efetuar o processo equivalente a esse no domínio espacial, que é a multiplicação da função contorno maior por um trem de impulsos de valor igual ao inverso do fator de redução (versão espacial da Eq. 2.17), reamostragem desta sequência de forma a eliminar os elementos zerados e cálculo da DFT dessa nova sequência. A diferença entre este procedimento e o "downsampling" é que neste último é necessária uma filtragem do tipo passa-baixas ideal antes da reamostragem para garantir que o critério de Nyquist seja satisfeito para a nova sequência. Essa filtragem elimina as altas frequências que podem ser introduzidas na reamostragem (devido à diferenças elevadas entre as novas amostras consecutivas). No procedimento indicado nesta seção para redução de objetos (e conseqüente redução do comprimento da transformada do seu contorno) os valores das amostras foram divididos pelo fator de redução, o que aproxima os pontos consecutivos e impede a geração de altas frequências (que permitem as variações bruscas) na reamostragem.

E, após isso, para que seja obtida invariância à escala deve-se reduzir todos os objetos a um tamanho padrão, o que pode ser feito dividindo todas as harmônicas pelo módulo da primeira, tornando-a sempre unitária.

2.3.3.6 Normalização completa

Para obter todas as invariâncias desejadas, como os critérios de normalização determinados pelas seções anteriores (2.3.3.1 a 2.3.3.5) não são conflitantes, deve-se considerá-los todos, ou seja:

- faz-se a convolução da DFT da função contorno com um trem de impulsos para reduzir seu comprimento a um tamanho padrão, divide-se todas as harmônicas pelo quadrado da relação entre os números de harmônicas e divide-se novamente todas as harmônicas pelo módulo da primeira ($F[1]$) para obter invariância à escala
- despreza-se $F[0]$ para obtenção da invariância ao deslocamento

- considera-se somente o módulo das harmônicas para haja invariância sob rotação, ponto inicial e reflexão



Figura 2.5a: objeto original



Figura 2.5b: objeto reconstruído com harmônicas 2 e $N - 2$ conjugadas



Figura 2.5c: objeto reconstruído com partes reais e imaginárias das harmônicas 2 e $N - 2$ trocadas entre si



Figura 2.5d: objeto reconstruído com partes imaginárias das harmônicas 2 e $N - 2$ nulas, mas mantendo os módulos



Figura 2.5e: objeto reconstruído com sinais das harmônicas 2 e $N - 2$ invertidos



Figura 2.5f: objeto reconstruído com harmônicas 2 e $N - 2$ conjugadas e partes reais e imaginárias trocadas entre si

Pelo fato de ser considerado apenas o módulo de cada componente do espectro passa-se de 2 parâmetros por harmônica para somente 1. Isto implica em perda de informação, perdeu-se exatamente o que não era desejado num descritor invariante, isto é, as informações de posição absoluta, posição angular, posição de início de varredura, tamanho e espelhamento. Mas não foram somente esses dados os perdidos. Esse processo de obtenção de invariância também produziria saídas iguais para, por exemplo, um objeto cuja DFT do contorno é igual a de outro para algumas

das harmônicas, mas é igual ao conjugado para outras. Esses objetos obviamente não são iguais, mas não podem ser distinguidos por esse método que utiliza apenas o módulo. Exemplos desse fenômeno podem ser vistos na Fig. 2.5, onde foram efetuadas várias operações matemáticas nas harmônicas 2 e N-2 (isto é, frequência positiva e negativa da segunda harmônica para evitar entrelaçamento das funções contorno $x[n]$ e $y[n]$ na reconstrução) que alteram seus valores mas mantém os módulos constantes. As operações são especificadas na legenda de cada figura. As diferenças entre as Figs. 2.5a a 2.5f são pequenas porque o valor da 2ª harmônica dos seus contornos é baixo, isto é, esta frequência tem pouca contribuição na composição desses objetos. Para limitar a perda de informação a somente aquelas indesejadas, é necessário outro método, os descritores elípticos de Fourier, definidos na próxima seção.

2.4 Descritores elípticos de Fourier

A utilização destes descritores foi proposta por Staib e Duncan (1992) para evitar a ambigüidade inerente dos descritores de Fourier mostrada na seção 2.3.3.6 e vem da observação que cada harmônica (exceto a zero), quando vista isoladamente, representa uma elipse no plano xy . Para verificar isso pode-se partir da Eq. (2.3) reescrita abaixo, que permite a reconstrução das seqüências $x[n]$ e $y[n]$ a partir de suas componentes harmônicas.

$$\begin{bmatrix} x[n] \\ y[n] \end{bmatrix} = \sum_{k=0}^{N-1} \begin{bmatrix} a_k & b_k \\ c_k & d_k \end{bmatrix} \begin{bmatrix} \cos(2k\pi n / N) \\ \sin(2k\pi n / N) \end{bmatrix} \quad (2.3)$$

Para simplificar os cálculos, pode-se supor que os coeficientes b_k e c_k são nulos (a generalização é feita abaixo), então a contribuição de cada harmônica na reconstrução das funções originais é dada por :

$$x_k[n] = a_k \cos(2k\pi n / N) \quad \text{e} \quad y_k[n] = d_k \sin(2k\pi n / N)$$

Dividindo as duas equações pelos coeficientes no caso em que eles não são nulos obtém-se:

$$\frac{x_k[n]}{a_k} = \cos(2k\pi n / N) \quad \text{e} \quad \frac{y_k[n]}{d_k} = \sin(2k\pi n / N)$$

Elevando ambos os lados de cada equação ao quadrado, vem :

$$\left(\frac{x_k[n]}{a_k}\right)^2 = \cos^2(2k\pi n / N) \quad \text{e} \quad \left(\frac{y_k[n]}{d_k}\right)^2 = \sin^2(2k\pi n / N)$$

Somando-as,

$$\left(\frac{x_k[n]}{a_k}\right)^2 + \left(\frac{y_k[n]}{d_k}\right)^2 = \cos^2(2k\pi n / N) + \sin^2(2k\pi n / N) = 1$$

Ou seja,

$$\left(\frac{x_k[n]}{a_k}\right)^2 + \left(\frac{y_k[n]}{d_k}\right)^2 = 1 \quad (2.21)$$

que é a equação da elipse com semi-eixos de comprimento a_k e d_k coincidentes com os eixos x e y , respectivamente. No caso geral, serão observadas elipses cujos semi-eixos formam ângulos com as retas $x=0$ e $y=0$. A observação acima é válida para qualquer harmônica (exceto a zero, que é apenas um ponto: o centróide do contorno). Os descritores elípticos utilizam essas elipses para representar os contornos realizando normalizações para obter apenas as invariâncias desejadas, sem outros efeitos colaterais prejudiciais ao reconhecimento e classificação.

2.4.1 Determinação dos descritores elípticos de Fourier

Como a transformada discreta da função contorno dos objetos é calculada pela Eq. (2.6) neste trabalho, a reconstrução da função no domínio espacial deve ser obtida pela Eq. (2.7), e a contribuição de cada harmônica k para esta reconstrução é dada por:

$$f_k[n] = \frac{1}{N} (F[k] \exp(j2k\pi n / N) + F[N-k] \exp(j2(N-k)\pi n / N))$$

$$= \frac{1}{N} (F[k](\cos(2k\pi n / N) + \sin(2k\pi n / N)) + F[N-k](\cos((2(N-k)\pi n / N) + j \sin(2(N-k)\pi n / N)))$$

Dada a periodicidade e simetria das funções seno e cosseno resulta:

$$f_k[n] = \frac{1}{N} (\cos(2k\pi n / N)(F[k] + F[N-k]) + j \sin(2k\pi n / N)(F[k] - F[N-k]))$$

Supondo que a DFT da sequência $x[n]$ seja $X[k] = a_k + jb_k$ e a DFT de $y[n]$ igual a $Y[k] = c_k + jd_k$ e utilizando as Eqs. (2.26) vem:

$$f_k[n] = \frac{1}{N} (2a_k \cos(2k\pi n / N) + j2c_k \cos(2k\pi n / N) - j2d_k \sin(2k\pi n / N) - 2b_k \sin(2k\pi n / N))$$

Portanto,

$$x_k[n] = \frac{2}{N} (a_k \cos(2k\pi n / N) - b_k \sin(2k\pi n / N))$$

e

$$y_k[n] = \frac{2}{N} (c_k \cos(2k\pi n / N) - d_k \sin(2k\pi n / N))$$

ou,

$$x_k[n] = \frac{2}{N} \sqrt{a_k^2 + b_k^2} \cos\left(2k\pi n / N + \tan^{-1} \frac{b_k}{a_k}\right) \quad (2.22)$$

e

$$y_k[n] = \frac{-2}{N} \sqrt{c_k^2 + d_k^2} \sin\left(2k\pi n / N - \tan^{-1} \frac{c_k}{d_k}\right) \quad (2.23)$$

Definindo

$$d_k^2[n] = x_k^2[n] + y_k^2[n] \quad (2.24)$$

e derivando esta Eq. (2.24) em relação a n pode-se encontrar os n 's em que ocorrem as distâncias máximas e mínimas. As soluções são dadas por:

$$n = \frac{N}{2k\pi} \left(a \tan \left(\frac{\sqrt{a_k^2 + b_k^2} \sin(2 \tan^{-1} b_k / a_k) + \sqrt{c_k^2 + d_k^2} \sin(2 \tan^{-1} c_k / d_k)}{\sqrt{c_k^2 + d_k^2} \cos(2 \tan^{-1} c_k / d_k) - \sqrt{a_k^2 + b_k^2} \cos(2 \tan^{-1} b_k / a_k)} \right) - \lambda \pi \right) \quad (2.25)$$

com $\lambda=0,1$

A demonstração desse resultado não é mostrada por ser bastante longa e não apresentar dificuldade teórica. E substituindo esse n 's na Eq. (2.24), pode-se calcular os semi-eixos maior (M_k) e menor (m_k), e utilizando os mesmos n 's nas Eqs (2.22) e (2.23) determina-se o ângulo θ_k , que é definido como o menor (em módulo) dos ângulos que o semi-eixo maior forma com o eixo das abscissas. Usando $n=0$ nas Eqs. (2.22) e (2.23) pode-se calcular o ângulo φ_k , definido como o ângulo formado entre o segmento de reta que passa pela origem e pelo ponto $(x[0],y[0])$ e o eixo das abscissas. Estes parâmetros podem ser identificados na Fig. 2.6, e formam outro conjunto não redundante de $4N/2=2N$ valores (a divisão por 2 é realizada porque a geração de cada elipse inclui as frequências positivas e negativas).

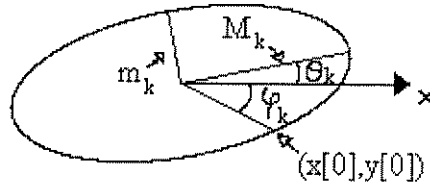


Figura 2.6: ilustração dos parâmetros dos descritores elípticos de Fourier

Na reconstrução da curva através das harmônicas, estas elipses somam-se vetorialmente, numa espécie de sistema planetário onde cada elipse move seu centro sobre a curva formada pelas anteriores com velocidade (determinada pela frequência) crescente, e os pontos da curva final são obtidos pelo somatório vetorial (para cada instante) de todos os fasores que geram as elipses, como ilustrado na Fig. 2.7. Uma aproximação da curva pode ser obtida pela interrupção (em qualquer ordem de harmônica) do somatório vetorial ordenado pela frequência, e pelas características da transformada de Fourier, essa reconstrução é a melhor possível para cada número de elipses.

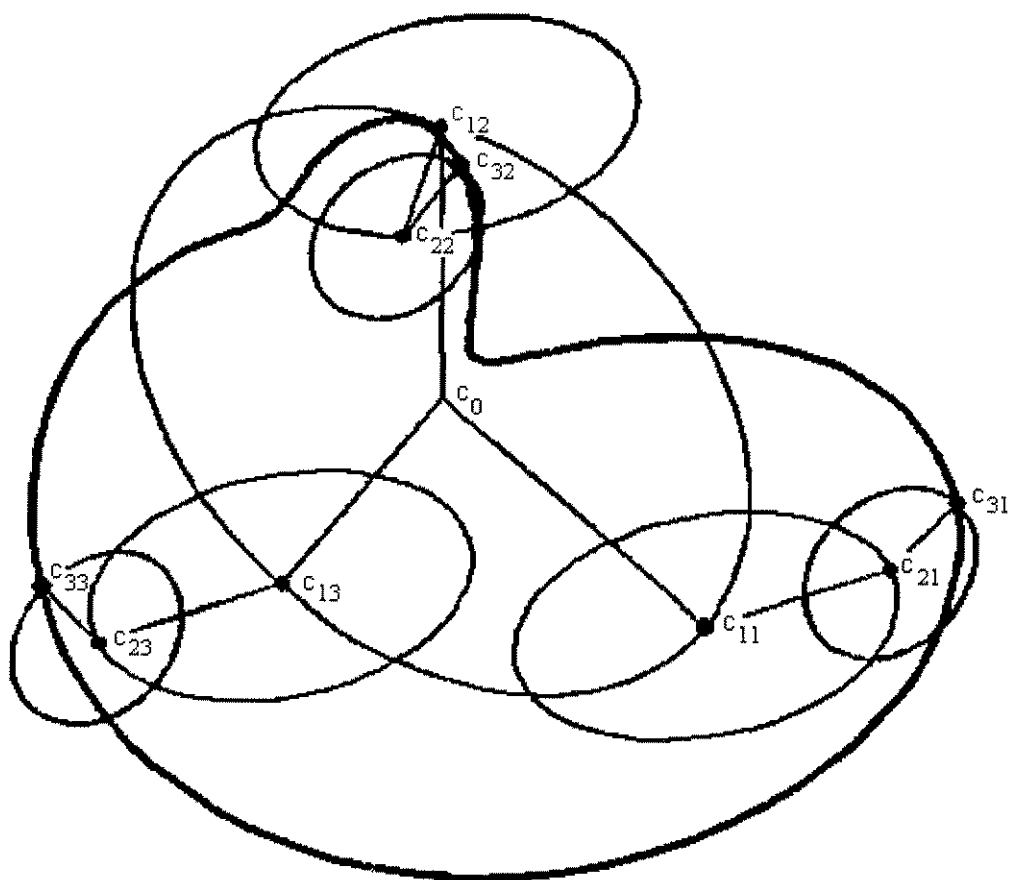


Figura 2.7: ilustração da reconstrução de um objeto a partir de suas elipses. A curva mais grossa é composta por três elipses indicadas pelos seus centros em três instantes diferentes. O primeiro índice indica a ordem da harmônica (sua frequência), e o segundo o instante em que ela foi traçada. Nesse caso os pontos com primeiro índice igual a 3 formam a curva original. Retirado de Staib e Duncan (1992).

2.4.1.1 Cálculo dos coeficientes

O cálculo dos coeficientes dos descritores elípticos de Fourier é feito a partir dos valores já conhecidos da DFT de $f[n]=x[n]+jy[n]$. Porém é necessário conhecer os valores das DFTs de $x[n]$ e de $y[n]$ isoladamente. Como a DFT é linear, eles podem ser recuperados da seguinte forma:

Como

$$F[k]=X[k]+jY[k] \quad \text{e} \quad X[k]=a_k + jb_k \quad \text{e} \quad Y[k]=c_k + jd_k$$

Então

$$F[k]=g_k + jh_k = a_k + jb_k + j(c_k + jd_k)$$

e portanto tem-se o sistema de equações:

$$\begin{cases} g_k = a_k - d_k \\ h_k = b_k + c_k \end{cases}$$

Igualmente,

$$\begin{cases} g_{N-k} = a_{N-k} - d_{N-k} \\ h_{N-k} = b_{N-k} + c_{N-k} \end{cases}$$

Mas como $x[n]$ e $y[n]$ são funções reais, a parte real de suas DFTs tem simetria par, isto é, $a_k = a_{N-k}$ e a parte imaginária tem simetria ímpar, isto é, $b_k = -b_{N-k}$. Portanto, o sistema de 2 equações e 4 incógnitas definido acima pode ser resolvido:

$$a_k = (g_k + g_{N-k})/2 \quad (2.26a)$$

$$b_k = (h_k - h_{N-k})/2 \quad (2.26b)$$

$$c_k = (h_k + h_{N-k})/2 \quad (2.26c)$$

$$d_k = (g_{N-k} - g_k)/2 \quad (2.26d)$$

Apesar de serem necessários os valores das DFTs isoladas de $x[n]$ e de $y[n]$ agora, seu cálculo em conjunto foi proveitoso porque agora não são realizadas mais operações com números complexos, apenas somas, subtrações e divisões por 2, todas reais, e como será visto mais adiante, isso não precisa ser feito para todas as harmônicas, tendo um lucro líquido em quantidade de cálculos e tempo.

2.4.2 Normalizações

Da mesma que foi feita para os descritores de Fourier nas seções 2.3.3.1 a 2.3.3.6, para que o reconhecimento através dos descritores elípticos de Fourier seja possível mesmo com todas as variações na imagem de um mesmo objeto (devidas ao processo de digitalização) mencionadas na introdução deve-se utilizar descritores

invariantes (ou insensíveis) a todas essas mudanças. Isto é obtido realizando operações matemáticas nos descritores que produzem normalizações. Estas operações são deduzidas nas seções seguintes.

2.4.2.1 Normalização quanto à translação

Como demonstrado para os descritores de Fourier na seção 2.3.3.1, a informação de posição absoluta (coordenadas do centróide) está na componente de $k=0$ do espectro, as outras harmônicas contém a informação sobre a forma do objeto. E como os descritores elípticos de Fourier também são obtidos a partir da DFT, para normalização quanto à translação basta desprezar a componente zero (ou DC) da transformada.

2.4.2.2 Normalização quanto à rotação

Se um objeto é rotacionado em torno do seu centróide de um ângulo α , todas as suas componentes elípticas são rodadas do mesmo ângulo. Matematicamente, isso pode ser verificado revendo a Eq. (2.3):

$$\begin{bmatrix} x[n] \\ y[n] \end{bmatrix} = \sum_{k=0}^{N-1} \begin{bmatrix} a_k & b_k \\ c_k & d_k \end{bmatrix} \begin{bmatrix} \cos(2k\pi n / N) \\ \sin(2k\pi n / N) \end{bmatrix}$$

A rotação de um ângulo α , como já referido na seção 2.3.1 (Eq. 2.9), corresponde à pré-multiplicação das coordenadas dos pontos pela matriz:

$$\begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$$

Pré-multiplicando-se os dois lados da equação de síntese dos descritores de Fourier pela matriz de rotação e, sendo essa matriz uma constante para a variável k , ela pode ser inserida no somatório, rotacionando portanto cada harmônica do mesmo ângulo. Embora isso tenha sido indicado para os descritores de Fourier, e não os

elípticos, estes últimos são calculados a partir dos primeiros, cada harmônica isoladamente, portanto, se os descritores de Fourier foram rotacionados de um ângulo α , os elípticos também o serão. Esse ângulo de rotação α será adicionado ao ângulo de orientação θ_k de cada elipse (e também aos φ_k , definidos na seção 2.4.1), não alterando as posições angulares relativas entre as elipses, mas apenas o ângulo que as mesmas formam com os eixos x e y (ângulo absoluto).

Para normalizar os descritores elípticos quanto à rotação, basta portanto considerar apenas os ângulos relativos entre as elipses. Isso pode ser feito rotacionando arbitrariamente todas as componentes elípticas do ângulo $-\theta_1$, que é o menor dos ângulos que o semi-eixo maior da primeira elipse forma com o eixo x, fazendo, dessa forma, com que os semi-eixos da primeira elipse componente de todos os contornos coincidam com as retas $x = 0$ e $y = 0$. Isto corresponde no processo de normalização a definir os ângulos $\theta'_k = \theta_k - \theta_1$ e $\varphi'_k = \varphi_k - \theta_1$ que devem ser utilizados como parâmetros de descrição.

2.4.2.3 Normalização quanto ao ponto inicial

A variação do ponto inicial modifica o ângulo de ponto inicial φ (definido na seção 2.4.1) de cada elipse. Para normalizar os descritores sobre esse aspecto, o procedimento é semelhante ao da rotação: altera-se φ_1 para 0 e reajusta-se todos os outros φ_k de modo que a diferença relativa se mantenha, isto é, não deformando o objeto. Diferentemente de θ_k , a modificação do ângulo φ_k depende do índice da harmônica (k). Como pôde ser calculado a partir da equação de síntese, para n variando de 0 a $N-1$, cada harmônica gerará uma elipse, mas essa elipse será percorrida k vezes, devido a frequência crescente das funções seno e cosseno e sua periodicidade. Em outras palavras, para uma variação de n correspondente ao percurso por um arco de ângulo ψ na primeira elipse, em cada uma das outras elipses será atingido o ângulo $k\psi$. Portanto, para a normalização, define-se o parâmetro invariante $\varphi''_k = \varphi'_k - k\varphi'_1$

2.4.2.4 Normalização quanto à reflexão

Como indicado na seção 2.3.3.4, a reflexão de um objeto pode ser obtida pela combinação da reflexão em torno do eixo jy , uma translação e inversão da ordem dos pontos da função contorno. Pela Fig. 2.4 da mesma seção e Fig. 2.6, que ilustra os descritores elípticos, pode-se perceber (realizando esse processo para uma elipse) que isso simplesmente reflete os ângulos θ_k e φ_k de todas as harmônicas em torno do eixo jy , isto é, um objeto, se espelhado será composto por elipses cujos ângulos θ'_k são iguais ao negativo dos valores dos mesmos ângulos para a imagem original e os ângulos φ'_k são iguais $\pi - \varphi_k$ se $0 \leq \varphi_k \leq \pi$ e $-\pi - \varphi_k$ se $-\pi \leq \varphi_k < 0$. Para normalização quanto à reflexão deve-se eleger um dos ângulos θ_k como referência tornando-o positivo (se já não for) para todos os objetos e alterando também todos os outros ângulos θ_k e φ_k desses objetos.

Para este trabalho foi escolhido o parâmetro θ'_2 como referência porque ele é o mais significativo da representação já que θ'_1 foi tornado nulo para obtenção da invariância quanto à rotação e as elipses de maior ordem têm contribuição sucessivamente menor (de modo geral) na descrição da forma dos objetos.

2.4.2.5 Normalização quanto à escala

Como mostrado na seção 2.4.1, os descritores elípticos derivam dos descritores de Fourier, portanto a forma de obtenção da invariância com a escala é semelhante à indicada na seção 2.3.3.5, isto é, deve-se efetuar uma convolução para redução do comprimento da DFT da função contorno para um tamanho padrão e após dividir todas as harmônicas pelo quadrado da relação entre os comprimentos, e somente então calcular os parâmetros das elipses. Feito isto deve-se alterar todas as elipses de forma que os semi-eixos da primeira tenham comprimento unitário, isto é, chamando os semi-eixos maiores de M_k e os menores de m_k , faz-se $M'_k = M_k / M_1$ e $m'_k = m_k / m_1$ para normalizar os tamanhos dos objetos descritos.

Isso corresponde a reescalonar o objeto todo para um formato padrão que permita a comparação, degenerando a primeira elipse para uma circunferência.

Outro processo alternativo seria definir $M'_k = M_k / M_1$ e $m'_k = m_k / M_1$, o que mantém a relação entre os comprimentos dos eixos de cada elipse constante, apenas altera seus tamanhos de forma que o semi-eixo maior da primeira elipse tenha valor unitário.

2.4.2.6 Normalização completa

Como os procedimentos deduzidos nas seções 2.4.2.1 a 2.4.2.5 não são conflitantes, a normalização completa é obtida fazendo todas as normalizações indicadas, isto é,

- realizando a convolução e divisão pelo fator de comprimentos na DFT
- desconsiderando a harmônica zero para obtenção da invariância ao deslocamento
- calculando $\theta'_k = \theta_k - \theta_1$ e $\varphi'_k = \varphi_k - \theta_1$ para que haja invariância sob rotação
- definindo $\varphi''_k = \varphi'_k - k\varphi_1$ para obtenção da invariância quanto ao ponto inicial
- calculando $M'_k = M_k / M_1$ e $m'_k = m_k / m_1$ (ou $M'_k = M_k / M_1$ e $m'_k = m_k / M_1$) para invariância à escala
- alterando o sinal de todos os θ'_k e refletindo os φ''_k em torno do eixo jy se θ'_2 for negativo para ignorar reflexões

2.5 Descritores cossenoidais

A geração desses descritores vem do fato que sempre pode-se calcular a DCT (transformada cosseno discreta) de uma sequência de números e é desejável pela propriedade de maior compactação de energia nas baixas frequências (em comparação com DFT) que esta transformada possui (ver Jain (1989)). A equação da DCT é :

$$V[k] = \alpha[k] \sum_{n=0}^{N-1} f[n] \cos\left(\frac{(2n+1)k\pi}{2N}\right) \quad \text{para } 0 \leq k < N \quad (2.27)$$

onde $\alpha[0] = \sqrt{\frac{1}{N}}$ e $\alpha[k] = \sqrt{\frac{2}{N}}$ para $0 < k < N$

Como foi demonstrado por Jain (1989), a DCT de N pontos pode ser calculada através da DFT, e portanto pelo algoritmo de FFT de N pontos segundo a Eq. (2.28).

$$V[k] = \text{Re}\left\{\alpha[k] \exp\left(\frac{-jk\pi}{2N}\right) \text{DFT}_N\{f[n]\}\right\} \quad (2.28)$$

Como os coeficientes $V[k]$ são reais, são necessários apenas dois coeficientes por harmônica (um devido a $x[n]$, e outro a $y[n]$), em vez dos quatro parâmetros dos descritores elípticos de Fourier para representar a função contorno dos objetos, e sem as ambigüidades mostradas na seção 2.3.3.6 que ocorrem com os descritores de Fourier. A comparação desses descritores cossenoidais exigiria, portanto, menos memória, cálculos e conseqüentemente, menos tempo para o reconhecimento.

A equação de síntese da DCT é :

$$f[n] = \sum_{k=0}^{N-1} \alpha[k] V[k] \cos\left(\frac{(2n+1)k\pi}{2N}\right) \quad \text{para } 0 \leq n < N \quad (2.29)$$

com os mesmos $\alpha[k]$ da Eq. (2.27).

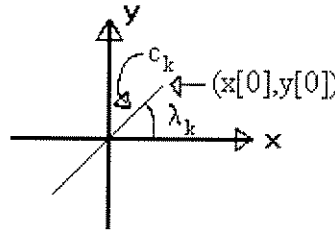


Figura 2.8: ilustração dos parâmetros dos descritores cossenoidais

Pela Eq. (2.29) pode-se perceber que a contribuição de cada harmônica para a reconstrução de $f[n]$ é uma função cosseno. A reconstituição da função contorno do

objeto, quando plotada no plano xy , é um segmento de reta que passa pela origem. Pode-se então associar a esse segmento duas parametrizações: o seu comprimento e o ângulo que seu ponto inicial faz com o eixo x (ver Fig. 2.8). Como esse remapeamento não altera o número de parâmetros e não é redundante, não foi perdida informação.

2.5.1 Cálculo dos coeficientes

Como mencionado na seção anterior, os coeficientes da DCT são calculados pela equação (2.28), e a partir deles obtém-se os descritores cossenoidais c_k e λ_k , onde

$$c_k = 2\alpha[k]\sqrt{V_x^2[k] + V_y^2[k]} \quad (2.30)$$

e

$$\lambda_k = \tan^{-1}\left(\frac{V_y[k]}{V_x[k]}\right) \quad (2.31)$$

onde $V_x[k]$ é a DCT da sequência $x[n]$, e $V_y[k]$ é a DCT de $y[n]$.

2.5.2 Normalizações

Da mesma forma feita para os outros descritores, devem ser feitas operações matemáticas que produzam invariâncias à translação, rotação, ponto inicial, reflexão e escala para que os descritores cossenoidais possam ser utilizados para identificação de objetos digitalizados. As normalizações para os descritores cossenoidais são (em parte) semelhantes àquelas já deduzidas para os outros descritores, sendo as demonstrações apresentadas nas seções seguintes.

2.5.2.1 Normalização quanto à translação

Analogamente à transformada de Fourier, uma translação do objeto é realizada pela adição de uma constante à função contorno, o que implicará numa variação

somente da componente de $k=0$ da transformada cosseno. Usando a equação da definição da DCT (Eq. 2.27) resulta:

$$\begin{aligned} V'[k] &= \alpha[k] \sum_{n=0}^{N-1} (f[n] + \Delta x) \cos\left(\frac{(2n+1)k\pi}{2N}\right) \quad \text{para } 0 \leq k < N \\ &= \alpha[k] \sum_{n=0}^{N-1} f[n] \cos\left(\frac{(2n+1)k\pi}{2N}\right) + \alpha[k] \sum_{n=0}^{N-1} \Delta x \cos\left(\frac{(2n+1)k\pi}{2N}\right) \end{aligned}$$

E substituindo a primeira parcela pela DCT de $f[n]$, conforme Eq. (2.27) e deslocando o fator Δx para fora do segundo somatório resulta:

$$V'[k] = V[k] + \alpha[k] \Delta x \sum_{n=0}^{N-1} \cos\left(\frac{(2n+1)k\pi}{2N}\right)$$

E sabendo que $\sum_{n=0}^{N-1} \cos\left(\frac{(2n+1)k\pi}{2N}\right) = N\delta[k]$ obtém-se:

$$V'[k] = V[k] + \alpha[k] \Delta x N \delta[k]$$

Portanto, como ocorrido antes, exceto pela harmônica de $k=0$, a transformada cosseno discreta não se altera com a translação do objeto, portanto, para obter invariância ao deslocamento deve-se desconsiderar a harmônica zero.

2.5.2.2 Normalização quanto à rotação

Como demonstrado a seguir, uma rotação irá girar todos os ângulos λ_k do mesmo valor:

$$\begin{aligned} \begin{bmatrix} V'_x[k] \\ V'_y[k] \end{bmatrix} &= \alpha[k] \sum_{n=0}^{N-1} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x[n] \\ y[n] \end{bmatrix} \cos\left(\frac{(2n+1)k\pi}{2N}\right) \quad \text{para } 0 \leq k < N \\ &= \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \alpha[k] \sum_{n=0}^{N-1} \begin{bmatrix} x[n] \\ y[n] \end{bmatrix} \cos\left(\frac{(2n+1)k\pi}{2N}\right) \\ &= \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} V_x[k] \\ V_y[k] \end{bmatrix} \end{aligned}$$

Ou seja, o vetor formado por $V_x[k]$ e $V_y[k]$ foi rotacionado de θ radianos, e como λ_k é a fase desse vetor, também foi girada do mesmo ângulo.

Portanto, da mesma forma feita para os descritores elípticos, para normalizar os cossenoidais quanto à rotação deve-se considerar apenas os ângulos relativos entre os segmentos de reta. Isso pode ser feito rotacionando arbitrariamente todas os componentes do valor $-\lambda_1$, que é o negativo do ângulo que o ponto inicial do primeiro segmento forma com o eixo x, fazendo, dessa forma, com que os primeiros segmentos de todos os contornos coincidam com a reta $y=0$. Isto é, define-se os parâmetros invariantes $\lambda'_k = \lambda_k - \lambda_1$

2.5.2.3 Normalização quanto ao ponto inicial

A normalização quanto ao ponto inicial é a mais complexa delas. Pode ser demonstrado que a transformada cosseno de uma seqüência não pode ser obtida diretamente a partir da transformada de qualquer um dos deslocamentos circulares da seqüência dentro de um período (isto é, sem utilizar os valores da seqüência, ver Eq. (2.32) e demonstração simplificada em Rao e Yip (1990)). Uma solução possível para a normalização é descobrir qual deslocamento fornece o maior valor para c_1 , mas isso implicaria no cálculo de N transformadas cossenos (ou pelo menos a 1ª harmônica das transformadas dos N deslocamentos circulares da seqüência). Contudo, pode-se alterar um algoritmo de FFT para que este forneça somente estas N primeiras harmônicas de uma só vez por um custo computacional semelhante ao de uma FFT de N elementos. O fluxograma em malha de ramos direcionados desta variação pode ser visto na Fig. 2.9, onde já estão incluídos os cálculos para os coeficientes c_{1d} , onde o sub-índice d representa o valor do deslocamento circular. O fator $2\alpha[1]$ não foi incluído porque é uma constante para todos os coeficientes, portanto não altera suas relações. Este algoritmo efetua um cálculo eficiente dos valores da Eq. (2.32) para $k=1$ e $0 \leq d < N$ e o custo desta operação é de $8N$ multiplicações reais e $2N(\log_2 N + 1) - 2$ adições reais para calcular todos os $V[1]_d$. Como ilustração, um algoritmo de FFT requer $2N \log_2 N$ multiplicações reais e $4N \log_2 N$ adições reais. Ambas as especificações desprezam o cálculo dos "twiddle factors" e multiplicações por -1.

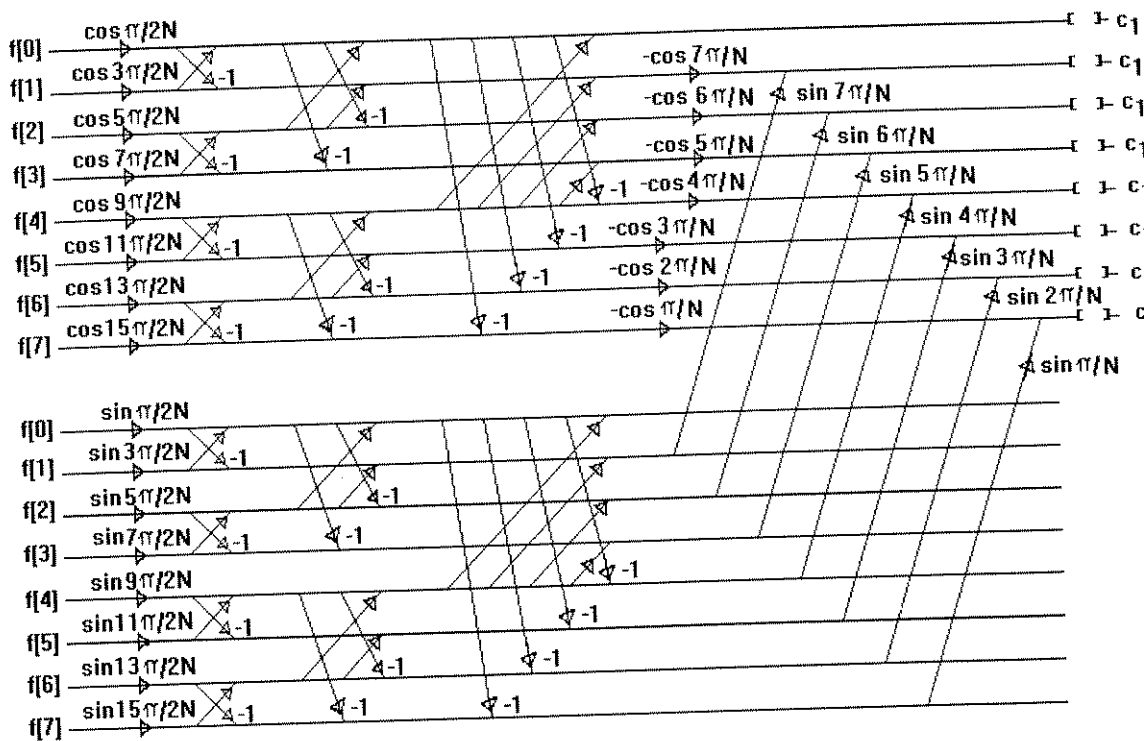


Figura 2.9: exemplo com $N=8$ do fluxograma do algoritmo de cálculo do comprimento do primeiro segmento de reta dos descritores cossenoidais para os N deslocamentos circulares da sequência de entrada. Os colchetes representam a operação módulo. São executados $8N$ multiplicações reais e $2N(\log_2 N + 1) - 2$ adições reais, sem contar as operações de módulo.

A equação para o cálculo dos valores da DCT para qualquer deslocamento circular da sequência de entrada é:

$$V_d[k] = \cos(dk / N) \sum_{n=0}^{N-1} f[n] (1 - 2u_{-1}[n - N + d])^k \cos\left(\frac{(2n+1)k\pi}{2N}\right) \quad (2.32)$$

onde d é o índice do deslocamento circular, e $u_{-1}[n]$ é uma função que vale 0 para $n < 0$ e 1 para os demais valores de seu argumento (também chamada de função salto ou degrau unitário).

2.5.2.4 Normalização quanto à reflexão

Definindo a sequência de pontos do objeto refletido como $f[n] = -f^*[N-n]$, isto é, realizando apenas a reflexão em torno do eixo jy da seção 2.4.2.4 e desprezando a translação (já que ela somente altera a componente $k=0$ da DCT, como demonstrado em 2.5.2.1), a transformada cosseno discreta do objeto refletido é:

$$V'[k] = \alpha[k] \sum_{n=0}^{N-1} (-f^*[N-n]) \cos\left(\frac{(2n+1)k\pi}{2N}\right)$$

Sendo $V'[k] = V'_x[k] + jV'_y[k]$, e após algumas manipulações algébricas chega-se a:

$$\begin{aligned} V'[k] &= -(-1)^k \alpha[k] \left(\sum_{n=0}^{N-1} f[n] \left(\cos\left(\frac{(2n+1)k\pi}{2N}\right) \cos(k\pi/N) - \sin\left(\frac{(2n+1)k\pi}{2N}\right) \sin(k\pi/N) \right) \right)^* \\ &= -(-1)^k \alpha[k] \left(V[k] \cos(k\pi/N) - \sin(k\pi/N) \sum_{n=0}^{N-1} f[n] \sin\left(\frac{(2n+1)k\pi}{2N}\right) \right)^* \end{aligned}$$

E aplicando a Eq. (2.31) para cálculo dos ângulos λ_k vem:

$$\lambda'_k = \tan^{-1} \left(\frac{V_y[k] \cos(k\pi/N) - \sin(k\pi/N) \sum_{n=0}^{N-1} y[n] \sin\left(\frac{(2n+1)k\pi}{2N}\right)}{-V_x[k] \cos(k\pi/N) + \sin(k\pi/N) \sum_{n=0}^{N-1} x[n] \sin\left(\frac{(2n+1)k\pi}{2N}\right)} \right) \quad (2.33)$$

E reconhecendo os somatórios da Eq. (2.33) como a transformada seno discreta das seqüências dos pontos do objeto e chamando-as de $W[k]$, resulta:

$$\lambda'_k = \tan^{-1} \left(\frac{V_y[k] \cos(k\pi/N) - \sin(k\pi/N) W_y[k]}{-V_x[k] \cos(k\pi/N) + \sin(k\pi/N) W_x[k]} \right)$$

Ou seja, esses ângulos não podem ser calculados somente a partir dos valores da DCT do objeto não refletido, são necessários os valores das funções espaciais (ou a transformada seno discreta delas). Para obter invariância à reflexão pode-se eleger um dos ângulos λ_k como referência tornando-o o mais próximo possível de zero através da reflexão (ou não) do objeto no domínio espacial. Mas este procedimento é conflitante com a invariância ao ponto inicial já considerada, e como esta última é

mais importante no conjunto de normalizações, a invariância à reflexão não é aplicada aos descritores cossenoidais neste trabalho.

2.5.2.5 Normalização quanto à escala

Como o teorema da convolução circular não pode ser aplicado à transformada cosseno discreta (porque suas funções base são diferentes da transformada de Fourier discreta) deve-se reduzir o comprimento das seqüências contorno ao tamanho padrão no domínio espacial (através do procedimento explicado na seção 2.3.3.5) e depois calcular a DCT da seqüência reamostrada. A invariância à escala dos descritores cossenoidais é então realizada escalando-se todos os comprimentos dos segmentos de reta pelo mesmo fator, não alterando assim seus valores relativos. Isto é, realiza-se um reescalamento no objeto todo para um formato padrão (possibilitando a comparação) de forma a alterar o primeiro segmento de reta para um comprimento unitário. Isto é, determina-se os descritores invariantes $c'_k = c_k / c_1$.

2.5.2.6 Normalização completa

Não sendo conflitantes os procedimentos para obtenção de cada invariância isoladamente, a normalização completa é obtida fazendo-se simultaneamente todas as normalizações apresentadas nas seções 2.5.2.1 a 2.5.2.5, ou seja:

- aplicando o procedimento especificado na seção 2.3.3.5 (na versão domínio espaço) para reduzir a seqüência de pontos do contorno a um comprimento padrão.
- aplicando o algoritmo especificado na seção 2.5.2.3 para determinação do deslocamento circular da seqüência que produz o maior c_1 e calculando a DCT somente da seqüência com este ordenamento para obtenção da invariância quanto ao ponto inicial
- desconsiderando a harmônica zero para obtenção da invariância ao deslocamento
- calculando $\lambda'_k = \lambda_k - \lambda_1$ para que haja invariância sob rotação

- determinando $c'_k = c_k / c_l$ para invariância à escala

2.6 Momentos

Os momentos estão entre os mais populares descritores de objetos, principalmente por serem conhecidos há décadas e por terem uma interpretação física (Gonzalez e Woods (1993), Pratt (1991), Rosenfeld e Kak (1976) e Schalkoff (1989)). Por situar-se dentro de uma imagem que é finita por natureza, um objeto possui momentos de todas as ordens, que são definidos por:

$$m_{pq} = \sum_x \sum_y x^p y^q f(x, y) \quad (2.34)$$

onde x e y pertencem à imagem e $f(x, y)$ é a intensidade ou nível de cinza do "pixel" de coordenadas (x, y) .

Para obter-se invariância à translação deve-se utilizar os momentos centrais, que são:

$$\mu_{pq} = \sum_x \sum_y (x - x_c)^p (y - y_c)^q f(x, y) \quad (2.35)$$

onde x_c e y_c são as coordenadas do centróide do objeto, isto é, $x_c = m_{10} / m_{00}$ e $y_c = m_{01} / m_{00}$.

Para invariância à escala utiliza-se os momentos centrais normalizados, através de:

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^\gamma} \quad (2.36a)$$

$$\text{onde } \gamma = \frac{p+q}{2} + 1 \quad (2.36b)$$

Os momentos centrais normalizados definidos nas Eqs. (2.36) não são invariantes com a rotação; para eliminar essa dependência Bell (1965) e Hu (1962) propuseram sete momentos invariantes definidos por:

$$\phi_1 = \eta_{20} + \eta_{02} \quad (2.37a)$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \quad (2.37b)$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \quad (2.37c)$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \quad (2.37d)$$

$$\phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad (2.37e)$$

$$\phi_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \quad (2.37f)$$

$$\phi_7 = (3\eta_{21} - \eta_{30})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ + (3\eta_{21} - \eta_{30})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad (2.37g)$$

Porém, mesmo estes 7 momentos invariantes dependem das condições de iluminação na aquisição das imagens do objeto (determinadas por $f(x,y)$). A solução trivial para esta limitação (quando se deseja descrever somente a forma dos objetos) é a binarização da imagem, ou seja, tornar $f(x,y)=1$ para todos os "pixels" do objeto, e zero para os outros.

Outro conjunto de momentos invariantes, estes calculados a partir da seqüência unidimensional de contorno normalizada, foi proposto por Gupta e Srinath (1987) como:

$$\text{Sendo } z_i = \sqrt{(x_i - x_c)^2 + (y_i - y_c)^2} \quad \text{para } 0 \leq i < N \quad (2.38)$$

onde (x_i, y_i) é a seqüência ordenada dos N pontos de contorno, e

$$m_r = \frac{1}{N} \sum_{i=0}^{N-1} z_i^r \quad (2.39)$$

$$\mu_r = \frac{1}{N} \sum_{i=0}^{N-1} (z_i - m_1)^r \quad (2.40)$$

os momentos invariantes da sequência de contorno são dados por:

$$\phi_r = \frac{m_r}{\mu_2^{r/2}} \quad (2.41)$$

$$\varphi_r = \frac{\mu_r}{\mu_2^{r/2}} \quad (2.42)$$

Porém, como estes momentos ainda são muito sensíveis à ruídos, os mesmos autores propõem no mesmo artigo outro conjunto (de apenas 4 elementos) que reduz essa deficiência e são definidos por:

$$F_1 = \frac{\sqrt{\mu_2}}{m_1} \quad (2.43a)$$

$$F_2 = \frac{\mu_3}{\mu_2^{3/2}} \quad (2.43b)$$

$$F_3 = \frac{\mu_4}{\mu_2^2} \quad (2.43c)$$

$$F_4 = \varphi_5 \quad (2.43d)$$

Sendo os valores de m_1 , μ_r e φ_5 das Eqs. (2.43) calculados pelas Eqs. (2.39), (2.40) e (2.42), respectivamente.

CAPÍTULO 3

Algoritmo para geração de funções discriminantes

No processo de reconhecimento de objetos, a aplicação de descritores quantitativos (isto é, descritores que atribuem um ou mais valores numéricos às suas funções de entrada) gera uma representação de cada objeto por um vetor de características. O conjunto de todos os possíveis vetores de características forma um espaço (de dimensão igual ao número de elementos do vetor) que é chamado de espaço de padrões ou espaço de características, onde cada vetor assume a forma de um ponto. Os descritores devem ser escolhidos de forma que as distribuições do conjunto de amostras (pontos dados pelos vetores de características) de cada classe possuam o mínimo possível de sobreposição no espaço de padrões. Ou seja, os descritores devem medir alguma característica que permita a separação das classes nesse espaço.

A classificação consiste em determinar o número de classes existentes no conjunto de amostras (se isso já não for conhecido) e suas posições no espaço de padrões sabendo ou não o rótulo de cada elemento do conjunto de amostras. Um dos métodos de classificação mais comuns é a análise de "clusters", que divide o espaço de padrões em regiões (separando as classes) a partir de alguma medida de similaridade entre os vetores de características. As classes são geradas agrupando os elementos mais próximos (segundo essa medida de similaridade) entre si. Os algoritmos K-means, ISODATA, K-NN (Bow (1984)) e CNN (Hart (1968)) são exemplos desse tipo de método.

O processo de identificação é posterior à etapa de classificação, sendo que existem métodos paramétricos e não paramétricos. Os métodos paramétricos determinam os parâmetros de alguma função analítica que modela a distribuição (ou

dispersão) das classes e permite a atribuição de um valor numérico para a certeza de que cada elemento pertença a cada classe. Como exemplos de funções paramétricas podem ser citadas a função perda, função discriminante de Bayes, função de máxima probabilidade (Andrews (1972)) que utilizam informações estatísticas dos dados, sua distribuição ou densidade de probabilidade. Porém, a eficiência desses métodos é comprometida se essas informações estatísticas não forem conhecidas "a priori" ou forem mal estimadas.

Os métodos não paramétricos utilizam uma função de distância para obtenção de um valor numérico para a similaridade entre o vetor de características do objeto a ser reconhecido e as classes consideradas, sendo atribuída a esse objeto a denominação da classe que estiver mais próxima dele. As funções mais comuns para este caso são: a inércia, pertinência fuzzy, distância euclidiana, distância de Mahalanobis e coeficiente de Tanimoto (Andrews (1972)). Estes métodos não paramétricos apresentam limitações devido a tendência dessas medidas de distância serem radiais, tornando as classes convexas e aproximadamente esféricas. Mesmo no caso da distância de Mahalanobis, que usa a matriz de covariância dos dados, e a distância euclidiana ponderada pela direção, haverá falhas em alguns casos extremos como as classes inter-penetrantes da Fig. 3.1. Além disso, estes métodos somente são bem sucedidos se as distâncias entre classes forem maiores que as intra-classes no espaço de padrões considerado.

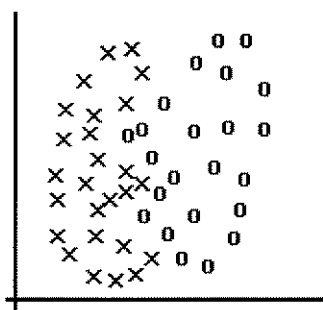


Figura 3.1: caso extremo de espaço de padrões com classes inter-penetrantes

Também podem ser encontrados na literatura métodos não paramétricos que utilizam distribuições gaussianas para definir as classes, como por exemplo o descrito em Fukunaga e Olsen (1970), que apresentam baixa confiabilidade quando a função densidade de probabilidade difere consideravelmente de uma distribuição gaussiana.

Tanto para os métodos paramétricos como não paramétricos, o que se procura é a determinação de funções discriminantes, que são superfícies de separação, ou

superfícies de decisão, que determinam os limites entre duas ou mais classes. Essas superfícies têm dimensão $n - 1$, onde n é a dimensão do vetor de características. As limitações ao uso desse tipo de superfície são a dificuldade de determiná-las e o seu grau. A forma matemática mais simples de separação é dada por superfícies discriminantes lineares, definidas pela Eq. (3.1).

$$f(X) = \sum_{i=1}^n a_i x_i + a_0 \quad (3.1)$$

Embora esta forma seja bastante eficiente e utilizada, algumas classes podem não ser linearmente separáveis, ou seja, podem ser necessárias superfícies quadráticas, cúbicas, etc para sua separação. A utilização de funções discriminantes lineares por partes pode eliminar esta dificuldade, mas em compensação gera outras, tais como a determinação do número de partes e seus intervalos de validade.

Com o objetivo de contribuir para a solução da questão colocada, neste capítulo é proposto um algoritmo rápido e automático que gera funções discriminantes lineares por partes mesmo quando as distribuições de probabilidades das classes não podem ser corretamente modeladas por funções paramétricas conhecidas. Para demonstrar a validade do algoritmo proposto, exemplos com dados sintéticos e reais são apresentados e seus resultados avaliados.

3.1 Descrição do algoritmo proposto

O fluxograma do algoritmo proposto é apresentado na Fig. 3.2, e seu funcionamento é descrito nessa seção. Sendo um algoritmo que necessita de treinamento, este exige uma amostra na qual elementos de cada classe devem ser fornecidos com suas respectivas classificações (não necessariamente todas corretas).

Como primeiro passo do algoritmo são determinados os exemplares limitantes exteriores de cada classe segundo o método rotacional descrito na seção 3.1.1. Exemplares limitantes exteriores de cada classe são aqueles elementos que, quando ligados em ordem, formam o perímetro externo da distribuição espacial que o conjunto de amostras de cada classe apresenta no espaço de padrões.

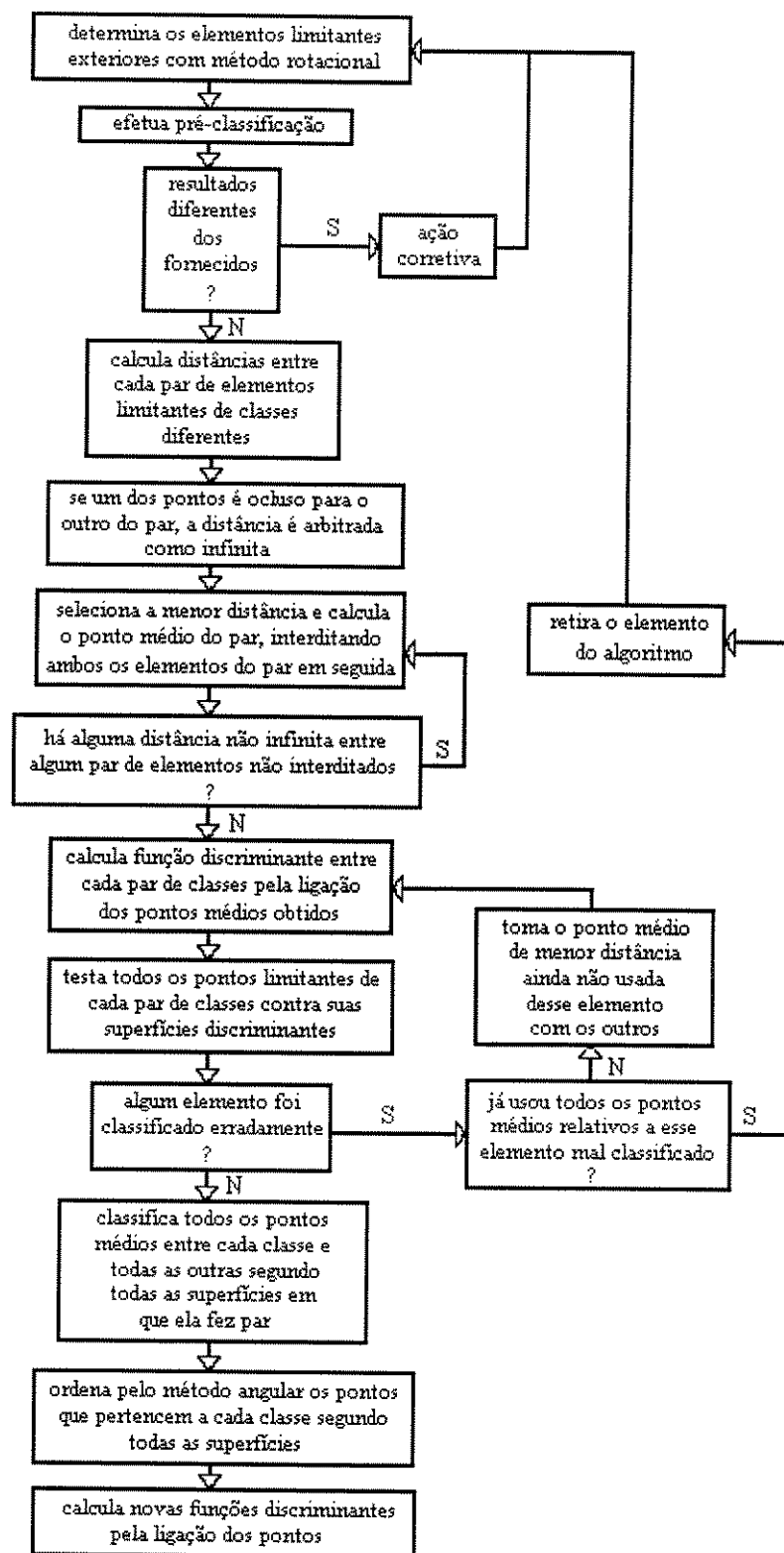


Figura 3.2: fluxograma do algoritmo para geração de funções discriminantes lineares por partes proposto

No segundo passo do algoritmo é efetuada uma pré-classificação com base nos elementos limitantes (descrita com detalhes na seção 3.1.2). Se o resultado obtido

não coincidir com o fornecido inicialmente, ações corretivas são tomadas segundo a Tab. 3.1, e volta-se ao primeiro passo enquanto não houver concordância entre os resultados.

No terceiro passo são calculadas as distâncias entre todos os elementos limitantes exteriores de cada par de grupos. Essa distância é arbitrada como infinita se um dos elementos for ocluso sob o ponto de vista do outro (o teste de oclusão é descrito na seção 3.1.3).

É selecionada a menor das distâncias entre elementos não interditados e calculado o ponto médio entre esses elementos. Os elementos utilizados são imediatamente definidos como interditados, para sua não reutilização. Esta interdição é realizada para não polarizar a superfície discriminante em torno desses elementos. Na sequência é verificado se existe alguma distância finita entre pares de elementos não interditados. Se houver, retorna-se ao passo anterior, caso contrário o algoritmo prossegue para o passo seguinte.

Os pontos médios obtidos são boas estimativas para a especificação das superfícies de separação pois, contendo exatamente os pontos que estão a igual distância entre cada duas classes, elas permitem igual variação para ambas as classes. Supondo que os exemplares fornecidos ao algoritmo determinam uma boa aproximação da função de distribuição das classes (seja ela de que tipo for), este método resulta numa estimativa quase ótima da função discriminante entre cada par de classes. Essa função é então gerada pela ligação entre os pontos médios obtidos, e será constituída por retas para um vetor de características bidimensional ($n=2$), planos para vetores tridimensionais ($n=3$), e hiperplanos para vetores de dimensão maior.

No passo seguinte é feita a verificação da validade dessa função discriminante utilizando o método especificado na seção 3.1.4. O teste deve ser realizado somente com os elementos limitantes do par de classes em questão, pois, por serem os mais exteriores, a função discriminante falhará para eles caso seja inválida. Se algum elemento limitante foi classificado erradamente, a função discriminante deve ser recalculada incluindo mais pontos médios para sua formação. Esse fato ocorre quando há poucos elementos em uma das classes, o que pode induzir a uma estimativa errada da superfície de separação. O ponto médio extra a ser usado é aquele entre o elemento mal classificado e seu vizinho mais próximo da classe oposta se ainda existe algum ponto médio relativo a esse elemento que não foi usado. Caso não exista,

provavelmente a classificação fornecida para o elemento limitante é errada, ele deve ser eliminado da amostra e o algoritmo reiniciado.

No passo seguinte, para cada classe, todos os seus pontos médios, que formam as superfícies de decisão com todas as outras classes, são classificados segundo todas as superfícies discriminantes relativas à classe em consideração. São eliminados os pontos que não pertencem à classe segundo todas as superfícies discriminantes. Os pontos que não foram eliminados com certeza são limitantes da classe, portanto são usados para geração da função discriminante linear por partes. Para isso, eles são ordenados segundo sua posição angular em relação ao centróide da classe, sem alterar a ordem dos pontos que pertencem a uma mesma dicotomia.

Este método não garante que todos os pontos do espaço de padrões estarão contidos em alguma classe, isto é, pode haver regiões de indeterminação (ver Fig. 3.6), entretanto os espaços interiores das classes são corretamente estabelecidos.

3.1.1 Método rotacional

Este passo do algoritmo seleciona e ordena os exemplares limitantes exteriores de cada classe. Na sequência é descrita sua aplicação para o caso bidimensional, sendo possível sua extensão para espaços de maior dimensão. O processo toma como ponto inicial o elemento da classe mais distante do centróide. Todos os pontos da classe são girados (em torno do centróide) de um ângulo θ de forma a colocar esse ponto sobre a vertical que passa pelo centróide, e tornando sua ordenada maior que a do centróide. Dentre os pontos com abscissa maior que a do centróide, o de maior ordenada é selecionado como próximo elemento e colocado numa lista.

Este processo é repetido para cada novo ponto até voltar ao inicial. Pela descrição, percebe-se que o algoritmo ordena os dados segundo o sentido horário, todavia, se houver uma fronteira de pequena concavidade, ao chegar ao ponto mais interior desta, a seleção avança para o vértice (final) dessa região. Como solução para esta restrição, o algoritmo é efetuado uma vez no sentido horário e outra no sentido anti-horário, e seus resultados são reordenados de forma a gerar a sequência correta.

Outra restrição ocorre quando há fronteiras côncavas tão longas que o centróide situa-se fora do contorno limitante da classe, como se observa na imagem de uma lua crescente ou minguante. Nesse caso, o algoritmo como proposto passaria do início para o final da região côncava desconsiderando todos os elementos intermediários. A ocorrência dessa restrição é identificada por $\theta \geq 90^\circ$, portanto, para esses valores, é usado um algoritmo alternativo para o método rotacional, chamado de método de varredura. Este algoritmo alternativo consiste em: para todo o elemento com ordenada maior que a do centróide e menor que a do último elemento processado pelo método rotacional, determinar a equação da reta que passa por esse elemento e o último selecionado. Para o caso em que a abscissa do centróide é maior que zero, é tomado como próximo elemento aquele cuja reta associada intercepta o eixo das ordenadas com menor valor. Caso contrário (abscissa negativa), é selecionado o elemento cuja intersecção ocorre com maior valor. Este processo pode ser visualizado imaginando-se um pêndulo fixo no último elemento processado, de comprimento inicial igual à distância entre este e o centróide, movendo-se para a esquerda até tocar algum outro elemento, daí originando a denominação "método de varredura".

3.1.2 Pré-classificação

Após a determinação dos elementos limitantes de cada classe, é gerado, no espaço bidimensional, um polígono com vértices nesses pontos, ou um hiperpolígono em espaços de maior dimensão. Na sequência, o seguinte algoritmo é executado:

para cada elemento da amostra

para cada classe i

verificar a localização do ponto a que o elemento está associado no espaço de padrões. Se ele estiver dentro ou no perímetro do polígono definido pelos elementos limitantes da classe i então o elemento pertence à classe i .

Essa verificação é feita usando um algoritmo de ponto no interior de polígonos de Rogers (1988) e outro de ponto sobre o perímetro do polígono.

Devido à forma de geração dos polígonos limitantes das classes, esse procedimento associa todos os elementos da amostra à alguma classe.

Se não houver coincidência entre a classificação obtida e a fornecida com a amostra é tomada uma das ações definidas na Tab. 3.1, senão prossegue o algoritmo calculando as distâncias entre elementos limitantes de classes diferentes.

número de elementos da classe A classificados como sendo de B	número de elementos da classe B classificados como sendo de A	ação a ser tomada
1	0	retirar da amostra o elemento classificado erradamente da classe A
> 1	0	retirar da amostra o elemento da classe B mais próximo do centróide de A
1	1	retirar da amostra o elemento classificado erradamente que estiver mais distante do centróide da sua classe
> 1	1	retirar da amostra o elemento da classe B
> 1	> 1	retirar da amostra o elemento classificado erradamente que estiver mais próximo do centróide da outra classe

Tabela 3.1: ações corretivas a serem tomadas na etapa de pré-classificação

As regras da Tab. 3.1 foram definidas empiricamente a partir de constatações realizadas para cada caso e são uma extensão das utilizadas no sistema S-MURICS proposto em Ichino (1979). Através delas são eliminadas as superposições e os elementos dúbios da amostra, aumentando a confiabilidade do algoritmo.

3.1.3 Geração dos pontos médios

Para determinação da superfície discriminante entre cada par de classes são necessários apenas os elementos de fronteira não oclusos em relação a outra classe.

Para isso, é gerada uma matriz com as distâncias entre os elementos limitantes do par de classes consideradas, e traçada uma reta paramétrica entre os pares de pontos (com $0 \leq t \leq 1$). Sendo p_{ik} o i -ésimo elemento da classe k e p_{jm} o j -ésimo elemento da classe m , para o caso bidimensional:

$$\begin{aligned} y &= a_{ij}t + y(p_{ik}) \\ x &= b_{ij}t + x(p_{ik}) \end{aligned} \quad (3.2a)$$

com

$$\begin{aligned} a_{ij} &= y(p_{jm}) - y(p_{ik}) \\ b_{ij} &= x(p_{jm}) - x(p_{ik}) \end{aligned} \quad (3.2b)$$

onde $k \neq m$, e $y(p)$ e $x(p)$ representam a operação de extração da ordenada e abscissa do ponto p , respectivamente. A extensão para o caso geral é direta.

A seguir são calculadas as intersecções dessa reta paramétrica com os hiperplanos que contém as faces hiperpoligonais da distribuição espacial das duas classes. Se houver intersecção para valores do parâmetro t entre $0 < t < 1$ sobre os hiperpolígonos, um dos pontos é ocluso, e a distância entre os pontos é arbitrada como infinita.

Para o caso bidimensional ($n=2$), o teste pode ser realizado da seguinte forma: são calculadas as retas paramétricas que ligam os elementos de fronteira consecutivos de cada classe, isto é :

$$\begin{aligned} y &= c_{ik}s + y(p_{ik}) \\ x &= d_{ik}s + x(p_{ik}) \end{aligned} \quad (3.3a)$$

com

$$\begin{aligned} c_{ik} &= y(p_{(i+1)k}) - y(p_{ik}) \\ d_{ik} &= x(p_{(i+1)k}) - x(p_{ik}) \end{aligned} \quad (3.3b)$$

sendo p_{ik} o i -ésimo elemento da classe k , e são determinados os valores dos parâmetros t e s das intersecções das retas definidas pelas Eqs. (3.2) com as retas dadas pelas Eqs. (3.3) da classe k e da classe m . Se ocorrer $0 < t < 1$ e $0 < s < 1$ para qualquer das intersecções, então um dos pontos é ocluso e a distância entre eles é definida como infinita.

Após isso, as distâncias são ordenadas de forma crescente numa lista, desprezando aquelas entre elementos já usados em distâncias menores, e são calculados os pontos médios entre os elementos desta lista.

3.1.4 Classificação

A etapa de classificação é necessária no processo de determinação das superfícies discriminantes para eliminar os pontos médios que não pertencem às mesmas. Seja o caso de três classes circulares com desvio padrão igual a 0.1 e centróides sobre o eixo das abscissas, tendo coordenadas $C1=(1,0)$, $C2=(2,0)$ e $C3=(3,0)$. A superfície discriminante da classe 1 deve passar pelo ponto médio entre os centróides das classes 1 e 2, portanto o ponto médio entre as classes 1 e 3 deve ser eliminado.

As superfícies discriminantes separam o espaço em duas regiões. A avaliação da sua função matemática para os pontos sobre a superfície resulta zero, e nos outros resulta em valores maiores ou menores que zero.

A partir do elemento da classe que gerou o ponto médio por onde a superfície discriminante passa, é avaliado o resultado da função e determinado o sinal (positivo ou negativo) atribuído para os elementos da classe, e contra esse sinal devem ser comparados os resultados da avaliação de elementos desconhecidos.

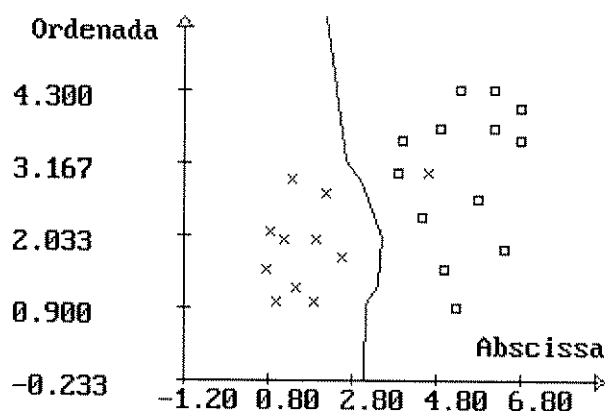


Figura 3.3: resultado da aplicação do algoritmo em dados sintéticos.

No caso de uma superfície linear por partes, deve ser determinado contra qual componente da superfície discriminante o elemento deve ser testado, ou seja, no intervalo de validade de qual das componentes o elemento desconhecido se situa. Para isso usa-se o conceito de superfícies de partição.

As superfícies de partição têm a mesma dimensão das superfícies discriminantes e determinam a região de validade de cada uma das componentes. São definidas como a bissetriz do ângulo formado entre cada par consecutivo de superfícies de decisão, ou seja, para cada superfícies de decisão há duas de partição, exceto para a primeira e última superfícies de decisão (se essas não se unirem formando um "cluster" fechado), porque elas são ilimitadas. A primeira superfície de partição de uma superfície discriminante e a segunda da superfície discriminante anterior coincidem. As avaliações das duas superfícies de partição para um ponto resultam em valores não negativos se o ponto está no intervalo de validade da superfície discriminante correspondente. Dessa forma, um elemento desconhecido somente é testado contra a superfície de decisão cujas superfícies de partição o aceitam.

Se mais de um par de superfícies de partição aceita um novo elemento, e suas superfícies de decisão fornecem classificações contraditórias, então o elemento está ocluso por outras superfícies discriminantes. Nesse caso é necessário realizar o teste de oclusão para todas as superfícies de decisão cujas superfícies de partição aceitam o ponto, e rejeitar a classificação daquelas que apresentam oclusão. Isso é determinado calculando-se as equações paramétricas da reta que parte do elemento desconhecido e alcança o ponto médio da superfície discriminante sob teste com o valor do parâmetro igual a um, e verificando se há intersecção com outras superfícies de decisão nos seus respectivos intervalos de validade com $0 < t < 1$. Isto é :

$$x_i = x_i(p) + a_{im}t \quad (3.4a)$$

com

$$a_{im} = \frac{\sum (x_i(pf_m) + x_i(pi_m))}{2(n-1)} - x_i(p) \quad (3.4b)$$

onde $x_i(pf_m)$ representa a coordenada i do ponto final da superfície de decisão m , e $x_i(pi_m)$ representa a coordenada i do ponto inicial da mesma superfície no caso da

dimensão do vetor de características ser igual a 2. Para dimensão maior, os pontos iniciais e finais formam um hiperpolígono delimitante da superfície discriminante. Se m não for a primeira nem a última superfície de decisão é utilizada a Eq. (3.4b); se for, é utilizado um ponto qualquer da superfície (que esteja no interior do seu intervalo de validade) no lugar do ponto médio que é calculado na primeira parcela da Eq. (3.4b).

O valor do parâmetro para o ponto de intersecção da reta definida pelas Eqs. (3.4) com a superfície discriminante k é obtido por:

$$t = \frac{-a_{0k} - \sum_{i=1}^n a_{ik} x_i(p)}{\sum_{i=1}^n a_{ik} a_{im}} \quad (3.5)$$

com $1 \leq k \leq n^\circ$ de superfícies de decisão, e $k \neq m$.

Se houver intersecção da reta definida nas Eqs. (3.4) com a superfície k para valores de t no intervalo $0 < t < 1$ e no intervalo de validade definido pelos pontos limitantes dessa superfície, então a classificação fornecida por m não é válida.

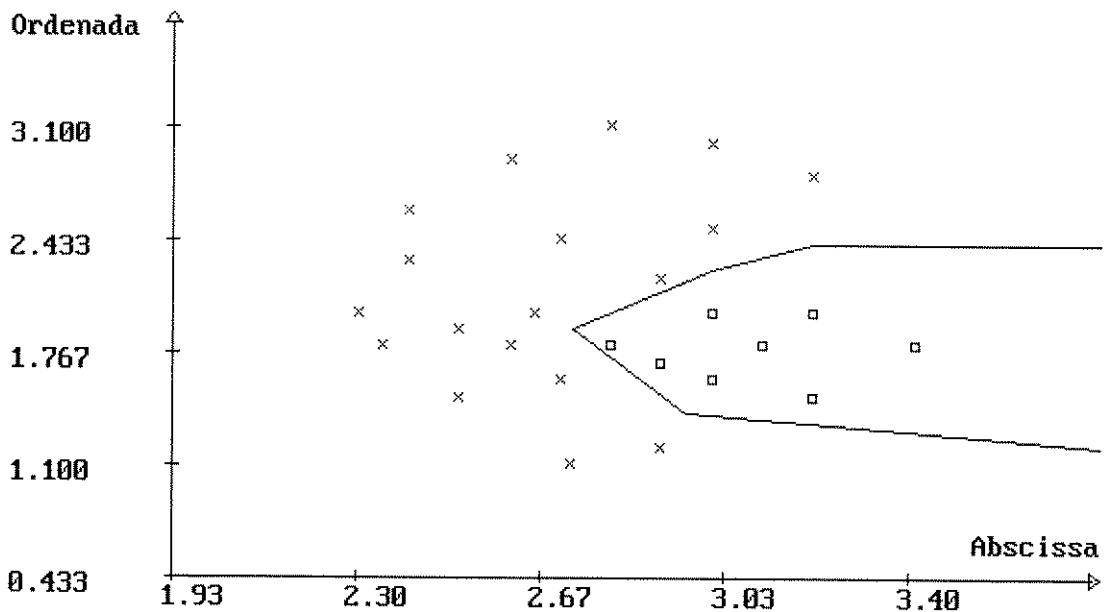


Figura 3.4: resultado obtido a partir de dados sintéticos formando um "cluster" convexo inserido num côncavo

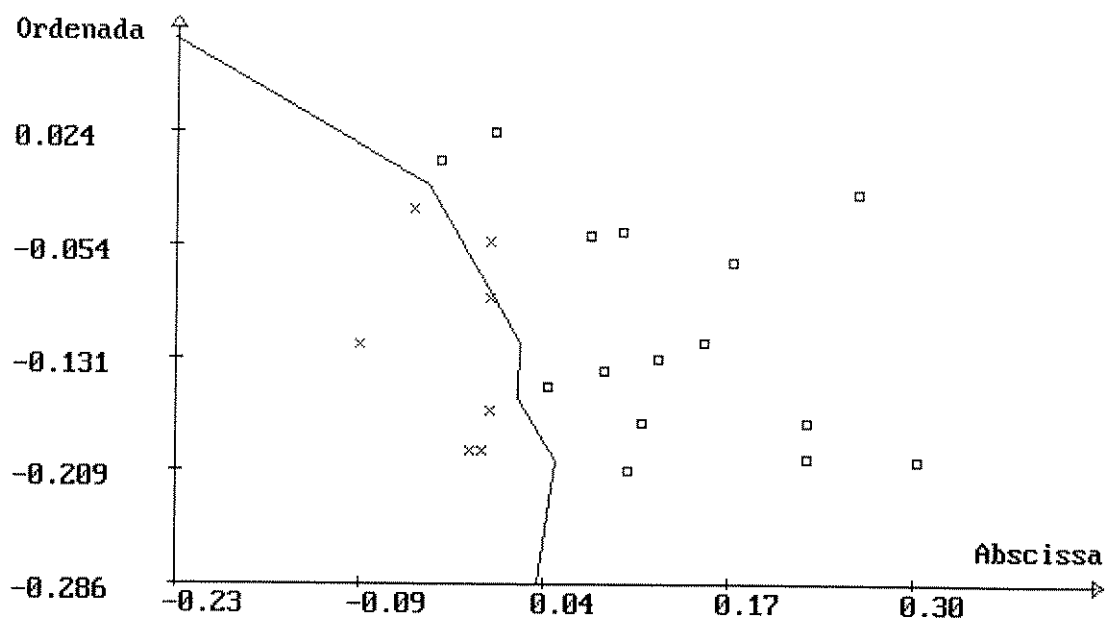


Figura 3.5: resultado obtido a partir de dados reais: descritores de Fourier de folhas das espécies *lonchocarpus guilleminianus* (letras X) e *lonchocarpus muehlbergianus* (quadrados).

3.2 Resultados

O algoritmo descrito foi implementado em linguagem C para o caso bidimensional. Dados sintéticos e reais foram considerados na avaliação do seu desempenho. As Figs. 3.3 a 3.6 mostram resultados da separação entre classes. No caso da Fig. 3.3 foi fornecida propositalmente errada a classificação de um elemento da amostra (há um X no interior da classe dos quadrados). Esse erro foi detectado pelo algoritmo, que eliminou esse elemento da amostra para efeito de determinação da superfície discriminante. Com isso, dentre os 24 elementos da amostra, somente este elemento foi classificado erradamente, ou seja, o índice de erros foi de 4,17%. Para os dados mostrados na Fig. 3.4 não houve erro de classificação. Na Fig. 3.5 pode ser observada a classificação errada de um único elemento dentre 22, o que representa um percentual de erros igual a 4,55%. Esse elemento não foi corretamente identificado porque está dentro da região determinada pelos exemplares limitantes exteriores da classe dos quadrados, portanto foi retirado do processo de geração das funções discriminantes (segundo ação determinada pela Tab. 1) na etapa de pré-classificação. Para os dados apresentados na Fig. 3.6 houve dois erros em 35 amostras, tendo o algoritmo apresentado um índice de erro igual a 5,71%. Uma

classificação intuitiva (proveniente da pergunta: dentro de qual "cluster" do espaço de padrões eles se situam ?) desses dois elementos mal rotulados coincide com a realizada pelo algoritmo. Para o caso mostrado nesta última figura ocorreu uma pequena região triangular de indeterminação.

Da análise realizada, foi observada para todos esses casos que a percentagem de erros de classificação foi baixa, demonstrando a eficiência do método proposto.

Apesar de seus vários passos, sub-algoritmos, uso de números em ponto flutuante com precisão dupla e de ter sido efetuado num computador lento para os padrões atuais (IBM PC 80386 com coprocessador aritmético), a execução do algoritmo foi efetuada em menos de 1 segundo para todos os exemplos mostrados, o que possibilita inclusive o refinamento das superfícies discriminantes após a adição de cada novo elemento classificado.

3.3 Comparações

Nas Tabs. 3.2 a 3.5 podem ser vistos e comparados os índices de erros de quatro algoritmos de classificação aplicados aos dados mostrados nas Figs. 3.3 a 3.6.

método	Ta	Tc	K	nº de erros	percentagem de erros
	1	0	-	0	0 %
Ichino	1	1	-	14	58,33 %
	2	0	-	0	0 %
	2	1	-	4	16,67 %
K-means	-	-	número de classes	1	4,17 %
K-NN	-	-	1	3	12,5 %
	-	-	2	3	12,5 %
	-	-	3	1	4,17 %
	-	-	4	1	4,17 %
	-	-	5	1	4,17 %
proposto	-	-	-	1	4,17 %

Tabela 3.2: índices de erros de quatro algoritmos para os dados mostrados na Fig. 3.3 (24 elementos considerados)

método	Ta	Tc	K	nº de erros	percentagem de erros
Ichino	1	0	-	0	0 %
	1	1	-	21	80,77 %
	2	0	-	0	0 %
	2	1	-	20	76,92 %
K-means	-	-	número de classes	8	30,77 %
K-NN	-	-	1	2	7,69 %
	-	-	2	2	7,69 %
	-	-	3	3	11,54 %
	-	-	4	2	7,69 %
	-	-	5	2	7,69 %
proposto	-	-	-	0	0 %

Tabela 3.3: índices de erros de quatro algoritmos para os dados mostrados na Fig. 3.4 (26 elementos considerados)

método	Ta	Tc	K	nº de erros	percentagem de erros
Ichino	1	0	-	0	0 %
	1	1	-	17	77,273 %
	2	0	-	0	0 %
	2	1	-	0	0 %
K-means	-	-	número de classes	5	22,73 %
K-NN	-	-	1	2	9,09 %
	-	-	2	2	9,09 %
	-	-	3	5	22,73 %
	-	-	4	3	13,64 %
	-	-	5	3	13,64 %
proposto	-	-	-	1	4,55 %

Tabela 3.4: índices de erros de quatro algoritmos para os dados mostrados na Fig. 3.5 (22 elementos considerados)

O algoritmo de Ichino gera os "clusters" com a forma de hiper-retângulos (no espaço de padrões n-dimensional, e retângulos para n=2) que contêm os elementos do conjunto de amostras. O método pode produzir mais de um hiper-retângulo para cada classe e o resultado do algoritmo é dependente da ordem com que os elementos da amostra são fornecidos e de dois parâmetros: Ta e Tc. O primeiro parâmetro controla a geração e crescimento dos hiper-retângulos, e o segundo indica o número máximo de componentes do vetor de características de um elemento que podem situar-se fora

dos hiper-retângulos de uma classe para que o elemento pertença a esta classe. O parâmetro T_a pode variar dentro dos limites $1 < T_a \leq n$ e o T_c entre $0 \leq T_c < n$, devendo os valores ótimos serem escolhidos manualmente após a execução do algoritmo com todas as combinações possíveis. A escolha $T_c=0$ limita a capacidade de abstração que o algoritmo deve extrair do conjunto de amostras, pois esse valor para T_c somente permite que um elemento desconhecido seja identificado se ele estiver no interior de algum dos hiper-retângulos de alguma classe (e somente de uma). Como os hiper-retângulos não abrangem necessariamente todo o espaço de padrões (ver Figs. A-14 a A-17 no apêndice A), haverá muitos intervalos de indeterminação, apesar das respostas a essas indeterminações serem óbvias para uma classificação intuitiva, isto é, a resposta à pergunta: dentro de qual "cluster" do espaço de padrões esses pontos se situam ?

método	T_a	T_c	K	nº de erros	percentagem de erros
Ichino	1	0	-	0	0 %
	1	1	-	21	60 %
	2	0	-	0	0 %
	2	1	-	0	0 %
K-means	-	-	número de classes	8	22,86 %
K-NN	-	-	1	9	25,71 %
	-	-	2	9	25,71 %
	-	-	3	8	22,86 %
	-	-	4	7	20,00 %
	-	-	5	7	20,00 %
proposto	-	-	-	2	5,71 %

Tabela 3.5: índices de erros de quatro algoritmos para os dados mostrados na Fig. 3.6 (35 elementos considerados)

Além disso, o algoritmo de Ichino (1979) possui a tendência a gerar hiper-retângulos a partir de um único elemento, isto é, o intervalo de variação de cada componente do espaço de padrões é zero, e o volume do hiper-retângulo é nulo. Por essas características o algoritmo de Ichino apresenta bons resultados de classificação para o conjunto de amostras fornecido (quase sempre 0% de erros, conforme indicado nas Tabs. 3.2 a 3.5), mas possui um desempenho fraco quando utilizado no reconhecimento de elementos novos.

Como também mostrado nas Tabs. 3.2 a 2.5, o algoritmo proposto neste capítulo apresenta um desempenho igual (para o caso trivial esboçado na Fig. 12,

onde a distância entre classes é maior que a intra-classes), ou melhor ao dos métodos K-means e K-NN. O algoritmo proposto produz menos erros de classificação que o K-NN devido à sua capacidade de rejeição da classificação fornecida com as amostras (exercida na etapa de pré-classificação). Quanto ao algoritmo de Ichino, uma simples avaliação visual dos seus hiper-retângulos (mostrados nas Figs. A-14 a A-17) prova que o ótimo índice de erros do método é enganoso quando este é aplicado ao próprio conjunto de amostras, e que seria bastante maior ao do método proposto neste capítulo na classificação de elementos diferentes dos fornecidos na amostra de treinamento (pelos motivos já citados nos parágrafos anteriores).

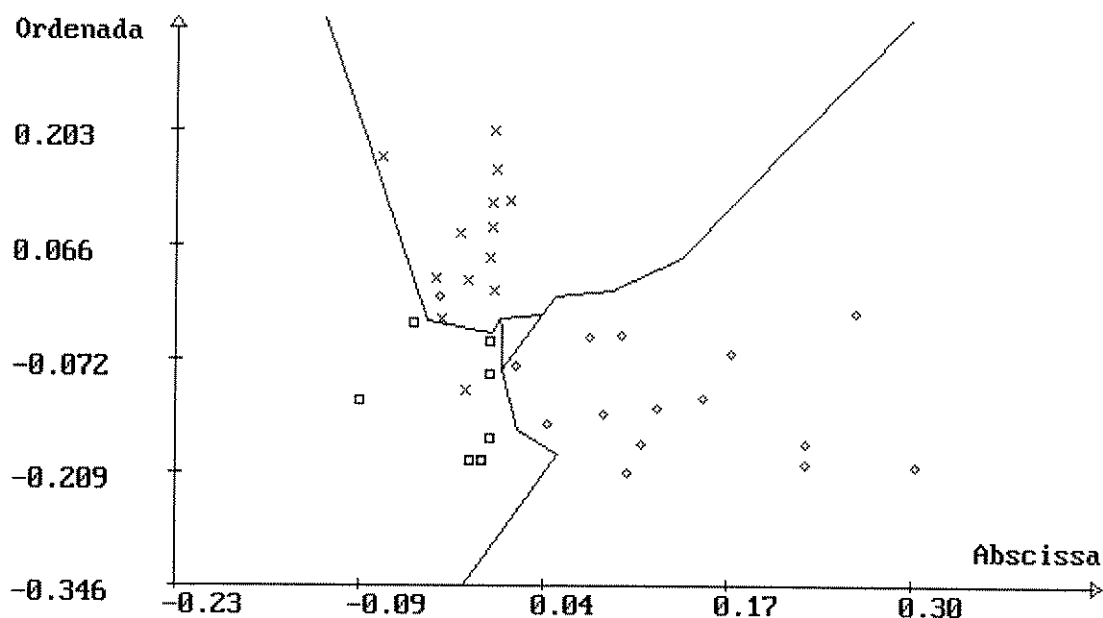


Figura 3.6: resultado obtido a partir de dados reais: descritores de Fourier de folhas das espécies *lonchocarpus campestris* (letras X), *lonchocarpus guilleminianus* (quadrados) e *lonchocarpus muehlbergianus* (losangos).

3.4 Conclusão

Os resultados obtidos e a comparação realizada na seção anterior mostram a validade, rapidez e eficiência do algoritmo proposto. Este algoritmo não depende da seleção experimental de valores de limiar, como o apresentado em Ichino (1979), ou da pré-definição do número de componentes das funções discriminantes lineares por partes nem requer um longo treinamento para definição dos hiperplanos como o método proposto em Foroutan e Sklansky (1987). Também não apresenta uma

convergência lenta e dependente de parâmetros arbitrados como ocorre em Spivak (1989).

Abordagem semelhante à apresentada neste capítulo é encontrada no caso de redes neurais, que também podem ser usadas para classificação. Os perceptrons que compõem a rede podem ser representados por uma função discriminante linear seguida de uma binarização ou função sigmóide, portanto pode-se obter as superfícies de decisão a partir de redes neurais, como demonstrado em Bobrowski (1991). Entretanto sua convergência é normalmente lenta e dependente do número de neurônios e arquitetura utilizada.

CAPÍTULO 4

Aplicação

Para testar a eficácia dos métodos de reconhecimento estudados, isto é, a representação de objetos por descritores de Fourier e a identificação através de funções discriminantes lineares por partes geradas pelo algoritmo proposto, esses métodos foram aplicados à classificação de quatro espécies de plantas do gênero *lonchocarpus*: *lonchocarpus campestris*, *lonchocarpus guilleminianus*, *lonchocarpus latifolius* e *lonchocarpus muehlbergianus* a partir da forma de seus folíolos. Foram utilizadas as definições de folha e folíolos ilustradas na Fig. 4.1. Todos os descritores especificados no capítulo 2 foram aplicados na representação dos folíolos basais, sub-terminais e terminais das espécies citadas, sendo almejada a identificação da espécie a partir dos descritores do seu contorno sabendo qual posição o folíolo ocupa na folha, isto é, foram feitas análises isoladas para os três tipos de folíolos.

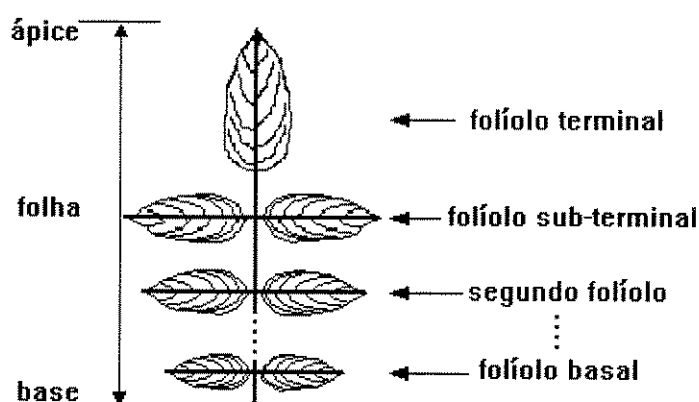


Figura 4.1: ilustração dos conceitos de folha e folíolo

No apêndice D podem ser observadas amostras de folíolos das quatro espécies de plantas utilizadas.

As folhas propiciam um teste difícil ao processo de descrição e reconhecimento porque as amostras não são exatamente iguais, como no caso de objetos construídos em série pelo homem, e podem apresentar uma grande variação intra-classe devido às diferentes condições de crescimento, alimentação e competição. Por pertencerem ao mesmo gênero, estas espécies devem possuir alguma similaridade entre si, o que significa que a distância entre classes no espaço de padrões não deve ser tão grande que permita uma separação óbvia. Isto é, não deve haver uma variação intra-classe tão grande que torne nula a distância entre classes, porque isto inviabiliza a separação das distribuições no espaço formado pelos vetores de características, nem muito pequena porque as amostras não são objetos idênticos.

4.1 Obtenção das imagens

Foram digitalizadas por um "scanner" amostras dos folíolos das espécies citadas distribuídas na seguinte forma: 13 *campestris* basais, 12 *campestris* sub-terminais, 13 *campestris* terminais, 8 *guilleminianus* basais, 9 *guilleminianus* sub-terminais, 7 *guilleminianus* terminais, 6 *latifolius* basais, 6 *latifolius* sub-terminais, 6 *latifolius* terminais, 4 *muehlbergianus* basais, 5 *muehlbergianus* sub-terminais e 15 *muehlbergianus* terminais. As imagens dos folíolos foram geradas com 256 tons de cinza numa resolução de 250 pontos por polegada. Estas imagens apresentavam dimensões variadas (dependendo do tamanho dos folíolos), mas seus histogramas eram tipicamente bimodais, isto é, possuíam duas concentrações facilmente distinguíveis. Uma centrada num valor baixo e devida ao folíolo propriamente dito, e a outra centrada num valor alto e proveniente do fundo claro da imagem. A binarização das imagens foi realizada através do próprio "software" do "scanner" e quase sempre com o valor de limiar indicado por esse. Após a binarização, os pontos da imagem que pertencem à folha apresentam valor 0, e os do fundo 1. Neste trabalho foi definido que o fundo é identificado pelo valor 0, portanto todas as imagens foram invertidas para o processamento subsequente. Como padronização, todos os folíolos foram digitalizados com a base na parte inferior da imagem, e o peciólulo (haste que une o folíolo ao resto da folha) foi cortado segundo o prolongamento da curva da base antes da digitalização. As imagens foram gravadas no formato "windows bitmap" sem compressão (BMP) e todo o processamento foi realizado por programas escritos em linguagem C.

4.1.1 Rastreamento do contorno

Foi extraído um contorno 8-conectado das imagens binárias (com somente dois valores) usando um algoritmo de rastreamento de curvas com inércia (Sonka et al. (1993), Ferson et al. (1985)). Esse algoritmo parte de um dado ponto inicial pertencente ao contorno e, para cada "pixel", procura (no sentido anti-horário) numa vizinhança de 3x3 "pixels" centrada no inicial, o último (no sentido da varredura) "pixel" de valor unitário (isto é, que pertence ao objeto) num intervalo de $[-180^\circ, 45^\circ]$ em relação a direção anterior de movimento. Armazena as coordenadas do "pixel" presente e move-se para o próximo, atualizando a direção tomada. A inércia polariza a busca do próximo "pixel" em torno da direção tomada pela passagem do "pixel" anterior para o presente. Ela é necessária para evitar que o rastreio siga para o interior do objeto (e se perca) devido à ruídos ou mesmo por traços característicos do objeto, e para que o rastreio não volte sobre si mesmo. O algoritmo termina quando volta ao primeiro ponto. Foi dada arbitrariamente uma inércia inicial para a esquerda no algoritmo.

O ponto inicial do rastreamento é determinado por uma varredura feita na imagem no sentido de baixo para cima e da esquerda para a direita até encontrar o primeiro "pixel" não nulo. Esse "pixel" será o ponto mais baixo do contorno. Devido a isso e à inércia inicial, o contorno do objeto é sempre registrado no sentido horário. Contudo, pode ocorrer que não seja encontrado um "pixel" do contorno, mas algum ruído devido à digitalização ou binarização da imagem. Nesse caso é rastreado o contorno do ruído e desprezado o objeto. Para eliminar esse problema é realizada uma pré-filtragem na imagem antes do rastreamento da borda. Essa filtragem é detalhada na seção seguinte.

4.1.2 Pré-filtragem da imagem

Como mencionado na seção anterior, é necessário retirar os ruídos da imagem, ou pelo menos aqueles abaixo do ponto de menor ordenada da borda do objeto. Isso é feito levando-se em conta o fato de que os ruídos normalmente são pontos ou pequenas aglomerações de pontos nas imagens binárias. Portanto, um filtro que apague um conjunto de pontos de área pequena e deixe intocados os grandes conjuntos (para

não deformar o objeto) seria o ideal. Um filtro com essas características não é linear (por depender da área do objeto), e portanto essa filtragem não pode ser realizada por convolução (que é uma operação linear). Neste trabalho foi utilizada morfologia matemática (Gonzalez e Woods (1993) e Serra (1982)) para efetuar essa filtragem. Uma das soluções possíveis é a aplicação, na seqüência, de uma erosão com um elemento estruturante planar extenso, e uma dilatação geodésica com um elemento estruturante 3x3 composto somente de uns, e tomando como base a imagem original. O elemento estruturante é uma máscara semelhante a da convolução, mas binária nesse caso. Algumas definições básicas utilizadas pela morfologia matemática são dadas a seguir.

A translação de A pelo valor $x=(x_1, x_2)$ é simbolizada por $(A)_x$ e definida como:

$$(A)_x = \{c | c = a + x, \text{ para } a \in A\} \quad (4.1)$$

A reflexão de B é simbolizada por $\tilde{B}$ e dada por:

$$\tilde{B} = \{x | x = -b, b \in B\} \quad (4.2)$$

O complemento de A é definido como:

$$A' = \{x | x \notin A\} \quad (4.3)$$

A erosão é uma função que realiza uma varredura na imagem de entrada A com um elemento estruturante B e, através da operação definida na Eq. (4.4) gera uma outra imagem como saída. Os "pixels" dessa nova imagem têm valor 1 se o elemento estruturante está contido na imagem de entrada tendo como referência a posição do "pixel" que está sendo calculado na nova imagem, e 0 caso contrário. A erosão de A por B é denotada como $A-B$ e pode ser obtida por:

$$A - B = \{x | (B)_x \subseteq A\} \quad (4.4)$$

A dilatação é a operação dual da erosão. Também tem como entrada uma imagem e o elemento estruturante, e a saída é outra imagem, sendo que cada "pixel" da imagem de saída tem valor 1 se a intersecção do elemento estruturante refletido (segundo Eq. 4.2) com a imagem de entrada na posição do "pixel" que está sendo calculado é não nula, e 0 caso contrário. Sendo $\emptyset$ o conjunto vazio, a dilatação de A por B é denotada como $A \oplus B$ e pode ser obtida por:

$$A \oplus B = \{x \mid (\tilde{B})_x \cap A \neq \emptyset\} \quad (4.5)$$

A dilatação geodésica (ou dilatação condicional) tem como entrada duas imagens e o elemento estruturante, e sua saída é outra imagem. Uma das imagens de entrada é chamada de base. O procedimento é semelhante ao da dilatação, mas é feita além da dilatação convencional, uma comparação com o "pixel" da imagem base na mesma posição do que o que está sendo calculado na saída. É feita uma operação lógica E (ou intersecção) entre estes últimos, o que só vai permitir que os "pixels" da saída tenham valor não nulo se eles também o tiverem na imagem base, ou seja, a saída está contida na imagem base. A dilatação geodésica realiza esse processo até a idempotência, isto é, até que uma nova iteração não altere a imagem de saída. A dilatação geodésica de A por B tendo como imagem base C pode ser obtida pela seguinte expressão iterativa:

$$X_k = (X_{k-1} \oplus B) \cap C \quad (4.6)$$

com $k=1,2,3,\dots$ e $X_0=A$

Neste trabalho a erosão é executada com um elemento estruturante 7x7 com todos os "pixels" com valor unitário. Isso elimina todas as aglomerações de "pixels" da imagem que não contenham essa área (os ruídos) e deforma o objeto. A dilatação geodésica executada na sequência reconstrói completamente o objeto (pois supõem-se que a imagem do objeto contenha um quadrado de 7x7 "pixels") mas não pode regenerar os ruídos, pois eles não deixam "sementes" para serem dilatadas. Desta forma, esse processo realiza a filtragem desejada.

Na prática não foi implementada uma dilatação geodésica porque sua convergência é muito lenta (pela própria natureza do algoritmo). É aplicado no seu lugar o algoritmo Blob-coloring modificado e com vizinhança 8-conectada. Esse método é mais rápido porque executa as alterações diretamente na imagem de entrada e não precisa compará-la com um elemento estruturante. Normalmente sua convergência é atingida com duas iterações, mas em casos especiais pode exigir mais. É composto por duas etapas que se alternam. Na primeira é realizada uma varredura na imagem de entrada no sentido de cima para baixo e da esquerda para a direita, sendo ao "pixel" presente atribuído o valor 1 se um dos seu vizinhos (da esquerda, acima e da diagonal acima, que já foram processados) possuir valor 1 e o "pixel" da imagem base da mesma posição também tiver valor 1. Esta operação é representada pela Eq. (4.7).

$$f(x, y) = \{f(x-1, y) + f(x-1, y-1) + f(x, y-1)\} \cdot g(x, y) \quad (4.7)$$

onde $g(x, y)$ é a imagem base, o símbolo $+$ (quando fora do argumento das funções) representa o operador lógico OU e o símbolo $\cdot$ é o operador lógico E

Na segunda etapa é feita a varredura no sentido contrário, isto é, de baixo para cima e da direita para a esquerda, sendo considerados os vizinhos da direita, de baixo e da diagonal para baixo. Sua expressão matemática é :

$$f(x, y) = \{f(x+1, y) + f(x+1, y+1) + f(x, y+1)\} \cdot g(x, y) \quad (4.8)$$

onde $g(x, y)$ é a imagem base, o símbolo $+$ (quando fora do argumento das funções) representa o operador lógico OU e o símbolo $\cdot$ é o operador lógico E

Volta-se à primeira etapa se houve alguma alteração na última, garantindo a idempotência. Por levar em conta as alterações feitas durante a execução da iteração corrente, esse algoritmo converge muito mais rapidamente e seu efeito final é exatamente o mesmo da dilatação geodésica com elemento estruturante 3×3 composto somente por valores unitários.

Para diminuir o tempo gasto, a dilatação geodésica (ou mais exatamente, a aplicação do algoritmo Blob-coloring) é feita somente até uma linha acima do primeiro "pixel" não nulo da imagem erodida, e as linhas acima desta são copiadas da imagem original. Isso pode ser feito porque o que é realmente necessário eliminar são os ruídos abaixo do ponto de menor ordenada do objeto. Os outros são naturalmente desprezados pelo algoritmo de rastreamento de contorno. A erosão também é interrompida de forma semelhante. Essas otimizações rendem uma razoável economia de tempo (dependendo da imagem, pode reduzir para um quinto o tempo consumido nessa operação).

4.2 Determinação dos descritores

Como os folíolos de cada planta não são exatamente iguais (porque existe uma variação intra-classe), não há sentido em comparar os descritores de todas as

harmônicas, mas como a transformada de Fourier decompõem sua função de entrada em funções harmônicas de diferentes velocidades de variação espacial, examinar apenas as harmônicas de menor ordem é equivalente a fazer uma pré-filtragem na função contorno com um filtro passa-baixas ideal, eliminando os ruídos introduzidos pela digitalização, limiarização e as pequenas diferenças que existem entre folíolos da mesma espécie (que são informações contidas nas frequências mais altas por serem detalhes pequenos). Portanto faz-se necessário determinar a frequência de corte a ser utilizada, ou equivalentemente, o número de harmônicas a ser considerado para a representação dos contornos. Uma forma de definição da frequência de corte é a verificação da exatidão da reconstrução parcial que ela induz. Para efeito de classificação deve-se ter uma representação de cada objeto precisa o suficiente para distingui-lo de outros mas não tão exata que diferencie várias imagens do mesmo objeto.

Um método que permite a avaliação da precisão da reconstrução parcial é o cálculo do erro quadrático médio entre o objeto original e sua reconstituição (sendo o erro definido como a distância euclidiana entre cada "pixel" correspondente dos contornos dos objetos). Pelo fato da distância ser medida em "pixels", o erro quadrático médio também é expresso nessa unidade. E por ser uma medida de erro absoluto, seu valor varia em função das dimensões da imagem (e do objeto), de forma que não pode ser utilizado como termo de comparação se aplicado a imagens de tamanhos muito diferentes. Para que sejam medidas válidas de comparação estes erros devem ser expressos numa forma relativa, isto é, dependente do tamanho da imagem (ou objeto). Um método de descrição de tamanho é obtido pelas distâncias características (Thiran et al. (1994)) e sua média. As distâncias características são as distâncias entre o centro de gravidade do objeto e seu contorno segundo várias direções, e a partir delas pode ser definida uma distância média (ou raio médio) que pode ser utilizada como valor de referência para tamanho do objeto. Expressando-se o erro quadrático médio como percentagem do raio médio tem-se um valor para a exatidão da reconstrução que é independente do tamanho das imagens. Esta fórmula reduz-se ao erro quadrático médio normalizado se forem utilizados todos os elementos do contorno no cálculo do raio médio (o que foi feito).

4.2.1 Determinação dos descritores do contorno completo

Inicialmente foram calculados os descritores baseados na transformada discreta de Fourier do contorno completo dos folíolos. Como neste trabalho os contornos foram super-amostrados e a DFT tem a propriedade de grande compactação de energia nas baixas frequências, o valor das componentes de alta frequência (que a convolução somaria as baixas para poder comparar as DFTs) é muito pequeno (isto foi verificado experimentalmente e também pode ser percebido intuitivamente sabendo que os objetos utilizados têm formas bastante elípticas), e a $F[k]$ é aproximadamente igual à $G[k]$ (definidas na seção 2.3.3.5) para as harmônicas $0 \leq k < N/a = N'$ exceto por uma constante multiplicativa (o quadrado da relação entre os comprimentos $1/(a*a)$). Portanto, a convolução pode até ser desprezada, sendo necessária apenas uma normalização de ganho, o que foi feito dividindo todas as harmônicas pelo módulo da primeira, tornando esta unitária para todos os objetos e obtendo uma aproximação da invariância à escala. Para determinar o número de harmônicas a ser usado no reconhecimento foi calculado o erro quadrático médio entre a reconstrução usando somente as primeiras harmônicas (positivas e negativas) e o contorno original (ou reconstrução completa) para todas as amostras. Os valores médios dos erros quadráticos médios, assim como a média dos erros percentuais (medidos em relação ao raio médio de cada objeto) podem ser vistos na Tab. 4.1.

harmônicas usadas	média dos erros quadráticos médios (em "pixels")	média dos erros percentuais
1	13,9	10,1 %
1 2	13,0	9,5 %
1 2 3	5,6	4,1 %
1 2 3 4	5,1	3,7 %
1 2 3 4 5	3,2	2,3 %
1 2 3 4 5 6	3,0	2,2 %
1 2 3 4 5 6 7	2,1	1,6 %
1 2 3 4 5 6 7 8	2,1	1,5 %

Tabela 4.1: erros das reconstruções parciais do contorno completo dos folíolos

Analisando esses dados e levando em conta a relação custo/benefício, decidiu-se utilizar as primeiras cinco harmônicas, e foram aplicados os descritores de Fourier definidos na seção 2.3 e os descritores elípticos especificados na seção 2.4. Para reduzir a dimensionalidade do vetor de parâmetros para o reconhecimento foi aplicada

a transformada de Hotelling (na forma definida no apêndice C) a cada conjunto de descritores isoladamente, e tomadas somente as duas primeiras componentes desse novo espaço. Isso pode ser efetuado porque a transformada de Hotelling é ótima no sentido de minimização do erro quadrático médio das reconstruções parciais (ver demonstração em Gonzalez e Woods (1993) e Jain (1989)), isto é, apresenta a melhor compactação de energia nas componentes baixas, sendo comumente usada na compressão de dados e porque a razão entre o maior autovalor e os outros da transformadas de Hotelling é elevada (ver Tab. 4.2), o que significa que a primeira componente possui muito mais informação que as outras, e o erro quadrático médio introduzido por essa eliminação é pequeno se comparado ao erro que resulta do desprezo somente da primeira componente (ver apêndice C).

harmônicas usadas	autovalores em ordem decrescente (multiplicados por 10^6)				
1 2 3 4 5	3434	229	133	32	19
	12	8	6	1	-

Tabela 4.2: autovalores da transf. de Hotelling dos descritores de Fourier dos folíolos

Foram utilizadas somente as duas primeiras componentes da transformada de Hotelling para que fossem gerados gráficos bidimensionais das distribuições dos descritores, porém as distribuições das quatro espécies (para os folíolos terminais, sub-terminais e basais) apresentam-se sobrepostas nesse espaço de padrões, assim como no espaço dos descritores (ver Figs. 4.2 a 4.6).

Isso ocorre porque essas espécies não se distinguem pelo formato geral dos folíolos, mas pelos formatos da base e ápice desses. Portanto os descritores devem ser aplicados a essas extremidades dos folíolos para que seja possível a classificação.

4.2.2 Determinação dos descritores da base e ápice

A base e o ápice foram definidos como o intervalo do contorno centrado nos pontos de maior distância ao centróide do objeto (componente DC da transformada dividida por N) e de comprimento igual a um quarto do perímetro do contorno do folíolo (segundo Hickey (1973)), mas essas curvas são abertas e, se fosse aplicada diretamente a transformada de Fourier nelas, devido à descontinuidade entre o ponto inicial e o final, surgiria o fenômeno de Gibbs, deformando a representação da curva.

A solução usualmente aceita para esse fato é o retraço da curva sobre seus próprios pontos até o início, mas isso gera um objeto de espessura (e área) nula, cuja interpretação intuitiva da decomposição em elipses é dificultada. Em lugar disso foi realizada uma reflexão de cada ponto da curva em torno do ponto médio entre o primeiro e o último pontos da curva aberta. Esse processo gera um contorno fechado com uma simetria tal que todas as suas harmônicas pares (exceto a zero, que somente possui informação sobre o centróide do objeto) são nulas (ver demonstração no apêndice B), e o custo do cálculo dos novos pontos será de apenas uma subtração de números complexos por ponto, mais a soma complexa dos pontos inicial e final, isto é, a partir da curva aberta $f[n]$ é definida uma nova função contorno $f'[n]$ dada por:

$$f'[n] = \begin{cases} f[n] & \text{para } 0 \leq n < N \\ c - f[n - N] & \text{para } N \leq n < 2N \end{cases} \quad (4.9a)$$

sendo $c = f[0] + f[N - 1]$ (4.9b)

e $f[n] = x[n] + jy[n]$ para $0 \leq n < N$ (4.9c)

Com isso, somente é necessário utilizar (e calcular) as harmônicas pares da DFT de $f'[n]$.

As Tabs. 4.3 e 4.4 mostram as médias dos erros quadráticos médios e percentuais das reconstruções parciais das funções base e ápice simetrizadas (definidas pelas Eqs. 4.9) a partir dos seus espectros de Fourier.

harmônicas usadas	média dos erros quadráticos médios (em "pixels")	média dos erros percentuais
1	7,1	10,0 %
1 3	3,1	4,3 %
1 3 5	1,7	2,3 %
1 3 5 7	1,1	1,6 %
1 3 5 7 9	0,8	1,1 %
1 3 5 7 9 11	0,6	0,9 %

Tabela 4.3: erros das reconstruções parciais da base simetrizada dos folíolos

Analisando as Tabs. 4.3 e 4.4, decidiu-se calcular os descritores de Fourier definidos na seção 2.3 para as primeiras cinco harmônicas ímpares (pois o ganho

obtido pelo uso de mais frequências não compensaria seu custo), o que acarreta num erro percentual de reconstrução em torno de 1%. Foi utilizada esta versão dos descritores de Fourier porque apresenta menor custo computacional apesar das suas ambigüidades de representação (indicadas na seção 2.3.3.6). Novamente foi aplicada a transformada de Hotelling aos descritores da base e ápice isoladamente, e tomada somente a primeira componente de cada uma dessas transformadas, o que resulta num vetor de características bidimensional por folíolo. A transformada de Hotelling foi aplicada à cinco combinações dos descritores segundo a ordem crescente das harmônicas ímpares, isto é, aos descritores somente da 1ª harmônica, aos descritores da 1ª e 3ª, da 1ª, 3ª e 5ª, da 1ª, 3ª, 5ª e 7ª, e aos da 1ª, 3ª, 5ª, 7ª e 9ª harmônicas para avaliação da variação do erro de classificação com o exatidão da reconstrução. Foram utilizadas as frequências positivas e negativas dessas harmônicas, mas a frequência positiva da 1ª harmônica foi utilizada para a normalização, sendo tornada igual para todos os folíolos. Somente foram consideradas a primeira componente da transformada de Hotelling do ápice e da base, de forma a gerar um espaço de padrões bidimensional. Isso pode ser efetuado porque a razão entre o maior autovalor e o segundo maior autovalor de cada uma das transformadas de Hotelling é elevada, o que significa que a primeira componente possui muito mais informação que as outras (ver apêndice C). As Tabs. 4.5 e 4.6 mostram os autovalores das transformadas de Hotelling em função do número de harmônicas utilizadas.

harmônicas usadas	média dos erros quadráticos médios (em "pixels")	média dos erros percentuais
1	9,6	14,0 %
1 3	3,5	5,1 %
1 3 5	1,9	2,7 %
1 3 5 7	1,3	1,9 %
1 3 5 7 9	0,9	1,4 %
1 3 5 7 9 11	0,8	1,1 %

Tabela 4.4: erros das reconstruções parciais do ápice simetrizado dos folíolos

harmônicas usadas	autovalores em ordem decrescente (multiplicados por 10 ⁶)								
1 3	3222	165	35						
1 3 5	3225	167	57	26	4				
1 3 5 7	3229	168	57	27	9	3	2		
1 3 5 7 9	3230	170	57	28	11	6	3	2	1

Tabela 4.5: autovalores da transf. de Hotelling dos descritores de Fourier da base simetrizada dos folíolos

De forma semelhante a ocorrida para os descritores do contorno completo, as distribuições dos descritores da base e ápice se sobrepõem no espaço de padrões para os folíolos basal e sub-terminal. Mas, para os folíolos terminais pode-se definir três classes distintas, uma composta pelas espécies *campestris* e *latifolius*, outra somente com as amostras da espécie *guilleminianus* e a última com os folíolos de plantas *muehlbergianus*. A delimitação dessas classes foi realizada usando funções discriminantes lineares por partes calculadas pelo algoritmo descrito no capítulo 3 e em Silva e Tozzi (1996). Os gráficos com as funções discriminantes lineares por partes para as 5 combinações de harmônicas podem ser observados nas Figs. A-4 a A-8 no apêndice A, e neles percebe-se que a taxa de erros de classificação é a mesma para as combinações de duas ou mais harmônicas, sendo levemente superior quando somente a primeira harmônica é utilizada. Isto é, os descritores da 3ª harmônica e da frequência negativa da 1ª são suficientes para a representação e separação das espécies de folíolos consideradas com o mesmo índice de erros que ocorre para descrições mais precisas (que utilizam mais harmônicas). A distribuição espacial dos descritores da espécie *lonchocarpus latifolius* está completamente sobreposta na distribuição da espécie *lonchocarpus campestris* para todas as combinações (ver Figs. 4.7 a 4.11), por isso as amostras dessas duas espécies foram consideradas como pertencentes a uma mesma classe nos gráficos que mostram a separação (Figs. A-4 a A-8). Essa sobreposição deve-se ao fato dessas espécies se diferenciarem pelo tamanho (os folíolos da espécie *latifolius* são muito maiores que os da espécie *campestris*), e como os descritores são intencionalmente invariantes à escala (e conseqüentemente à resolução), não puderam absorver essa característica.

harmônicas usadas	autovalores em ordem decrescente (multiplicados por 10 ⁶)								
1 3	12906	290	91						
1 3 5	13041	380	100	77	21				
1 3 5 7	13045	387	102	78	21	9	4		
1 3 5 7 9	13047	387	103	78	23	10	5	5	1

Tabela 4.6: autovalores da transf. de Hotelling dos descritores de Fourier do ápice simetrizado dos folíolos

Portanto, para diferenciar folíolos *campestris* de *latifolius* é necessário incorporar mais um passo no processo de identificação, onde as amostras que se situam no "cluster" das *lonchocarpus campestris* e *latifolius* têm seu tamanho avaliado e, a partir de certo valor limite, são reconhecidas como *lonchocarpus*

latifolius. Na Fig. 4.12 estão as distribuições das espécies separadas pelas superfícies discriminantes lineares por partes geradas pelo algoritmo descrito no capítulo 3 (somente as harmônicas 1 e 3), onde percebe-se que o índice de erros foi de 4,88% para os dados utilizados. Na identificação usando somente a primeira harmônica (ver Fig. A-8) essa taxa de erros aumenta para 7,32% no algoritmo utilizado. É necessário calcular os descritores da base e do ápice e considerar ambos para diferenciar essas espécies. Isto pode ser observado pelos descritores da harmônica N-1 das Fig. 4.7 e 4.9 ou pelo gráfico bidimensional da Fig. 4.12. Nesta última percebe-se que uma reta vertical, que portanto só considera os descritores da base, separa a espécie *lonchocarpus muehlbergianus* das outras. Uma reta horizontal, agora somente utilizando informação do ápice, permite a diferenciação entre a classe formada pelas *lonchocarpus campestris e latifolius* e as outras espécies. Essas retas correspondem a valores de limiar nos histogramas dos descritores da base e ápice (Figs. 4.7 e 4.9).

O processo completo de identificação de um folíolo é apresentado no fluxograma da Fig. 4.15 (utilizando descritores de Fourier, no caso da aplicação de outro tipo de descritor basta substituir a DFT pela transformada correspondente e seu descritor). E na Fig. 4.16 são mostradas as imagens resultantes dos passos intermediários do algoritmo de reconhecimento para um folíolo.

Outra razão para não terem sido calculados os descritores elípticos de Fourier para a base e ápice é sugerida pelas Figs. 4.7 a 4.10. Nelas percebe-se que o fator discriminatório situa-se somente na 1ª harmônica, sendo que as outras contêm informação de variação intra-classe e ruído. Outra forma de verificação desse fato é obtida pela observação que o índice de erros de identificação não varia entre as Figs. A-4 a A-7, que utilizam quatro combinações sucessivamente menores de harmônicas.

Para os descritores de Fourier a frequência positiva da 1ª harmônica é tornada unitária para todas as amostras pela normalização, sendo sua informação perdida, mas a capacidade identificatória da frequência negativa é preservada. Para os descritores elípticos definidos na seção 2.4 são utilizadas a frequência positiva e negativa da 1ª harmônica para normalização, sendo destruída, portanto, sua informação e frustrada qualquer tentativa de classificação a partir desse tipo de descritor.

Pode-se eliminar essa deficiência dos descritores elípticos alterando o seu processo de normalização quanto à escala especificado na seção 2.4.2.5. Se o parâmetro invariante de descrição dos semi-eixos menores das elipses for definido

como $m'_k = m_k / M_1$ as elipses mantêm a relação entre seus eixos constante durante o processo de normalização, isto é, não são deformadas por ele processo. Esses novos descritores foram aplicados à representação da base e ápice simetrizados dos folíolos terminais das mesmas espécies (usando as harmônicas ímpares de ordem 1 a 7), mas, como pode ser observado na Fig. 4.13, eles não possibilitam a separação desejada. Porém, como os descritores elípticos são derivados dos descritores de Fourier, e para estes últimos a primeira harmônica é suficiente para separação das espécies, isso sugere que a classificação pode ser realizada somente pelo semi-eixo menor dos descritores elípticos de Fourier da base e ápice da 1ª harmônica (porque esses foram os únicos descritores elípticos da primeira harmônica que não foram reduzidos a um valor padrão pela normalização modificada). De fato isso pode ser observado na Fig. 4.14, portanto os descritores elípticos também podem realizar a classificação desejada.

Os resultados mostrados na Fig. 4.13 não permitem a separação das espécies porque representam apenas a primeira componente de cada transformada de Hotelling, e as razões entre os autovalores dessas transformadas não são elevadas como ocorreu com os outros descritores, isto é, não é possível desconsiderar as outras componentes dessas transformadas sem incorrer numa perda de informação razoável.

Também foram calculadas as transformadas de Hotelling dos descritores cossenoidais da base e ápice dos folíolos (Fig. A-12). Mas, como pode ser observado nessa figura, esses descritores não permitem a separação das espécies consideradas (apesar da razão entre os autovalores da transformada de Hotelling ser elevada nesse caso), provavelmente devido à inadequação das funções de base desses descritores para representação desse tipo de objeto (que são bastante elípticos). Foi verificada também a eficiência de classificação dos folíolos através dos 7 momentos invariantes do contorno completo (Fig. A-8) e da base e ápice (Fig. A-9), da mesma forma utilizando a transformada de Hotelling para redução de dimensionalidade e facilidade de visualização. Esses momentos não realizam a diferenciação desejada, de forma que foram calculados a seguir o conjunto de momentos invariantes para uma sequência unidimensional do contorno segundo indicado na seção 2.6 para o perímetro completo da folha (Fig. A-10) e para a base e ápice (Fig. A-11). Nenhum destes descritores permite uma separação clara entre as espécies, isto é, há superposição entre suas distribuições no espaço de padrões.

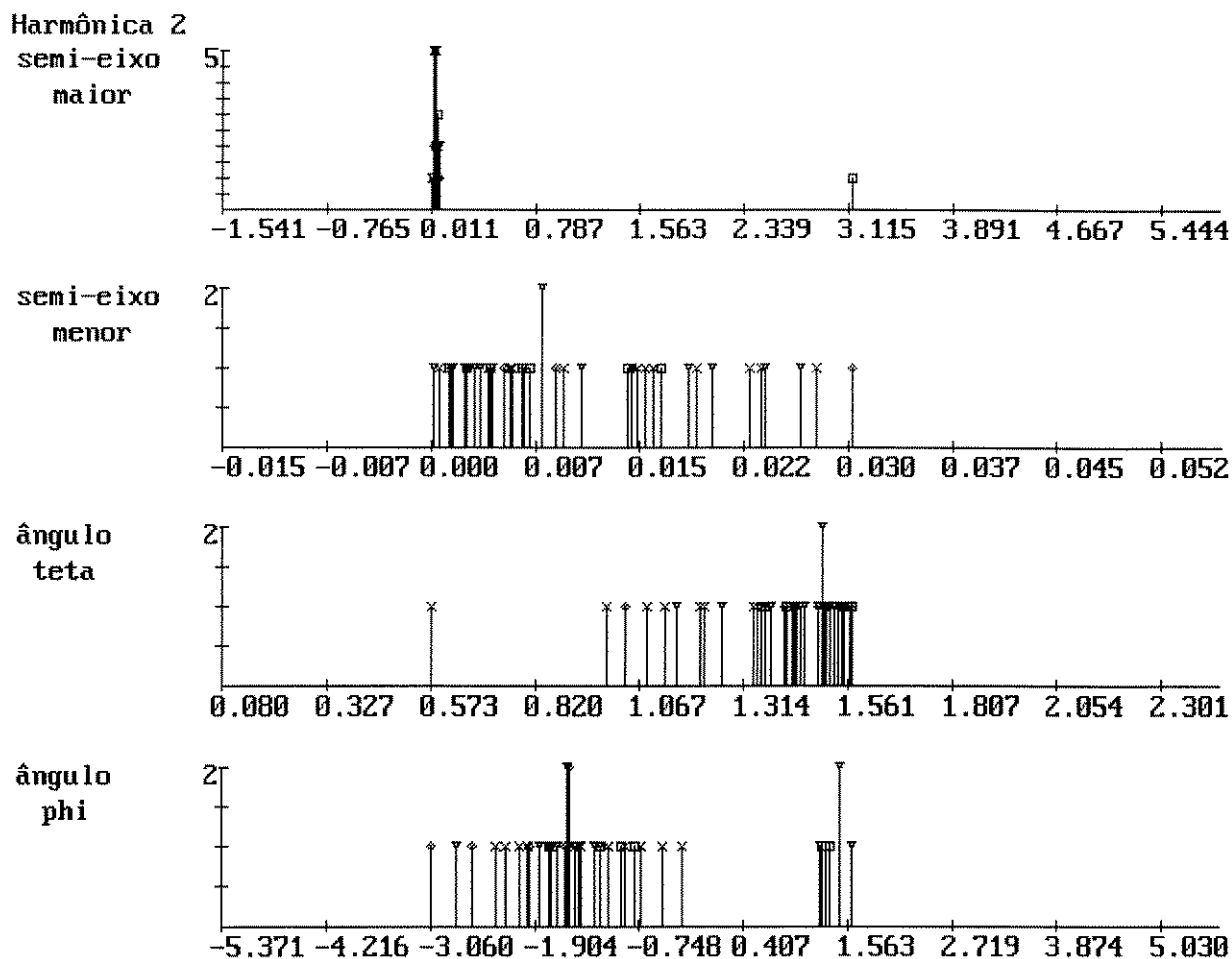


Figura 4.2: histogramas dos descritores elípticos da 2ª harmônica do contorno completo dos folíolos terminais das folhas das espécies *lonchocarpus campestris* (letras X), *lonchocarpus guilleminianus* (quadrados), *lonchocarpus latifolius* (losangos) e *lonchocarpus muehlbergianus* (triângulos invertidos).

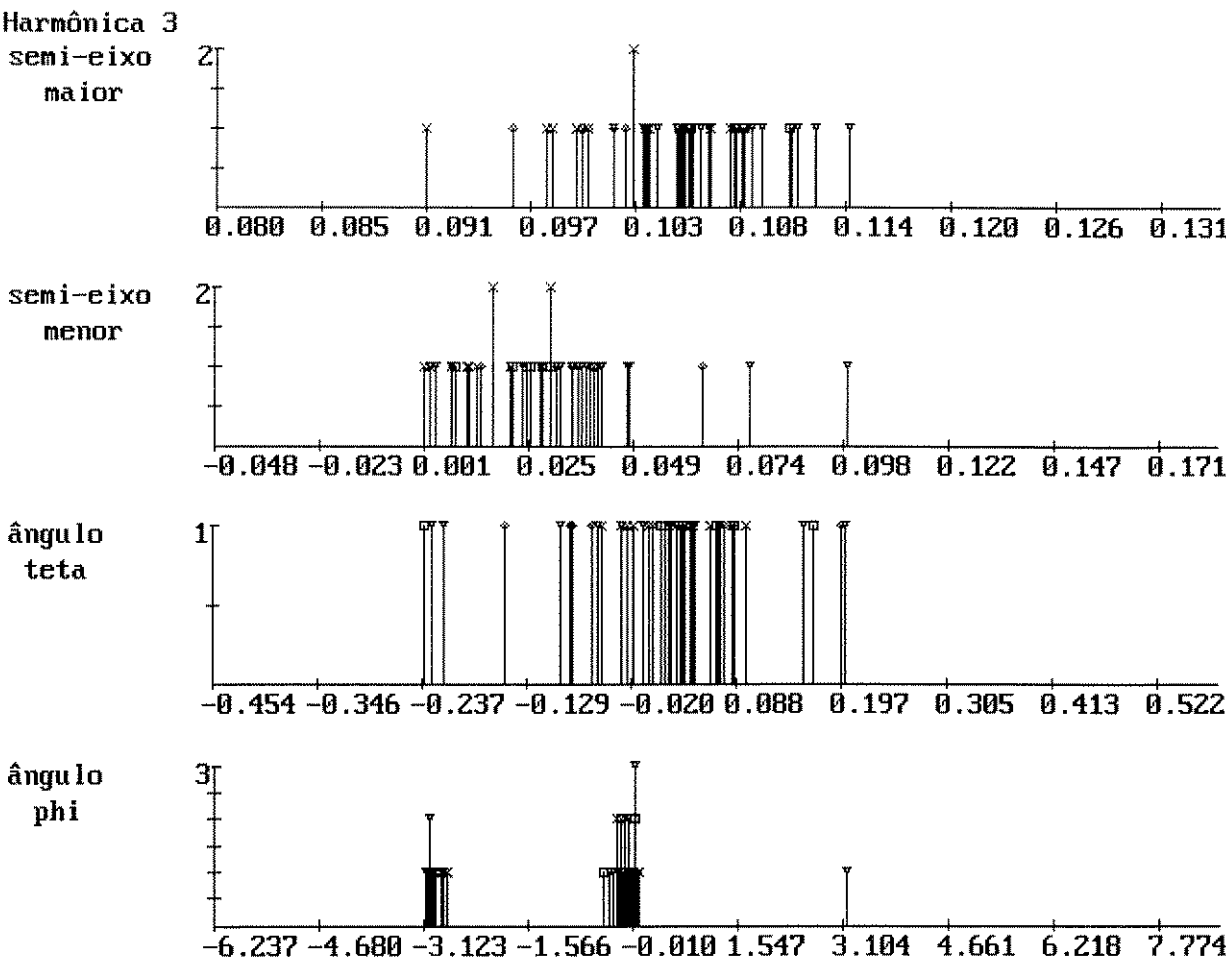


Figura 4.3: histogramas dos descritores elípticos da 3ª harmônica do contorno completo dos folíolos terminais das folhas das espécies *lonchocarpus campestris* (letras X), *lonchocarpus guilleminianus* (quadrados), *lonchocarpus latifolius* (losangos) e *lonchocarpus muehlbergianus* (triângulos invertidos).

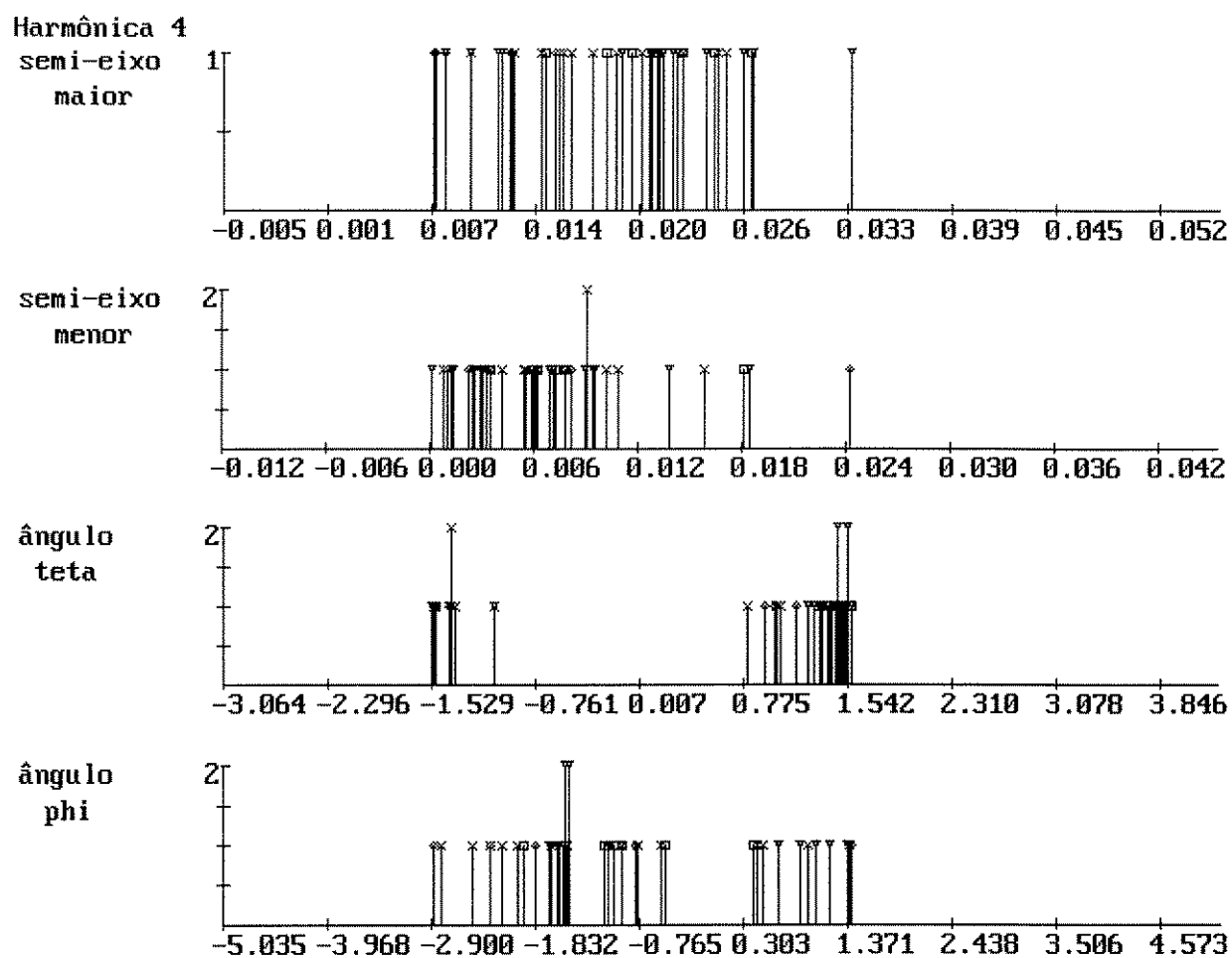


Figura 4.4: histogramas dos descritores elípticos da 4ª harmônica do contorno completo e folíolos terminais das folhas das espécies *lonchocarpus campestris* (letras X), *lonchocarpus guilleminianus* (quadrados), *lonchocarpus latifolius* (losangos) e *lonchocarpus muehlbergianus* (triângulos invertidos).

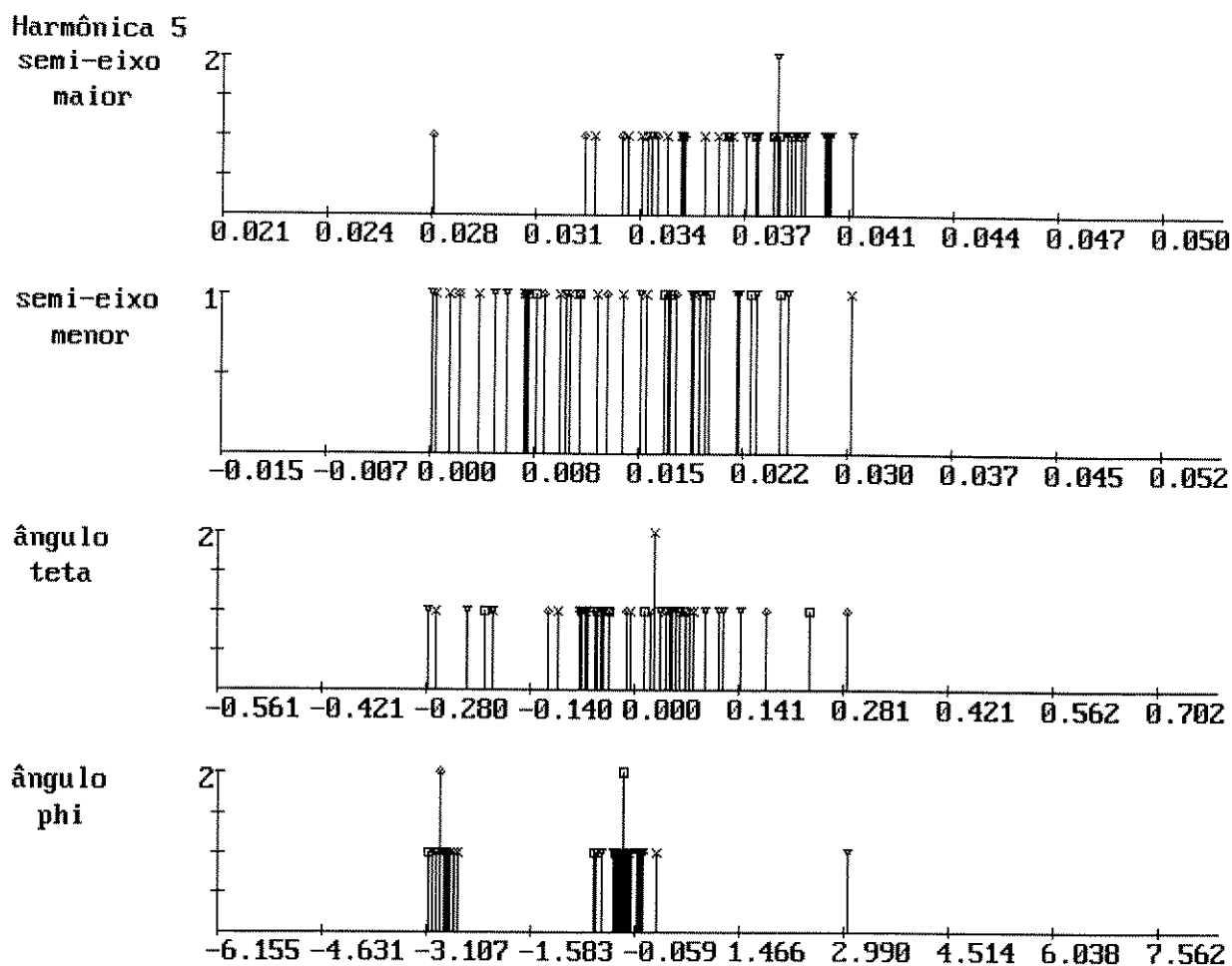


Figura 4.5: histogramas dos descritores elípticos da 5ª harmônica do contorno completo do folíolos terminais das folhas das espécies *lonchocarpus campestris* (letras X), *lonchocarpus guilleminianus* (quadrados), *lonchocarpus latifolius* (losangos) e *lonchocarpus muehlbergianus* (triângulos invertidos).

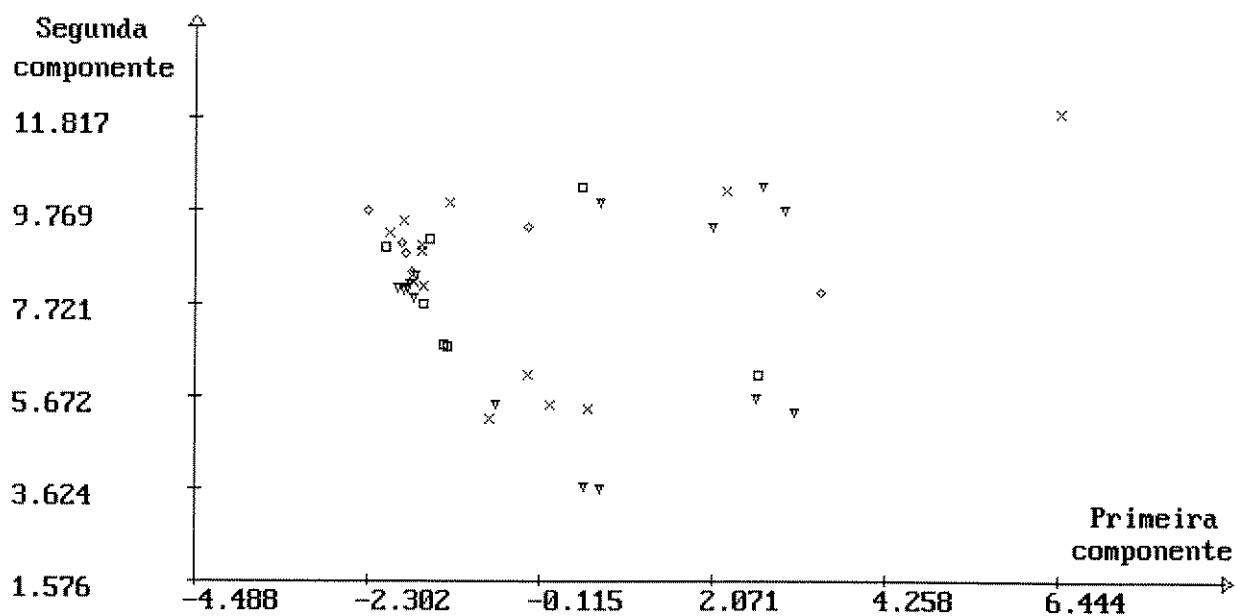


Figura 4.6: gráfico das duas primeiras componentes da transformada de Hotelling dos descritores elípticos das harmônicas 2 a 5 do contorno completo dos folíolos terminais das folhas das espécies *lonchocarpus campestris* (letras X), *lonchocarpus guilleminianus* (quadrados), *lonchocarpus latifolius* (losangos) e *lonchocarpus muehlbergianus* (triângulos invertidos).

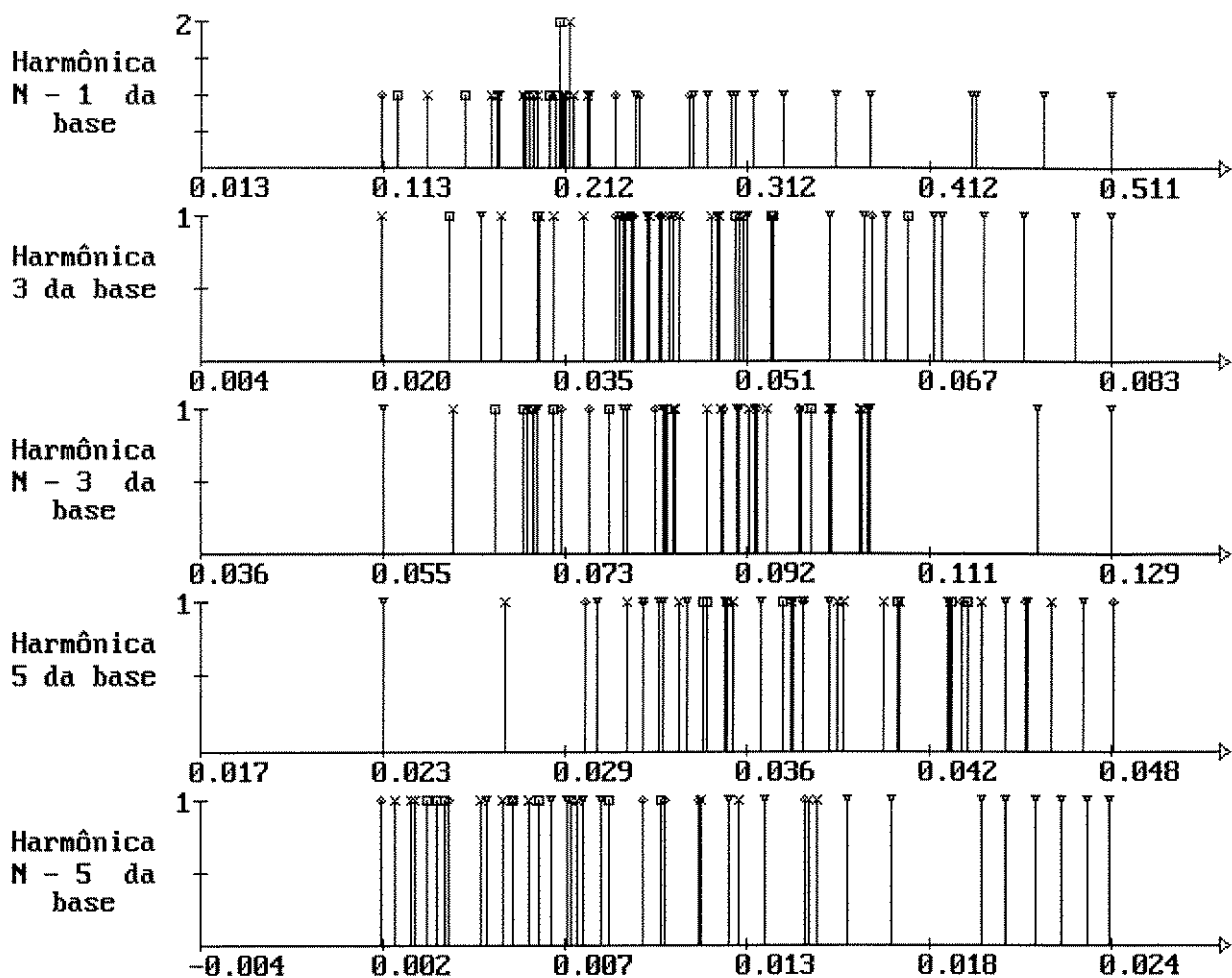


Figura 4.7: histograma dos descritores de Fourier das harmônicas ímpares de 1 a 5 da base simetrizada dos folíolos terminais das folhas das espécies *lonchocarpus campestris* (letras X) *lonchocarpus guilleminianus* (quadrados), *lonchocarpus latifolius* (losangos) e *lonchocarpus muehlbergianus* (triângulos invertidos).

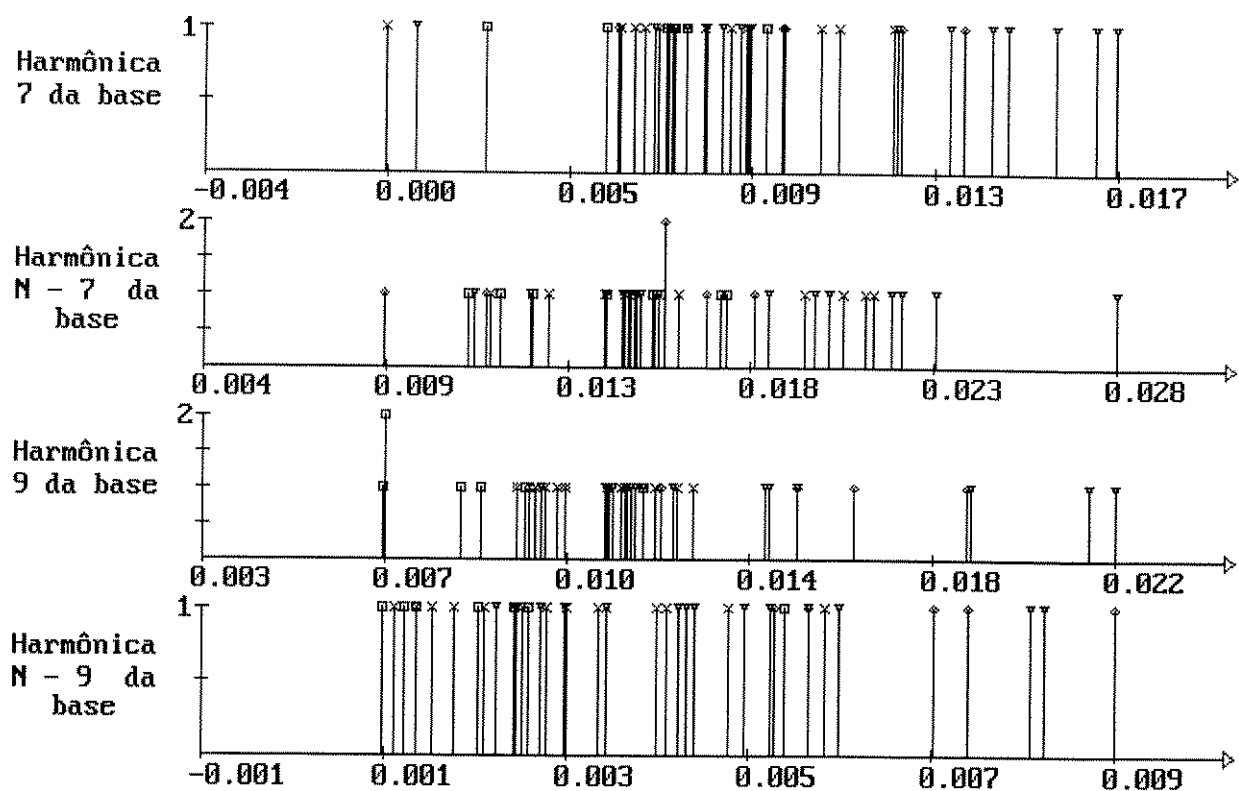


Figura 4.8: histograma dos descritores de Fourier das harmônicas 7 e 9 da base simetrizada do folíolos terminais das folhas das espécies *lonchocarpus campestris* (letras X), *lonchocarpus guilleminianus* (quadrados), *lonchocarpus latifolius* (losangos) e *lonchocarpus muehlbergianus* (triângulos invertidos).

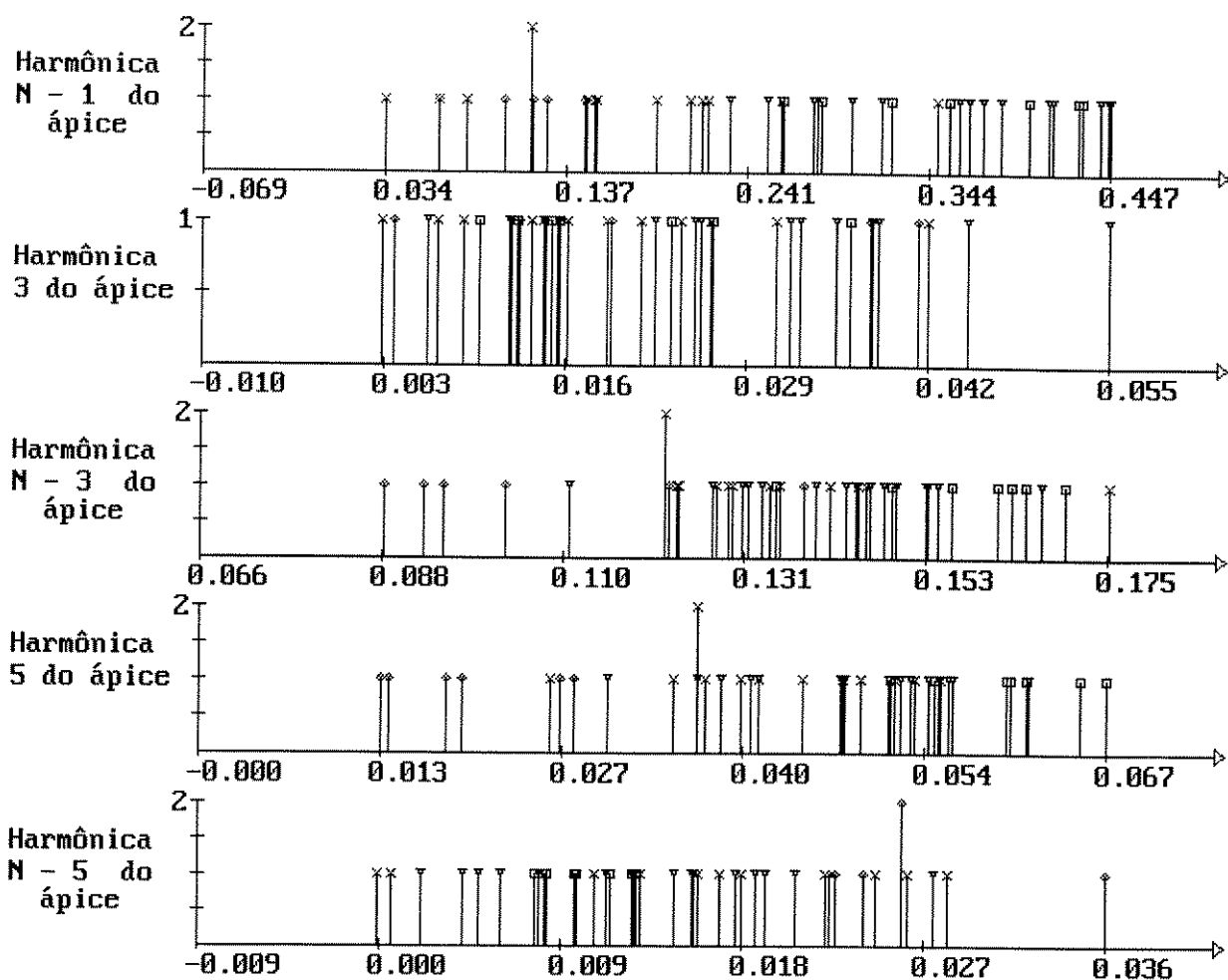


Figura 4.9: histograma dos descritores de Fourier das harmônicas ímpares de 1 a 5 do ápice simetrizado dos folíolos terminais das folhas das espécies *lonchocarpus campestris* (letras X) *lonchocarpus guilleminianus* (quadrados), *lonchocarpus latifolius* (losangos) e *lonchocarpus muehlbergianus* (triângulos invertidos).

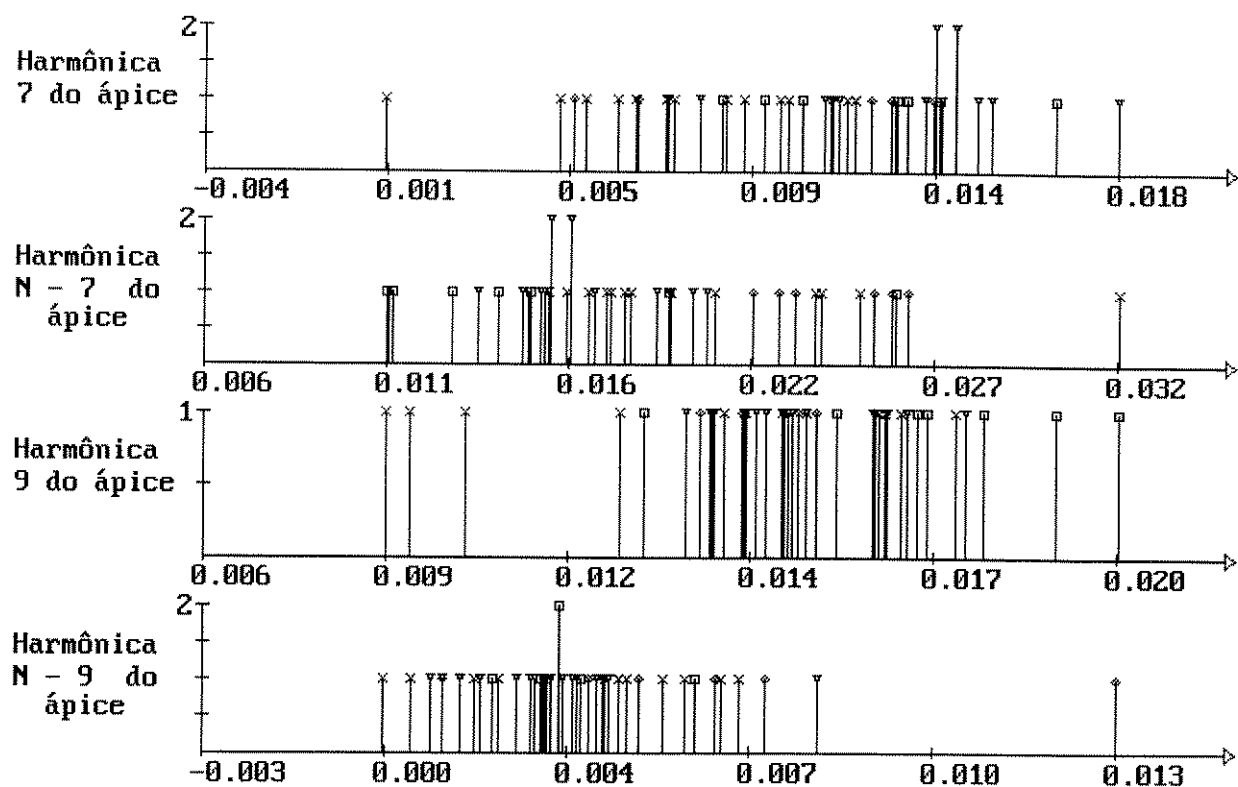


Figura 4.10: histograma dos descritores de Fourier das harmônicas 7 e 9 do ápice simetrizado do folíolos terminais das folhas das espécies *lonchocarpus campestris* (letras X), *lonchocarpus guilleminianus* (quadrados), *lonchocarpus latifolius* (losangos) e *lonchocarpus muehlbergianus* (triângulos invertidos).

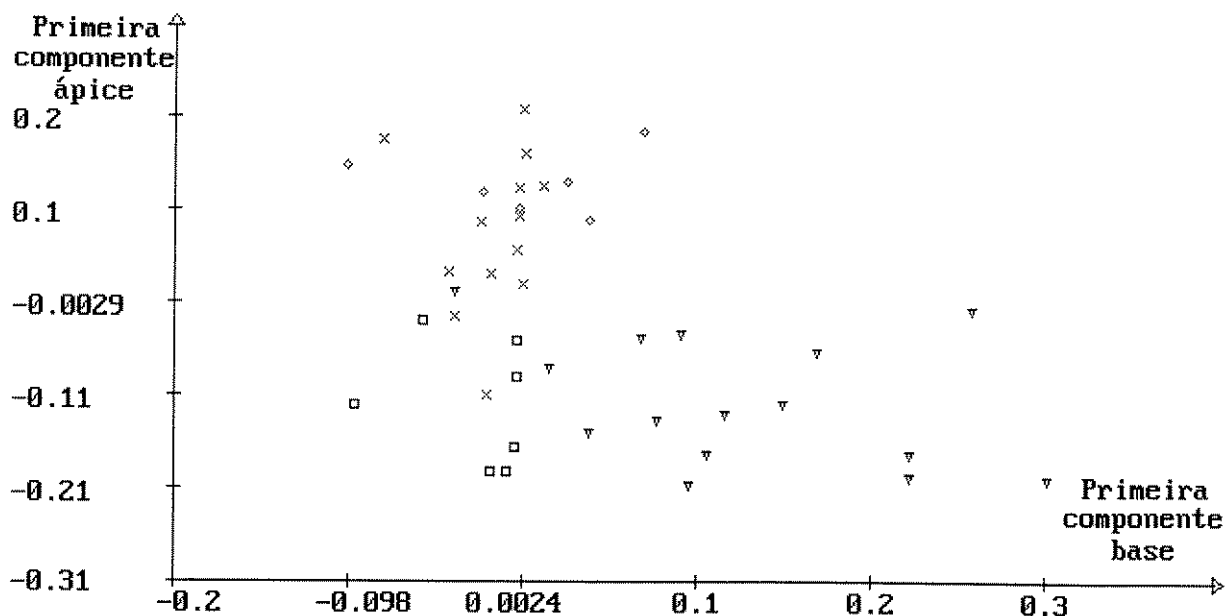


Figura 4.11: gráfico da primeira componente da transformada de Hotelling dos descritores de Fourier das harmônicas ímpares de 1 a 9 da base simetrizada versus a primeira componente da transformada de Hotelling dos descritores de Fourier das harmônicas ímpares de 1 a 9 do ápice simetrizado dos folíolos terminais das folhas das espécies *lonchocarpus campestris* (letras X), *lonchocarpus guilleminianus* (quadrados), *lonchocarpus latifolius* (losangos) e *lonchocarpus muehlbergianus* (triângulos invertidos).

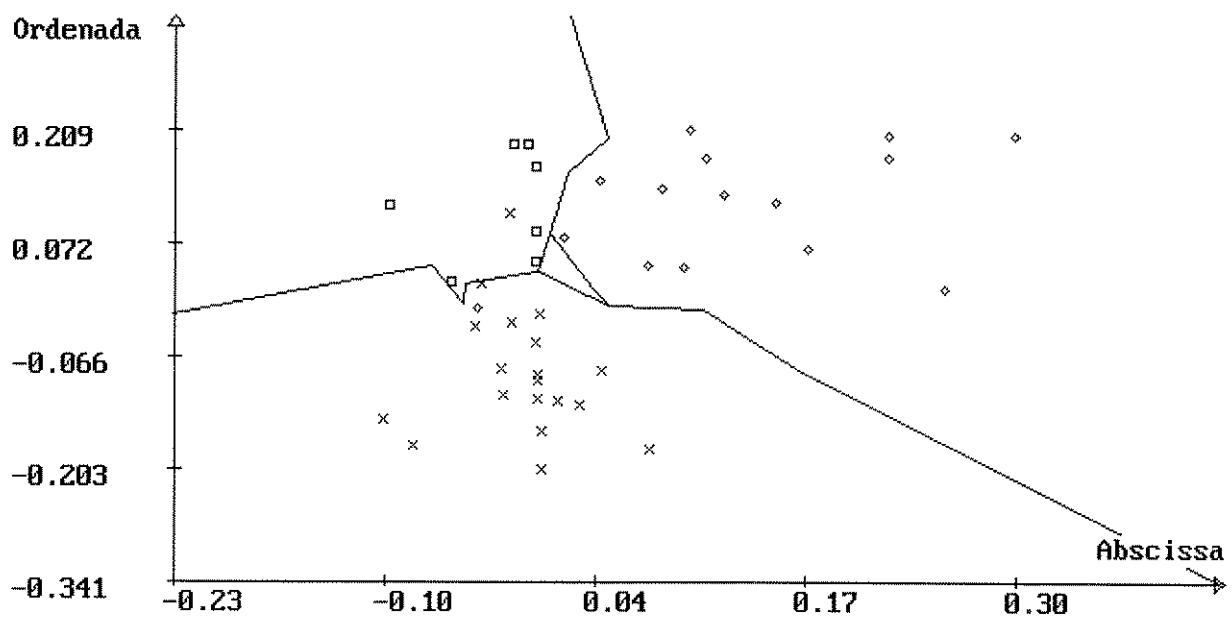


Figura 4.12: gráfico da primeira componente da transformada de Hotelling dos descritores de Fourier das harmônicas ímpares de 1 e 3 da base simetrizada versus a primeira componente da transformada de Hotelling dos descritores de Fourier das harmônicas ímpares de 1 e 3 do ápice simetrizado dos folíolos terminais das folhas das espécies *lonchocarpus campestris* e *latifolius* (letras X), *lonchocarpus guilleminianus* (quadrados) e *lonchocarpus muehlbergianus* (losangos). Também mostrando as funções discriminantes geradas pelo algoritmo proposto, que induzem dois erros.

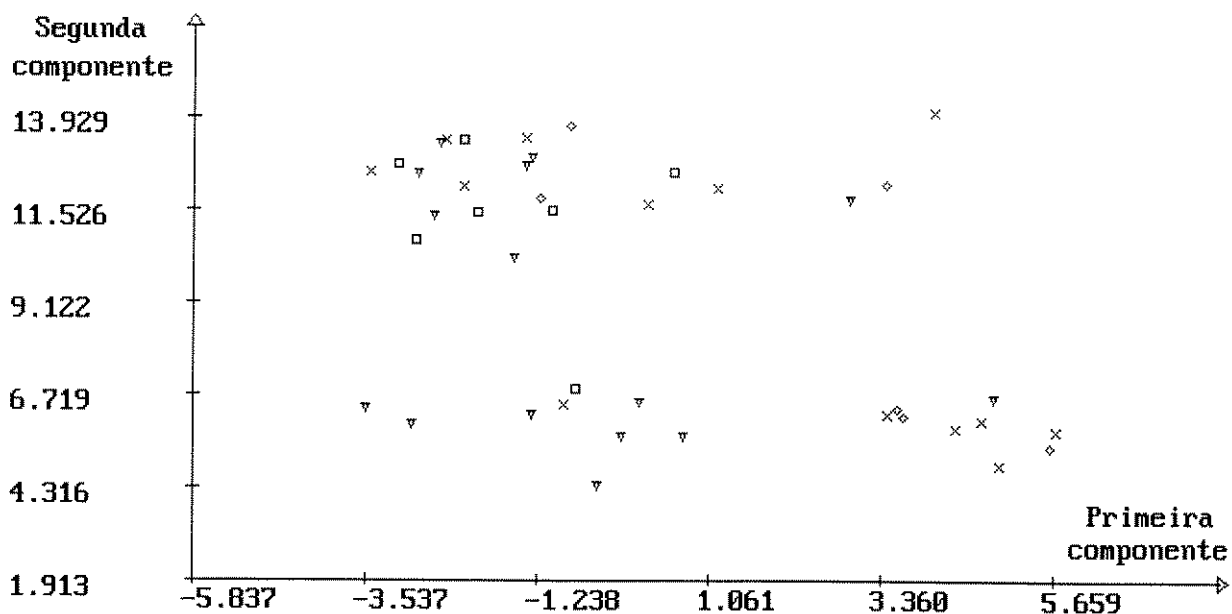


Figura 4.13: gráfico da primeira componente da transformada de Hotelling dos descritores elípticos de Fourier (com invariância à escala modificada) das harmônicas ímpares de 1 a 7 da base simetrizada versus a primeira componente da transformada de Hotelling dos descritores elípticos de Fourier (também com invariância à escala modificada) das harmônicas ímpares de 1 a 7 do ápice simetrizado dos folíolos terminais das folhas das espécies *lonchocarpus campestri* (letras X), *lonchocarpus guilleminianus* (quadrados), *lonchocarpus latifolius* (losangos) e *lonchocarpus muehlbergianus* (triângulos invertidos).

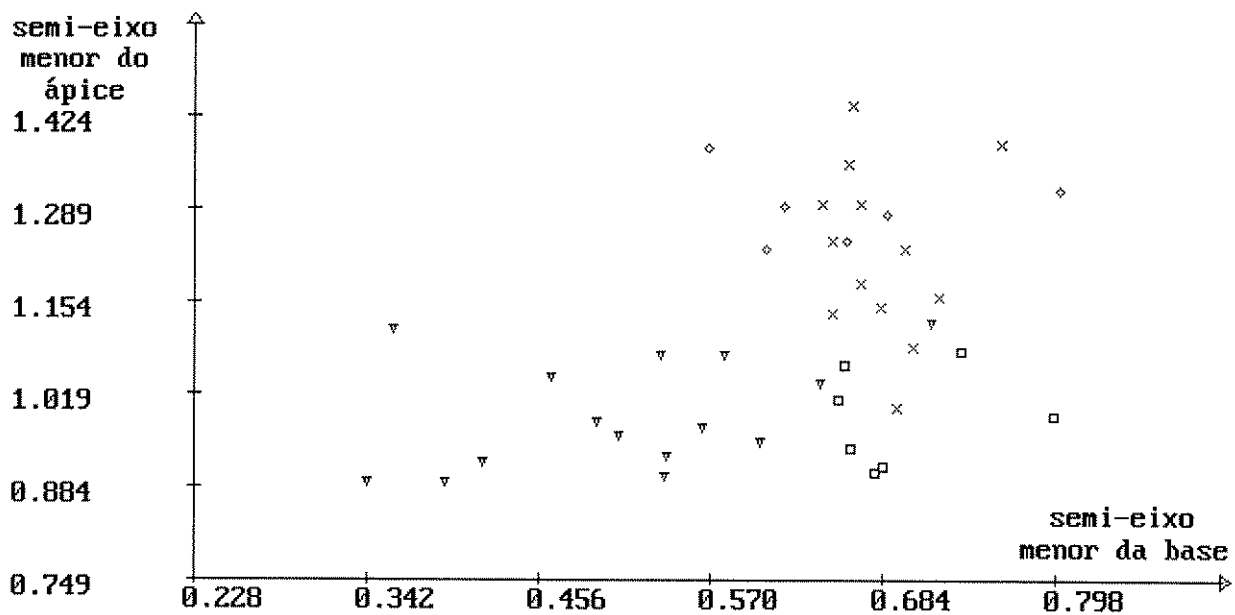


Figura 4.14: gráfico dos semi-eixos menores dos descritores elípticos de Fourier (com invariância à escala modificada) da primeira harmônica da base e ápice simetrizados dos folíolos terminais (com a mesma simbologia da figura anterior anterior)

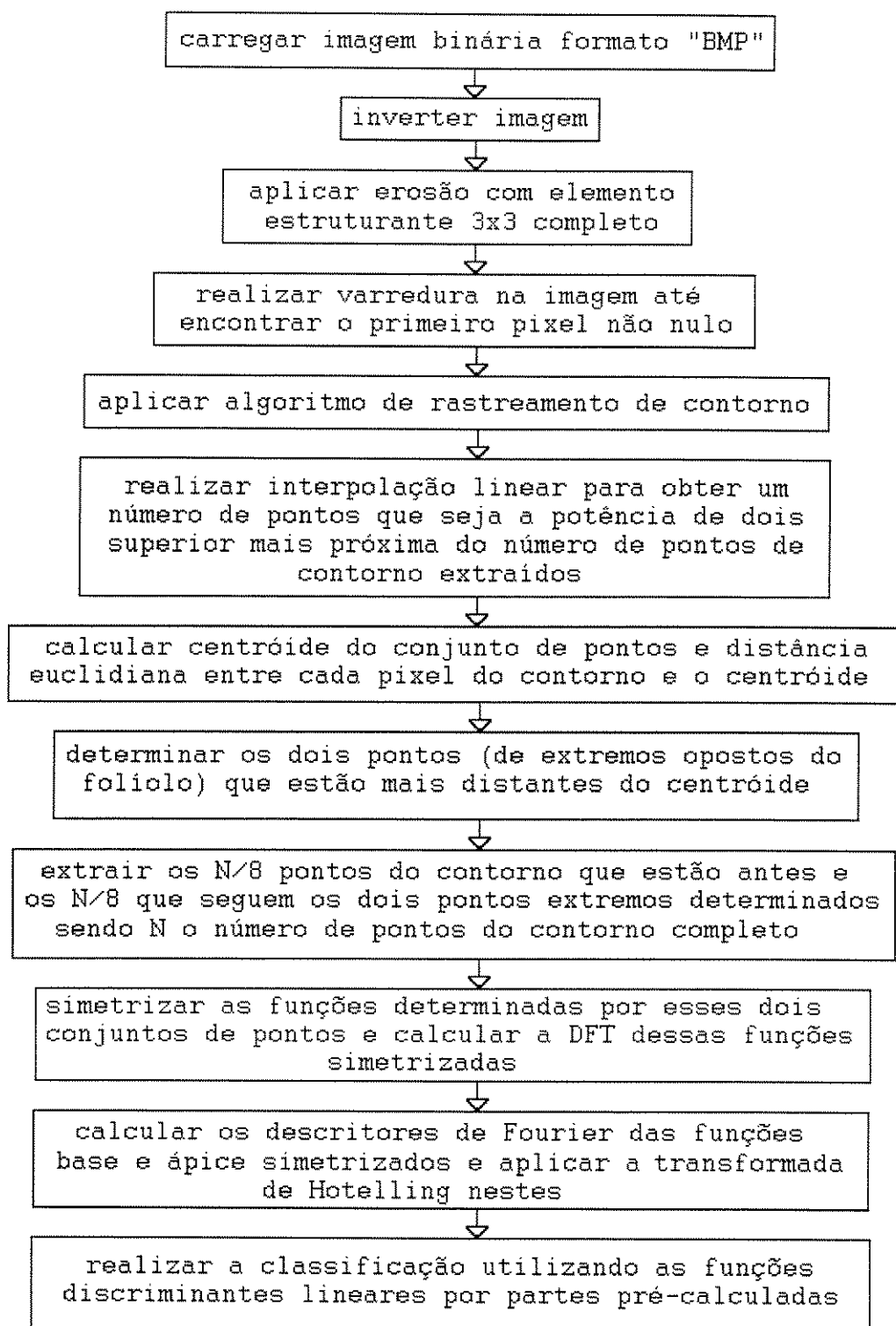


Figura 4.15: fluxograma do procedimento completo de identificação de um folíolo utilizando descritores de Fourier.

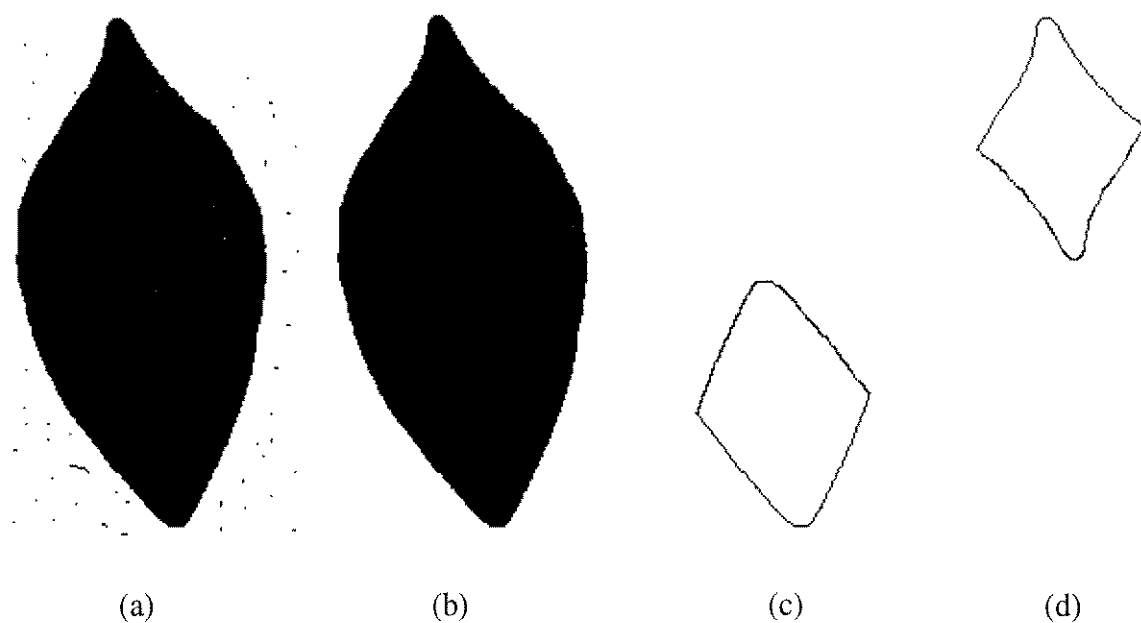


Figura 4.16: imagens dos passos intermediários do processo de identificação - (a) imagem original, (b) imagem filtrada, (c) contorno da base simetrizada e (d) contorno do ápice simetrizado.

CAPÍTULO 5

Conclusão

O objetivo deste trabalho foi estudar técnicas de reconhecimento de objetos através de descritores de contorno e a geração e utilização de funções discriminantes para sua classificação. E em particular o comportamento dos descritores de Fourier do contorno dos objetos como fator classificatório em comparação com outros descritores conhecidos. Foram deduzidos descritores cossenoidais a partir da transformada cosseno discreta (DCT) e foi demonstrada matematicamente a maneira correta de evitar "aliasing" e normalizar os descritores de Fourier em relação à escala. As técnicas estudadas foram aplicadas na classificação de quatro espécies de plantas segundo a forma de seus folíolos para avaliação de seus desempenhos .

Também foi proposto e aplicado um algoritmo para geração de funções discriminantes lineares por partes que permite a separação de classes no espaço de padrões de modo rápido, automático e eficiente para qualquer tipo de função de densidade de probabilidade. Este algoritmo não depende da seleção experimental de valores de limiar, como o apresentado em Ichino (1979), ou da pré-definição do número de componentes das funções discriminantes lineares por partes, nem requer um longo treinamento para definição dos hiperplanos como o método proposto em Foroutan e Sklansky (1987). Também não apresenta uma convergência lenta e dependente de parâmetros arbitrados como ocorre em Spivak (1989).

Somente a conjunção dos descritores de Fourier da forma simetrizada da base e ápice dos folíolos terminais permite uma distinção entre as espécies na aplicação efetuada. Todos os outros descritores considerados geram distribuições sobrepostas no espaço de padrões, mas para os descritores de Fourier são suficientes a 1ª e 3ª harmônicas para efetuar a separação com baixo índice de erros. O que leva a supor que estas espécies são identificadas justamente por essas partes dos folíolos.

Os descritores de Fourier têm melhor desempenho que os momentos porque utilizam o conceito de frequência, que permite a decomposição das formas segundo a velocidade de suas variações espaciais, e produz representações sucessivamente mais exatas, e portanto descrições separadas da forma básica e dos detalhes dos objetos. Os descritores cossenoidais não permitem a distinção na aplicação considerada porque suas funções de base não são adequadas para a representação dessa categoria de objetos, visto que eles são bastante elípticos, mas podem ser os mais eficientes em outras aplicações.

Os cálculos necessários no reconhecimento são realizados (em média) em 1,7 segundos num computador padrão IBM PC 80386 com co-processador aritmético, e deste tempo somente a filtragem das imagens consumiu em média 0,7 segundos. Nesta especificação de tempo não está incluída a digitalização dos folíolos, que é limitada pela velocidade do "scanner". Foi possível separar com baixa margem de erro duas espécies de plantas das outras através dos descritores de Fourier das bases e ápices dos folíolos terminais, sendo necessária a dimensão dos folíolos para distinguir outras duas entre si.

Referências bibliográficas

- Andrews, H. C. **Introduction to mathematical techniques in pattern recognition**. New York: Wiley-Interscience, 1972, 242 p.
- BoBrowski, L. Design of piecewise linear classifiers from formal neurons by a basis of exchange technique. **Pattern Recognition**, London, v. 24, n. 9, p. 863-870, setembro 1991.
- Bow, S. **Pattern Recognition**. New York: Marcel Dekker Inc, 1984, 323 p.
- Blum, H. A transformation for extracting new descriptors of shape. **Models for the Perception of Speech and Visual Form**, Cambridge: MIT Press, 1967.
- Bookstein, F. L., Strauss, R. E., Humphries, J. M., Chernoff, B., Elder, R. L. e Smith, G. R. A comment upon the uses of Fourier methods in systematics. **Syst. Zool**, v. 31, n. 1, p. 85-92, 1982.
- Ehrlich, R., Pharr, R. B. e Healy-Williams, N. Comments on the validity of Fourier descriptors in systematics: a reply to Bookstein et al. **Syst. Zool**, v. 32, n. 2, p. 202-206, 1983.
- Ferson, S., Rohlf F. J. e Koehn, R. K. Measuring shape variation of two-dimensional outlines. **Syst. Zool**, v. 34, n. 1, p. 59-68, 1985.
- Foroutan, I. e Sklansky, J. Feature selection for automatic classification of non-gaussian data. **IEEE Trans. Systems, Man, and Cybernetics**, New York, v. SMC-17, n. 2, p. 187-198, março/abril 1987.
- Fukunaga, K. e Olsen, D. R. Piecewise linear discriminant functions and classification errors for multiclass problems. **IEEE Trans. Inform. Theory**, New York, v. IT-16, n. 1, p. 99-100, janeiro 1970.

- Gonzalez, R. C. e Woods, R. E. **Digital image processing**. New York: Addison Wesley, 1993, 716 p.
- Granlund, G. H. Fourier preprocessing for hand print character recognition. **IEEE Trans. Computers**, New York, v. C-21, n. 2, p. 195-201, fevereiro 1972.
- Gupta, L. e Srinath, M. D. Contour sequence moments for the classification of closed planar shapes. **Pattern Recognition**, London, v. 20, n. 3, p. 267-272, 1987.
- Hart, P. E. The condensed nearest neighbor rule. **IEEE Trans. Inform. Theory**, New York, v. IT-14, n. 3, p. 515-516, maio 1968.
- Hickey, L. J. Classification of the architecture of dicotyledonous leaves. **American J. of Botany**, v. 60, n. 1, p. 17-33, 1973.
- Hu, M. K. Visual pattern recognition by moment invariants. **IRE Trans. Inform. Theory**, New York, v. IT-8, n. 2, p. 179-187, 1962.
- Ichino, M. A nonparametric multiclass pattern classifier. **IEEE Trans. Systems, Man, and Cybernetics**, New York, v. SMC-9, n. 6, p. 345-352, junho 1979.
- Kuhl, F. P. e Giardin, C. R. Elliptic Fourier features of a closed contour. **Comp. Graph. and Image Proc.**, New York, v. 18, n. 3, p. 236-258, 1982.
- Lin, C. C. e Chellapa R. Classification of partial 2D shapes using Fourier descriptors. **IEEE Trans. on Pattern Analysis and Machine Intelligence**, New York, v. PAMI-9, n. 5, p. 686-690, 1987.
- Lin, C. S. e Hwang, C. L. New forms of shape invariants from elliptic Fourier descriptors. **Pattern Recognition**, London, v. 20, n. 5, p. 535-545, 1987.
- Jain, A. K. **Fundamentals of digital image processing**. Englewood Cliffs: Prentice-Hall, Inc., 1989, 569 p.
- Mahmoud, Sabri A. Arabic character recognition using Fourier descriptors and character contour encoding. **Pattern Recognition**, London, v. 27, n. 6, p. 815-824, junho 1994.

- Oppenheim, A. V. e Schafer, R. W. **Discrete-time signal processing**. Englewood Cliffs: Prentice-Hall, 1989, 879 p.
- Paglieroni D. W. e Jain, A. K. Fast classification of discrete shape contours. **Pattern Recognition**, London, v. 20, n. 6, p. 583-598, 1987.
- _____ Control point transformation for shape representation and measurement. **Comp. Vision, Graph. and Image Proc.** Duluth, v. 42, n. 1, p. 87-111, abril 1988.
- Pavlidis, T. **Strutural pattern recognition**. Berlin: Springer, 1977, 302 p.
- Persoon, E. e Fu, K. Shape discrimination using Fourier descriptors. **IEEE Trans. on Pattern Analysis and Machine Intelligence**, New York, v. PAMI-8, n. 3, p. 388-397, maio 1986.
- Pratt, W. K. **Digital image processing**. 2. ed. New York: John Wiley, 1991, 698 p.
- Press, W. H., Flannery, B. P., Teukolsky, S. A. e Vetterling, W. T. **Numerical Recipes in C**. New York: Cambridge University Press, 1986, 994 p.
- Rao, K. R. e Yip, P. **Discret cosine transform**; algorithms, advantages, applications. New York: Academic Press, Inc., 1990, 490 p.
- Rogers, D. F. **Procedural elements for computer graphics**. 2. ed. Singapura: McGraw-Hill, 1988, 611 p.
- Rosenfeld, A. e Kak, A. C. **Digital picture processing**. New York: Academic Press, Inc., 1976, 457 p.
- Schalkoff, R. J. **Digital image processing and computer vision**. New York: John Wiley, 1989, 489 p.
- Serra, J. **Image analysis and mathematical morphology**. London: Academic Press, Inc., 1982, v. 1, 610 p.

- Silva, J. R. P. e Tozzi, C. L. Um algoritmo geométrico para geração de funções discriminantes. **IX Simpósio Brasileiro de Computação Gráfica e Processamento de Imagens**, 1996, Caxambu. Anais ..., p. 31-38, 1996.
- Sonka, M., Hlavac, V. e Boyle, R. **Image processing, analysis and machine vision**. London: Chapman & Hall, 1993, 555 p.
- Spivak, S. A multisurface method for pattern classification. **Pattern Recognition**, London, v. 22, n. 5, p. 587-591, 1989.
- Staib, L. H. e Duncan, J. S. Boundary finding with parametrically deformable models. **IEEE Trans. on Pattern Analysis and Machine Intelligence**, New York, v. PAMI-14, n. 11, p. 388-397, novembro 1992.
- Thiran, J., Macq, B. e Mairesse, J. Morphological classification of cancerous cells. **IEEE INTERNATIONAL CONFERENCE ON IMAGE PROCESSING**, 1994 **Proceedings**, Austin, novembro 1994.
- Tseng, Y. e Schenk, T. A least-square approach to matching lines with Fourier descriptors. **Report nº 418**, Ohio State University, p. 99-108, 1992.
- Wang, S., Chen, P. e Lin, W. Invariant pattern recognition by moment Fourier descriptor. **Pattern Recognition**, London, v. 27, n. 12, p. 1735-1742, dezembro 1994.
- Wood, Jeffrey Invariant pattern recognition: a review. **Pattern Recognition**, London, v. 29, n. 1, p. 1-17, janeiro 1996.
- Zhang, T. Y. e Suen, C. Y. A fast parallel algorithm for thinning digital patterns. **Comm. ACM**, vol. 27, n. 3, p. 236-239, 1984.

Apêndice A

Gráficos de descritores

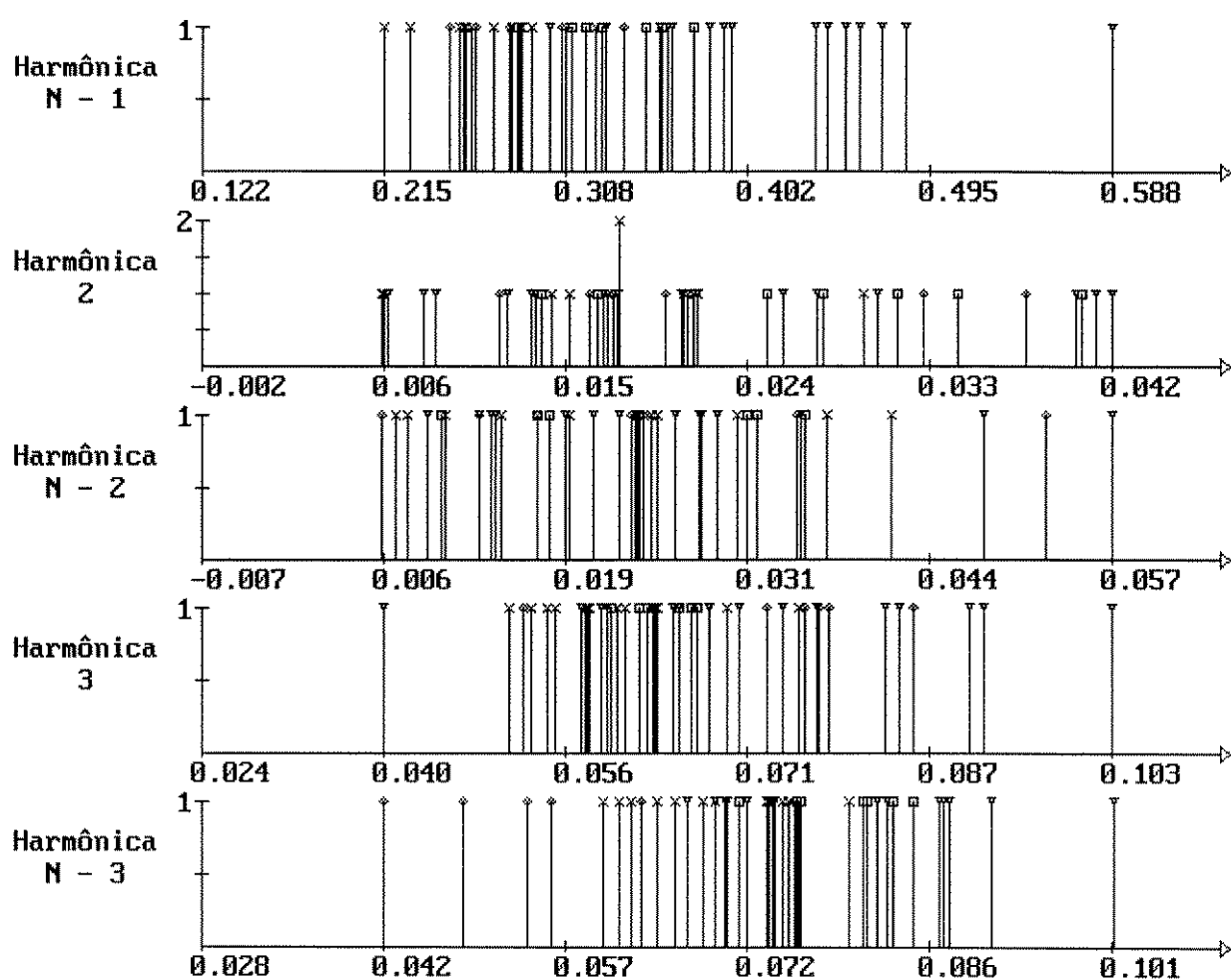


Figura A-1: histograma dos descritores de Fourier das harmônicas 1 a 3 do contorno completo do folíolo terminal das folhas das espécies *lonchocarpus campestris* (letras X), *lonchocarpus guilleminianus* (quadrados), *lonchocarpus latifolius* (losangos) e *lonchocarpus muehlbergianus* (triângulos invertidos).

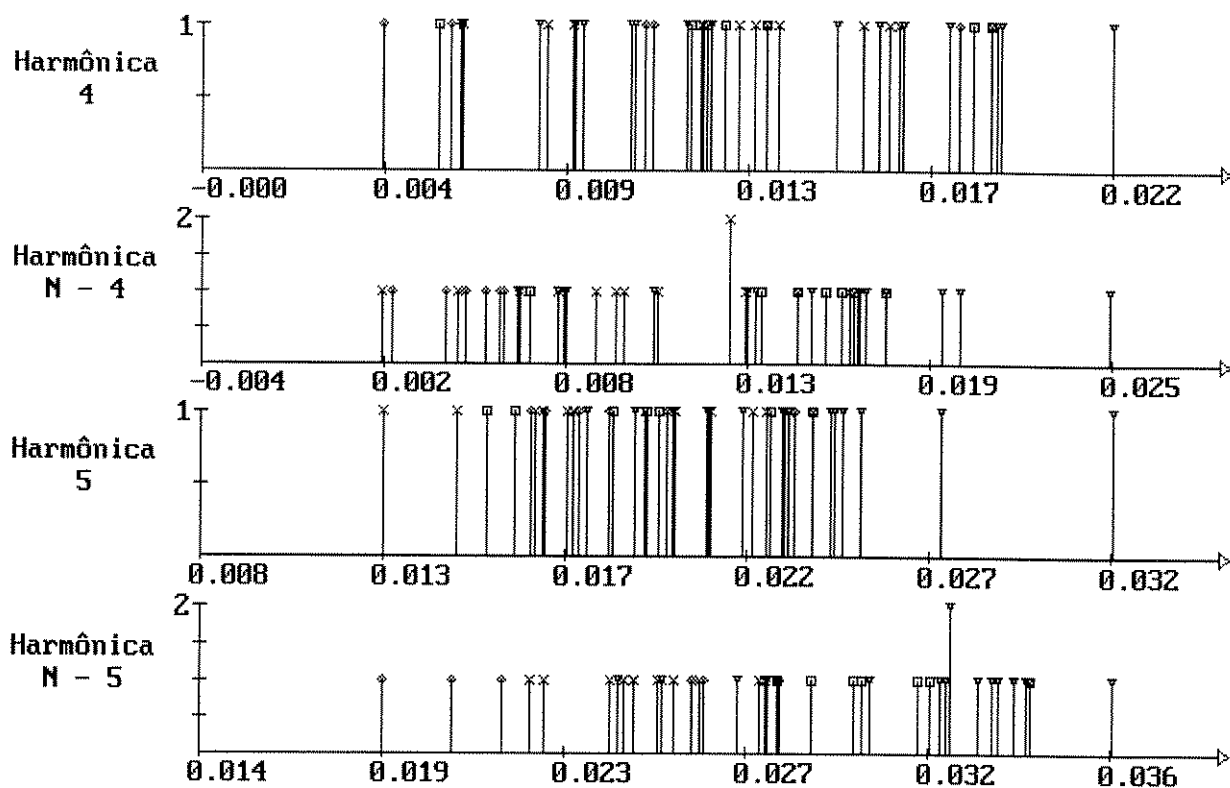


Figura A-2: histograma dos descritores de Fourier das harmônicas 4 e 5 do contorno completo dos folíolos terminais das folhas das espécies *lonchocarpus campestris* (letras X), *lonchocarpus guilleminianus* (quadrados), *lonchocarpus latifolius* (losangos) e *lonchocarpus muehlbergianus* (triângulos invertidos).

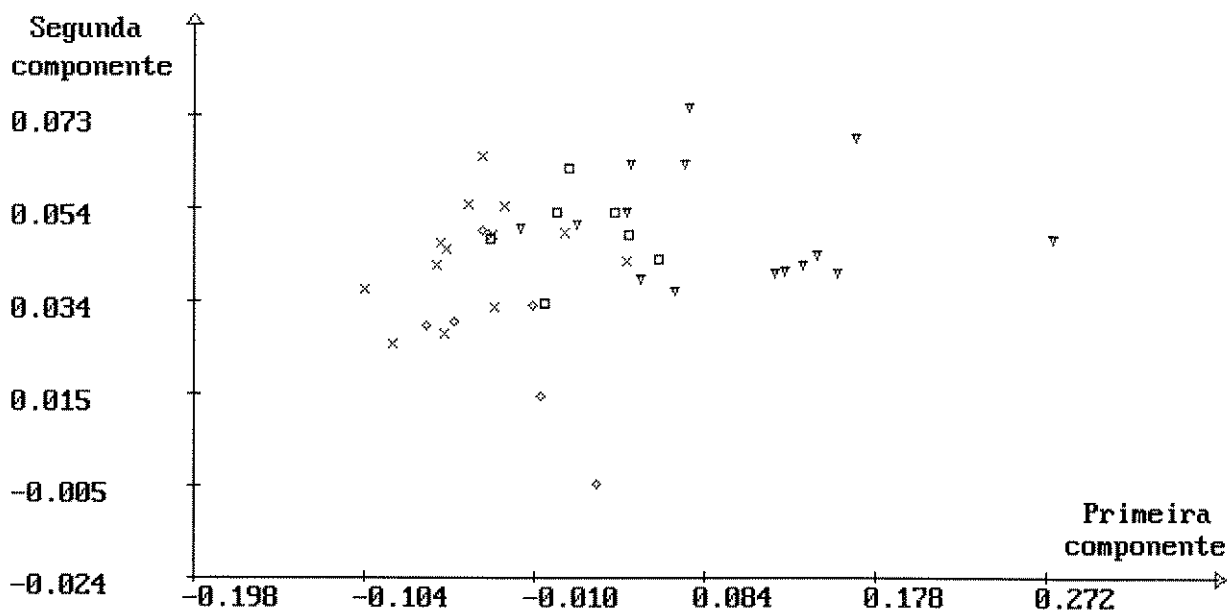


Figura A-3: gráfico das duas primeiras componentes da transformada de Hotelling dos descritores de Fourier das harmônicas 1 a 5 do contorno completo dos folíolos terminais das folhas das espécies *lonchocarpus campestris* (letras X), *lonchocarpus guilleminianus* (quadrados), *lonchocarpus latifolius* (losangos) e *lonchocarpus muehlbergianus* (triângulos invertidos).

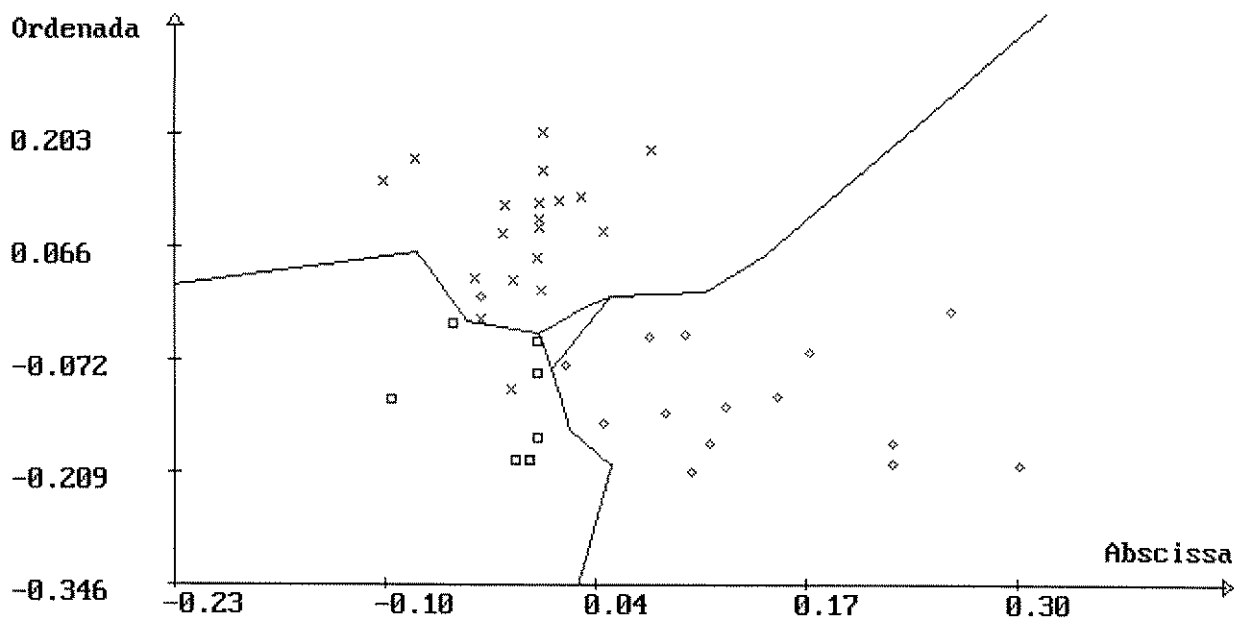


Figura A-4: gráfico da primeira componente da transformada de Hotelling dos descritores de Fourier das harmônicas ímpares de 1 a 9 da base simetrizada versus a primeira componente da transformada de Hotelling dos descritores de Fourier das harmônicas ímpares de 1 a 9 do ápice simetrizado dos folíolos terminais das folhas das espécies *lonchocarpus campestris* e *latifolius* (letras X), *lonchocarpus guilleminianus* (quadrados) e *lonchocarpus muehlbergianus* (losangos). Também mostrando as funções discriminantes geradas pelo algoritmo proposto.

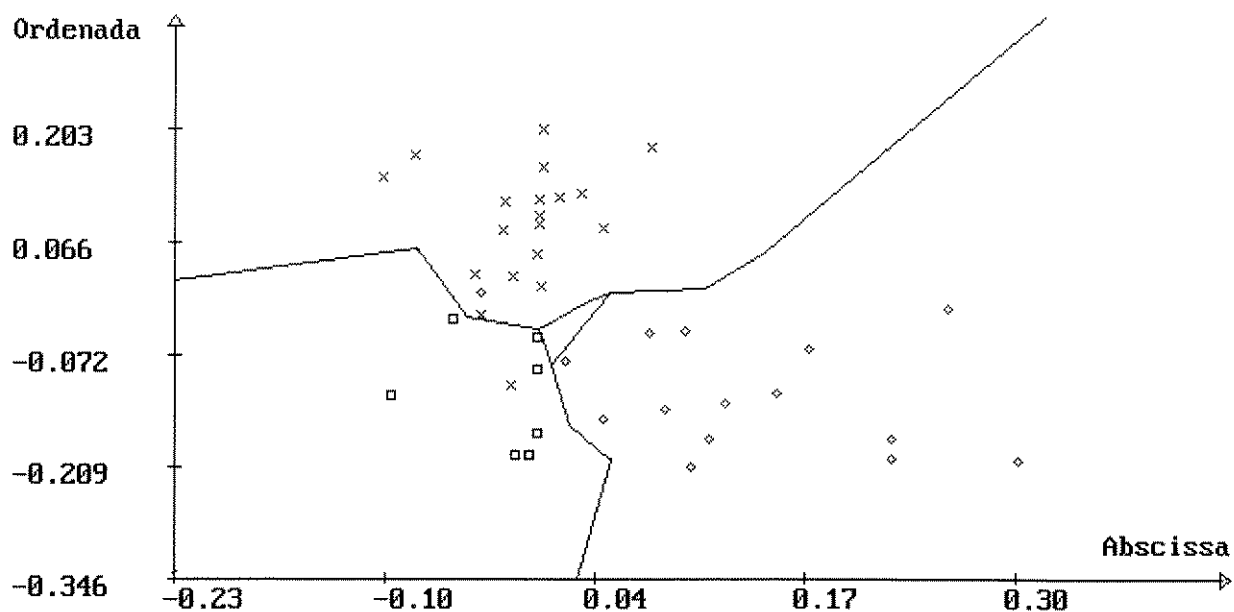


Figura A-5: gráfico da primeira componente da transformada de Hotelling das harmônicas ímpares de 1 a 7 dos descritores de Fourier da base simetrizada versus a primeira componente da transformada de Hotelling das harmônicas ímpares de 1 a 7 dos descritores de Fourier do ápice simetrizado dos folíolos terminais das folhas das espécies *lonchocarpus campestris* e *latifolius* (letras X), *lonchocarpus guilleminianus* (quadrados) e *lonchocarpus muehlbergianus* (losangos). Também mostrando as funções discriminantes geradas pelo algoritmo proposto.

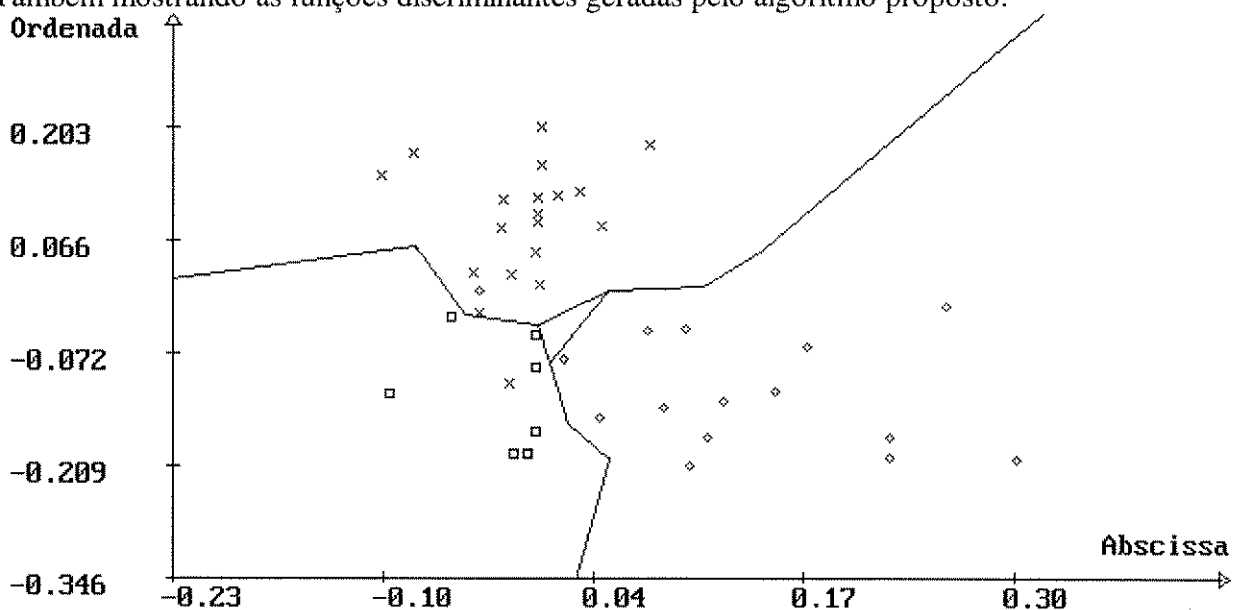


Figura A-6: gráfico da primeira componente da transformada de Hotelling dos descritores de Fourier das harmônicas ímpares de 1 a 5 da base simetrizada versus a primeira componente da transformada de Hotelling dos descritores de Fourier das harmônicas ímpares de 1 a 5 do ápice simetrizado dos folíolos terminais das folhas das espécies *lonchocarpus campestris* e *latifolius* (letras X), *lonchocarpus guilleminianus* (quadrados) e *lonchocarpus muehlbergianus* (losangos). Também mostrando as funções discriminantes geradas pelo algoritmo proposto.

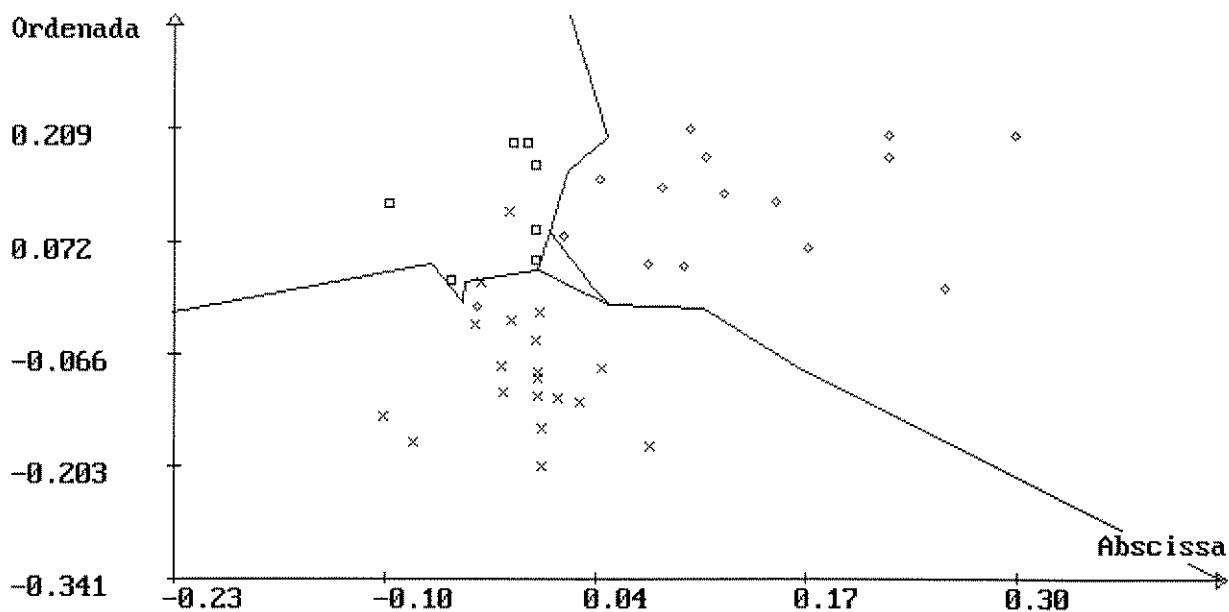


Figura A-7: gráfico da primeira componente da transformada de Hotelling dos descritores de Fourier das harmônicas ímpares de 1 e 3 da base simetrizada versus a primeira componente da transformada de Hotelling dos descritores de Fourier das harmônicas ímpares de 1 e 3 do ápice simetrizado dos folíolos terminais das folhas das espécies *lonchocarpus campestris* e *latifolius* (letras X), *lonchocarpus guilleminianus* (quadrados) e *lonchocarpus muehlbergianus* (losangos). Também mostrando as funções discriminantes geradas pelo algoritmo proposto.

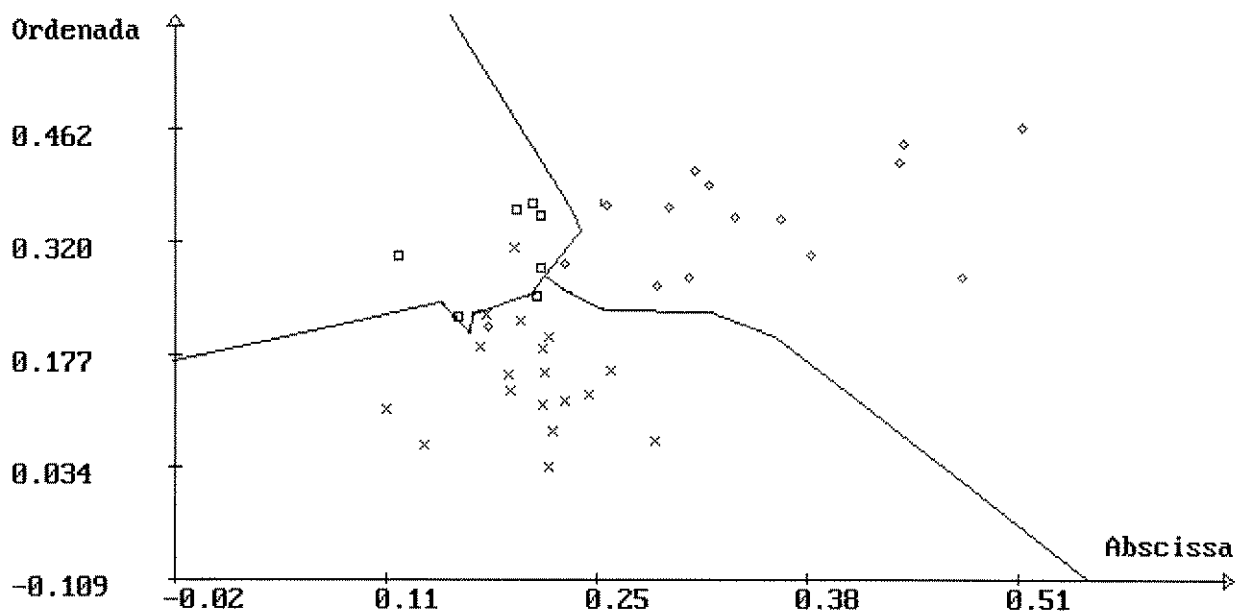


Figura A-8: gráfico do descritor de Fourier da primeira harmônica da base simetrizada versus o descritor de Fourier da primeira harmônica do ápice simetrizado dos folíolos terminais das folhas das espécies *lonchocarpus campestris* e *latifolius* (letras X), *lonchocarpus guilleminianus* (quadrados) e *lonchocarpus muehlbergianus* (losangos). Também mostrando as funções discriminantes geradas pelo algoritmo proposto.

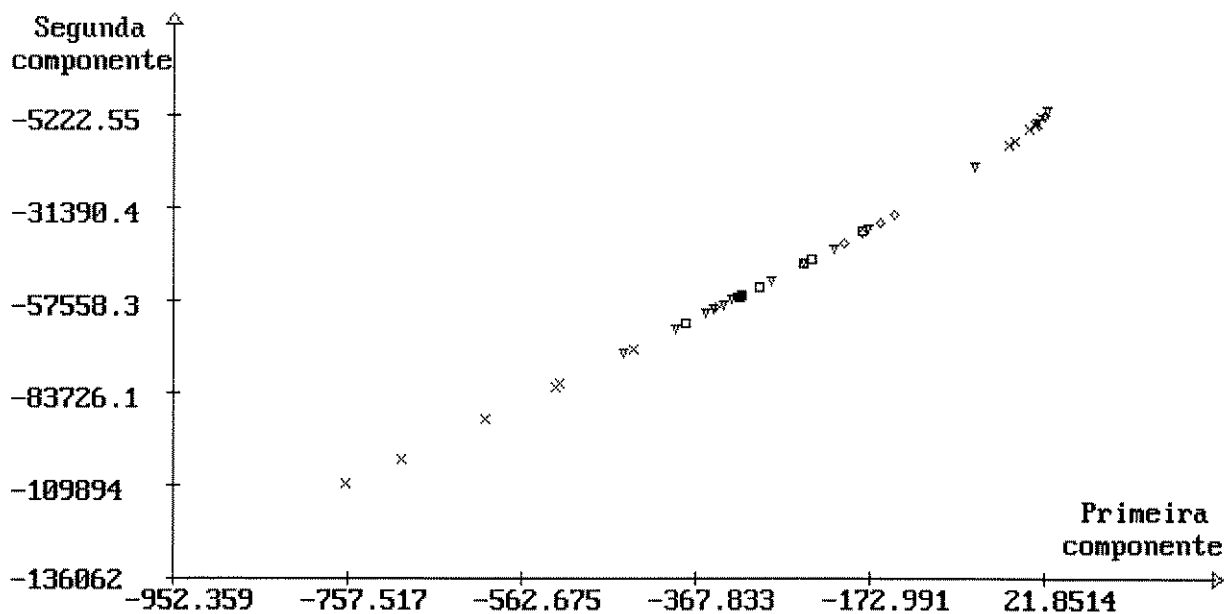


Figura A-9: gráfico das duas primeiras componentes da transformada de Hotelling dos 7 momentos centrais invariantes dos folíolos terminais das folhas das espécies *lonchocarpus campestris* (letras X), *lonchocarpus guilleminianus* (quadrados), *lonchocarpus latifolius* (losangos) e *lonchocarpus muehlbergianus* (triângulos invertidos).

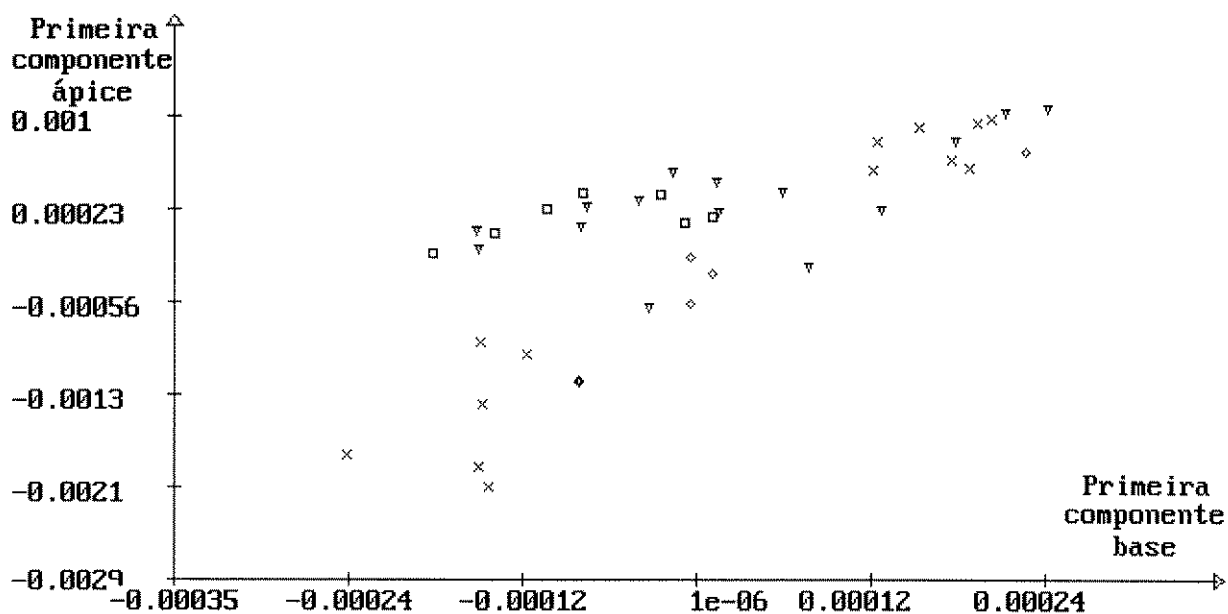


Figura A-10: gráfico da primeira componente da transformada de Hotelling dos 7 momentos centrais invariantes da base simetrizada versus a primeira componente da transformada de Hotelling dos 7 momentos centrais invariantes do ápice simetrizado dos folíolos terminais das folhas das espécies *lonchocarpus campestris* (letras X), *lonchocarpus guilleminianus* (quadrados), *lonchocarpus latifolius* (losangos) e *lonchocarpus muehlbergianus* (triângulos invertidos).

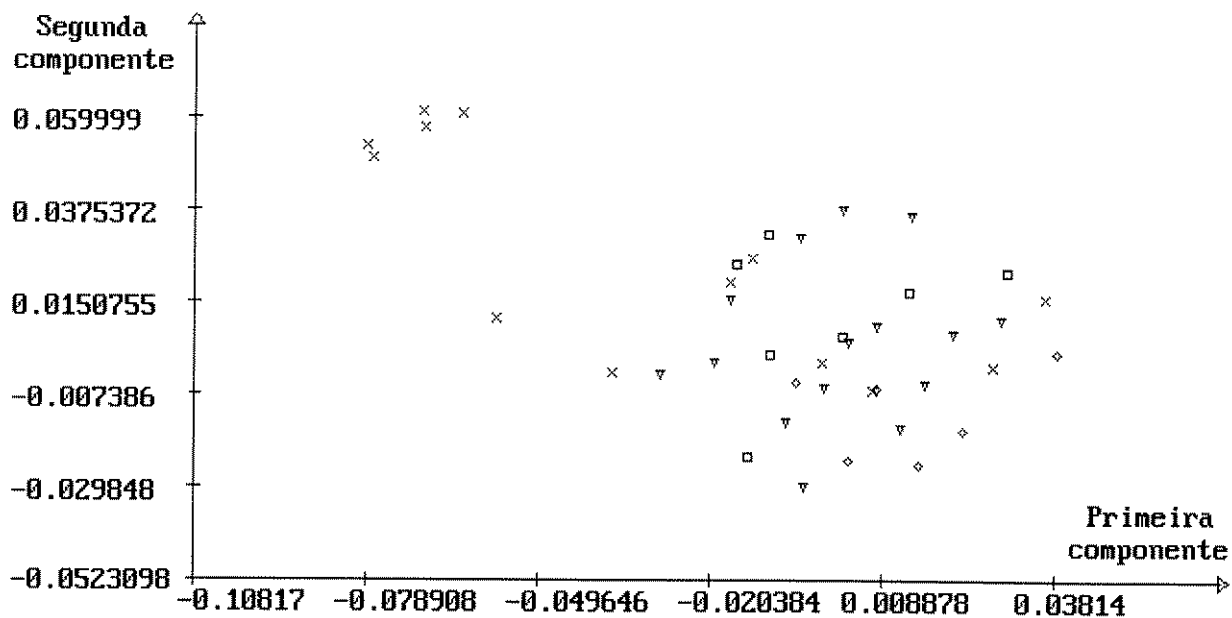


Figura A-11: gráfico das duas primeiras componentes da transformada de Hotelling dos 4 momentos da sequência unidimensional de contorno completo dos folíolos terminais das folhas das espécies *lonchocarpus campestris* (letras X), *lonchocarpus guilleminianus* (quadrados), *lonchocarpus latifolius* (losangos) e *lonchocarpus muehlbergianus* (triângulos invertidos).

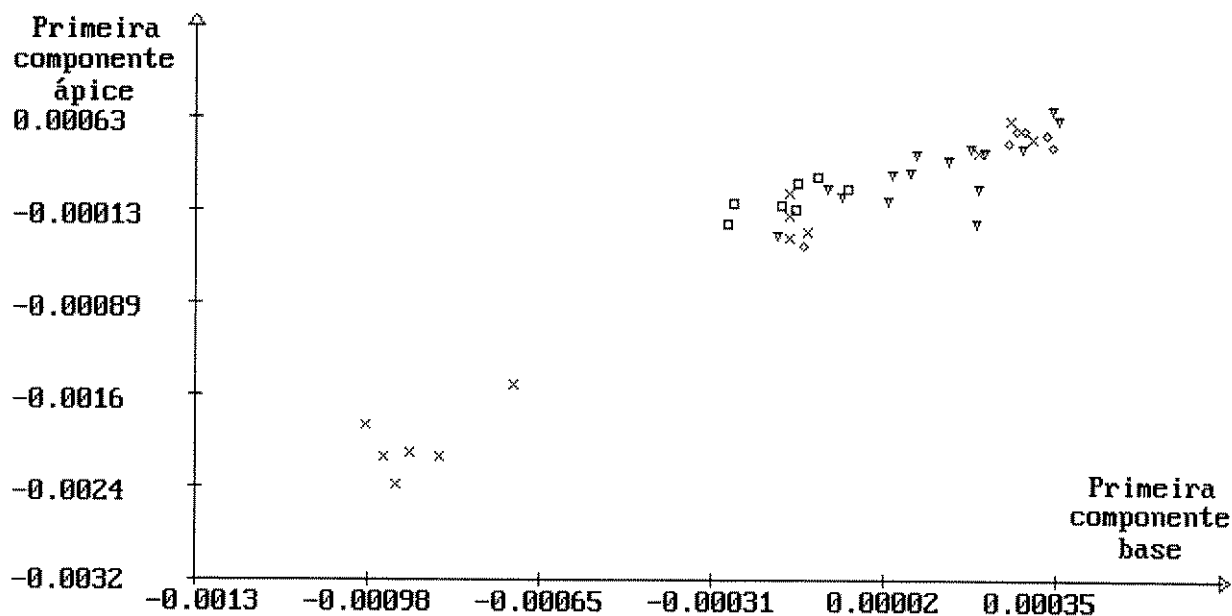


Figura A-12: gráfico da primeira componente da transformada de Hotelling dos 4 momentos da sequência unidimensional de contorno da base simetrizada versus a primeira componente da transformada de Hotelling dos 4 momentos da sequência unidimensional de contorno do ápice simetrizado dos folíolos terminais das folhas das espécies *lonchocarpus campestris* (letras X), *lonchocarpus guilleminianus* (quadrados), *lonchocarpus latifolius* (losangos) e *lonchocarpus muehlbergianus* (triângulos invertidos).

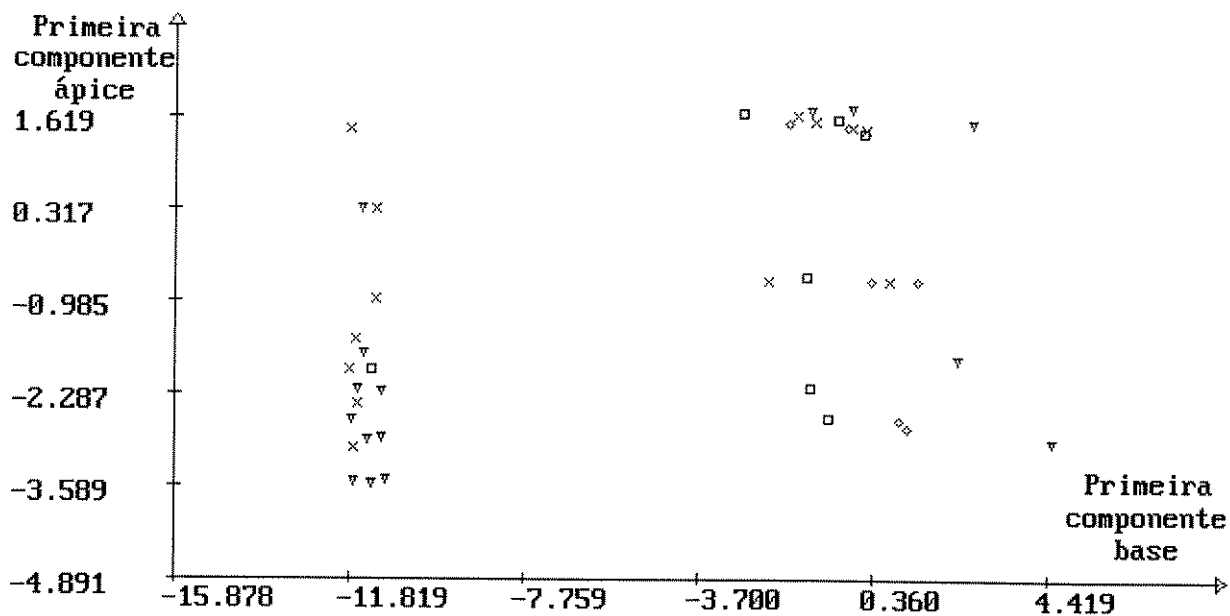


Figura A-13: gráfico da primeira componente da transformada de Hotelling dos descritores cossenoidais das harmônicas 2 a 5 da base simetrizada versus a primeira componente da transformada de Hotelling dos descritores cossenoidais das harmônicas 2 a 5 do ápice simetrizada dos folíolos terminais das folhas das espécies *lonchocarpus campestris* (letras X), *lonchocarpus guilleminianus* (quadrados), *lonchocarpus latifolius* (losangos) e *lonchocarpus muehlbergianus* (triângulos invertidos).

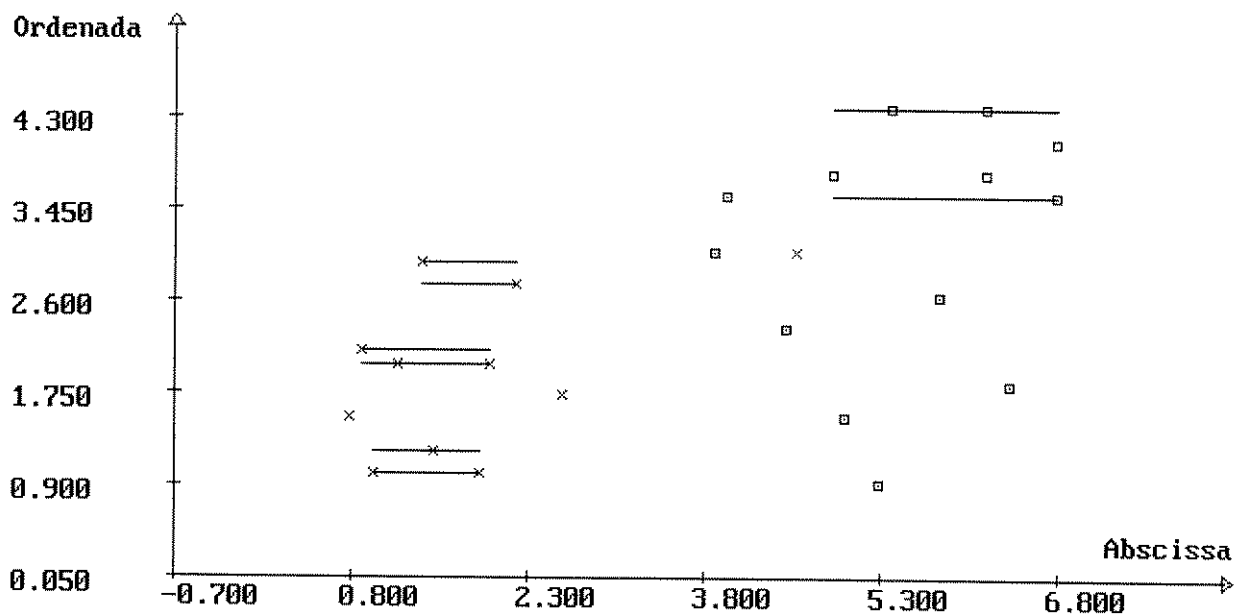


Figura A-14: gráfico do espaço de padrões com os mesmos dados da Fig. 3.3 e hiper-retângulos gerados pelo algoritmo de Ichino com $T_a=2$ e $T_c=0$.

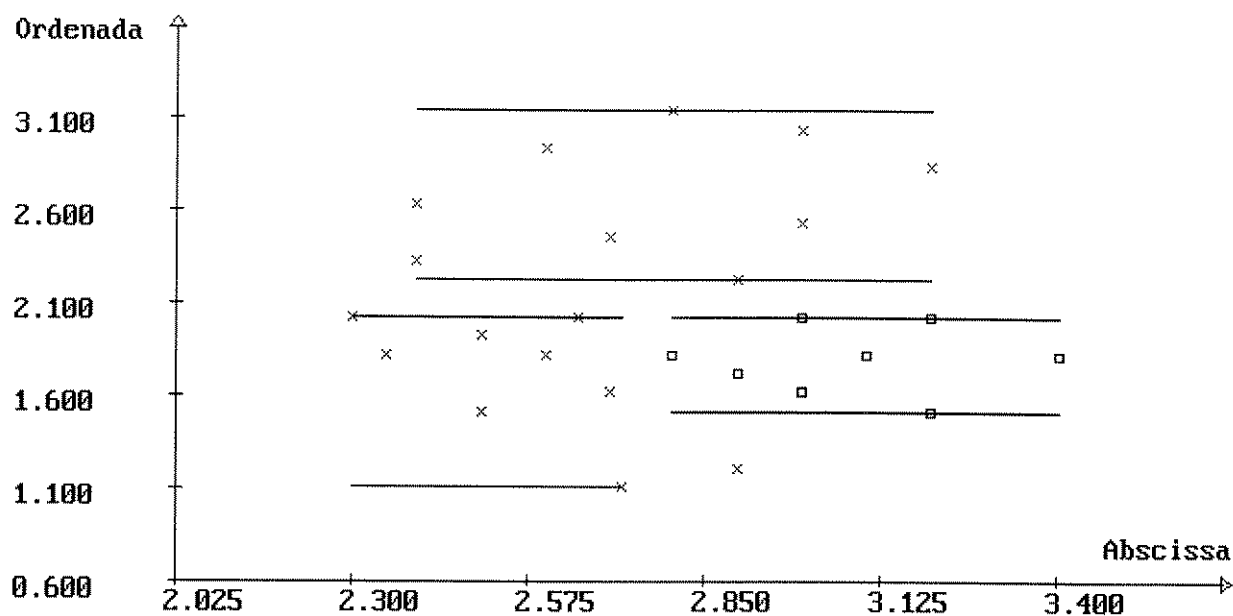


Figura A-15: gráfico do espaço de padrões com os mesmos dados da Fig. 3.4 e hiper-retângulos gerados pelo algoritmo de Ichino com $T_a=1$ e $T_c=0$.

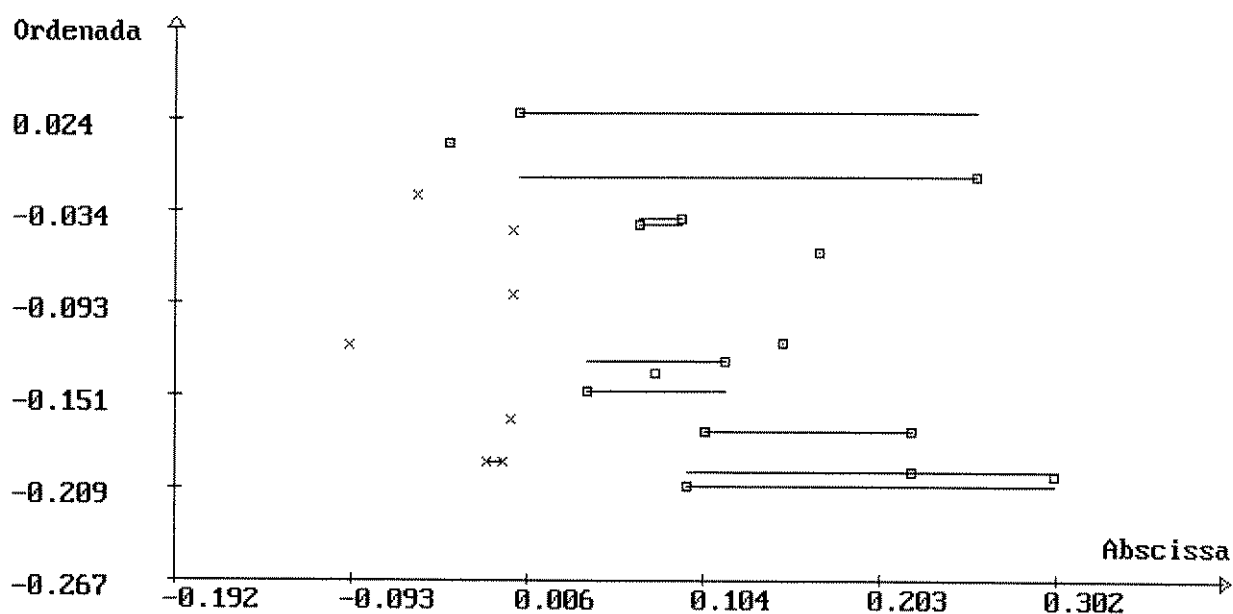


Figura A-16: gráfico do espaço de padrões com os mesmos dados da Fig. 3.5 e hiper-retângulos gerados pelo algoritmo de Ichino com $T_a=2$ e $T_c=1$.

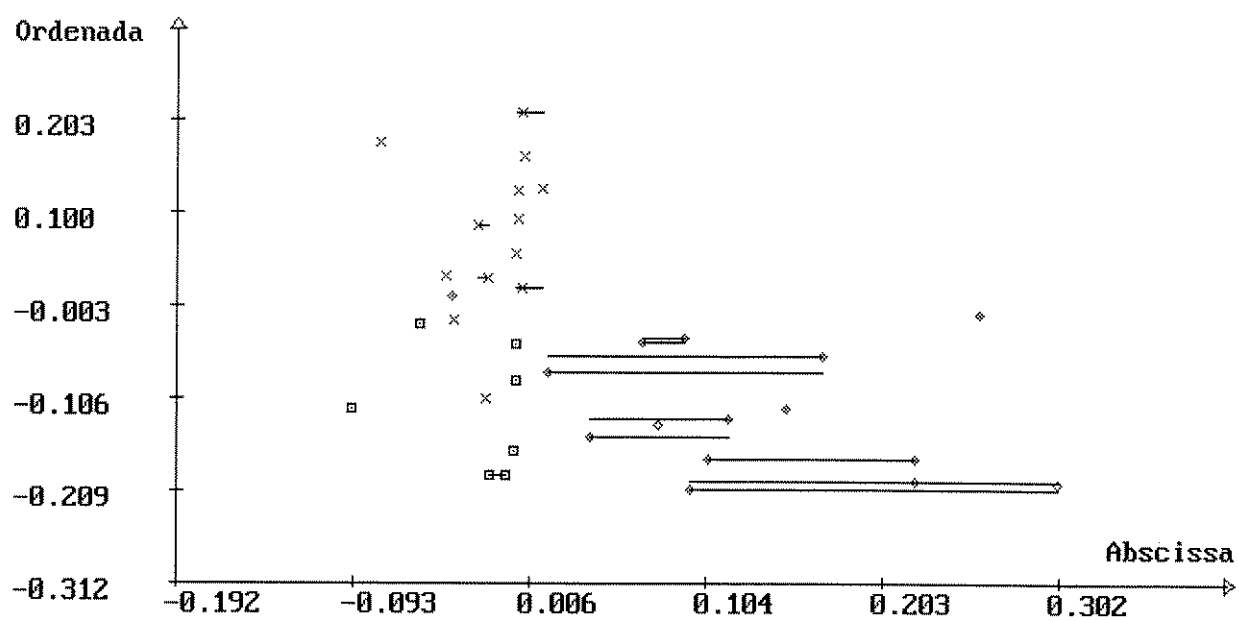


Figura A-17: gráfico do espaço de padrões com os mesmos dados da Fig. 3.6 e hiper-retângulo gerados pelo algoritmo de Ichino com $T_a=2$ e $T_c=1$.

Apêndice B

Este apêndice apresenta a demonstração matemática que uma função complexa aberta, quando tornada fechada e simétrica segundo as Eqs. (4.9), tem uma transformada discreta de Fourier cujas harmônicas pares (exceto a zero) são identicamente nulas.

Dada uma função contorno aberto $f[n]=x[n]+jy[n]$ para $0 \leq n < N$ define-se um função $f'[n]$ fechada com $0 \leq n < 2N$ como:

$$f'[n] = \begin{cases} f[n] & \text{para } 0 \leq n < N \\ c - f[n - N] & \text{para } N \leq n < 2N \end{cases}$$

sendo o número complexo c determinado por $c = f[0] + f[N-1]$

Então a transformada discreta de Fourier da seqüência $f'[n]$ é obtida por:

$$\begin{aligned} F'[k] &= \sum_{n=0}^{2N-1} f'[n] \exp(-j2k\pi n / (2N)) \quad \text{para } 0 \leq k < 2N \\ &= \sum_{n=0}^{N-1} f'[n] \exp(-j2k\pi n / (2N)) + \sum_{n=N}^{2N-1} f'[n] \exp(-j2k\pi n / (2N)) \\ &= \sum_{n=0}^{N-1} f[n] \exp(-j2k\pi n / (2N)) + \sum_{n=N}^{2N-1} (c - f[n - N]) \exp(-j2k\pi n / (2N)) \end{aligned}$$

Isolando o termo constante no segundo somatório vem:

$$\begin{aligned} F'[k] &= c \sum_{n=N}^{2N-1} \exp(-j2k\pi n / (2N)) + \sum_{n=0}^{N-1} f[n] \exp(-j2k\pi n / (2N)) \\ &\quad - \sum_{n=N}^{2N-1} f[n - N] \exp(-j2k\pi n / (2N)) \end{aligned}$$

Pode-se agora separar essa função em duas partes para facilitar sua avaliação , isto é, definindo:

$$S_1[k] = c \sum_{n=N}^{2N-1} \exp(-j2k\pi n / (2N)) \quad (B1)$$

e

$$S_2[k] = \sum_{n=0}^{N-1} f[n] \exp(-j2k\pi n / (2N)) - \sum_{n=N}^{2N-1} f[n - N] \exp(-j2k\pi n / (2N)) \quad (B2)$$

Sabendo que a equação que fornece o resultado do somatório de uma progressão geométrica de razão constante é:

$$\sum_{n=a}^b \alpha^n = \frac{\alpha^a - \alpha^{b+1}}{1 - \alpha} \quad \text{para } b \geq a$$

obtém-se para a Eq. (B1):

$$S_1[k] = c \frac{\exp(-j2k\pi / 2) - \exp(-j2k\pi)}{1 - \exp(-j2k\pi / (2N))} = c \frac{\exp(-jk\pi) - \exp(-j2k\pi)}{1 - \exp(-jk\pi / N)}$$

E colocando os fatores apropriados em evidência resulta:

$$\begin{aligned} S_1[k] &= c \frac{\exp(-j3k\pi / 2)(\exp(jk\pi / 2) - \exp(-jk\pi / 2))}{\exp(-jk\pi / (2N))(\exp(jk\pi / (2N)) - \exp(-jk\pi / (2N)))} \\ &= c \frac{\exp(-jk\pi(3 - 1/N) / 2) \sin(k\pi / 2)}{\sin(k\pi / (2N))} \end{aligned}$$

Mas como $\sin(k\pi/2) = 0$ para k par, $S_1[k]$ resulta zero para valores pares de k com $k \neq 0$. Quando k vale zero há uma indeterminação em $S_1[k]$, que deve ser resolvida usando a regra de l'Hôpital:

$$S_1[k] = c \frac{\exp(-jk(3 - 1/N) / 2) \frac{d}{dk} \sin(k\pi / 2)}{\frac{d}{dk} \sin(k\pi / (2N))} = c \frac{\exp(-jk(3 - 1/N) / 2) \cos(k\pi / 2) \pi / 2}{\cos(k\pi / (2N)) \pi / (2N)}$$

para $k = 0$, e portanto:

$$S_1[0] = c \frac{\exp(-j0) \cos(0) N}{\cos(0)} = cN$$

Voltando à Eq. (B2) e fazendo $p = n - N$ no segundo somatório de $S_2[k]$ vem:

$$\begin{aligned}
S_2[k] &= \sum_{n=0}^{N-1} f[n] \exp(-j2k\pi n / (2N)) - \sum_{p=0}^{N-1} f[p] \exp(-j2k\pi(p+N) / (2N)) \\
&= \sum_{n=0}^{N-1} f[n] \exp(-j2k\pi n / (2N)) - \sum_{p=0}^{N-1} f[p] \exp(-jk\pi) \exp(-j2k\pi p / (2N))
\end{aligned}$$

Substituindo $\exp(-jk\pi)$ por $(-1)^k$ e fazendo $n = p$ pode-se unir os dois somatórios em:

$$S_2[k] = \sum_{n=0}^{N-1} f[n] (1 - (-1)^k) \exp(-j2k\pi n / (2N))$$

Mas como $(1 - (-1)^k) = 0$ para k par, então $S_2[k] = 0$ para k par.

E como $F[k] = S_1[k] + S_2[k]$ por formação, resulta $F[k] = 0$ para valores pares de k e $k \neq 0$.

Apêndice C

Transformada de Hotelling

Enquanto as outras transformadas decompõem as seqüências de dados de entrada em famílias de funções base, a transformada de Hotelling, também conhecida como transformada Karhunen-Loève discreta e método das componentes principais, baseia-se em propriedades estatísticas dos dados para descorrelacioná-los. Portanto é dependente destes últimos e não pode ser obtida por um algoritmo rápido no caso geral. Mas é ótima no sentido de minimização do erro quadrático médio da reconstrução parcial (ou compactação de energia nas componentes baixas) e produz uma seqüência de saída com média nula e elementos descorrelacionados. É reversível, isto é, não induz perda de informação e é utilizada para compressão de dados e avaliação do desempenho de outros algoritmos. É obtida por:

$$Y = A(X - M_x) \quad (C1)$$

e sua inversa por:

$$X = A^{-1}Y + M_x \quad (C2)$$

onde X é um vetor coluna formado pelos dados, $M_x = E\{X\}$ (esperança matemática dos dados), $R = E\{(X - M_x)(X - M_x)^T\}$ (matriz de covariância dos dados), A é uma matriz cujas linhas são os autovetores normalizados da matriz R ordenados na forma decrescente de seus respectivos autovalores, e Λ é uma matriz diagonal cujos elementos são os autovalores de R em ordem crescente, de forma que:

$$AR = \Lambda A \quad (C3)$$

Para efeitos práticos, pode-se utilizar as seguintes aproximações:

$$M_x \cong \frac{1}{N} \sum_{i=0}^{N-1} X_i \quad (C4)$$

onde cada X_i é um vetor coluna do conjunto de amostras, N é o número de amostras e

$$R \cong \frac{1}{N} \sum_{i=0}^{N-1} (X_i - M_x)(X_i - M_x)^T \quad (C5)$$

onde X_i e M_x são vetores coluna e o símbolo T significa transposto.

Os autovalores são iguais às variâncias dos elementos da transformada da sequência de entrada na direção dos seus respectivos autovetores, e pode ser demonstrado que o erro quadrático médio entre o conjunto de dados inicial e sua reconstrução a partir da transformada com alguns elementos desprezados é igual à soma dos autovalores relativos a esses elementos omitidos (ver Gonzalez e Woods (1993) e Jain (1989)). Isto é, quanto menor um autovalor, menor importância tem o elemento da transformada relativo a ele porque o erro quadrático médio introduzido na reconstrução parcial devido à omissão desse elemento é igual ao seu autovalor. Este é o critério de decisão utilizado para a seleção das componentes significativas da transformada de Hotelling.

Apêndice D

Neste apêndice são apresentadas amostras de folíolos terminais de plantas das quatro espécies utilizadas.

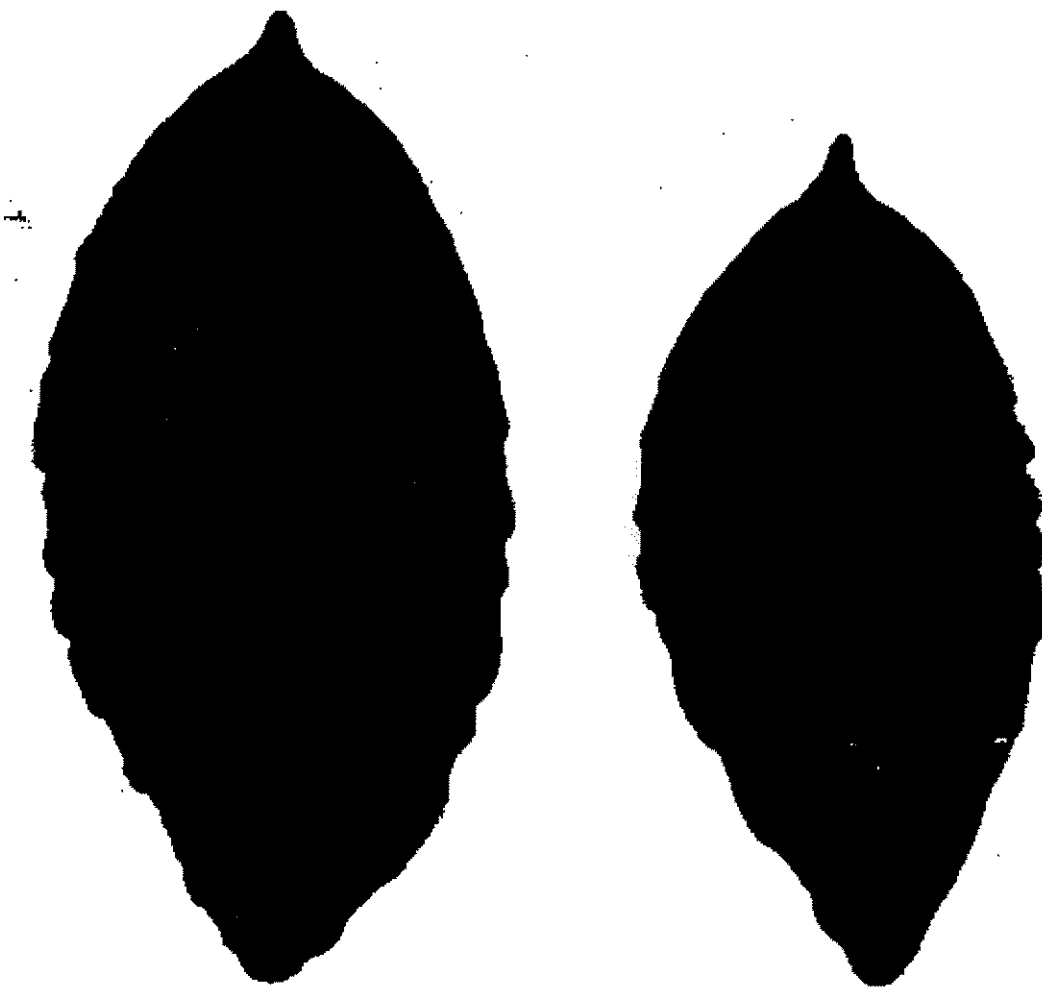


Figura D-1: dois folíolos terminais de plantas da espécie *Lonchocarpus Latifolius*.



Figura D-2: três folíolos terminais de plantas da espécie *Lonchocarpus Campestris*.

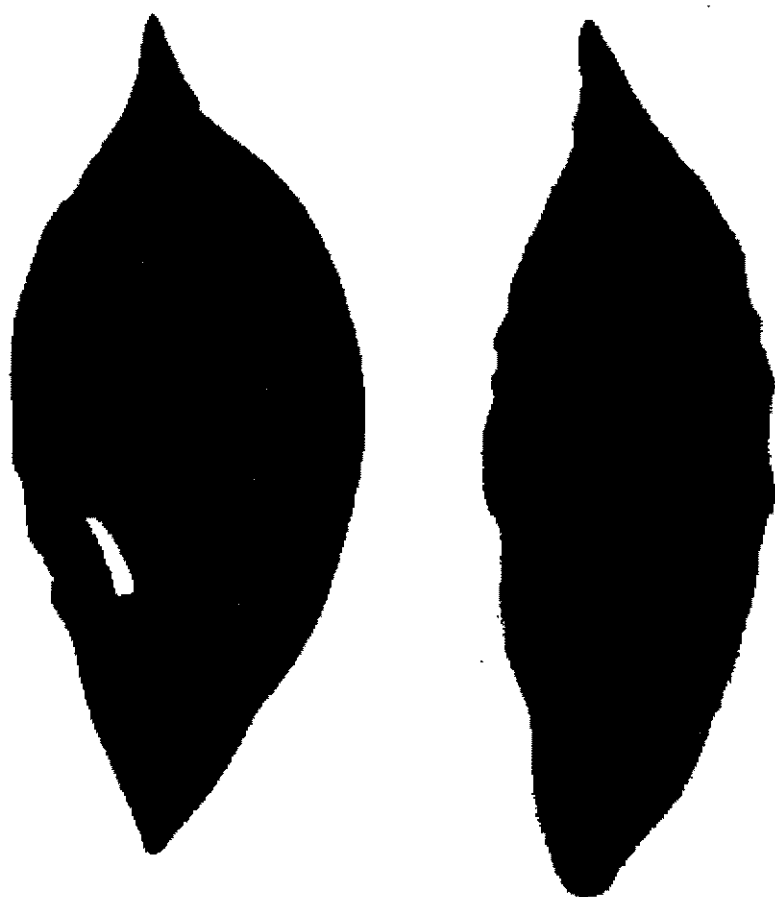


Figura D-3: dois folíolos terminais de plantas da espécie *Lonchocarpus Muehlbergianus*.



Figura D-4: dois folíolos terminais de plantas da espécie *Lonchocarpus Guilleminianus*.