

**CARLOS RAMON P. DIONISIO**

**CARACTERÍSTICAS DE FORMAS PLANAS  
INVARIANTES SOB TRANSFORMAÇÕES DE  
SIMILARIDADE E AFIM**

Tese apresentada à Escola Politécnica da  
Universidade de São Paulo para obtenção do  
Título de Doutor em Engenharia.

São Paulo  
2005

**CARLOS RAMON P. DIONISIO**

**CARACTERÍSTICAS DE FORMAS PLANAS  
INVARIANTES SOB TRANSFORMAÇÕES DE  
SIMILARIDADE E AFIM**

Tese apresentada à Escola Politécnica da  
Universidade de São Paulo para obtenção do  
Título de Doutor em Engenharia.

Área de Concentração:  
Sistemas Eletrônicos

Orientador:  
Prof. Dr. Hae Yong Kim

São Paulo  
2005

## FICHA CATALOGRÁFICA

Dionisio, Carlos Ramon Pantaleon

Características de Formas Planas Invariantes sob Transformações de Similaridade e Afim/ C. R. P. Dionisio. São Paulo, 2005.

93 p.

Tese (Doutorado) — Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Sistemas Eletrônicos.

1. Reconhecimento de Formas. 2. Características Invariante. 3. Parametrização Invariante. I. Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Sistemas Eletrônicos. II. t.

# AGRADECIMENTOS

Primeiramente, ao Deus Eterno, porque Ele sempre está ao meu lado nos bons e maus momentos, dando-me força e perseverança de continuar no bom caminho.

Ao Prof. Dr. Hae Yong Kim, pela paciência e excelente orientação demonstrado durante a realização deste trabalho.

Aos meus pais Felipe e Felipa; meus irmãos José, Alex, Heber e Yaella; muito obrigado pelo apoio e carinho.

À minha família, Silvanira, Katherine e Kymberly.

Aos meus amigos do laboratório, pela amizade e apoio demonstrados durante a realização deste trabalho. Em especial para Harold e Javier pela amizade demonstrada.

Aos professores e funcionários do IME-USP e EP-USP, pela formação e serviços prestados.

Finalmente, a todos que de alguma forma contribuíram a realização deste trabalho.

# RESUMO

Esta tese contribui para a solução do problema de reconhecimento de formas planas através do uso de características invariantes locais e globais. Especificamente, propomos um conjunto de características invariantes sob transformações de similaridade e um outro conjunto de características invariantes sob transformações afins.

As características do primeiro conjunto, invariantes sob transformações de similaridade, são simples descritores de formas porém eficazes no reconhecimento. As transformações de similaridade englobam as transformações de rotação, translação e escala uniforme. Estas características são baseadas na curvatura euclidiana e em propriedades do contorno da forma (área, perímetro e distância). O classificador de distância mínima usa estas características para o reconhecimento de formas invariante sob transformações de similaridade. Uma técnica baseada em aproximação poligonal é proposta e usada para diminuir o número de pontos do contorno da forma, com a finalidade de reduzir o tempo de processamento.

As características do segundo conjunto, invariantes sob transformações afins, são a nossa principal contribuição. As transformações afins englobam as transformações de rotação, translação, escala (uniforme e não-uniforme) e cisalhamento (*shear*). Sob certas condições, a mudança do ponto de vista da câmera pode ser aproximada por uma transformação afim. Este conjunto de características permite reconhecer objetos mesmo se eles são observados de diferentes pontos de vista. Estas características são baseadas nas áreas dos triângulos formados por três pontos pertencentes ao contorno do objeto. O algoritmo de  $k$ -vizinhos mais próximos usa estas características para o reconhecimento de objetos invariante sob transformações afins. Para este problema, parametrizamos o contorno dos objetos através do parâmetro comprimento de arco afim.

Também estudamos a importância da correta parametrização do contorno dos objetos sobre a invariância das características. Especificamente, mostramos a importância da parametrização sobre o reconhecimento de objetos invariante sob transformações afins.

Para poder medir o poder de discriminação e a robustez dos dois conjuntos de características, aplicamo-los ao problema de reconhecimento supervisionado de gestos estáticos de mão.

# ABSTRACT

This thesis contributes to the solution of the planar shape recognition problem using local and global invariant features. Specifically, we propose a set of features invariant under similarity transformations, and another invariant under affine transformations.

The features of the first set, invariant under similarity transformations, are simple shape descriptors but efficient for recognition. The similarity transformations include rotation, translation and uniform scaling. These features are based on the Euclidean curvature and simple properties of the shape (area, perimeter and distance). The minimum distance classifier uses these features for the recognition of shapes invariant under similarity transformations. A technique based on polygonal approximation is proposed and used to reduce the number of points of the shape contour with the purpose of reducing the processing time.

The features of the second set, invariant under affine transformation, are our main contribution. Affine transformations include rotation, translation, scaling (uniform and no uniform) and shearing. Under certain conditions, the change of the camera's point of view can be approached by an affine transformation. This set of features allows recognizing objects even if they are observed from different points of view. These features are based on the areas of the triangles formed by three points of the contour of the object. The k-nearest neighbors algorithm uses these features for object recognition invariant under affine transformation. For this problem, we parameterize the shape contour of objects using the affine length parameter.

We also discuss the importance of the correct parameterization on the invariance of the features. Specifically, we show the importance of the parameterization on the object recognition invariant under affine transformations.

In order to measure the discrimination power and the robustness of the two sets of features, we apply them to the problem of supervised recognition of static hand gestures.

# SUMÁRIO

## Lista de Figuras

## Lista de Tabelas

<b>1</b>	<b>Introdução</b>	<b>14</b>
1.1	Exposição do Cenário . . . . .	14
1.2	Contexto Científico . . . . .	16
1.3	Objetivos . . . . .	18
1.4	Abordagem Proposta . . . . .	18
1.5	Contribuições . . . . .	19
1.6	Publicações Associadas a esta Tese . . . . .	20
1.7	Conteúdo dos Capítulos . . . . .	20
<b>2</b>	<b>Descrição Invariante das Formas Planas sob Movimento no Espaço <math>\mathbb{R}^3</math></b>	<b>22</b>
2.1	Introdução . . . . .	22
2.2	Descrição das Formas . . . . .	24
2.2.1	Definição de Forma . . . . .	24
2.2.2	Contorno de Formas . . . . .	24
2.3	Representação de Curvas no Plano . . . . .	25
2.3.1	Representação Paramétrica das Curvas . . . . .	25

2.3.1.1	Descrição Paramétrica de uma Curva Plana . . . . .	25
2.3.1.2	Equivalência Geométrica de Duas Curvas Paramétricas . . . . .	25
2.3.1.3	Curva Paramétrica Fechada . . . . .	26
2.4	Transformação Geométrica através do Grupo Afim Geral . . . . .	26
2.4.1	Hipótese de Aproximação . . . . .	27
2.4.2	Definição Topológica de uma Transformação Afim Geral . . . . .	29
2.4.2.1	Grupo Afim Geral: $AG(2)$ . . . . .	29
2.4.2.2	Grupo Afim Especial: $AE(2)$ . . . . .	29
2.4.2.3	Decomposição da Matriz Afim no Caso do Grupo $AE(2)$ . . . . .	29
2.4.2.4	Transformação Euclidiana . . . . .	31
2.4.2.5	Transformação sob o Grupo de Similaridades . . . . .	31
2.4.2.6	Decomposição da Matriz Afim no Caso do Grupo $AG(2)$ . . . . .	31
2.4.2.7	Resumo sobre a Transformação Afim . . . . .	32
2.5	Parametrização Invariante . . . . .	33
2.5.1	Derivadas Discretas . . . . .	34
2.5.1.1	Diferenças Progressivas . . . . .	34
2.5.1.2	Diferenças Atrasadas . . . . .	35
2.5.1.3	Diferenças Centrais . . . . .	35
2.5.2	Mudança da Parametrização de uma Curva . . . . .	35



2.5.3	Parametrização pelo Comprimento de Arco Normalizado . . .	36
2.5.4	Parametrização pelo Comprimento de Arco Afim Normalizado	37
2.5.5	Implementação da Parametrização Invariante . . . . .	38
2.6	Características Invariantes . . . . .	40
2.6.1	Características Invariantes Globais . . . . .	41
2.6.2	Características Invariantes Locais . . . . .	41
2.7	Seleção de Características . . . . .	43
2.8	Conclusão . . . . .	44
<b>3</b>	<b>Reconhecimento de Formas Invariantes sob Transformações de Similari-</b>	
	<b>dade</b>	<b>45</b>
3.1	Introdução . . . . .	45
3.2	O Método . . . . .	47
3.2.1	Características . . . . .	51
3.2.2	Algoritmo de Treinamento . . . . .	52
3.2.3	Algoritmo de Classificação . . . . .	52
3.3	Resultados Experimentais . . . . .	53
3.4	Conclusões . . . . .	55
<b>4</b>	<b>Reconhecimento de Formas Invariantes sob Transformações Afins</b>	<b>59</b>
4.1	Introdução . . . . .	59
4.2	Conceitos Fundamentais . . . . .	61
4.2.1	Transformação Afim . . . . .	62

4.2.2	Área e Transformação Afim . . . . .	62
4.2.3	Parametrização Afim . . . . .	64
4.3	Extração de Informações Locais e Globais de uma Matriz de Áreas . .	65
4.3.1	Matriz de Áreas . . . . .	67
4.3.2	Matriz de Áreas Normalizada . . . . .	69
4.3.3	Extração de Informações Locais e Globais . . . . .	69
4.4	Características e Algoritmo de Classificação . . . . .	71
4.4.1	Características Propostas . . . . .	71
4.4.2	Análise das Características . . . . .	73
4.4.3	Algoritmo de Treinamento . . . . .	76
4.4.4	Algoritmo de Classificação . . . . .	77
4.5	Resultados Experimentais . . . . .	77
4.5.1	Formas Deformadas . . . . .	78
4.5.2	Formas Não-Deformadas . . . . .	80
4.6	Conclusões . . . . .	82
<b>5</b>	<b>Conclusões</b>	<b>84</b>
	<b>Referências</b>	<b>86</b>
	<b>Apêndice A - Área de um Polígono</b>	<b>91</b>
A.1	Área de um Triângulo . . . . .	91
A.2	Área de um Polígono . . . . .	91



# LISTA DE FIGURAS

1	Representação cartesiana do contorno de uma forma. . . . .	27
2	(a) Câmera de Furo; (b) Modelo da Câmera de Furo. . . . .	28
3	Alongamento de uma forma de $\mathbb{R}^2$ . . . . .	30
4	Transformações afins do contorno de uma mão. . . . .	32
5	Divisão do Grupo Afim Geral. . . . .	32
6	Exemplos de transformação afim. . . . .	33
7	Parte do contorno de uma forma. . . . .	39
8	Parâmetros invariantes divididos em partes iguais. . . . .	40
9	Duas re-parametrizações do contorno ilustrado na Figura 7. . . . .	41
10	Algumas re-amostragens de uma curva: (a) contorno original com parametrização arbitrária (89 pontos), (b) parametrização pelo comprimento de arco normalizado (200 pontos), e (c) parametrização pelo comprimento de arco afim normalizado (200 pontos). . . . .	42
11	Curvatura do contorno ilustrado na Figura 12 e os limites superior e inferior para o cálculo dos pontos significantes. . . . .	50
12	Os pontos sobre o contorno pertencem a aproximação poligonal determinados pelo método proposto. . . . .	50
13	Contornos de gestos de mão das classes A, B, C, D, E e F. . . . .	53

14	Quatro contornos de consulta manualmente etiquetados como sendo da classe A. Os gestos 1, 2 e 3 foram classificados por nosso algoritmo como A. O gesto 4 foi classificado como C. . . . .	55
15	Quatro contornos de consulta manualmente etiquetados como sendo da classe B. Os gestos 1, 2 e 3 foram classificados por nosso algoritmo como B. O gesto 4 foi classificado como C. . . . .	56
16	Quatro contornos de consulta manualmente etiquetados como sendo da classe C. Os gestos 1, 2 e 3 foram classificados por nosso algoritmo como C. O gesto 4 foi classificado como B. . . . .	56
17	Quatro contornos de consulta manualmente etiquetados como sendo da classe D. Todos eles foram classificados por nosso algoritmo como D.	57
18	Quatro contornos de consulta manualmente etiquetados como sendo da classe E. Os gestos 1, 2 e 3 foram classificados por nosso algoritmo como E. O gesto 4 foi classificado como B. . . . .	58
19	Quatro contornos de consulta manualmente etiquetados como sendo da classe F. Os gestos 1, 2 e 3 foram classificados por nosso algoritmo como F. O gesto 4 foi classificado como E. . . . .	58
20	Cálculo da matriz de áreas. . . . .	68
21	Implementação de uma curva gaussiana. . . . .	70
22	Configuração das gaussianas $g_e, g_c$ e $g_d$ para uma aplicação em particular.	71
23	A parametrização da forma é robusta à escolha do ponto inicial. . . .	73
24	Vetores $\psi_e, \psi_c$ e $\psi_d$ de uma mesma forma com parametrizações diferentes. . . . .	74
25	Vetores $\psi_e, \psi_c$ e $\psi_d$ de duas formas similares com parametrização pelo comprimento de arco afim. . . . .	75

26	Vetores $\psi_e$ , $\psi_c$ e $\psi_d$ sob transformação de cisalhamento. . . . .	76
27	Gestos de mão das classes A, B, C, D, E e F. . . . .	78
28	Vetores $\psi_e$ , $\psi_c$ e $\psi_d$ para cada uma das seis classes de gestos da mão. .	79
29	O gesto na parte superior esquerda pertence à classe F. Os outros são as formas deformadas com $\kappa = 1$ . . . . .	80
30	O gesto na parte superior esquerda pertence à classe F. Os outros são as formas deformadas com $\kappa = 2$ . . . . .	81
31	O gesto na parte superior esquerda pertence à classe F. Os outros são as formas deformadas com $\kappa = 3$ . . . . .	81
32	Alguns contornos de consultas não-deformados e as suas classificações.	83

# LISTA DE TABELAS

1	As transformações contidas nos grupos Euclidiano, Similaridade e Afim Geral; e as invariantes sob estes grupos. . . . .	33
2	Classificação dos dados de consulta. . . . .	54
3	Classificação dos dados de treinamento. . . . .	54
4	Configuração das três gaussianas ( $m$ é o número de colunas de $S$ ). . .	71
5	As características são robustas à escolha do ponto inicial. . . . .	73
6	Classificação dos contornos de consulta. . . . .	82

# 1 INTRODUÇÃO

Este capítulo faz uma exposição do tema e motivação desta tese. É feita uma breve apresentação do contexto científico, dos objetivos a serem atingidos, da metodologia a ser considerada, das contribuições deste trabalho e das publicações associadas a esta tese. Finalmente, apresenta-se a estrutura geral da tese.

## 1.1 Exposição do Cenário

O trabalho apresentado nesta tese pertence ao domínio do reconhecimento de objetos por computador. Trata-se de um problema fundamental e muito amplo da visão computacional.

Este domínio compreende outros problemas muito variados que vão desde a aquisição de imagens, segmentação, descrição e reconhecimento de objetos, até a interpretação da imagem.

A visão computacional não pretende diretamente reproduzir a visão humana em um computador, mas antes procura encontrar algoritmos que permitam chegar a resultados próximos ao humano. Desta forma, a visão computacional tem como objetivo final usar o computador para emular a visão humana, incluindo a fase do aprendizado e sendo capaz de fazer inferências e tomar ações baseadas nas entradas visuais.

A visão computacional pode ser dividida em três tipos de processos computacionais: processo de nível baixo, intermediário e alto ([GONZALEZ; WOODS, 2002](#)).



**1) Nível Baixo:** este tipo de processo compreende aquisição e pré-processamento de imagens.

A aquisição permite representar as imagens de maneira interna no computador, através de uma matriz de valores, chamados pixels: cada pixel codifica a intensidade das cores (para imagens coloridas) ou a intensidade luminosa (para imagens em níveis de cinza).

Esta informação é o resultado de um processo óptico e eletrônico cujo objetivo é transformar o sinal óptico, captado pelo ponto de vista da câmera ou do sistema óptico do scanner, numa imagem digitalizada, registrado nessa matriz de valores (chamado daqui a diante de imagem).

O pré-processamento de imagens pode consistir, por exemplo, de redução de ruído ou borramento, realce do contraste e aumento da nitidez da imagem.

**2) Nível Intermediário:** este tipo de processo compreende a segmentação, representação, descrição e reconhecimento.

A segmentação consiste na partição da imagem em regiões ou objetos. A representação consiste na forma como representamos os objetos contidos na imagem. Neste trabalho, representamos os objetos através do seu contorno. Desta forma, reduzimos consideravelmente a informação presente na imagem.

A descrição consiste no cálculo das características dos objetos observados. O reconhecimento é o processo que atribui um rótulo a um objeto presente na imagem, baseado na informação fornecida pelas suas características.

O processo de nível intermediário é caracterizado pelo fato que as entradas geralmente são imagens e as saídas são atributos extraídos das imagens (por exemplo, o contorno e a identidade de cada objeto).

**3) Nível Alto:** este tipo de processo compreende a interpretação da imagem.

A interpretação consiste em atribuir um significado a uma imagem, numa linguagem verbal e não geométrica ou matemática. Este processo, também chamado de compreensão de imagens ou análise de cenas, usa toda a informação gerada pelos dois processos anteriores em uma tentativa de interpretar o conteúdo de uma imagem.

## 1.2 Contexto Científico

Devido ao fato de que a forma é uma propriedade fundamental de um objeto, uma peça chave para a descrição do objeto é encontrar um conjunto de características da forma o qual descreva e discrimine o objeto da melhor forma possível. As aplicações de métodos de descrição de forma podem ser encontradas em muitas áreas, tais como reconhecimento, recuperação e indexação de imagens.

Os métodos de descrição de forma podem ser divididos em três categorias principais: características baseadas em contorno, área e esqueleto ([LATECKI; LAKÄMPER; WOLTER, 2003](#)).

**1) Características baseadas em contorno:** as características são extraídas a partir da representação do contorno da forma.

Características baseadas em contorno da forma exploram somente a informação do contorno e não a informação do interior da forma. Como exemplos temos os descritores de Fourier ([ARBTER; SNYDER; BURKHARDT, 1990](#)), espaço-escala da curvatura ([MOKHTARIAN; ABBASI, 2001](#); [MOKHTARIAN; ABBASI, 2002](#)) e os descritores de momentos ([FLUSSER; SUK, 1993](#); [HUANG; COHEN, 1994](#); [ZHAO; CHEN, 1997](#)).

**2) Características baseadas em área:** as características são extraídas considerando todos os pixels contidos na forma.

Geralmente, os métodos baseados em área usam descritores de momento para descrever a forma. Estes métodos incluem momentos geométricos (MARTINEZ; THOMAS, 2002), momentos de Legendre (AHMADIAN; FARAMARZI; SAYADIAN, 2003), momentos de Zernike (KHOTANZAD; HONG, 1990; KIM et al., 2000), pseudo-momentos de Zernike (HADDADNIA; AHMADI; FAEZ, 2002) e um método baseado em grade (CHAKRABARTI et al., 2000). Estes métodos de descrição da forma exploram a informação contida na forma e não apenas a informação contida no contorno da forma.

**3) Características baseadas em esqueleto:** as características são baseadas no esqueleto da forma. O esqueleto é um grafo que descreve a informação topológica da forma. O esqueleto é um descritor atrativo porque captura a noção das partes e dos componentes da forma. A similaridade de duas formas pode ser calculada através de algoritmos de casamento de grafos (GOLD; RANGARAJAN, 1996). Alguns exemplos destes métodos encontram-se em (LIN; KUNG, 1996; SIDDIQI et al., 1999; SUNDAR et al., 2003).

Gostaríamos de saber a qual dos três métodos de descrição da forma, mencionados acima, pertence o descritor com melhor desempenho sobre um mesmo conjunto de dados de entrada.

Para o caso de descritores invariantes sob transformações de similaridade, o trabalho de Latecki (LATECKI; LAKÄMPER; ECKHARDT, 2000) compara os melhores descritores de cada uma das três categorias sobre um mesmo conjunto de dados. Como resultado, dois descritores baseados em contorno ( (MOKHTARIAN; ABBASI; KITTLER, 1997) e (LATECKI; LAKÄMPER, 2000) ) obtiveram os melhores desempenhos. Recentemente, o descritor chamado de BAS (Beam Angle Statistics) (ARICA; VURAL, 2003) mostrou ter o melhor desempenho sobre o mesmo conjunto de dados.

Considerando os resultados sobre o desempenho dos métodos de descrição da

forma, o método proposto nesta tese pertence à primeira categoria (características baseadas em contorno). Para esta categoria, uma boa descrição dos objetos depende fortemente da escolha das características invariantes.

### 1.3 Objetivos

O objetivo desta tese é contribuir na solução do problema de reconhecimento de objetos através do uso de características invariantes locais e globais. Especificamente, propomos dois conjuntos de características invariantes sob transformação de similaridade e transformação afim, respectivamente. Os objetivos do presente trabalho são:

1. Descrever alguns fundamentos teóricos da Geometria Diferencial Euclidiana e Afim, relevantes para o trabalho.
2. Desenvolver características invariantes sob transformação de similaridade e transformação afim, para o reconhecimento de formas.
3. Estudar a importância da escolha da parametrização do contorno dos objetos sobre a invariância das características.
4. Implementar os algoritmos para o reconhecimento de objetos invariante sob transformação de similaridade e transformação afim.

### 1.4 Abordagem Proposta

O trabalho descrito nesta tese situa-se no processo de visão computacional de nível intermediário. Especificamente, tratamos da representação, descrição e reconhecimento dos objetos.

O objetivo deste trabalho é poder realizar tarefas como reconhecimento, recuperação ou indexação de imagens baseados no contorno dos objetos.

A nossa abordagem assume uma segmentação prévia das imagens em contornos dos objetos. Os objetos podem estar sujeitos a transformações de similaridade ou a transformação afim (os objetos podem ser observados de diferentes pontos de vista). Também consideramos a presença de ruído no contorno dos objetos.

A abordagem que detalhamos na tese, para o problema de reconhecimento de objetos invariante sob transformação de similaridade e afim, descansa sobre dois conjuntos de características local e global invariante sob transformação de similaridade e afim, respectivamente. Estes conjuntos de características servirão de medida de semelhança entre os objetos.

## 1.5 Contribuições

Esta tese contém as seguintes contribuições principais:

- Propomos um conjunto de características invariantes sob transformações de similaridade e um método para a aproximação poligonal do contorno do objeto (DIONISIO; KIM, 2002). Estas características são baseadas em propriedades do contorno da forma (O'HIGGINS, 1997; COSTA; CESAR, 2001). A aproximação poligonal do contorno da forma é usada para diminuir o número de pontos do contorno, com a finalidade de reduzir o tempo de processamento.
- Propomos um conjunto de características invariantes sob transformações afins (DIONISIO; KIM, 2004a; DIONISIO; KIM, 2004b). Uma das características está relacionada com a deficiência convexa e as outras são extraídas a partir da matriz de áreas. A matriz de áreas foi usada por Shen (SHEN; WONG; IP, 1999) para o casamento por similaridade na recuperação de imagens. Diferentemente do trabalho de Shen, re-parametrizamos o contorno das formas usando um parâmetro invariante afim. Isto torna as nossas características robustas a parametrização da forma, enquanto que o trabalho de Shen não possui esta propriedade.

- Para cada conjunto de características mencionado acima, realizamos uma aplicação ao problema de reconhecimento supervisionado de gestos estáticos de mão.

## 1.6 Publicações Associadas a esta Tese

Esta tese está baseada no resultado das seguintes publicações:

1. No artigo (DIONISIO; KIM, 2002) definimos um conjunto de características para o problema de reconhecimento de formas invariante sob transformações de similaridade e um método para a aproximação poligonal do contorno do objeto. Este conjunto de características foi escolhido considerando as características descritas em (O'HIGGINS, 1997; COSTA; CESAR, 2001). Os resultados experimentais em reconhecimento de gestos estáticos de mão mostram o desempenho do método proposto.
2. Nos artigos (DIONISIO; KIM, 2004a; DIONISIO; KIM, 2004b) definimos um conjunto de características para o problema de reconhecimento de formas invariante sob transformações afins. Uma das características definidas está relacionada com a deficiência convexa e as outras são extraídas da matriz de áreas. A matriz de áreas foi usada por Shen (SHEN; WONG; IP, 1999) para o casamento por similaridade na recuperação de imagens. Os resultados experimentais indicam que as nossas características podem classificar corretamente formas deformadas e ruidosas usando um pequeno conjunto de treinamento.

## 1.7 Conteúdo dos Capítulos

Esta tese é composta de cinco capítulos.

No Capítulo 2, reunimos as considerações teóricas para a descrição invariante das curvas planas. Mostramos como o movimento de um objeto plano pode ser aproximado

através de transformações de similaridade ou afim. Também, descrevemos a descrição invariante das formas planas sob estas transformações.

No Capítulo 3, propomos um conjunto de características invariantes sob transformação de similaridade. Uma técnica baseada na aproximação poligonal do contorno do objeto também é proposto. Este conjunto de características é usado para o reconhecimento de gestos estáticos de mão e os resultados experimentais são apresentados.

No Capítulo 4, propomos um conjunto de características invariantes sob transformação afim. Estas características são baseadas na área do triângulo formado por três pontos pertencentes ao contorno do objeto. Este conjunto de características também é usado para o reconhecimento de gestos estáticos de mão e os resultados experimentais são apresentados.

No Capítulo 5, apresentamos as nossas conclusões.

## 2 DESCRIÇÃO INVARIANTE DAS FORMAS PLANAS SOB MOVIMENTO NO ESPAÇO $\mathbb{R}^3$

Neste capítulo, mostramos como modelar o movimento das formas planas observados no plano de uma câmera. Especificamente, estamos interessados em modelar tipos de movimentos que possam ser aproximados pelas seguintes transformações: transformações de similaridade e transformações afins. O grupo das transformações de similaridade é um subgrupo do grupo das transformações afins.

### 2.1 Introdução

As transformações de similaridade consistem de translações, rotações e mudança de escala uniforme. Este tipo de transformação descreve o movimento das formas observadas num plano paralelo ao plano do movimento.

As transformações afins consistem de translações, rotações, mudanças de escala (uniformes e não-uniformes) e cisalhamento (shear). Este tipo de transformação descreve o movimento tridimensional de um objeto rígido planar, observado numa projeção em eixos paralelos. Sob certas condições, este tipo de movimento pode ser interpretado como mudanças do ponto de vista da câmera.

Nesta tese, as formas são representadas pelo seu contorno. Especificamente, as formas são representadas através de curvas paramétricas no plano.

As transformações de similaridade ou afim atuam sobre as curvas paramétricas mudando, assim, a descrição geométrica das curvas. Porém, os objetos continuam



sendo os mesmos. Estas curvas podem ser descritas através de características invariantes locais e globais.

O objetivo deste capítulo, que é mostrar como modelar o movimento das formas planas, repousa sobre a descrição invariante, paramétrica, local e global das curvas planas sob uma deformação, no plano de observação de uma câmera. Esta deformação das curvas pode ser descrita por uma transformação de similaridade ou afim. Para cumprir este objetivo, utilizaremos a abordagem baseada na geometria diferencial. Especificamente, para o caso das transformações de similaridade (transformações afins) utilizaremos as ferramentas da geometria diferencial euclidiana (geometria diferencial afim).

Portanto, neste capítulo reunimos as considerações teóricas para a descrição invariante das curvas planas, o qual constitui a base desta tese. A maioria das definições e conceitos contidos neste capítulo foram extraídas da tese de M. Stanislaw ([STANISLAW, 1999](#)).

Este capítulo se compõe de sete seções.

Nas Seções 2.2 e 2.3 definimos as formas planas, observadas por uma câmera, como uma função paramétrica de  $\mathbb{R}^2$ .

Na Seção 2.4 definimos as transformações lineares de  $\mathbb{R}^2$  e mostramos as condições da aplicação de estas para a aproximação do tipo de movimento dos objetos.

Nas Seções 2.5 e 2.6, apresentamos a parametrização e descrição invariante de curvas planas em função de um parâmetro invariante.

Na Seção 2.7 apresentamos algumas considerações relacionadas ao processo de seleção de características.

Na Seção 2.8 apresentamos as conclusões deste capítulo.

## 2.2 Descrição das Formas

### 2.2.1 Definição de Forma

A noção de forma foi introduzida por vários autores de acordo com objetivos diferentes (ARBTER, 1989; GAUTHIER; BORNARD; SILBERMANN, 1991; MOKHTARIAN; MACKWORTH, 1992). No caso de uma análise geométrica, a forma de um objeto pode ser caracterizada por pontos, linhas, curvas, planos e assim por diante. Especificamente, a forma de um objeto descreve a sua característica independente do ponto de vista da câmera ou da transformação geométrica aplicada a esta forma. Em outras palavras, se dois objetos possuem a mesma forma então existe uma aplicação que transforma uma forma na outra.

Do ponto de vista matemático, a noção de forma é definida em relação a um grupo de transformações (GHORBEL, 1992).

**Definição 1.** *Seja  $G$  um grupo de transformações que agem sobre o espaço dos objetos  $\mathbb{O}$  e sejam  $O_1$  e  $O_2$  dois elementos de  $\mathbb{O}$ . Diz-se que  $O_1$  e  $O_2$  têm a mesma  $G$ -forma ( $O_1 \equiv_G O_2$ ) se e somente se existe uma transformação  $F \in G$  tal que  $O_2 = F(O_1)$ .*

Como consequência da definição anterior, a relação definida entre todos os objetos de  $\mathbb{O}$  é uma relação de equivalência. Portanto, os elementos de  $\mathbb{O}$  admitem uma partição formada pelas suas classes de equivalência.

Por conseguinte, uma forma pertencente ao conjunto  $\mathbb{O}$  é uma classe de equivalência relativa ao grupo de transformações  $G$  e o conjunto quociente  $\mathbb{O}/G$  representa o conjunto das formas.

### 2.2.2 Contorno de Formas

Em algumas aplicações, particularmente em processamento de imagens, a informação que representa qualquer objeto se resume muitas vezes no seu contorno. Por-

tanto, o contorno dos objetos ocupam um lugar muito importante na percepção e interpretação das imagens através do sistema visual humano.

Do ponto de vista matemático, podemos definir o contorno de um objeto como sendo uma fronteira entre duas regiões de níveis de cinza diferentes e relativamente homogêneos.

## 2.3 Representação de Curvas no Plano

Neste trabalho, qualquer objeto é representado pelo seu contorno e descrito por uma função paramétrica de  $\mathbb{R}^2$ .

### 2.3.1 Representação Paramétrica das Curvas

Representamos as formas através do seu contorno. Ao longo deste trabalho consideramos somente os contornos planos, fechados e simples, com exceção do Capítulo 3 onde é considerado contornos abertos. Mostramos abaixo algumas definições matemáticas relacionadas com a representação das formas.

#### 2.3.1.1 Descrição Paramétrica de uma Curva Plana

**Definição 2.** *Uma curva paramétrica plana é uma aplicação  $\mathbf{u} : I \rightarrow \mathbb{R}^2$  onde  $I$  é um intervalo de  $\mathbb{R}$ .*

Vamos supor que a curva paramétrica  $\mathbf{u}$  é de classe  $C^2$  (a primeira e segunda derivada de  $\mathbf{u}$  são contínuas).

#### 2.3.1.2 Equivalência Geométrica de Duas Curvas Paramétricas

**Definição 3.** *Duas curvas paramétricas  $\mathbf{u} : I \rightarrow \mathbb{R}^2$  e  $\bar{\mathbf{u}} : \bar{I} \rightarrow \mathbb{R}^2$  são ditas geometricamente equivalentes se existe duas funções  $\phi : I \rightarrow \bar{I}$  e  $\psi : \bar{I} \rightarrow I$  tal que  $\mathbf{u} = \bar{\mathbf{u}} \circ \phi$  e  $\bar{\mathbf{u}} = \mathbf{u} \circ \psi$ .*

A relação anterior define uma relação de equivalência. Denotamos por  $[\mathbf{u}]$  a classe de equivalência de  $\mathbf{u}$  para esta relação. Como consequência da definição anterior, uma mesma curva pode ter diferentes parametrizações. Mais ainda, estas parametrizações mudam sob transformações lineares.

Portanto, é necessário escolher uma parametrização apropriada para resolver satisfatoriamente um problema de reconhecimento de formas invariante a um grupo de transformações  $\mathbf{G}$ . A idéia seria descrever as curvas através de um mesmo tipo de parametrização invariante ao grupo de transformações  $\mathbf{G}$ .

### 2.3.1.3 Curva Paramétrica Fechada

**Definição 4.** *Uma curva paramétrica  $\mathbf{u} : I = [a, b] \rightarrow \mathbb{R}^2$  é fechada se  $\mathbf{u}(a) = \mathbf{u}(b)$ .*

Para representar qualquer curva paramétrica de  $\mathbb{R}^2$ , escolhemos a representação cartesiana por ser o caso mais geral.

Se  $\mathbf{u} : [0, T] \rightarrow \mathbb{R}^2$  é uma curva paramétrica fechada, de período  $T$ , então:

$$\mathbf{u}(t) = (x(t), y(t)) \quad \text{para } t \in [0, T].$$

A Figura 1 ilustra a construção da representação paramétrica de  $\mathbf{u}$  nas coordenadas cartesianas. No resto deste capítulo, consideramos somente curvas fechadas e simples.

## 2.4 Transformação Geométrica através do Grupo Afim Geral

O movimento tridimensional de um objeto plano pode ser aproximado através de transformações afins do objeto. A seguir, definimos a hipótese para poder realizar esta aproximação.

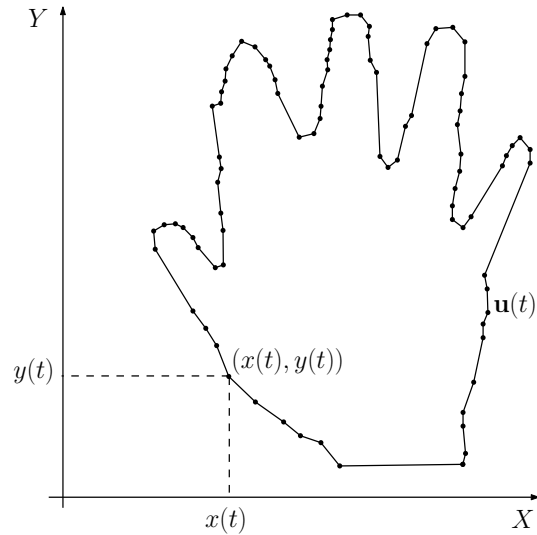


Figura 1: Representação cartesiana do contorno de uma forma.

### 2.4.1 Hipótese de Aproximação

Se o objeto plano e a câmera estiverem suficientemente distantes então as transformações afins modelam bem o movimento tridimensional do objeto observado no plano da câmera. Esta aproximação tem como base o modelo da Câmera de Furo (*pinhole camera*) (BALLARD; BROWN, 1982; HORN, 1986; NALWA, 1993) que está ilustrado na Figura 2b. O contorno visível de um objeto visto pela câmera é determinado pelas linhas tangentes à superfície do objeto que tem início no centro óptico da câmera. Na Figura 2b,  $d$  representa a distância focal do lente ao objetivo da câmera;  $(X, Y, Z)$  são as coordenadas do objeto filmado no espaço  $\mathbb{R}^3$  e  $(x, y)$  são as coordenadas do objeto ligadas à câmera no espaço  $\mathbb{R}^2$ .

No modelo da Câmera de Furo, o plano da câmera (plano  $xy$ ) situa-se entre o centro da projeção  $O$  e o objeto filmado (esta representação é simbólica e permite não ter a imagem invertida do objeto).

Para representar o objeto observado no plano da câmera em função das suas coordenadas  $(X, Y, Z)$ , ajustamos o centro  $o$  do plano da câmera com o centro da projeção  $O$  tal que  $Z = d$ .

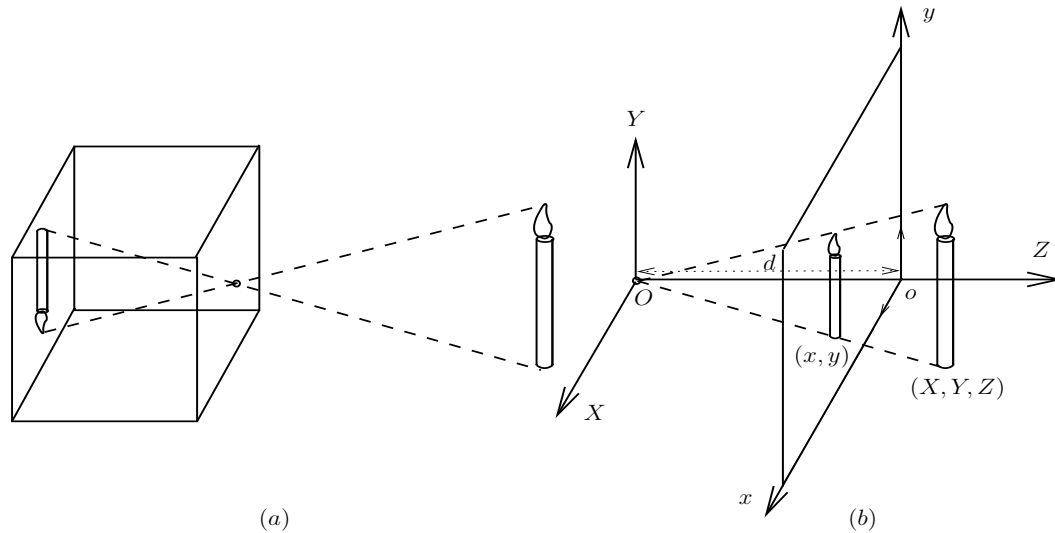


Figura 2: (a) Câmera de Furo; (b) Modelo da Câmera de Furo.

Assim, no plano da imagem (plano da câmera), podemos escrever as coordenadas  $(x, y)$  como:

$$x = \frac{d}{Z}X, \quad y = \frac{d}{Z}Y.$$

Se o objeto é plano e pequeno em comparação com a distância entre o objeto e a câmera, ou se for observado perpendicularmente em relação ao plano do objeto, então a mudança na coordenada  $Z$  dos pontos do contorno será insignificante. Neste caso, a coordenada  $Z$  é interpretada como uma constante. Escrevendo  $x = mX$  e  $y = mY$ , com  $m = \frac{d}{Z}$ , obtemos o modelo da transformação afim.

Se  $\mathbf{u}$  e  $\bar{\mathbf{u}}$  são as projeções de uma mesma forma em dois instantes diferentes então elas são vinculadas através de uma transformação afim. Isto nos leva a existência de uma matriz não singular  $\mathbf{A}$  e de um vetor de translação  $\mathbf{T}$  os quais definem os parâmetros da transformação. As novas coordenadas do contorno são descritas em função das antigas coordenadas pela seguinte expressão:

$$\bar{\mathbf{u}} = \mathbf{A}\mathbf{u} + \mathbf{T}.$$

Esta aproximação é válida unicamente quando a câmera está suficientemente distante em relação à dimensão do objeto (PAUWELS et al., 1995).

Do ponto de vista geométrico, a transformação afim em  $\mathbb{R}^2$  transforma círculos em elipses e retângulos em paralelogramos; somente o triângulo permanece sendo um triângulo.

## 2.4.2 Definição Topológica de uma Transformação Afim Geral

A transformação afim geral em  $\mathbb{R}^2$  é representada por uma matriz  $\mathbf{A}$   $2 \times 2$  inversível e um vetor de translação  $\mathbf{T}$ . Daqui em diante chamaremos a tal matriz  $\mathbf{A}$  de *matriz afim*.

**Definição 5.** *Seja  $\mathbf{u}(t) = (x(t), y(t))$ ,  $t \in [0, T]$ , uma curva paramétrica em  $\mathbb{R}^2$ . Uma transformação afim  $\mathbf{F} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  atuando sobre  $\mathbf{u}$  é uma aplicação tal que:*

$$\mathbf{F}(\mathbf{u}) = \mathbf{A}\mathbf{u} + \mathbf{T} \quad (2.1)$$

onde  $\mathbf{A}$  é uma matriz  $2 \times 2$  inversível e  $\mathbf{T} \in \mathbb{R}^2$  é o vetor de translação.

### 2.4.2.1 Grupo Afim Geral: $AG(2)$

O Grupo Afim Geral  $AG(2)$  é o conjunto de todas as transformações afins em  $\mathbb{R}^2$ .

### 2.4.2.2 Grupo Afim Especial: $AE(2)$

O Grupo Afim Especial  $AE(2)$  está formado por todas as transformações afins pertencentes ao grupo  $AG(2)$  com determinante da matriz afim associada igual a um. As transformações deste grupo não alteram a área da forma (veja a Subseção 4.2.2).

### 2.4.2.3 Decomposição da Matriz Afim no Caso do Grupo $AE(2)$

A matriz afim  $\mathbf{A}$  de uma transformação  $\mathbf{F} \in AE(2)$  pode ser decomposta como um produto matricial de duas transformações simples:

$$\mathbf{A} = \mathbf{R}_\theta \mathbf{M}_{ab},$$

onde  $\mathbf{R}_\theta$  e  $\mathbf{M}_{ab}$  denotam uma matriz de rotação e uma matriz de alongamento (*stretching*), respectivamente. Especificamente, estas matrizes se definem como:

$$\mathbf{R}_\theta = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \quad \mathbf{M}_{ab} = \begin{bmatrix} a & b \\ 0 & 1/a \end{bmatrix}$$

onde  $\theta \in [0, 2\pi[$ ,  $a \in \mathbb{R}_+$  e  $b \in \mathbb{R}$ .

Esta decomposição não é única ( $R_\theta M_{ab} = M_{cd} R_\gamma$ ), mas pode tornar-se como tal escolhendo a ordem das transformações simples. Pode-se supor, por exemplo, que as transformações do grupo  $AE(2)$  se compõem de uma rotação seguida de alongamento.

A matriz de cisalhamento (*shearing*) é um caso particular da matriz de alongamento quando  $a = 1$ . Esta matriz é definida da seguinte forma:

$$\mathbf{A}_\kappa = \begin{bmatrix} 1 & \kappa \\ 0 & 1 \end{bmatrix} \quad (2.2)$$

onde o parâmetro  $\kappa$ , chamado de taxa de deformação (shear ratio), controla a medida da deformação.

A Figura 3 ilustra um exemplo de alongamento na direção  $X$ . Os contornos (b) e (c) são os resultados da aplicação da matriz  $\mathbf{M}_{11}$  e  $\mathbf{M}_{12}$ , respectivamente, sobre o contorno (a).

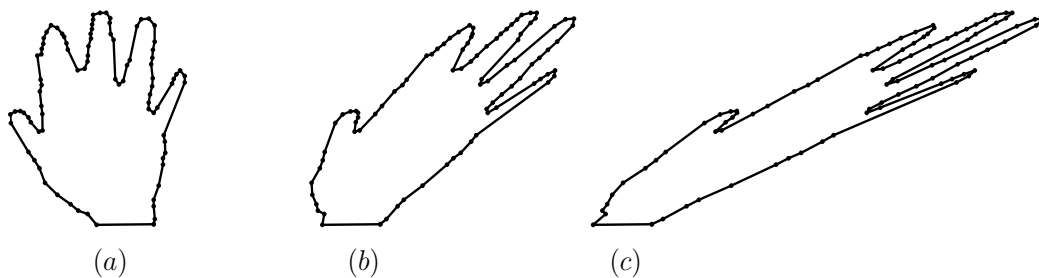


Figura 3: Alongamento de uma forma de  $\mathbb{R}^2$ .



#### 2.4.2.4 Transformação Euclidiana

Se a matriz afim for uma matriz de rotação ( $a = 1, b = 0$ ) então a transformação pertence ao grupo das transformações euclidianas em  $\mathbb{R}^2$ . As transformações euclidianas preservam o comprimento e o ângulo entre os pontos da forma.

#### 2.4.2.5 Transformação sob o Grupo de Similaridades

Se a matriz afim é composta de uma matriz de rotação e uma matriz de homotetia (com  $\alpha > 0$ ) então a transformação pertence ao grupo de similaridades em  $\mathbb{R}^2$ . A matriz de homotetia  $H_\alpha$  é definida como:

$$H_\alpha = \begin{bmatrix} \alpha & 1 \\ 1 & \alpha \end{bmatrix}$$

onde  $\alpha \neq 0$  é o fator de homotetia.

#### 2.4.2.6 Decomposição da Matriz Afim no Caso do Grupo $AG(2)$

As transformações do Grupo  $AG(2)$  constitui o caso geral das transformações lineares no plano. Neste caso, a decomposição da matriz afim é a seguinte:

$$\mathbf{A} = \mathbf{R}_\theta \mathbf{M}_{ab} \mathbf{H}_\alpha$$

onde  $\mathbf{H}_\alpha$  é uma matriz de homotetia. Esta decomposição é única, supondo que a matriz de rotação seja seguida pela matriz de alongamento.

A Figura 4 ilustra um exemplo de transformação sob o grupo  $AG(2)$ . Os contornos (b) e (c) são os resultados da aplicação da transformação afim com matrizes  $\mathbf{A}_1$  e  $\mathbf{A}_2$ , respectivamente, sobre o contorno (a).

$$\mathbf{A}_1 = \begin{bmatrix} 1.6 & -0.16 \\ 0.06 & 0.61 \end{bmatrix} \quad \mathbf{A}_2 = \begin{bmatrix} 1.2 & 0.5 \\ 0 & 0.833 \end{bmatrix}$$

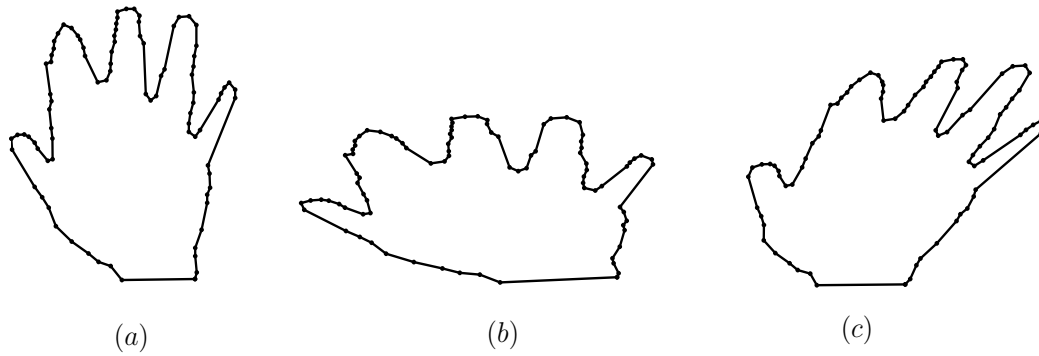


Figura 4: Transformações afins do contorno de uma mão.

#### 2.4.2.7 Resumo sobre a Transformação Afim

Em seguida, descrevemos as relações entre os diferentes grupos de transformações apresentados.

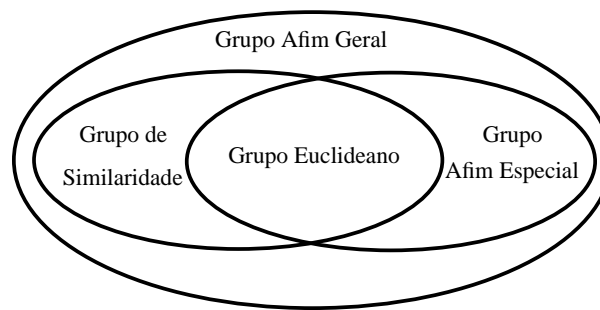


Figura 5: Divisão do Grupo Afim Geral.

O Grupo Euclidiano (composto de rotações e translações) está incluído não somente no Grupo Afim Especial (composto de rotações, translações e alongamento, com determinante da matriz afim igual a um) mas também no Grupo de Similaridades (composto de rotações, translações e homotetias).

Como ilustrado na Figura 5, estes três grupos pertencem ao Grupo Afim Geral, o qual compreende todas as transformações lineares no plano. Estas transformações são representadas por uma matriz afim inversível (composta de rotações, alongamento e homotetias) e um vetor de translação. A Figura 6 ilustra alguns exemplos de transformação afim.

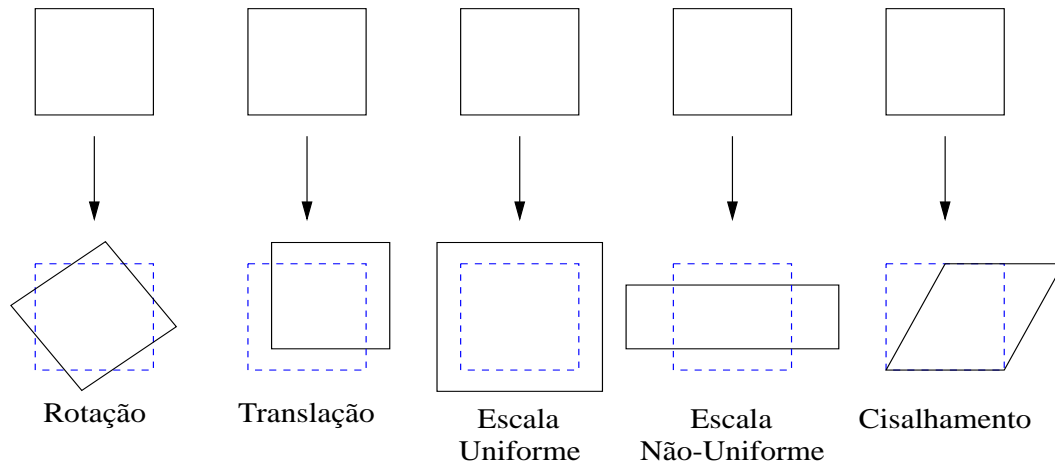


Figura 6: Exemplos de transformação afim.

A Tabela 1 mostra as transformações contidas nos grupos euclidiano, de similaridade e afim. As invariantes sobre cada grupo também são mostrados.

Tabela 1: As transformações contidas nos grupos Euclidiano, Similaridade e Afim Geral; e as invariantes sob estes grupos.

Transformações e Invariantes	Euclidiano	Similaridade	Afim Geral
Transformações			
rotação	✓	✓	✓
translação	✓	✓	✓
escala uniforme		✓	✓
escala não-uniforme			✓
cisalhamento			✓
Invariantes			
comprimento	✓		
ângulo	✓	✓	
razão de comprimentos	✓	✓	
paralelismo	✓	✓	✓

## 2.5 Parametrização Invariante

Como consequência da Definição 3, uma mesma curva pode ter diferentes parametrizações. Mais ainda, duas parametrizações diferentes de uma mesma curva podem ter valores diferentes da primeira derivada (velocidade), segunda derivada (aceleração), etc. Além disso, esta parametrização pode ser alterada pela ação de qualquer

transformação.

Portanto, devemos re-parametrizar as curvas por um mesmo tipo de parametrização que dependa somente da curva e seja invariante à ação de qualquer transformação.

**Definição 6.** *Seja  $\mathbf{G}$  um subgrupo de  $AG(2)$  e  $\mathbf{u} : I \rightarrow \mathbb{R}^2$  uma curva paramétrica. O conjunto  $\{\mathbf{F}(\mathbf{u}), \mathbf{F} \in \mathbf{G}\}$  representa uma família de curvas com a mesma  $\mathbf{G}$ -forma.*

## 2.5.1 Derivadas Discretas

Nesta subseção, mostramos brevemente três formas de calcular numericamente uma aproximação da primeira e da segunda derivada de uma função. Uma referência mais detalhada e completa encontrasse em (FORTUNA, 2000).

Estas aproximações são baseadas nas equações de diferenças finitas da função os quais são obtidas a partir da expansão em série de Taylor da função a ser aproximada.

Seja  $\mathbf{u} : [0, T] \rightarrow \mathbb{R}^2$  é uma curva paramétrica fechada. Lembremos que temos somente a seqüência de pontos  $\mathbf{u}(0), \mathbf{u}(1), \dots, \mathbf{u}(T)$  a partir dos quais devemos calcular as derivadas.

### 2.5.1.1 Diferenças Progressivas

A primeira aproximação para a primeira e segunda derivada de  $\mathbf{u}$  é obtida utilizando-se diferenças progressivas (forward differences):

$$\begin{aligned}\dot{\mathbf{u}}(t) &\approx \frac{\mathbf{u}(t+h) - \mathbf{u}(t)}{h} \\ \ddot{\mathbf{u}}(t) &\approx \frac{\mathbf{u}(t+2h) - 2\mathbf{u}(t+h) + \mathbf{u}(t)}{h^2}\end{aligned}$$

onde o valor de  $h$  é constante e na prática considera-se  $h = 1$ .

### 2.5.1.2 Diferenças Atrasadas

A segunda aproximação para a primeira e segunda derivada de  $\mathbf{u}$  é obtida utilizando-se diferenças atrasadas (backward differences):

$$\begin{aligned}\dot{\mathbf{u}}(t) &\approx \frac{\mathbf{u}(t) - \mathbf{u}(t - h)}{h} \\ \ddot{\mathbf{u}}(t) &\approx \frac{\mathbf{u}(t) - 2\mathbf{u}(t - h) + \mathbf{u}(t - 2h)}{h^2}\end{aligned}$$

onde o valor de  $h$  é constante e na prática considera-se  $h = 1$ .

### 2.5.1.3 Diferenças Centrais

A terceira aproximação para a primeira e segunda derivada de  $\mathbf{u}$  é obtida utilizando-se diferenças centrais (central differences):

$$\begin{aligned}\dot{\mathbf{u}}(t) &\approx \frac{\mathbf{u}(t + h) - \mathbf{u}(t - h)}{2h} \\ \ddot{\mathbf{u}}(t) &\approx \frac{\mathbf{u}(t + h) - 2\mathbf{u}(t) + \mathbf{u}(t - h)}{h^2}\end{aligned}$$

onde o valor de  $h$  é constante e na prática considera-se  $h = 1$ .

## 2.5.2 Mudança da Parametrização de uma Curva

Seja  $\mathbf{u} : I \rightarrow \mathbb{R}^2$  uma curva paramétrica e  $\mathbf{G}$  um grupo de transformações.

Seja  $s_{\mathbf{u}} : I \rightarrow I_0$  o novo parâmetro inversível através do qual a curva  $\mathbf{u}$  será re-parametrizada. Esta mudança de parâmetro deve satisfazer o seguinte:

- A curva re-parametrizada  $\mathbf{u} \circ s_{\mathbf{u}}^{-1} : I_0 \rightarrow \mathbb{R}^2$  depende somente da classe de equivalência  $[\mathbf{u}]$  (Definição 3). Isto traduz a invariância da representação geométrica da curva por mudança de parâmetro.
- Se  $\mathbf{F} \in \mathbf{G}$  então os parâmetros  $s_{\mathbf{u}}$  e  $s_{\mathbf{F}(\mathbf{u})}$  coincidem. Isto traduz a invariância do parâmetro  $s_{\mathbf{u}}$  pela ação da transformação  $\mathbf{F}$  sobre a curva  $\mathbf{u}$ .

### 2.5.3 Parametrização pelo Comprimento de Arco Normalizado

Como sabemos, as transformações euclidianas conservam o comprimento entre os pontos da curva. Baseado nesta propriedade, pode-se construir uma parametrização invariante sob transformação de similaridade a partir do comprimento de arco da curva.

Seja  $\mathbf{u}(t) = (x(t), y(t))$ ,  $t \in [0, T]$ , uma curva paramétrica, de período  $T$ , em  $\mathbb{R}^2$ .

O comprimento de arco entre os pontos  $\mathbf{u}(0)$  e  $\mathbf{u}(p)$  pode ser calculado, a partir da parametrização inicial  $t$ , pela seguinte fórmula:

$$s(p) = \int_0^p |\dot{\mathbf{u}}(t)| dt = \int_0^p (\dot{x}^2(t) + \dot{y}^2(t))^{\frac{1}{2}} dt \quad (2.3)$$

onde  $\dot{\mathbf{u}}$ ,  $\dot{x}$  e  $\dot{y}$  denotam a primeira derivada de  $\mathbf{u}$ ,  $x$  e  $y$ , respectivamente, em relação ao parâmetro  $t$ . O comprimento total da curva é  $L = s(T)$ . O parâmetro  $s$  é invariante sob rotação e translação da curva  $\mathbf{u}$ . Portanto, para obter invariância em relação à mudança de escala uniforme, definimos o comprimento de arco normalizado como  $\bar{s}(p) = \frac{s(p)}{L}$  (MOKHTARIAN; MACKWORTH, 1992).

A seguinte proposição define uma curva re-parametrizada através do parâmetro comprimento de arco normalizado.

**Proposição 1.** *Para qualquer forma, com uma parametrização inicial arbitrária, podemos extrair o parâmetro comprimento de arco normalizado  $\bar{s}$ . Seja  $\mathbf{u}(t)$  a descrição paramétrica de uma forma parametrizada pelo parâmetro  $t \in [0, T]$ . A descrição paramétrica  $\mu(\bar{s}) : [0, 1] \rightarrow \mathbb{R}^2$  definida pela equação  $\mu(\bar{s}) = \mathbf{u}(\bar{s}^{-1})$  é a representação de  $\mathbf{u}(t)$  re-parametrizada pelo parâmetro  $\bar{s}$ .*

Esta re-parametrização da curva pode ser vista como uma amostragem equidistante dos pontos da curva, no sentido euclidiano.

Após a re-parametrização da forma através do parâmetro  $\bar{s}$ , temos que:

- As imagens de  $\mathbf{u}(t)$  e  $\mu(\bar{s})$  coincidem.

- O parâmetro  $\bar{s}$  é invariante em relação às transformações do grupo de similaridades.
- A curva re-parametrizada pelo parâmetro  $\bar{s}$  verifica a equação  $\left| \frac{d\mu}{d\bar{s}} \right| = L$ .

**Significado Físico:** Suponhamos que o parâmetro  $\bar{s}$  representa o tempo e que  $\mu(\bar{s})$  é a trajetória percorrida por um ponto material. Então  $\left| \frac{d\mu}{d\bar{s}} \right| = L$  significa que o módulo do vetor que representa a velocidade é constante sobre toda a trajetória.

#### 2.5.4 Parametrização pelo Comprimento de Arco Afim Normalizado

Na geometria diferencial afim, pode-se verificar que o comprimento da reta é nulo (porque sobre a ação do grupo  $AE(2)$  a reta pode ser transformada num ponto). Conseqüentemente, não podemos utilizar o comprimento de arco como parâmetro invariante sob o grupo  $AG(2)$ .

Denotaremos por  $\dot{\mathbf{u}} = (\dot{x}, \dot{y})$  e  $\ddot{\mathbf{u}} = (\ddot{x}, \ddot{y})$  como sendo a primeira e segunda derivada de  $\mathbf{u}$ , respectivamente.

O comprimento de arco afim pode ser calculado, a partir da parametrização inicial  $t$ , pela seguinte expressão:

$$\begin{aligned} s_a(p) &= \int_0^p |\text{Det}(\dot{\mathbf{u}}(t), \ddot{\mathbf{u}}(t))|^{\frac{1}{3}} dt \\ &= \int_0^p |\dot{x}(t)\ddot{y}(t) - \dot{y}(t)\ddot{x}(t)|^{\frac{1}{3}} dt \end{aligned} \quad (2.4)$$

O comprimento de arco afim total da curva é  $L_a = s_a(T)$ . O comprimento de arco afim normalizado  $\bar{s}_a(p)$  é definido como  $\bar{s}_a(p) = \frac{s_a(p)}{L_a}$  (FLUSSER; SUK, 1993; ZHAO; CHEN, 1997; SAPIRO; TANNENBAUM, 1993).

A seguinte proposição define uma curva re-parametrizada através do parâmetro comprimento de arco afim normalizado.

**Proposição 2.** *Para qualquer forma, com uma parametrização inicial arbitrária, podemos extrair o parâmetro comprimento de arco afim normalizado  $\bar{s}_a$ . Seja  $\mathbf{u}(t)$  a descrição paramétrica de uma forma parametrizada pelo parâmetro  $t \in [0, T]$ . A descrição paramétrica  $\mu(\bar{s}_a) : [0, 1] \rightarrow \mathbb{R}^2$  definida pela equação  $\mu(\bar{s}_a) = \mathbf{u}(\bar{s}_a^{-1})$  é a representação de  $\mathbf{u}(t)$  re-parametrizada pelo parâmetro  $\bar{s}_a$ .*

Esta re-parametrização pode ser visto como uma amostragem equidistante dos pontos da curva, no sentido afim.

Após a re-parametrização da forma através do parâmetro  $\bar{s}_a$ , temos que:

- A curva não altera o seu aspecto externo: as imagens de  $\mathbf{u}(t)$  e  $\mu(\bar{s}_a)$  coincidem.
- O parâmetro  $\bar{s}_a$  é invariante em relação ao grupo  $AG(2)$ .
- A curva re-parametrizada pelo parâmetro  $\bar{s}_a$  verifica a seguinte equação:  

$$|\text{Det}(\dot{\mu}(\bar{s}_a), \ddot{\mu}(\bar{s}_a))| = L_a^3.$$

**Significado físico:** Suponha que o parâmetro  $\bar{s}_a$  representa o tempo e que  $\mu(\bar{s}_a)$  é a trajetória percorrida por um ponto material. Então  $|\text{Det}(\dot{\mu}(\bar{s}_a), \ddot{\mu}(\bar{s}_a))| = L_a^3$  significa que a área do paralelogramo entre o vetor velocidade e o vetor aceleração é constante sobre toda a trajetória. No caso discreto da curva  $\mu$ , esta propriedade significa que a área do triângulo formado por três pontos consecutivos pertencentes a  $\mu$  é constante.

### 2.5.5 Implementação da Parametrização Invariante

Nesta subseção, descrevemos os passos principais para a implementação da parametrização invariante de formas. Especificamente, estes passos se resumem em:

1. dado o contorno  $\mathbf{u}(t)$ , calcular o parâmetro invariante  $s(p)$ ;
2. dado o número de pontos para a re-parametrização, divida o parâmetro  $s(p)$  em  $\Delta s$  partes iguais;



3. interpolar  $p$  para obter  $\bar{p}$  tal que  $s(\bar{p})$  seja uma função dividida em  $\Delta s$  partes iguais;
4. interpolar o contorno  $\mathbf{u}(t)$  nos pontos  $s(\bar{p})$  para obter a re-parametrização do contorno.

Na interpolação, pode-se usar qualquer método conhecido: linear, cúbico, spline, etc. Neste trabalho escolhemos o método linear.

Observe que o contorno re-parametrizado adquire uma propriedade baseada na propriedade do parâmetro  $s(\bar{p})$ . No caso do parâmetro comprimento de arco normalizado, os pontos consecutivos sobre o contorno são equidistantes. No caso do parâmetro comprimento de arco afim normalizado, a área do triângulo formado por três pontos consecutivos sobre o contorno é constante.

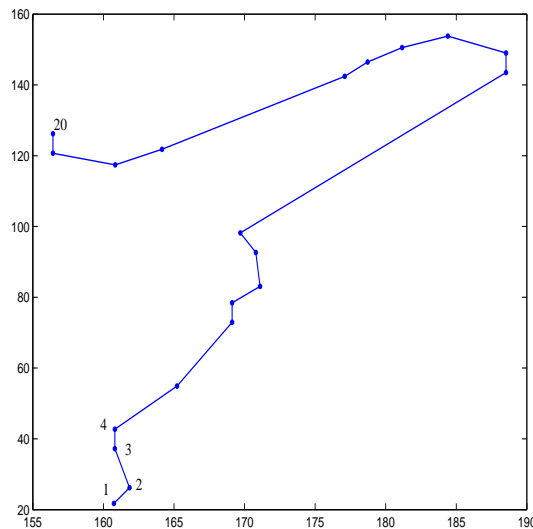


Figura 7: Parte do contorno de uma forma.

Considere o contorno ilustrado na Figura 7 como o contorno a ser re-parametrizado pelos parâmetros comprimento de arco normalizado e comprimento de arco afim normalizado. Este contorno consiste de 20 pontos ( $t = 1, 2, \dots, 20$ ). A Figura 8 ilustra estes parâmetros divididos em partes iguais.

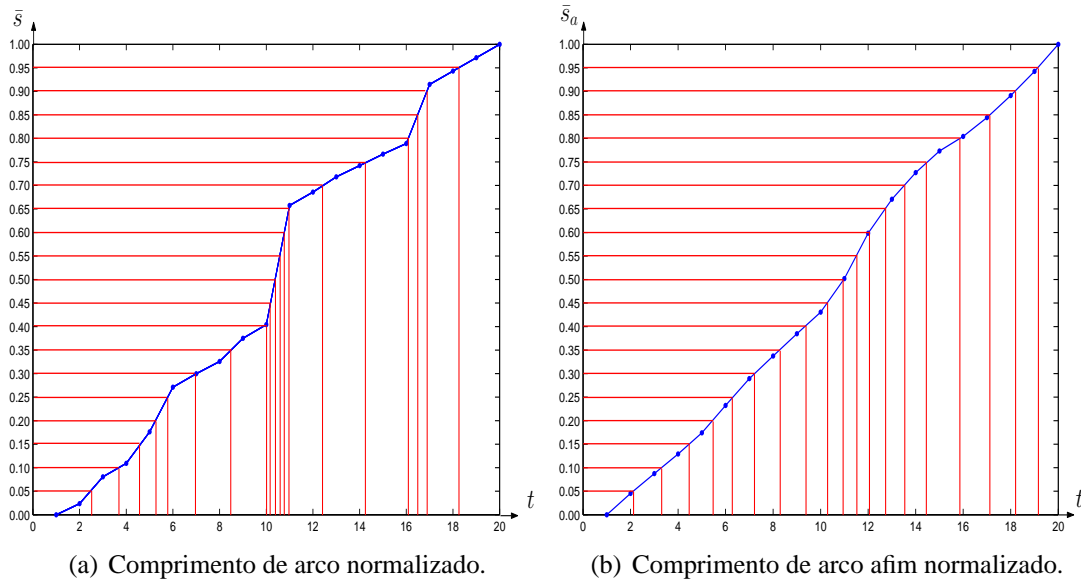


Figura 8: Parâmetros invariantes divididos em partes iguais.

A Figura 9 ilustra as re-parametrizações do contorno da Figura 7. Especificamente, a Figura 9(a) (Figura 9(b)) ilustra a re-parametrização do contorno pelo parâmetro comprimento de arco normalizado (comprimento de arco afim normalizado). No caso da re-parametrização pelo comprimento de arco normalizado, os pontos sobre o contorno são bastante equidistantes. Porém, o contorno perdeu informação sobre detalhes pequenos. No caso da re-parametrização pelo comprimento de arco afim normalizado, a propriedade do triângulo com área constante não se cumpre bem para todos os pontos. Isto é devido a que o cálculo do parâmetro invariante afim envolve o uso de derivadas de segundo ordem e como conseqüência o ruído é maior. Porém, o contorno não perdeu informação relevante.

Para mostrar o resultado da implementação da parametrização sobre um contorno maior, a Figura 10 ilustra as re-parametrizações de um contorno de gesto da mão.

## 2.6 Características Invariantes

É óbvio que quando um objeto muda de posição e de orientação ele mantém a sua forma. Geralmente, a noção de invariância pode ser definida em relação a um grupo

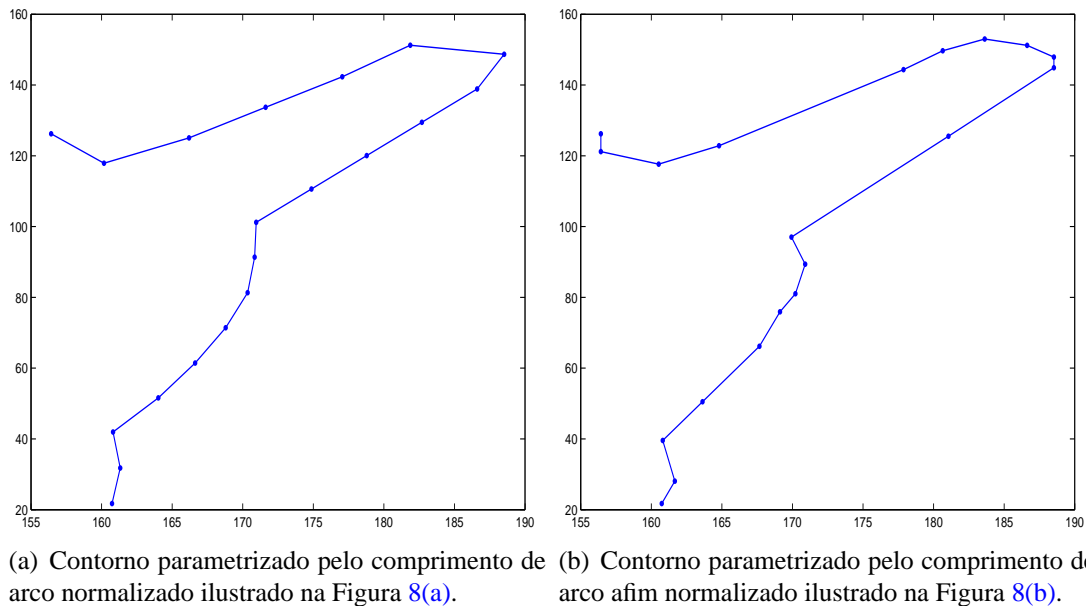


Figura 9: Duas re-parametrizações do contorno ilustrado na Figura 7.

de transformações (MOONS et al., 1995; WOOD, 1996).

As características invariantes podem ser classificadas em dois tipos: características invariantes globais e locais.

### 2.6.1 Características Invariantes Globais

Este tipo de características descrevem o contorno do objeto em forma global (eles são apropriados para contornos fechados). Por exemplo, no espaço de frequências temos os descritores de Fourier (ARBTER, 1989; KWOK; POON, 1996; YU; LO, 2001; CHAKER; BANNOUR; GHORBEL, 2003). Estas características são calculadas usando integrais ou somas. Também estas características podem ser baseadas sobre momentos (ZHAO; CHEN, 1997; YINAN; WEIJUN; YUECHAO, 2003).

### 2.6.2 Características Invariantes Locais

Também conhecidos como características diferenciais. Este tipo de características contém informação local do contorno da forma e permitem estabelecer uma descrição

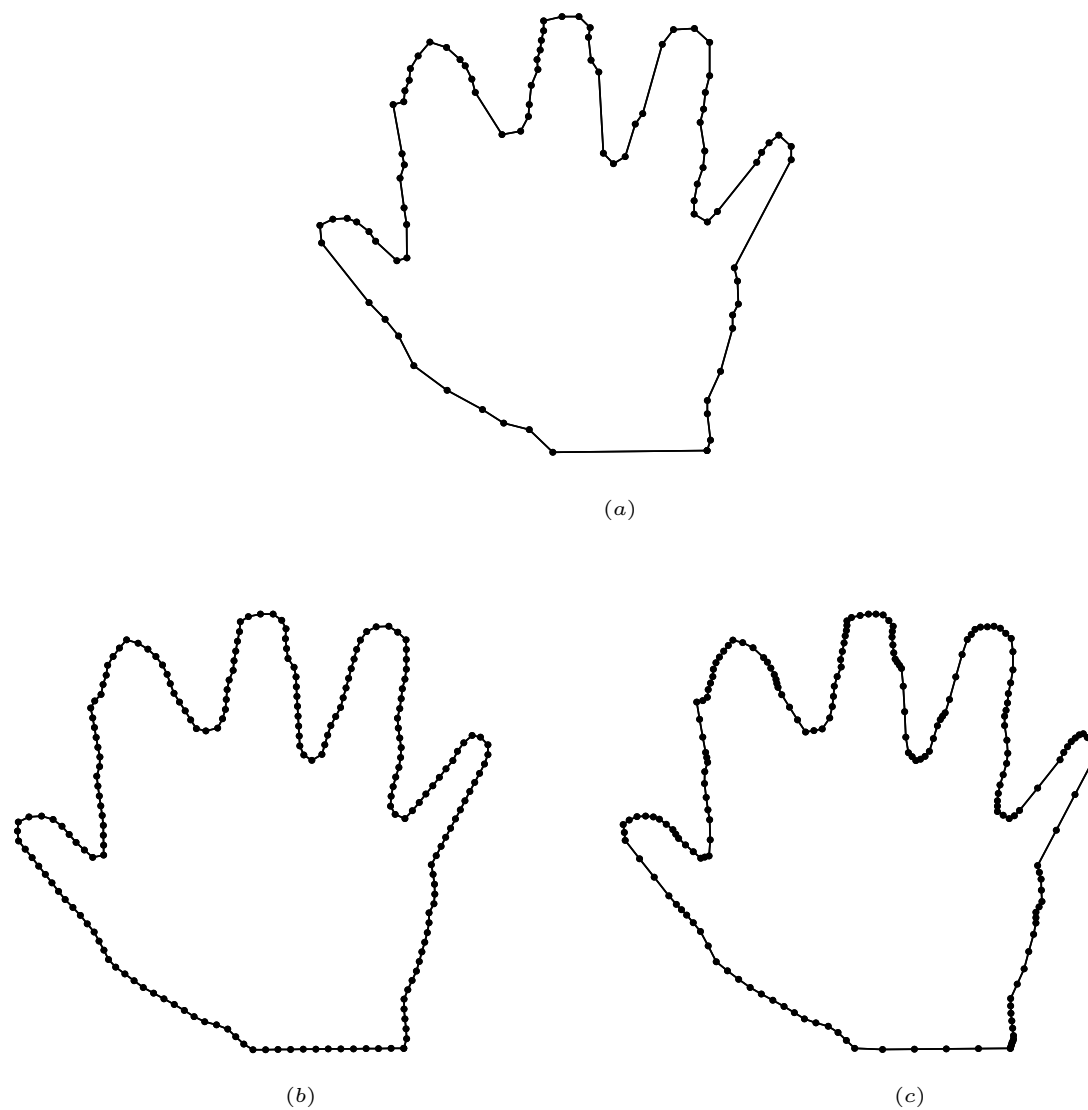


Figura 10: Algumas re-amostragens de uma curva: (a) contorno original com parametrização arbitrária (89 pontos), (b) parametrização pelo comprimento de arco normalizado (200 pontos), e (c) parametrização pelo comprimento de arco afim normalizado (200 pontos).

por partes. Portanto, eles são apropriados também para as curvas abertas. Estas invariantes são calculadas utilizando as derivadas. As características invariantes locais podem ser baseadas em comprimentos, ângulos ou curvaturas e são calculados através da primeira, segunda ou terceira derivada.

## 2.7 Seleção de Características

A escolha das características é particularmente crítica, visto que isto pode extremamente impactar o resultado final da classificação. Certamente, o processo de selecionar características apropriadas foi frequentemente identificado ([RIPLEY, 1996](#)) como sendo mesmo mais crítico do que os algoritmos de classificação. Embora nenhuma regra definitiva esteja disponível para definir que características usar numa situação específica, existem poucas direções gerais que podem ajudar tal processo, incluindo ([COSTA; CESAR, 2001](#)):

- Procurar características muito discriminativas a respeito dos objetos sob consideração. Observe que o conhecimento prévio sobre os objetos pode ser muito valioso.
- Evitar características muito correlacionadas. As medidas correlacionadas tendem a ser redundantes, implicando cálculos adicionais (tempo e armazenamento).
- Manter o número das características tão pequeno quanto possível. Além de implicar um custo computacional mais alto, um grande número de características faz a visualização e a exploração automatizada do espaço das características mais difícil, e exige também esquemas mais eficazes e mais poderosos de classificação. Certamente, se um número excessivo de características for usado, a similaridade entre os objetos calculados em termos de tais características tende a tornar-se menor, implicando menos poder discriminativo.

- Considerar as características que sejam invariantes sob transformações geométricas.
- Usar características que possam ser medidas objetivamente por métodos que não envolvam muitos parâmetros.
- A escolha das características adequadas torna-se mais natural e simples conforme o usuário adquire progressivamente mais familiaridade e experiência com a área da classificação e os problemas específicos.
- Dedicar uma atenção especial sobre os objetos que não parecem ser típicos de suas classes respectivas, visto que eles causam freqüentemente efeitos prejudiciais durante a classificação, incluindo sobreposição no espaço das características.
- Ter familiaridade com o maior número de características possíveis, seu poder discriminativo e seu custo computacional.

Neste trabalho, seguiremos as considerações anteriores para escolher características invariantes sob transformações de similaridade e afim.

## 2.8 Conclusão

Neste capítulo, especificamos as noções fundamentais relativas ao modelo da descrição invariante de um objeto plano, descrito pelo seu contorno, sob transformações de similaridade e afim.

Primeiramente, mostramos a ação das transformações do grupo de similaridades e afim sobre o contorno dos objetos descritos por curvas paramétricas. Posteriormente, parametrizamos o contorno das formas através dos parâmetros comprimento de arco normalizado euclidiano e afim. Após a re-parametrização do contorno das formas, as características invariantes são extraídas.

### **3 RECONHECIMENTO DE FORMAS INVARIANTES SOB TRANSFORMAÇÕES DE SIMILARIDADE**

Neste capítulo, propomos um conjunto de características invariantes sob transformação de similaridade. Uma técnica baseada na aproximação poligonal do contorno do objeto também é proposta. Os pontos da aproximação poligonal são formados por pontos de alta curvatura do contorno do objeto. As características são computadas a partir da aproximação poligonal e então o classificador de distância mínima é usado para reconhecer o objeto. A técnica proposta é rápida e invariante sob transformação de similaridade. Os resultados experimentais em reconhecimento de gestos estáticos de mão mostram o desempenho da técnica proposta.

#### **3.1 Introdução**

A representação de um objeto pode ser baseada no seu contorno ou na sua região. A caracterização do objeto pelo seu contorno é uma escolha mais natural pois isto parece ser a forma que nós humanos representamos mentalmente um objeto. Nosso sistema visual se focaliza sobre os contornos e ignora regiões uniformes. Esta habilidade é “embutida” em nossa retina: duas camadas de neurônios que executam uma operação similar ao Laplaciano são diretamente conectadas aos bastonetes e cones da retina. Esta operação é chamada de inibição lateral e ajuda a extrair contornos e bordas (RUSS, 1999).

Neste capítulo, simplificamos o contorno de um objeto pela sua aproximação po-

ligonal. A aproximação poligonal do contorno de um objeto consiste em encontrar pontos significantes ao longo do contorno do objeto de forma que estes pontos constituam uma boa aproximação do contorno original.

Uma boa abordagem para resolver este problema é considerar os pontos de alto valor absoluto da curvatura como pontos significantes. Estes pontos significantes são melhor computados numa abordagem multi-escala (PINHEIRO; IZQUIERDO; GHANHARI, 2000), porém, o custo computacional é alto. A nossa abordagem é a seguinte: a partir da curvatura do contorno, determinamos um limite inferior e superior para extrair os pontos de alta curvatura.

A curvatura é um dos mais importantes atributos que podem ser extraídos de um contorno. Por exemplo, podemos citar dois experimentos psicofísicos executados por Attneave (ATTNEAVE, 1954). No primeiro experimento, uma série de objetos planos foram mostradas as pessoas, os quais deviam representar cada contorno usando um conjunto de 10 pontos. Os resultados mostraram claramente que a maioria das pessoas particularmente usaram os pontos de alta curvatura para representar cada objeto. No segundo experimento, Attneave desenhou o perfil de um gato identificando os pontos de alta curvatura conectados por linhas retas, demonstrando que a maioria da informação na imagem original está concentrada nos pontos de alta curvatura. Similares experimentos e resultados foram reportados também por Fischler and Wolf (FISCHLER; WOLF, 1994). Conseqüentemente, a aproximação de curvas por linhas retas conectando pontos de alta curvatura guardam a informação necessária para o correto reconhecimento de formas. Este fato pode ser explicado considerando que o nosso sistema visual focaliza sobre singularidades e ignora curvas suaves usando o mecanismo de inibição lateral.

A representação do objeto através de pontos de alta curvatura diminui consideravelmente o número de pontos do contorno, enquanto que mantém a informação necessária para o reconhecimento da forma. Além disso, os pontos de alta curvatura são



descritores robustos do objeto, no sentido que eles são invariantes sob transformação de similaridade (TEH; CHIN, 1989).

Na literatura, existem muitos métodos para determinar a curvatura do contorno de um objeto (por exemplo, (BEUS; TIU, 1987; LIU; SRINATH, 1990; ROSENFELD; WESZKA, 1975; KOPLOWITZ; PLANTE, 1995)). Neste capítulo, usamos uma abordagem descrita por Costa e Cesar (COSTA; CESAR, 2001). Esta abordagem utiliza a derivada da transformada de Fourier para determinar a curvatura. Escolhemos este método devido a sua simplicidade.

Depois de aproximar o contorno dos objetos através de polígonos, um conjunto de características são calculadas para cada objeto. Estas características serão usadas para um problema de reconhecimento supervisionado de formas. Definimos e usamos oito características, todas invariantes sob transformação de similaridade. Na Seção 3.2 descrevemos estas características. Logo que todas as características são calculadas, calculamos o centróide (média dos vetores) para cada classe. Então, dado um vetor de consulta, o classificador de distância mínima será usado para identificar a sua classe.

Aplicamos o método descrito acima para o problema de reconhecimento supervisionado de gestos estáticos de mão. Resultados experimentais demonstram a eficiência das nossas características e da nossa técnica. O resto do capítulo está organizado como segue. Na Seção 3.2, o método, as características e o algoritmo de reconhecimento são descritos. Os resultados experimentais são apresentados na Seção 3.3. Finalmente, na Seção 3.4 apresentamos as nossas conclusões.

## 3.2 O Método

O método proposto para o reconhecimento de formas sob transformação de similaridade está baseado no cálculo rápido da aproximação poligonal do objeto e o uso de características da forma para a fase da classificação. Esta aproximação poligonal

está baseada no cálculo da curvatura do contorno do objeto no espaço de frequências, através do uso da transformada de Fourier. Dado a curvatura do contorno, propomos uma nova técnica para estimar os pontos de alta curvatura.

Não é fácil encontrar um conjunto de pontos de alta curvatura que constituam uma boa aproximação poligonal do contorno do objeto. Usualmente, o contorno obtido computacionalmente é áspero e existem muitos pontos de alta curvatura nesta aspereza que não pertencem a uma boa aproximação poligonal. Existem algumas técnicas que podem reduzir este problema (por exemplo, interpolação ou filtragem), mas eles aumentam o tempo total do processamento. Para evitar este problema, removemos pontos vizinhos do contorno que sejam muito próximos. Este pré-processamento é suficiente para aliviar o problema da aspereza.

A seguir, descrevemos a técnica descrita por Costa e Cesar ([COSTA; CESAR, 2001](#)) para o cálculo da curvatura. Seja  $(x(t), y(t))$  a representação paramétrica do contorno aberto de um objeto e  $\mathbf{u}(t) = x(t) + jy(t)$  sua representação no plano complexo. A curvatura  $k(t)$  de  $(x(t), y(t))$  é definida como:

$$k(t) = \frac{\dot{x}(t)\ddot{y}(t) - \ddot{x}(t)\dot{y}(t)}{(\dot{x}(t)^2 + \dot{y}(t)^2)^{\frac{3}{2}}} \quad (3.1)$$

A primeira e segunda derivada de  $\mathbf{u}(t)$ , denotado como  $\dot{\mathbf{u}}(t)$  e  $\ddot{\mathbf{u}}(t)$ , são definidas como  $\dot{\mathbf{u}}(t) = \dot{x}(t) + j\dot{y}(t)$  e  $\ddot{\mathbf{u}}(t) = \ddot{x}(t) + j\ddot{y}(t)$ , respectivamente. O conjugado de  $\ddot{\mathbf{u}}(t)$  é definido como  $\ddot{\mathbf{u}}^*(t) = \ddot{x}(t) - j\ddot{y}(t)$ . Considerando as seguintes relações

$$\begin{aligned} \dot{\mathbf{u}}(t)\ddot{\mathbf{u}}^*(t) &= \dot{x}(t)\dot{y}(t) + \ddot{x}(t)\dot{y}(t) - j(\dot{x}(t)\ddot{y}(t) - \ddot{x}(t)\dot{y}(t)) \\ |\dot{\mathbf{u}}(t)|^3 &= (\dot{x}(t)^2 + \dot{y}(t)^2)^{3/2} \end{aligned}$$

e a Equação 3.1, obtemos a seguinte representação da curvatura  $k(t)$  em termos de  $\mathbf{u}(t)$ :

$$k(t) = \frac{-\text{Im}\{\dot{\mathbf{u}}(t)\ddot{\mathbf{u}}(t)^*\}}{|\dot{\mathbf{u}}(t)|^3}. \quad (3.2)$$

onde  $\text{Im}$  denota a parte imaginária de um número complexo.

Seja  $U(f)$  a transformada de Fourier de  $\mathbf{u}(t)$ . Este fato será denotado como  $\mathbf{u}(t) \longleftrightarrow U(f)$ . As derivadas de  $\mathbf{u}(t)$  e sua transformada de Fourier possuem as seguintes propriedades:

$$\begin{aligned}\dot{\mathbf{u}}(t) &\longleftrightarrow (j2\pi f) U(f) \\ \ddot{\mathbf{u}}(t) &\longleftrightarrow (j2\pi f)^2 U(f)\end{aligned}$$

Assim,  $U(f)$  pode ser usado para estimar  $\dot{\mathbf{u}}(t)$ ,  $\ddot{\mathbf{u}}(t)$  e conseqüentemente a curvatura  $k(t)$ .

Logo que a curvatura é calculada, ela é normalizada pelo seu perímetro. Um ponto do contorno  $\mathbf{u}(t)$  é classificado como ponto de alta curvatura se sua curvatura satisfaz o seguinte sistema:

$$\begin{cases} k(t) \geq \delta_p, & k(t) \geq 0 \\ k(t) \leq \delta_n, & k(t) < 0 \end{cases} \quad (3.3)$$

onde  $\delta_p$  e  $\delta_n$  são os valores médios das partes positivas e negativas da curvatura, respectivamente. Estes pontos de alta curvatura são os pontos da aproximação poligonal do contorno. Num contorno aberto, é claro que os pontos extremos deveriam fazer parte da aproximação poligonal do contorno. Isto não sempre acontece se considerarmos o sistema anterior, devido a pouca informação para estimar a curvatura dos pontos extremos. Portanto, incluímos os dois pontos extremos do contorno aberto na aproximação poligonal.

A Figura 11 ilustra a curvatura e os limites para o cálculo dos pontos significantes do contorno da Figura 12, onde também os pontos que fazem parte da sua aproximação poligonal estão ilustrados. Note que esta aproximação pode não ser a melhor, porém, é bastante boa.

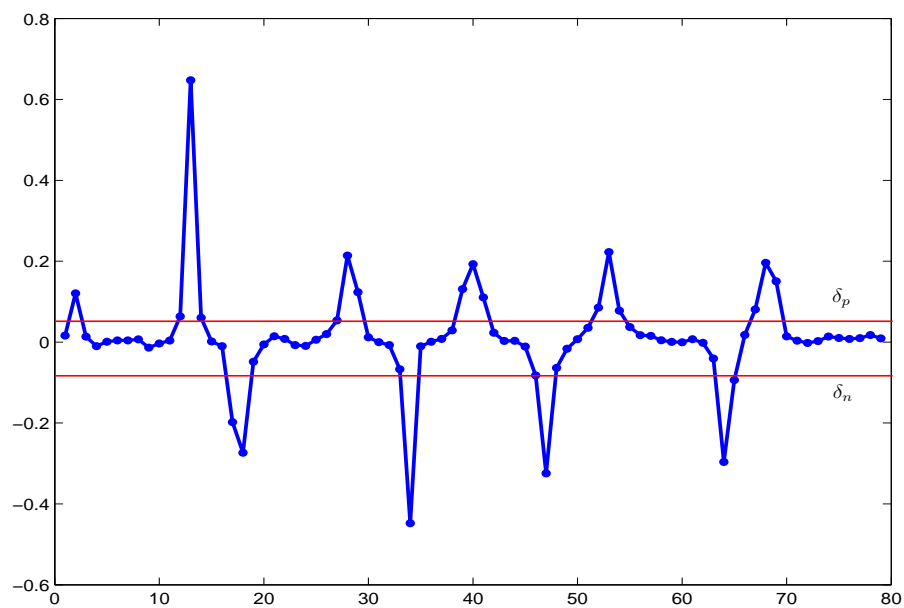


Figura 11: Curvatura do contorno ilustrado na Figura 12 e os limites superior e inferior para o cálculo dos pontos significantes.

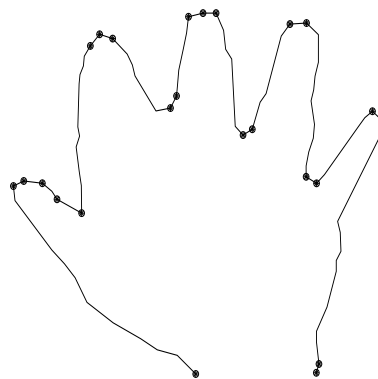


Figura 12: Os pontos sobre o contorno pertencem a aproximação poligonal determinados pelo método proposto.

### 3.2.1 Características

As características são calculadas a partir da aproximação poligonal do contorno do objeto. Considerando as características descritas em (O'HIGGINS, 1997; COSTA; CESAR, 2001), escolhemos oito de acordo aos critérios de seleção de características (Seção 2.7).

Seja  $u$  a aproximação poligonal do contorno de um objeto. Denotamos o perímetro e a área de  $u$  como  $P$  e  $A$ , respectivamente. A variável  $F$  denotará a área do fecho convexo de  $u$ . Seja  $D_{\max}$  e  $D_{\text{mean}}$ , respectivamente, a máxima e média distância entre o centróide de  $u$  e os pontos de  $u$ . Seja  $H$  o perímetro do fecho convexo de  $u$ . Para o cálculo do fecho convexo de qualquer polígono simples, usamos o Algoritmo de Graham (PREPARATA; SHAMOS, 1985). No Apêndice A, mostramos como calcular a área de qualquer polígono simples.

O vetor de características será composto pelas seguintes oito características:

$$F_1. \text{ Circularidade: } \frac{P^2}{A}.$$

$$F_2. \frac{P - \sqrt{P^2 - 4\pi A}}{P + \sqrt{P^2 - 4\pi A}}.$$

$$F_3. (\log_2(\frac{2P}{P-H}))^{-1}.$$

$F_4.$  A média dos valores da curvatura de  $u$  dividido pelo perímetro de  $u$ .

$F_5.$  A deficiência convexa dividido pela área de  $u$ , isto é,  $\frac{F-A}{A}$ .

$$F_6. \frac{A}{F}.$$

$F_7.$  A mediana dos valores da curvatura de  $u$ .

$$F_8. \frac{D_{\max}}{D_{\text{mean}}}.$$

Estas características são invariantes sob transformações de similaridade. A curvatura, área, perímetro e distância são medidas invariantes sob transformações de rotação e

translação. A invariância das características sob transformações de escala uniforme podem ser provadas para o caso discreto. A invariância é útil para obter uma classificação robusta. Para a classificação de um contorno de consulta, usamos o classificador de distância mínima.

### 3.2.2 Algoritmo de Treinamento

O algoritmo de treinamento pode ser resumido como segue. A entrada do algoritmo são os contornos dos objetos com suas respectivas etiquetas. Estas etiquetas certificam a correta classificação de cada contorno. A saída do algoritmo são os centróides de cada classe.

- Para cada contorno, remova os pontos que tem vizinhos muito próximos.
- Calcule a curvatura de cada ponto do contorno, usando a Equação 3.2.
- Determine os pontos de alta curvatura de cada contorno usando a Equação 3.3. Estes pontos, juntamente com os dois pontos extremos do contorno, definem a aproximação poligonal do contorno.
- Calcule as oito características para cada aproximação poligonal.
- Calcule o centróide de cada classe de objetos.

### 3.2.3 Algoritmo de Classificação

O algoritmo de classificação recebe como entrada um contorno de consulta com classificação desconhecida e calcula a sua classe.

- Dado um contorno de consulta com classificação desconhecida, calcule as oito características como descrito no algoritmo de treinamento. Denotamos o vetor de características obtido como  $x$ .

- O vetor  $x$  será classificado como pertencendo a classe que tem a menor distância euclidiana entre  $x$  e os centróides.

### 3.3 Resultados Experimentais

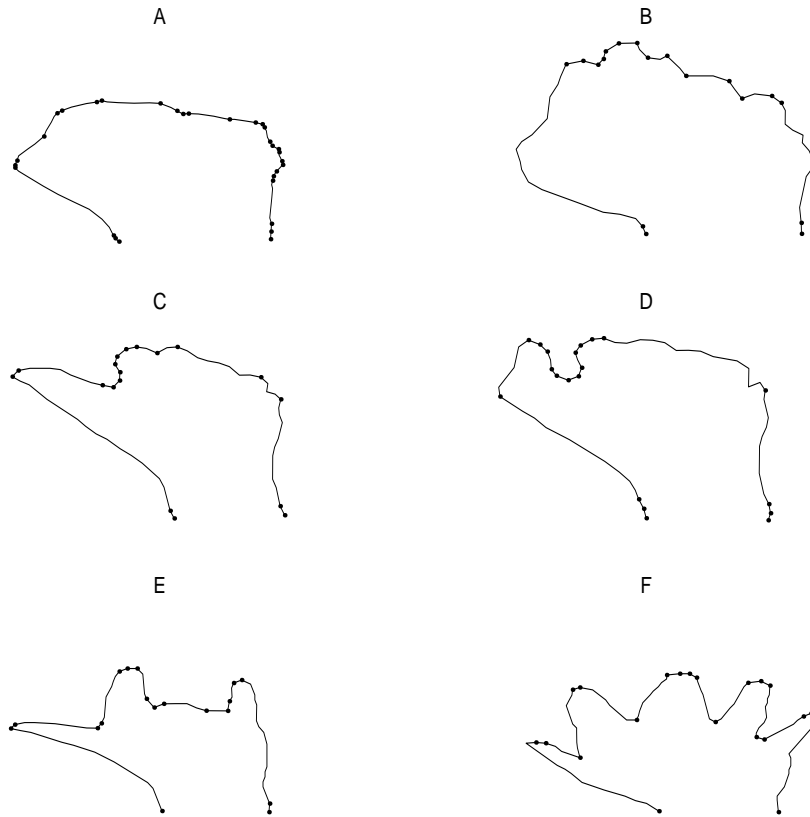


Figura 13: Contornos de gestos de mão das classes A, B, C, D, E e F.

Aplicamos as características e o método proposto para reconhecimento supervisionado de gestos estáticos de mão. Usamos parte dos contornos abertos de gestos estáticos de mão considerados por Milios e Petrakis em (MILIOS; PETRAKIS, 2000).

Etiquetamos manualmente 156 contornos abertos para o estágio de treinamento. Eles são divididos em seis classes, como ilustrado na Figura 13. Também etiquetamos 44 contornos de consulta. Os contornos de treinamento e consulta são disjuntos. Todos os programas necessários para o reconhecimento supervisionado de formas invariante sob transformação por similaridade, foram implementados em Matlab e se encontram

Tabela 2: Classificação dos dados de consulta.

	A	B	C	D	E	F
A	6	0	2	0	0	0
B	0	4	1	0	0	0
C	0	3	9	0	0	0
D	0	0	0	4	0	0
E	0	1	0	0	4	0
F	0	1	0	0	4	5

Tabela 3: Classificação dos dados de treinamento.

	A	B	C	D	E	F
A	22	0	0	0	0	0
B	0	13	3	0	0	0
C	0	2	24	1	0	0
D	0	0	0	22	0	0
E	0	0	1	0	15	0
F	0	0	0	0	2	51

no disquete incluído nesta tese.

A Tabela 2 resume os resultados da classificação dos contornos de consulta. Nesta tabela, o elemento  $(i, j)$  (fila  $i$ , coluna  $j$ ) corresponde ao número de gestos de mão manualmente etiquetados como  $i$  e classificados como  $j$  por nosso algoritmo. Por exemplo, a segunda linha expressa que existem cinco contornos de consulta da classe B. Quatro deles foram corretamente classificados por nosso algoritmo e um foi classificado como da classe C.

Também testamos nosso método sobre os contornos de treinamento. Os resultados são listados na Tabela 3. Devido a escolha do classificador de distância mínima (comparação com o centróide), os resultados não são ótimos sobre estes contornos.

As Figuras 14, 15, 16, 17, 18 e 19 ilustram alguns corretos e incorretos contornos de consulta. A maioria destes contornos apresentam bastante ruído, mesmo assim, as características e o método proposto conseguem classificar corretamente na maioria dos



casos.

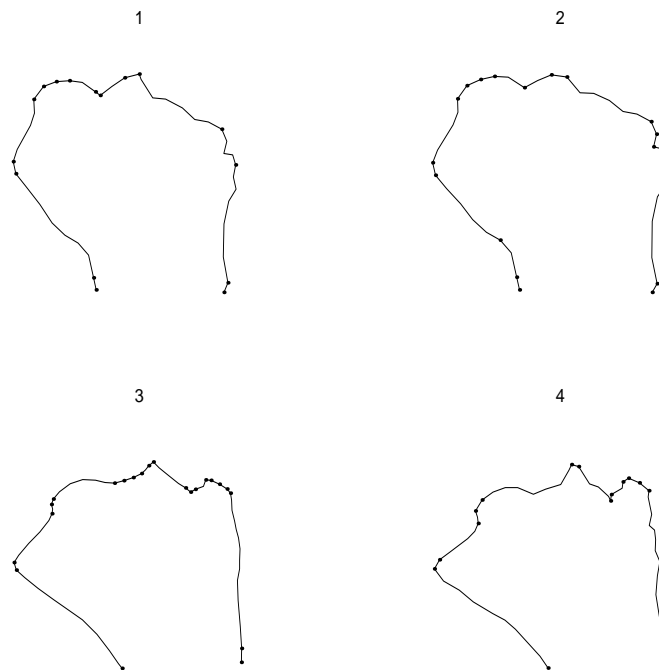


Figura 14: Quatro contornos de consulta manualmente etiquetados como sendo da classe A. Os gestos 1, 2 e 3 foram classificados por nosso algoritmo como A. O gesto 4 foi classificado como C.

Na Figura 15, observe que a aproximação poligonal dos gestos da classe B é muito pobre: existem muitos pontos de alta curvatura no parte superior do contorno. Isto é devido a aspereza de tal parte do contorno. O resto das aproximações poligonais tem uma aceitável qualidade.

### 3.4 Conclusões

Neste capítulo, propomos um conjunto de características invariantes sob transformação de similaridade e uma nova técnica para reconhecimento de formas baseada na aproximação poligonal do contorno do objeto. Esta técnica faz uso da transformada de Fourier para calcular os pontos de alta curvatura. Um conjunto de características são extraídas desde estes pontos de alta curvatura e o classificador de distância mínima é usado para a classificação. Aplicamos a técnica proposta para o reconhecimento

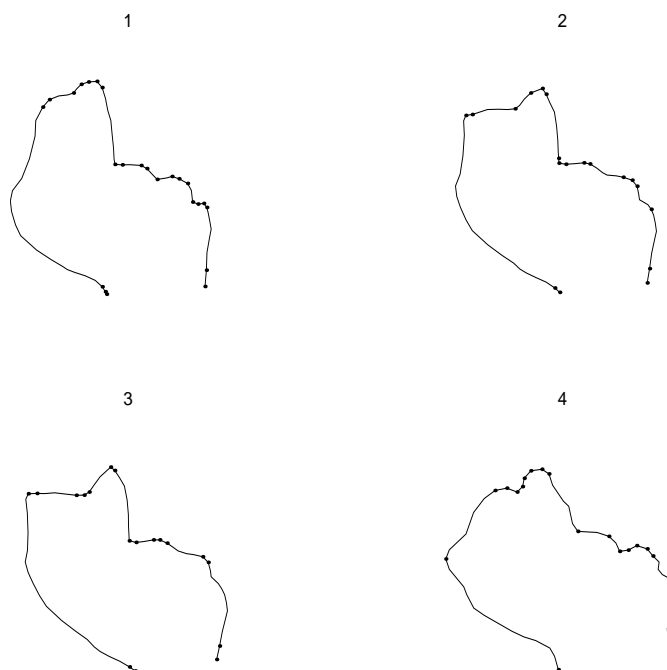


Figura 15: Quatro contornos de consulta manualmente etiquetados como sendo da classe B. Os gestos 1, 2 e 3 foram classificados por nosso algoritmo como B. O gesto 4 foi classificado como C.

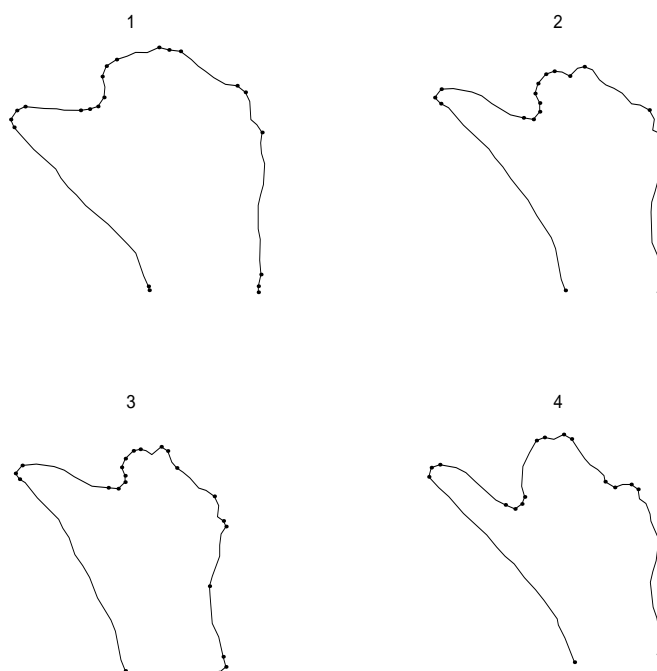


Figura 16: Quatro contornos de consulta manualmente etiquetados como sendo da classe C. Os gestos 1, 2 e 3 foram classificados por nosso algoritmo como C. O gesto 4 foi classificado como B.

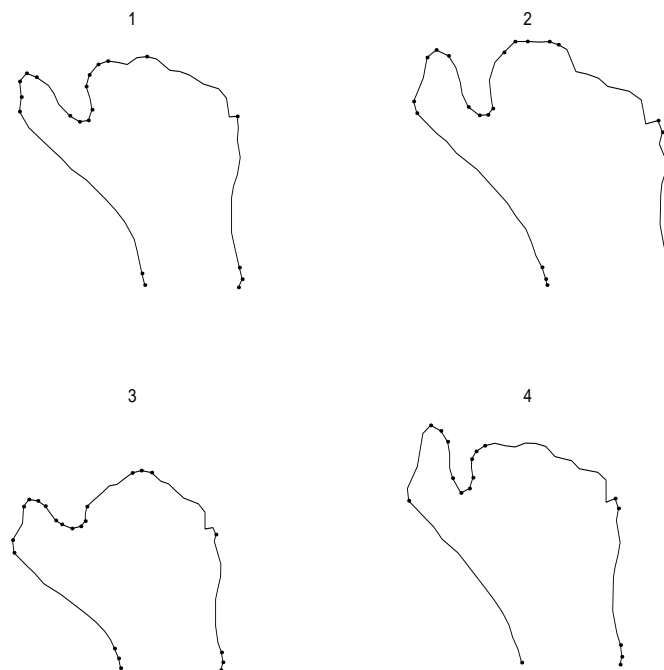


Figura 17: Quatro contornos de consulta manualmente etiquetados como sendo da classe D. Todos eles foram classificados por nosso algoritmo como D.

de gestos estáticos de mão e a alta taxa de reconhecimento obtida permite aplicações práticas da nossa técnica.

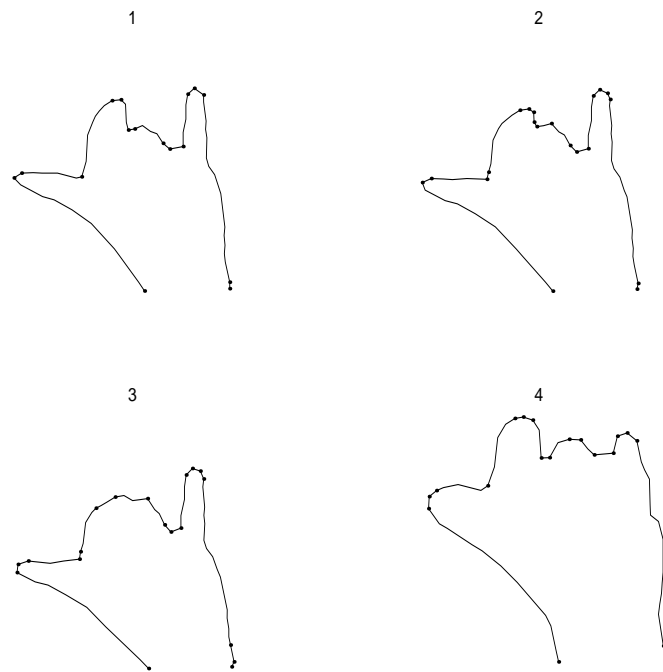


Figura 18: Quatro contornos de consulta manualmente etiquetados como sendo da classe E. Os gestos 1, 2 e 3 foram classificados por nosso algoritmo como E. O gesto 4 foi classificado como B.

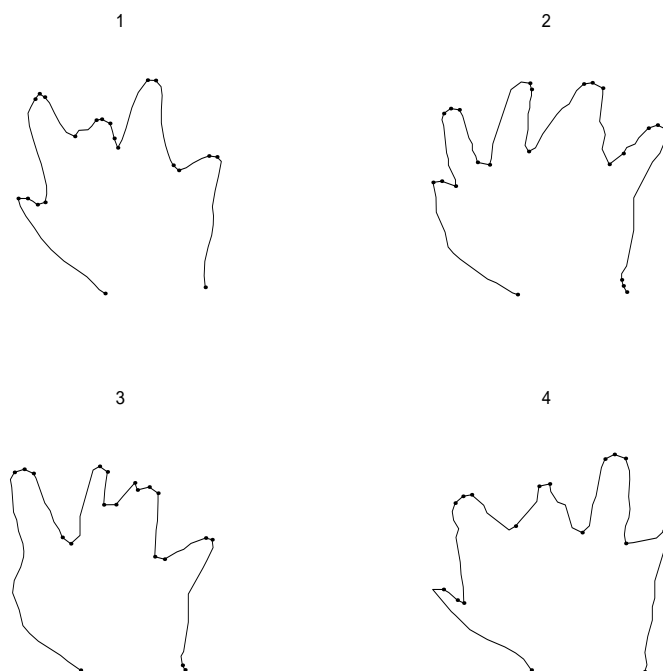


Figura 19: Quatro contornos de consulta manualmente etiquetados como sendo da classe F. Os gestos 1, 2 e 3 foram classificados por nosso algoritmo como F. O gesto 4 foi classificado como E.

## **4 RECONHECIMENTO DE FORMAS INVARIANTES SOB TRANSFORMAÇÕES AFINS**

Um objeto visto de diferentes pontos de vista resulta em imagens diferentemente deformadas. O reconhecimento de formas invariantes sob transformações afins deve classificar corretamente um objeto, independentemente do ponto de vista. Neste capítulo, propomos novas características locais e globais invariantes sob transformações afins. Estas características podem ser usadas para reconhecimento supervisionado ou não-supervisionado de formas, e para a recuperação e indexação de imagens baseadas no contorno da forma. Uma das características propostas está relacionada com a deficiência convexa e as outras são extraídas da matriz de áreas. A matriz de áreas foi usada por Shen ([SHEN; WONG; IP, 1999](#)) para o casamento por similaridade na recuperação de imagens. Porém, diferentemente do trabalho de Shen, parametrizamos o contorno das formas usando o parâmetro comprimento de arco afim. Isto torna as nossas características robustas a parametrização da forma, enquanto que o trabalho de Shen não possui esta propriedade. Os resultados experimentais indicam que o nosso método pode classificar corretamente formas deformadas e formas ruidosas usando um pequeno conjunto de treinamento.

### **4.1 Introdução**

Um problema importante na classificação de objetos é o fato de que um objeto pode ser visto desde diferentes pontos de vista da câmera, resultando em imagens diferentes.

Conseqüentemente, a invariância ao ponto de vista é uma propriedade desejável em muitos sistemas de reconhecimento de formas. Para objetos quase planos, estas deformações podem ser modeladas aproximadamente por transformações afins, se os pontos de vista estiverem suficientemente distantes do objeto (Subseção 2.4.1). Na literatura, existem muitos trabalhos sobre o reconhecimento de formas invariantes sob transformação afim. Alguns exemplos são os descritores de Fourier invariantes afins (ARBTER; SNYDER; BURKHARDT, 1990), os momentos invariantes afins (FLUSSER; SUK, 1993; HUANG; COHEN, 1994; ZHAO; CHEN, 1997) e o espaço-escala da curvatura afim (MOKHTARIAN; ABBASI, 2001; MOKHTARIAN; ABBASI, 2002). Nestes trabalhos, a idéia fundamental é usar uma parametrização afim, usualmente o comprimento de arco afim (ARBTER; SNYDER; BURKHARDT, 1990; FLUSSER; SUK, 1993; ZHAO; CHEN, 1997; SAPIRO; TANNENBAUM, 1993).

Usualmente, uma boa taxa de classificação pode ser obtida usando simultaneamente as características locais e globais. As características locais são sensíveis a ruídos e, na prática, a quantidade de informações locais disponíveis são usualmente insuficientes para uma classificação robusta. Alguns exemplos de características locais são os invariantes geométricos (RIVLIN; WEISS, 1995), cantos (corners) (RATTARANGSI; CHIN, 1992) e a curvatura (MOKHTARIAN, 1995). Por outro lado, as características globais são relativamente insensíveis a ruídos e duas formas similares normalmente possuem as características globais semelhantes. Dois exemplos simples de características globais são a área e a deficiência convexa.

A deficiência convexa é uma característica global que contém informação sobre as concavidades da forma. Esta característica é invariante sob transformação de translação, rotação e facilmente pode se tornar invariante à escala uniforme. Neste trabalho, mostramos que a deficiência convexa pode também se tornar invariante à transformação afim. As outras características propostas são extraídas da matriz de áreas. A matriz de áreas foi usada por Shen (SHEN; WONG; IP, 1999) para o casamento de formas por

similaridade em recuperação de imagens invariante afim. Shen mostra que esta matriz possui duas propriedades interessantes. A primeira é que ela pode se tornar invariante sob transformação afim. A segunda é que as suas colunas contém informações multi-escala: as colunas centrais contém informações globais ou em escala grossa e as colunas periféricas contém informações locais ou em escala fina. Diferentemente do trabalho de Shen, parametrizamos as formas usando o parâmetro comprimento de arco afim. Este procedimento torna a matriz de áreas robusta a parametrização da forma. A complexidade computacional do trabalho de Shen depende da dimensão da matriz de áreas. Nós, ao contrário, extraímos e processamos somente as informações relevantes, diminuindo a complexidade computacional. Na fase de reconhecimento, usamos o algoritmo dos  $k$ -vizinhos mais próximos.

O resto do capítulo está organizado como segue. A Seção 4.2 trata da área, da transformação afim e da parametrização afim. Na Seção 4.3, descrevemos a matriz de áreas, a sua normalização e a extração de informações relevantes. A Seção 4.4 trata das características propostas, do algoritmo de treinamento e do algoritmo de classificação. Os resultados experimentais estão expostos na Seção 4.5. Finalmente, na Seção 4.6 apresentamos as conclusões deste capítulo.

## 4.2 Conceitos Fundamentais

Seja  $\mathbf{u} : [0, T] \rightarrow \mathbb{R}^2$  uma curva paramétrica fechada, de período  $T$ , ou seja:

$$\mathbf{u}(t) = (x(t), y(t)) \quad \text{para } t \in [0, T].$$

Em geral, as curvas de uma aplicação podem ter sido parametrizadas com um número diferente de pontos  $T$ . Usualmente, todas as curvas são normalizadas para ter o mesmo número de pontos. Este número de pontos não pode ser muito pequeno (perda de informação na representação do contorno da forma) ou muito grande (aumento da complexidade computacional). Na prática, este número de pontos é calculado experi-

mentalmente para uma aplicação em particular.

A primeira e segunda derivadas de  $\mathbf{u}(t)$  são denotadas, respectivamente, como  $\dot{\mathbf{u}}(t) = (\dot{x}(t), \dot{y}(t))$  e  $\ddot{\mathbf{u}}(t) = (\ddot{x}(t), \ddot{y}(t))$ . Na Subseção 2.5.1 mostramos como aproximar estas derivadas.

### 4.2.1 Transformação Afim

De acordo a Definição 5, uma transformação afim geral  $\mathbf{F}$  em  $\mathbb{R}^2$ , atuando sobre a curva  $\mathbf{u}$ , é definida como:

$$\mathbf{F}(\mathbf{u}) = \mathbf{A}\mathbf{u} + \mathbf{T}$$

onde  $\mathbf{A}$  é a matriz afim e  $\mathbf{T}$  é o vetor translação. Veja a Subseção 2.4.2 para mais detalhes sobre transformação afim.

As transformações de rotação, translação e mudança de escala uniforme não alteram a descrição geométrica da forma. Por outro lado, as transformações de escala não-uniforme e cisalhamento deformam a forma. A transformação afim contém todos estes tipos de transformações.

Neste capítulo, propomos novas características invariantes sob transformações afins. Na fase experimental, a invariância e a robustez destas características são testadas somente para transformações de rotação e cisalhamento.

### 4.2.2 Área e Transformação Afim

Seja  $\triangle$  o triângulo formado pelos pontos  $(x_1, y_1)$ ,  $(x_2, y_2)$  e  $(x_3, y_3)$ . Então, a área do triângulo  $\triangle$  é dada pela seguinte fórmula:

$$\text{Area}(\triangle) = \frac{1}{2} \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} \quad (4.1)$$



onde  $|\cdot|$  denota a função determinante. Se  $(x_1, y_1)$ ,  $(x_2, y_2)$  e  $(x_3, y_3)$  são orientados no sentido anti-horário o valor da área é positivo, caso contrário é negativo. Se  $\mathbf{F}$  é uma transformação afim e  $\mathbf{A}$  a sua matriz afim associada, então:

$$\text{Area}(\mathbf{F}(\Delta)) = |\mathbf{A}|\text{Area}(\Delta) \quad (4.2)$$

onde  $|\mathbf{A}|$  é o determinante da matriz  $\mathbf{A}$ . Considerando a Equação 4.2 e o fato de que qualquer polígono simples 2-D  $\mathbf{u}$  pode ser decomposto como um conjunto de triângulos, podemos afirmar o seguinte:

$$\text{Area}(\mathbf{F}(\mathbf{u})) = |\mathbf{A}|\text{Area}(\mathbf{u}) \quad (4.3)$$

Neste trabalho, o fecho convexo de uma curva paramétrica fechada  $\mathbf{u}$  é definido como a seqüência de pontos que descrevem o menor polígono convexo contendo  $\mathbf{u}$ . Denotamos o fecho convexo da curva  $\mathbf{u}$  como  $\text{FC}(\mathbf{u})$ . Um dos algoritmos para o cálculo do fecho convexo de qualquer polígono simples é o Algoritmo de Graham (PREPARATA; SHAMOS, 1985). Neste algoritmo, a primitiva é o sinal da área do triângulo que verifica se um ponto está à esquerda de um vetor. Este sinal não muda sob transformações afins. Assim, podemos afirmar que um ponto pertence ao  $\text{FC}(\mathbf{u})$  se, e somente se, a sua transformada pertencer ao  $\text{FC}(\mathbf{F}(\mathbf{u}))$ . Conseqüentemente, podemos também afirmar o seguinte:

$$\text{Area}(\text{FC}(\mathbf{F}(\mathbf{u}))) = |\mathbf{A}|\text{Area}(\text{FC}(\mathbf{u})) \quad (4.4)$$

Baseado nas Equações 4.3 e 4.4, observamos que a deficiência convexa de  $\mathbf{u}$  é invariante sob transformação afim a menos do fator de proporcionalidade  $|\mathbf{A}|$ . Este fator pode ser simplificado dividindo a deficiência convexa, por exemplo, pela área de  $\mathbf{u}$  ou pela área do fecho convexo de  $\mathbf{u}$ . Neste trabalho, propomos a seguinte característica invariante afim:

$$F_1 = \frac{\text{Area}(\text{FC}(\mathbf{u})) - \text{Area}(\mathbf{u})}{\text{Area}(\text{FC}(\mathbf{u}))} \quad (4.5)$$

### 4.2.3 Parametrização Afim

Na Seção 2.5, descrevemos dois parâmetros invariantes usados para parametrizar curvas planas e úteis para resolver satisfatoriamente problemas de reconhecimento de formas. Estes parâmetros são o comprimento de arco e comprimento de arco afim.

O parâmetro comprimento de arco é invariante sob transformação de similaridade. Porém, ele não é invariante sob transformação afim, ao contrário do parâmetro comprimento de arco afim (FLUSSER; SUK, 1993; SAPIRO; TANNENBAUM, 1993; ZHAO; CHEN, 1997; MOKHTARIAN; ABBASI, 2001).

Na Subseção 2.5.3, o comprimento de arco normalizado é descrito com mais detalhe. Este parâmetro é definido como:

$$\bar{s}(p) = \frac{\int_0^p (\dot{x}^2(t) + \dot{y}^2(t))^{\frac{1}{2}} dt}{\int_0^T (\dot{x}^2(t) + \dot{y}^2(t))^{\frac{1}{2}} dt} \quad (4.6)$$

Na Subseção 2.5.4, o comprimento de arco afim normalizado é descrito com mais detalhe. Este parâmetro é definido como:

$$\bar{s}_a(p) = \frac{\int_0^p |\dot{x}(t)\ddot{y}(t) - \ddot{x}(t)\dot{y}(t)|^{\frac{1}{3}} dt}{\int_0^T |\dot{x}(t)\ddot{y}(t) - \ddot{x}(t)\dot{y}(t)|^{\frac{1}{3}} dt} \quad (4.7)$$

Na implementação da parametrização afim do contorno de uma forma, é desejável que o parâmetro  $\bar{s}_a$  seja uma função estritamente crescente (isto é, não existe  $p_i$  e  $p_j$  tal que  $\bar{s}_a(p_i) = \bar{s}_a(p_j)$ ). Para garantir isto, removemos todos os pontos da curva  $\mathbf{u}$  com a segunda derivada nula. Isto equivale a remover os pontos pertencentes a linhas retas.

Neste capítulo, normalizamos o número de pontos de todas as formas, pois as características propostas dependem da escolha deste número. Também reparametrizamos todas as formas usando o parâmetro comprimento de arco afim normalizado. Desta forma, as nossas características invariantes afins serão robustas a parametrização da forma pelo comprimento de arco afim normalizado.

### 4.3 Extração de Informações Locais e Globais de uma Matriz de Áreas

Antes de definir a matriz de áreas do contorno de uma forma, mostramos duas proposições que servirão para entender a natureza e as propriedades da matriz de áreas.

**Proposição 3.** *Seja  $\mathbf{u}(s) = (x(s), y(s))$  uma curva paramétrica no plano parametrizado pelo comprimento de arco afim normalizado  $s$ . Então, o determinante  $\text{Det}(\dot{\mathbf{u}}(s), \ddot{\mathbf{u}}(s))$  é relativamente invariante sob transformações afins.*

*Demonstração.* Seja  $\mu = \mathbf{A}\mathbf{u} + \mathbf{T}$  uma versão deformada da curva  $\mathbf{u}$  sob uma transformação afim, onde  $\mathbf{A}$  é a matriz afim e  $\mathbf{T}$  o vetor translação.

O determinante  $\text{Det}(\dot{\mu}(s), \ddot{\mu}(s))$  é relativamente invariante sob transformações afins se existe um valor constante  $K$  tal que:

$$\text{Det}(\dot{\mu}(s), \ddot{\mu}(s)) = K\text{Det}(\dot{\mathbf{u}}(s), \ddot{\mathbf{u}}(s)).$$

O determinante dos vetores  $\dot{\mathbf{u}}(s)$  e  $\ddot{\mathbf{u}}(s)$  é dado pela seguinte expressão:

$$\text{Det}(\dot{\mathbf{u}}(s), \ddot{\mathbf{u}}(s)) = \begin{vmatrix} \dot{x}(s) & \ddot{x}(s) \\ \dot{y}(s) & \ddot{y}(s) \end{vmatrix} \quad (4.8)$$

Se  $\mu(s) = (u(s), v(s))$  então podemos escrever as seguintes relações:

$$\begin{bmatrix} u(s) \\ v(s) \end{bmatrix} = \mathbf{A} \begin{bmatrix} x(s) \\ y(s) \end{bmatrix} + \mathbf{T} \quad (4.9)$$

A primeira e segunda derivada de  $\bar{\mathbf{u}}(s)$  são dados pelas seguintes equações:

$$\begin{bmatrix} \dot{u}(s) \\ \dot{v}(s) \end{bmatrix} = \mathbf{A} \begin{bmatrix} \dot{x}(s) \\ \dot{y}(s) \end{bmatrix} \quad (4.10)$$

$$\begin{bmatrix} \ddot{u}(s) \\ \ddot{v}(s) \end{bmatrix} = \mathbf{A} \begin{bmatrix} \ddot{x}(s) \\ \ddot{y}(s) \end{bmatrix} \quad (4.11)$$

Considerando as duas equações anteriores, podemos obter a seguinte igualdade:

$$\begin{vmatrix} \dot{u}(s) & \ddot{u}(s) \\ \dot{v}(s) & \ddot{v}(s) \end{vmatrix} = |\mathbf{A}| \begin{vmatrix} \dot{x}(s) & \ddot{x}(s) \\ \dot{y}(s) & \ddot{y}(s) \end{vmatrix} \quad (4.12)$$

A Equação 4.12 prova que o valor  $\text{Det}(\dot{\mathbf{u}}(s), \ddot{\mathbf{u}}(s))$  é relativamente invariante afim.  $\square$

**Proposição 4.** *Seja  $\mathbf{u}(t) = (x(t), y(t))$  uma curva paramétrica no plano com parametrização arbitrária  $t$ . O valor do determinante  $\text{Det}(\dot{\mathbf{u}}(t), \ddot{\mathbf{u}}(t))$  é proporcional à área do triângulo formado por três pontos consecutivos do contorno  $\mathbf{u}$ .*

*Demonstração.* Considerando as diferenças atrasadas como aproximação para a primeira e segunda derivada de  $\mathbf{u}$ , obtemos o seguinte:

$$\begin{aligned} \dot{x}(t) &= \frac{x(t) - x(t-h)}{h} & \ddot{x}(t) &= \frac{x(t) - 2x(t-h) + x(t-2h)}{h^2} \\ \dot{y}(t) &= \frac{y(t) - y(t-h)}{h} & \ddot{y}(t) &= \frac{y(t) - 2y(t-h) + y(t-2h)}{h^2} \end{aligned}$$

Desta forma, o valor do determinante dos vetores  $\dot{\mathbf{u}}(t)$  e  $\ddot{\mathbf{u}}(t)$  pode ser expressado como:

$$\begin{aligned} \text{Det}(\dot{\mathbf{u}}(t), \ddot{\mathbf{u}}(t)) &= \frac{1}{h^3} \begin{vmatrix} x(t-2h) & y(t-h) & 1 \\ x(t-h) & y(t-h) & 1 \\ x(t) & y(t) & 1 \end{vmatrix} \\ &= \frac{2}{h^3} \text{Area}(\mathbf{u}(t-2h), \mathbf{u}(t-h), \mathbf{u}(t)) \end{aligned} \quad (4.13)$$

Se  $h = 1$  então  $\text{Det}(\dot{\mathbf{u}}(t), \ddot{\mathbf{u}}(t))$  é proporcional a área do triângulo formado por três pontos consecutivos da forma. Isto prova a proposição.  $\square$

O sinal da curvatura euclideana de  $\mathbf{u}$  no ponto  $t$  é definido pelo sinal do  $\text{Det}(\dot{\mathbf{u}}(t), \ddot{\mathbf{u}}(t))$ . A Equação 4.13 mostra que este sinal é igual ao sinal da área do triângulo  $\text{Area}(\mathbf{u}(t-2), \mathbf{u}(t-1), \mathbf{u}(t))$ . O valor deste sinal nos indica as partes côncavas e convexas do contorno  $\mathbf{u}$ .

O determinante  $\text{Det}(\dot{\mathbf{u}}(t), \ddot{\mathbf{u}}(t))$  pode ser expresso como  $\dot{x}(t)\ddot{y}(t) - \ddot{x}(t)\dot{y}(t)$ . Considerando a Fórmula 3.1 para o cálculo da curvatura de  $\mathbf{u}$ , observamos que o valor do determinante é igual ao numerador da equação da curvatura de  $\mathbf{u}$ .

Seja  $\Psi_h(t) = \text{Area}(\mathbf{u}(t - 2h), \mathbf{u}(t - h), \mathbf{u}(t))$ . Considerando todos os pontos de  $\mathbf{u}(t), t = 0, 1, 2, \dots, T$ , a função  $\Psi_h$  representa uma curva contendo informação sobre as partes côncavas e convexas do contorno  $\mathbf{u}$ . Baseado nas duas proposições anteriores e a observação anterior, podemos concluir que a função  $\Psi_h$  é invariante sob transformação afim do contorno  $\mathbf{u}$  e proporcional ao numerador da equação da curvatura de  $\mathbf{u}$  (veja Fórmula 3.1).

A curvatura é um dos mais importantes atributos que podem ser extraídos de um contorno (ATTNEAVE, 1954; FISCHLER; WOLF, 1994). Isto nos motiva a explorar a função  $\Psi_h$  para  $h > 1$ . Fazendo  $h = 1, 2, 3, \dots$  teremos um conjunto de curvas  $\Psi_1, \Psi_2, \Psi_3 \dots$  que descrevem informações sobre as partes côncavas e convexas do contorno  $\mathbf{u}$  numa “escala”  $h$ .

Esta forma de descrever uma curva em diferentes “escalas” foi também usada por N. Arica (ARICA; VURAL, 2003) para descrever a curvatura euclideana. Considerando a curvatura em todas as “escalas” possíveis, ele definiu o descritor BAS (Beam Angle Statistics), invariante sob transformação de similaridade, com excelente desempenho sobre o Core Experiment CE-Shape-1 (LATECKI; LAKÄMPER; ECKHARDT, 2000).

### 4.3.1 Matriz de Áreas

Nesta subseção, definimos a matriz de áreas de uma curva  $\mathbf{u}$  como sendo a matriz que tem como vetores coluna as curvas  $\Psi_h$ .

Seja  $\mathbf{u}(t), t \in [0, T]$ , uma curva fechada de período  $T$  e seja  $m = (T - 1)/2$ . Seja  $S_{ti}$  a área do triângulo formado pelos pontos  $\mathbf{u}(t - i), \mathbf{u}(t)$  e  $\mathbf{u}(t + i)$ . Então, a matriz

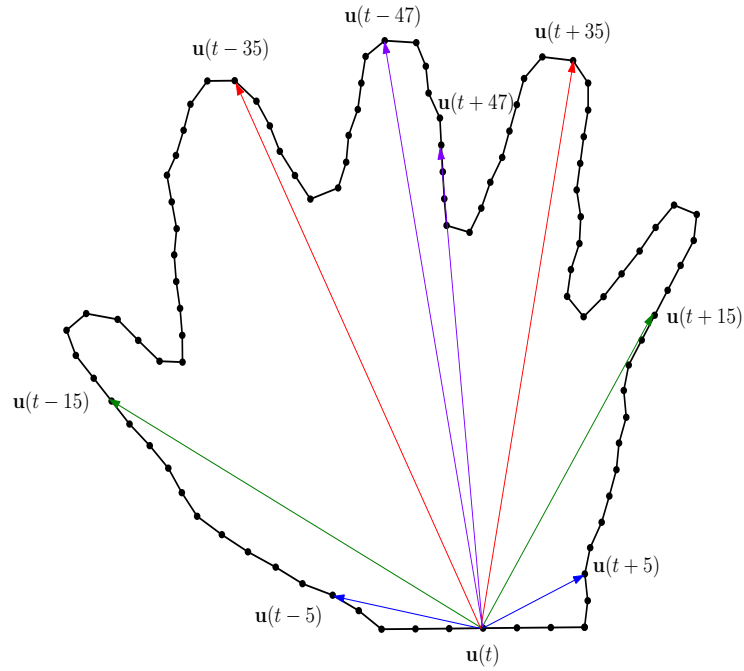


Figura 20: Cálculo da matriz de áreas.

de áreas de  $\mathbf{u}(t)$  é definida como:

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}_{01} & \mathbf{S}_{02} & \mathbf{S}_{03} & \cdots & \mathbf{S}_{0m} \\ \mathbf{S}_{11} & \mathbf{S}_{12} & \mathbf{S}_{13} & \cdots & \mathbf{S}_{1m} \\ \mathbf{S}_{21} & \mathbf{S}_{22} & \mathbf{S}_{23} & \cdots & \mathbf{S}_{2m} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{S}_{(T-1)1} & \mathbf{S}_{(T-1)2} & \mathbf{S}_{(T-1)3} & \cdots & \mathbf{S}_{(T-1)m} \end{bmatrix} \quad (4.14)$$

A Figura 20 ilustra o cálculo da matriz de áreas para o contorno de uma mão. Uma simples análise sobre a matriz de áreas nos mostra que ela contém informação local à esquerda, global no centro e local à direita:

$$\mathbf{S} = \left[ \begin{array}{ccc} \mathbf{S}_{t1} \cdots \mathbf{S}_{tk} & \mathbf{S}_{t(k+1)} \cdots \mathbf{S}_{tl} & \mathbf{S}_{t(l+1)} \cdots \mathbf{S}_{tm} \\ \text{informação local} & \text{informação global} & \text{informação local} \end{array} \right]_{t=0, \dots, T-1} \quad (4.15)$$

O  $i$ -ésimo vetor coluna de  $\mathbf{S}$  contém informações locais (escala fina) quando o índice  $i$  for pequeno ou próximo a  $m$  e contém informações globais (escala grossa) quando  $i$  estiver situado na região central de  $\mathbf{S}$ . De fato, cada vetor coluna da matriz de áreas

analisa a forma numa determinada escala, e a matriz de áreas completa contém as informações multi-escala da forma. Observemos também que a magnitude das informações locais são muito menores que a magnitude das informações globais.

### 4.3.2 Matriz de Áreas Normalizada

A matriz de áreas contém informações multi-escala. Para ser útil, devemos diminuir a dimensão desta matriz. A abordagem proposta neste trabalho consiste em reduzir a informação contida na matriz de áreas em três vetores coluna. Antes que a matriz de áreas seja reduzida, ela deve ser normalizada: tornar a matriz invariante sob transformação afim e ter todas as colunas com magnitudes similares.

Normalizamos a matriz de áreas  $\mathbf{S}$  em dois passos. Primeiro, tornamos a matriz  $\mathbf{S}$  invariante sob transformação afim dividindo cada elemento de  $\mathbf{S}$  pela soma de todos os elementos de  $\mathbf{S}$ . Segundo, fazemos todas as colunas terem magnitudes similares: Dividimos cada coluna  $i$  de  $\mathbf{S}$  pelo peso  $w(i)$ . O vetor de pesos  $w$  é calculado a partir de uma matriz de áreas invariante afim  $\mathbf{R}$  de uma forma representativa usando a seguinte equação:

$$w(i) = \max_{t=0, \dots, T-1} \{|\mathbf{R}_{t,i}|\} \quad i = 1, \dots, m \quad (4.16)$$

### 4.3.3 Extração de Informações Locais e Globais

Uma matriz de áreas contém informações em quantidade excessiva. Para ser útil, ela precisa ser reduzida e organizada de alguma forma. Extraímos as informações relevantes a partir de uma matriz de áreas normalizada  $\mathbf{S}$  e as armazenamos em três vetores  $\psi_e$ ,  $\psi_c$  e  $\psi_d$  que resumem, respectivamente, as informações das colunas situadas à esquerda, no centro e à direita da matriz  $\mathbf{S}$ :

$$\mathbf{S} = \left[ \begin{array}{ccc} \mathbf{S}_{t1} \dots \mathbf{S}_{tk} & \mathbf{S}_{t(k+1)} \dots \mathbf{S}_{tl} & \mathbf{S}_{t(l+1)} \dots \mathbf{S}_{tm} \\ \text{informação local} & \text{informação global} & \text{informação local} \\ \downarrow & \downarrow & \downarrow \\ \psi_e & \psi_c & \psi_d \end{array} \right]^{t=0, \dots, T-1} \quad (4.17)$$

O vetor  $\psi_e$  é calculado da seguinte forma:

$$\psi_e(t) = \frac{\sum_i g_e(i, \mu, \sigma) \mathbf{S}_{t,i}}{\sum_i g_e(i, \mu, \sigma)} \quad t = 0, \dots, T - 1 \quad (4.18)$$

onde  $g_e$  é uma curva gaussiana. O centro  $\mu$  desta gaussiana está situado no lado esquerdo de  $\mathbf{S}$ . Ele não pode estar muito próximo à primeira coluna, para reduzir os ruídos presentes em escala finas. O desvio-padrão  $\sigma$  é escolhido para extrair todas as informações presentes no lado esquerdo de  $\mathbf{S}$ , sem que haja mistura com as informações da região central de  $\mathbf{S}$ . Na implementação da curva gaussiana, consideramos que ela está definida somente no intervalo  $[\mu - \Delta m, \mu + \Delta m]$  como mostrado na Figura 21. Os vetores  $\psi_c$  e  $\psi_d$  são calculados de forma similar.

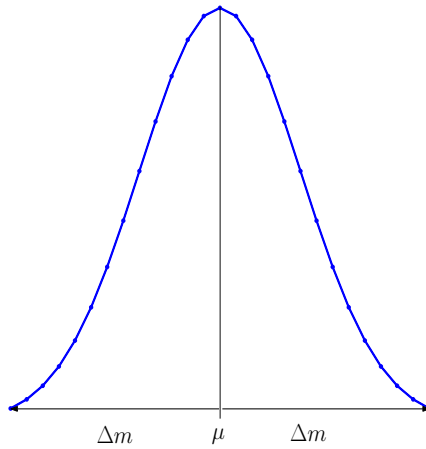


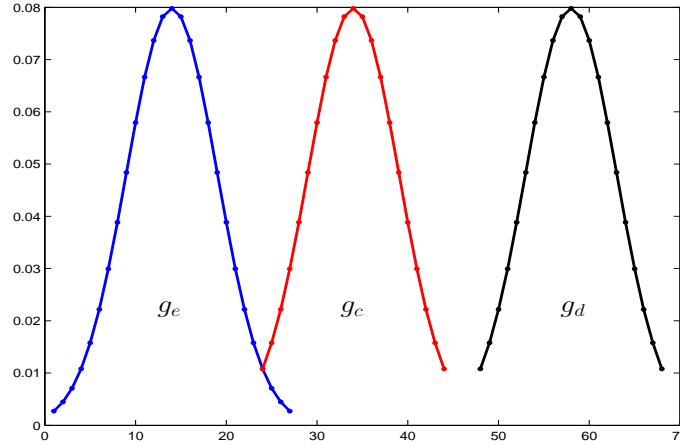
Figura 21: Implementação de uma curva gaussiana.

Os parâmetros que definem as três gaussianas são definidos experimentalmente para uma aplicação em particular. No entanto, esses parâmetros poderiam ser úteis também para outras aplicações. Neste trabalho, as três gaussianas foram definidas de acordo a Tabela 4. Se consideramos curvas fechadas com  $T = 140$  pontos, então a Figura 22 ilustra a configuração das três gaussianas (neste caso  $m = 69$ ).



Tabela 4: Configuração das três gaussianas ( $m$  é o número de colunas de  $\mathbf{S}$ ).

Gaussiana	$\sigma$	$\mu$	$\Delta m$
$g_e$	5	$0.20m$	$0.20m$
$g_c$	5	$0.50m$	$0.15m$
$g_d$	5	$0.85m$	$0.15m$

Figura 22: Configuração das gaussianas  $g_e$ ,  $g_c$  e  $g_d$  para uma aplicação em particular.

## 4.4 Características e Algoritmo de Classificação

A seguir, definimos as nossas características invariantes sob transformações afins. Elas são definidas a partir dos vetores  $\psi_e$ ,  $\psi_c$  e  $\psi_d$ . O algoritmo de treinamento e classificação também são apresentados.

### 4.4.1 Características Propostas

Os vetores  $\psi_e$ ,  $\psi_c$  e  $\psi_d$  foram definidos na Subseção 4.3.3. Usamos cinco características para classificar as formas. Estas características foram escolhidas de acordo aos critérios de seleção de características (Seção 2.7).

- A característica  $F_1$  foi definida na Equação 4.5:

$$F_1 = \frac{\text{Area}(\text{FC}(\mathbf{u})) - \text{Area}(\mathbf{u})}{\text{Area}(\text{FC}(\mathbf{u}))}$$

- A característica  $F_2$  é a soma dos valores negativos de  $\psi_e$ :

$$F_2 = \sum_{\psi_e(t) < 0} \psi_e(t)$$

- A característica  $F_3$  é a soma dos valores positivos de  $\psi_e$ :

$$F_3 = \sum_{\psi_e(t) > 0} \psi_e(t)$$

- A característica  $F_4$  é a soma de  $\psi_c$ :

$$F_4 = \sum \psi_c(t)$$

- A característica  $F_5$  é a soma de  $\psi_d$ :

$$F_5 = \sum \psi_d(t)$$

No início da escolha das características, começamos testando as somas das partes negativas e positivas dos três vetores sob ruído e deformação. Os resultados foram satisfatórios somente para o vetor  $\psi_e$ .

Observe que acabamos reduzindo a enorme quantidade de informação contida numa matriz de áreas em cinco características (vetor de características). Diferentemente do trabalho de Shen ([SHEN; WONG; IP, 1999](#)), que mede a similaridade de duas formas através do casamento de duas matrizes de áreas, vamos medir a similaridade de duas formas através da distância entre seus vetores de características. Isto torna muito mais rápido o processo de classificação. Além disso, parametrizamos o contorno das formas usando o parâmetro comprimento de arco afim. Isto torna as nossas características robustas a parametrização da forma, enquanto que o trabalho de Shen não possui esta propriedade.

A matriz de área normalizada e, por consequência, os vetores  $g_e, g_c$  e  $g_d$  são invariantes sob transformação afim. Assim, todas as características propostas são também invariantes sob transformação afim. Elas também são invariantes à escolha do ponto

inicial. Usamos o algoritmo dos  $k$ -vizinhos mais próximos para classificar as formas. Para poder usar este algoritmo, normalizamos as características para terem a média zero e o desvio-padrão um.

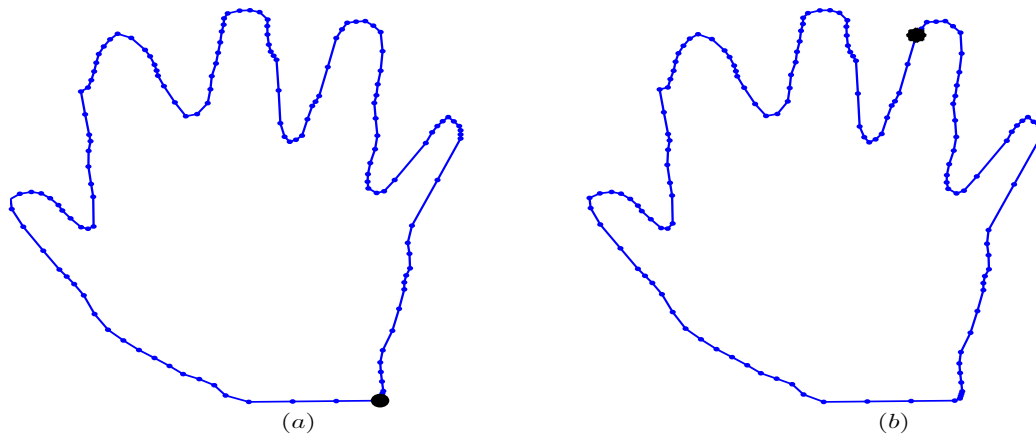


Figura 23: A parametrização da forma é robusta à escolha do ponto inicial.

A Figura 23 ilustra duas formas obtidas a partir da re-parametrização afim de uma mesma forma com diferentes pontos iniciais (representado por pontos maiores). Observamos que esta re-parametrização é robusta à escolha do ponto inicial. Portanto, as características são também robustas à escolha do ponto inicial, como mostrado na Tabela 5:

Tabela 5: As características são robustas à escolha do ponto inicial.

	$F_1$	$F_2$	$F_3$	$F_4$	$F_5$
Forma (a)	0.2064	-3.7753	24.9350	60.9692	53.1867
Forma (b)	0.2066	-3.7685	24.9252	60.9711	53.1849

#### 4.4.2 Análise das Características

Nesta subseção, apresentamos uma análise das características propostas. Esta análise é feita a partir dos vetores  $\psi_e$ ,  $\psi_c$  e  $\psi_d$ . Especificamente, estudamos a mudança dos vetores para diferentes parametrizações de uma mesma forma, a robustez dos vetores para formas similares e a invariância dos vetores sob transformação de cisalhamento.

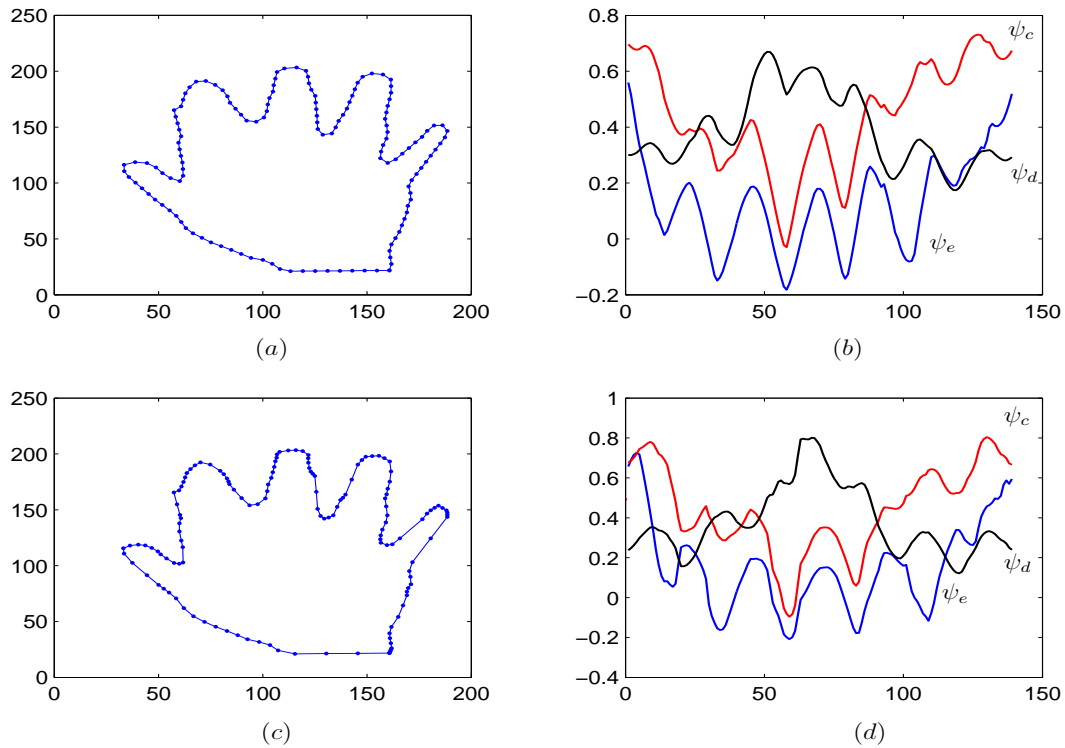


Figura 24: Vetores  $\psi_e$ ,  $\psi_c$  e  $\psi_d$  de uma mesma forma com parametrizações diferentes.

Os três vetores são baseados na soma de área de triângulos formados por três pontos pertencentes ao contorno da forma. Assim, eles são dependentes a parametrização da forma. A Figura 24 ilustra este fato para uma mesma forma parametrizada pelo parâmetro comprimento de arco (forma (a)) e comprimento de arco afim (forma (c)). Nesta figura, (b) e (d) são as curvas que representam os vetores  $\psi_e$ ,  $\psi_c$  e  $\psi_d$  para os contornos (a) e (c), respectivamente. Observe que a dependência da parametrização da forma é maior para  $\psi_c$  e  $\psi_d$ .

Para duas formas similares e parametrizadas através do parâmetro comprimento de arco afim, as curvas que representam os vetores  $\psi_e$ ,  $\psi_c$  e  $\psi_d$  também resultam similares. Isto é ilustrado na Figura 25. Nesta figura, (a) e (c) representam duas formas similares parametrizadas pelo parâmetro comprimento de arco afim, (b) e (d) são as curvas que representam os três vetores para os contornos (a) e (c), respectivamente. Observe que o vetor  $\psi_e$  é o mais robusto.

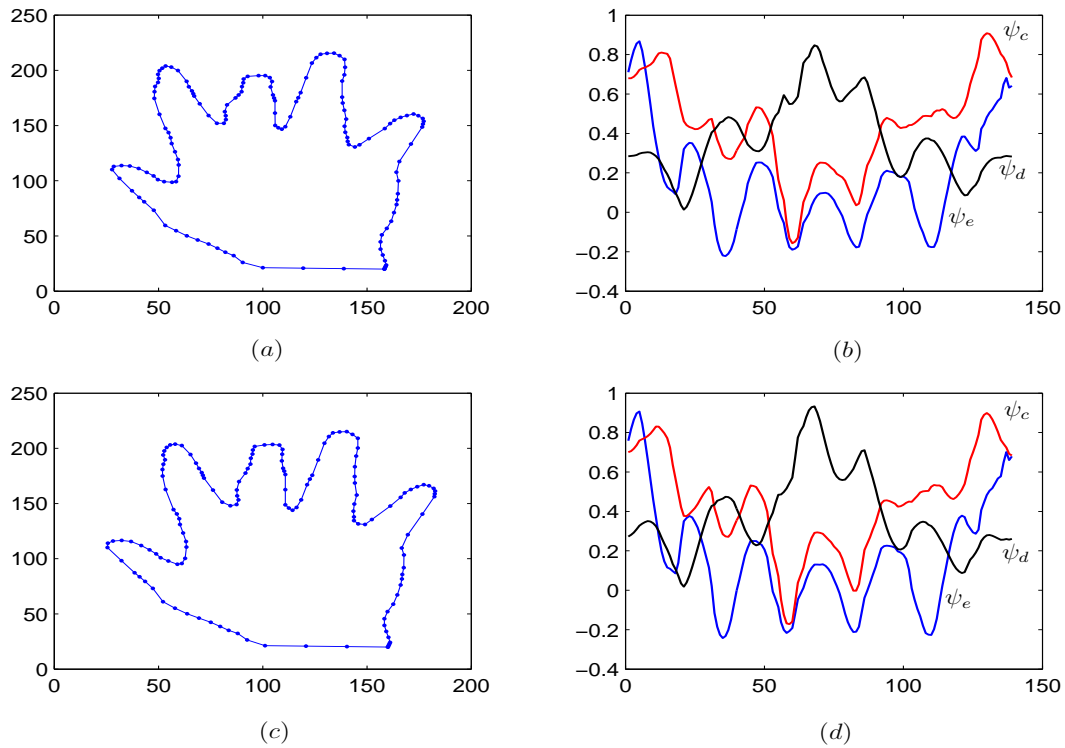


Figura 25: Vetores  $\psi_e$ ,  $\psi_c$  e  $\psi_d$  de duas formas similares com parametrização pelo comprimento de arco afim.

A invariância das curvas que representam os vetores  $\psi_e$ ,  $\psi_c$  e  $\psi_d$  sob transformação afim (especificamente sobre transformação de cisalhamento) é ilustrado na Figura 26. Os contornos deformados ((b), (c) e (d)) são obtidos a partir da aplicação de uma transformação de cisalhamento ( $\kappa = 1, 2$  e  $3$ ) sobre o contorno original (a). Antes de extrair os três vetores, parametrizamos o contorno original e os contornos deformados através do parâmetro comprimento de arco afim. Observe que os três vetores são os mesmos para o contorno original e deformados. A prova disto está no fato de que o parâmetro comprimento de arco afim e os vetores  $\psi_e$ ,  $\psi_c$  e  $\psi_d$  são invariantes sob transformações afins (veja a Subseção 2.5.4 para mais detalhes sobre parametrização afim).

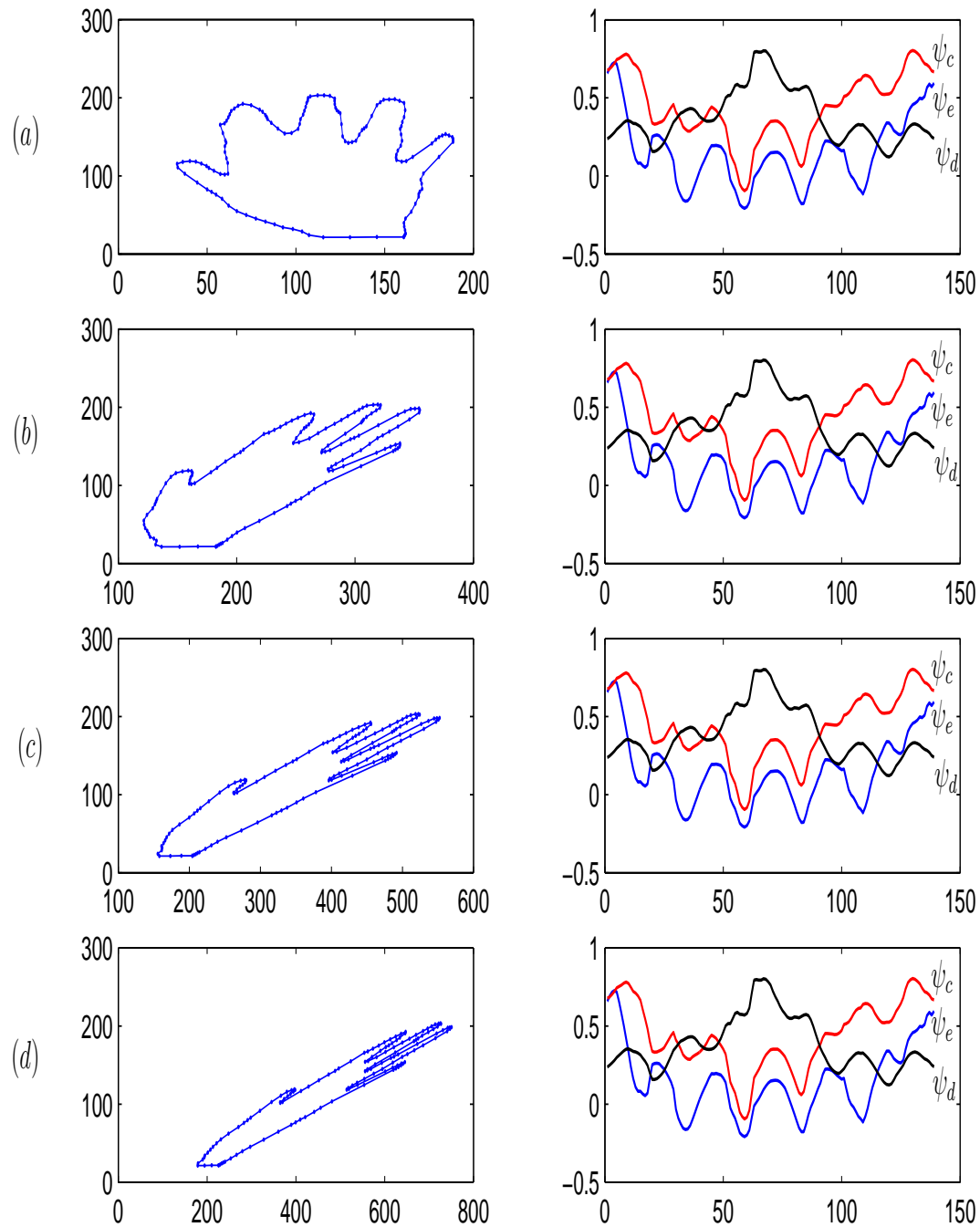


Figura 26: Vetores  $\psi_e$ ,  $\psi_c$  e  $\psi_d$  sob transformação de cisalhamento.

#### 4.4.3 Algoritmo de Treinamento

O algoritmo de treinamento consiste em calcular as cinco características e deve ser aplicado a todos os exemplos de treinamento. Cada exemplo de treinamento deve ser apresentado a este algoritmo junto com a sua correta classificação:

- Normalize o número de pontos da forma.
- Re-parametrize a forma usando o parâmetro comprimento de arco afim (Equação 4.7).
- Calcule e normalize a matriz de áreas.
- Determine os vetores  $\psi_e$ ,  $\psi_c$  e  $\psi_d$ .
- Calcule as cinco características da forma.

#### 4.4.4 Algoritmo de Classificação

O algoritmo de classificação recebe como entrada uma forma de consulta com classificação desconhecida e o número  $k$  a ser usado no algoritmo dos  $k$ -vizinhos mais próximos. A saída deste algoritmo é a classificação da forma.

- Dada uma forma de consulta, calcule as cinco características como descrito no algoritmo de treinamento. Denotemos o vetor de características obtido como  $x$ .
- Procure, no conjunto de treinamento,  $k$  formas que estão situadas mais próximas ao vetor  $x$ . A classificação será a moda das classificações das  $k$  formas de treinamento (isto é, a classificação mais freqüente).

## 4.5 Resultados Experimentais

Aplicamos o nosso método para o reconhecimento supervisionado dos gestos estáticos de mão. Usamos parte dos contornos dos gestos estáticos de mão considerados por Milios e Petrakis (MILIOS; PETRAKIS, 2000). Estes contornos e os programas (implementados em Matlab), necessários para o reconhecimento supervisionado de formas invariante sob transformação afim, se encontram no disquete incluído nesta tese.

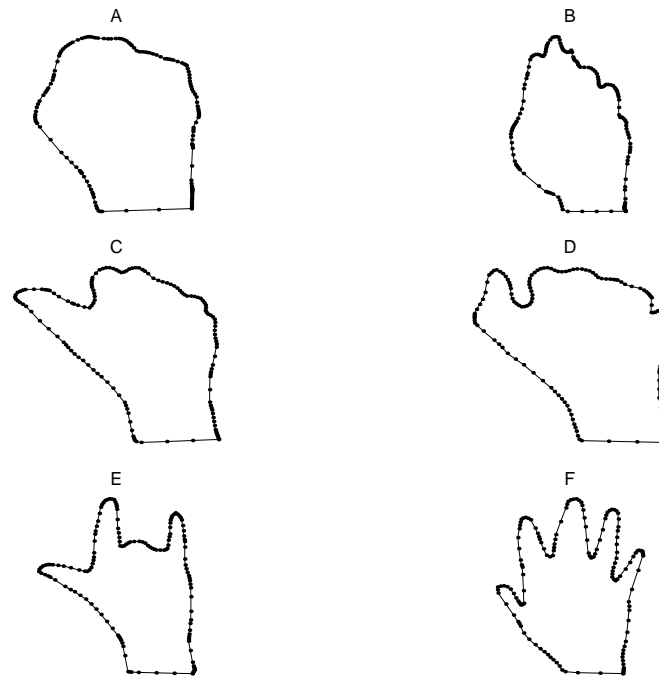


Figura 27: Gestos de mão das classes A, B, C, D, E e F.

Rotulamos manualmente 150 contornos fechados como exemplos de treinamento supervisionado. Eles foram divididos em seis classes, cada uma com 25 exemplos, conforme ilustrado na Figura 27. Todos os contornos foram re-parametrizados usando o parâmetro comprimento de arco afim e normalizados para terem 140 pontos. Este número de pontos foi escolhido experimentalmente de tal forma a não perder informação relevante do contorno da forma. A seguir, mostramos os resultados da classificação sobre formas deformadas e não-deformadas.

A Figura 28 ilustra os vetores  $\psi_e$ ,  $\psi_c$  e  $\psi_d$  para cada uma das seis classes de gestos da mão. Observe as diferenças dos vetores para classes diferentes.

#### 4.5.1 Formas Deformadas

Testamos a invariância afim das nossas características, especificamente para as transformações de rotação e deformação. Tomamos um contorno representativo de cada classe (que não pertence ao conjunto de treinamento) e geramos 9 versões rota-



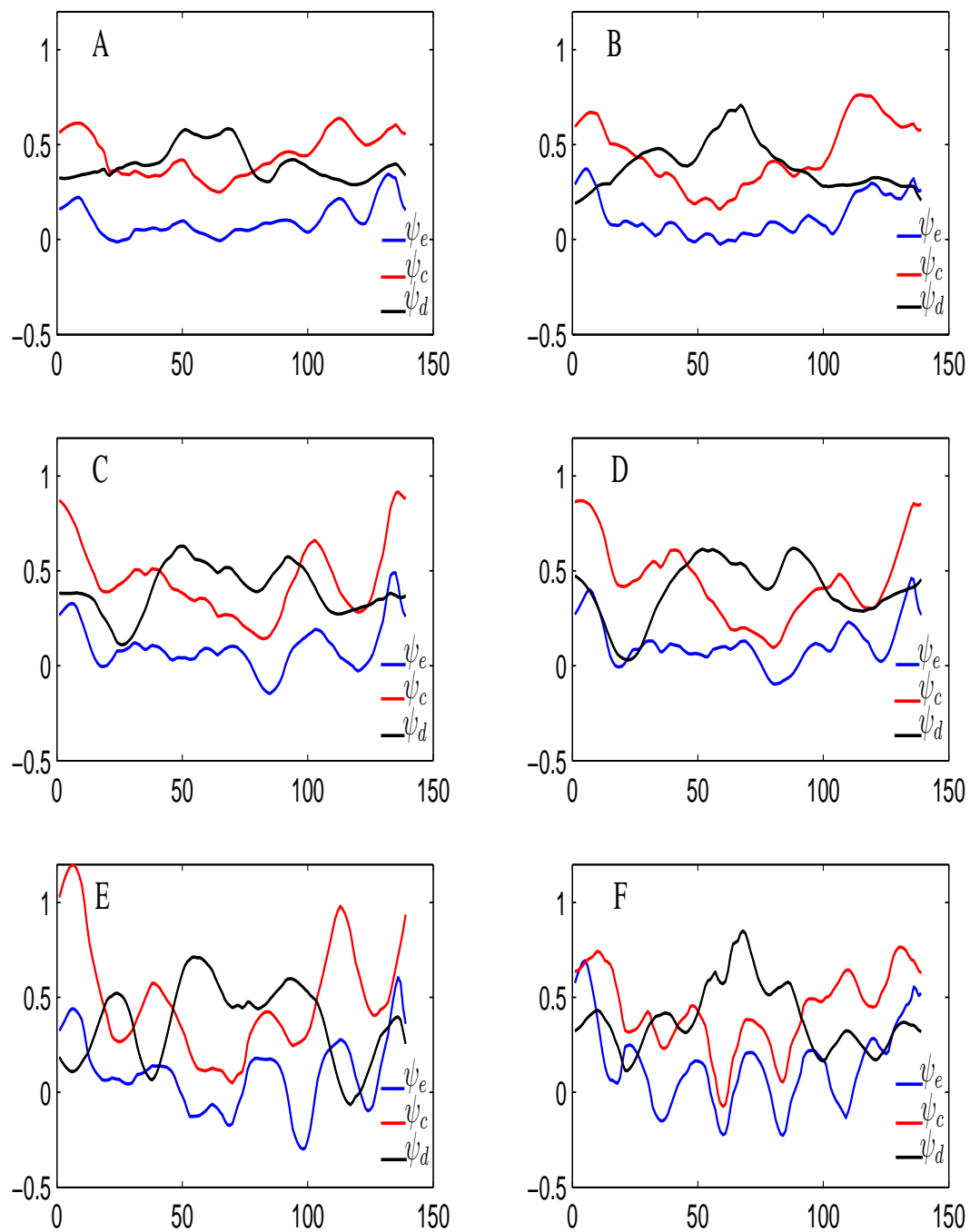


Figura 28: Vetores  $\psi_e$ ,  $\psi_c$  e  $\psi_d$  para cada uma das seis classes de gestos da mão.

cionadas, rodando o contorno original por  $20^\circ$ ,  $40^\circ$ , ...,  $180^\circ$ . Depois, deformamos as  $6 \times 9$  versões rotacionadas com as taxas de deformação  $\kappa = 1, 2$  e  $3$ , obtendo  $3 \times 6 \times 9$  formas deformadas. As Figuras 29, 30 e 31 ilustram a forma original e as versões deformadas com as taxas de deformação  $\kappa = 1, 2$  e  $3$ , respectivamente.

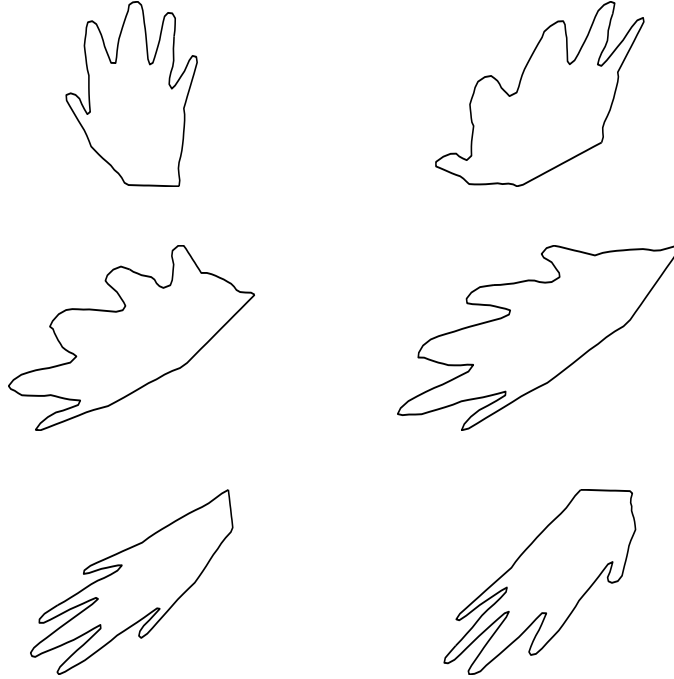


Figura 29: O gesto na parte superior esquerda pertence à classe F. Os outros são as formas deformadas com  $\kappa = 1$ .

Para a fase de reconhecimento, usamos o algoritmo dos  $k$ -vizinhos mais próximos. Testamos este algoritmo com  $k$  variando de 1 até 25. Todas as formas deformadas foram corretamente classificadas, para qualquer  $k \in \{1, \dots, 25\}$ . Isto mostra a robustez das características propostas sob transformação afim.

#### 4.5.2 Formas Não-Deformadas

Também rotulamos manualmente 150 contornos de consulta não-deformados, com 25 contornos em cada classe. O conjunto dos contornos de consulta é disjuncto do conjunto dos contornos de treinamento. A Tabela 6 resume a classificação dos gestos de mão de consulta para as formas não-deformadas, usando o algoritmo dos 5-vizinhos

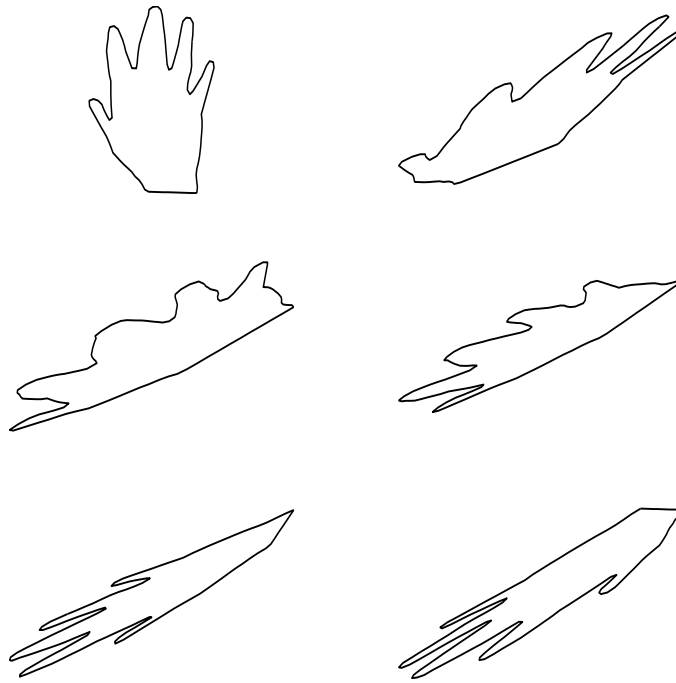


Figura 30: O gesto na parte superior esquerda pertence à classe F. Os outros são as formas deformadas com  $\kappa = 2$ .

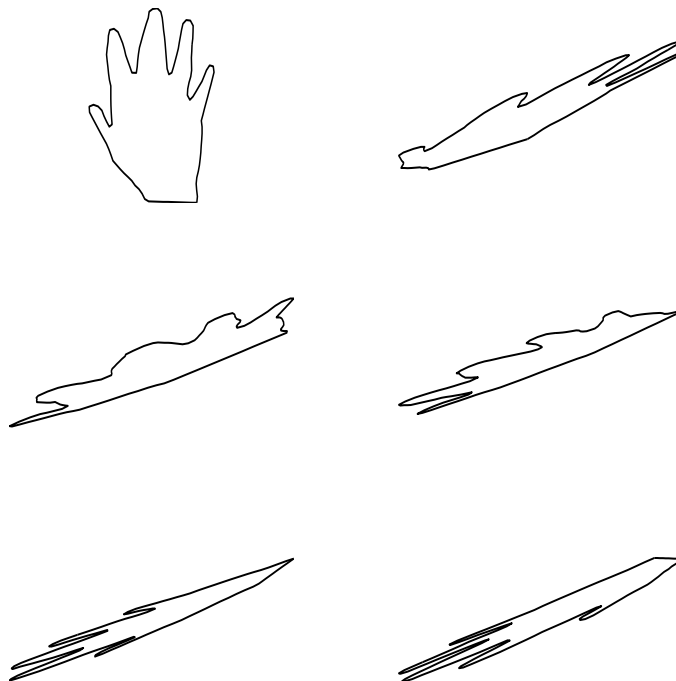


Figura 31: O gesto na parte superior esquerda pertence à classe F. Os outros são as formas deformadas com  $\kappa = 3$ .

Tabela 6: Classificação dos contornos de consulta.

	A	B	C	D	E	F
A	25	0	0	0	0	0
B	1	24	0	0	0	0
C	0	0	25	0	0	0
D	0	1	0	24	0	0
E	0	0	0	0	25	0
F	0	0	0	1	0	24

mais próximos. Nesta tabela, o elemento (linha  $i$ , coluna  $j$ ) corresponde ao número de gestos de mão manualmente rotulados como  $i$  e classificados pelo algoritmo como  $j$ . A taxa de erro é muito baixa, o que mostra a aptidão das características propostas para a classificação de formas. A Figura 32 ilustra alguns contornos de consultas não-deformados e as suas classificações. Nesta figura,  $X/Y$  denota que o contorno foi rotulado manualmente como sendo da classe  $X$  e classificado como sendo da classe  $Y$ . Devido a que as características propostas são versões simplificadas dos vetores  $\psi_e$ ,  $\psi_c$  e  $\psi_d$ , alguns contornos não-deformados não foram corretamente classificados.

## 4.6 Conclusões

Neste capítulo, propusemos um conjunto de novas características invariantes sob transformação afim. Utilizamos estas características para o reconhecimento supervisionado de formas, mas elas podem também ser usadas no reconhecimento não-supervisionado de formas e na indexação e recuperação de imagens baseado em contorno. Parametrizamos as formas usando o parâmetro comprimento de arco afim, para tornar as nossas características robusto a parametrização da forma. Os resultados experimentais mostraram que as nossas características são capazes de discriminar corretamente formas severamente deformadas sob transformação afim.

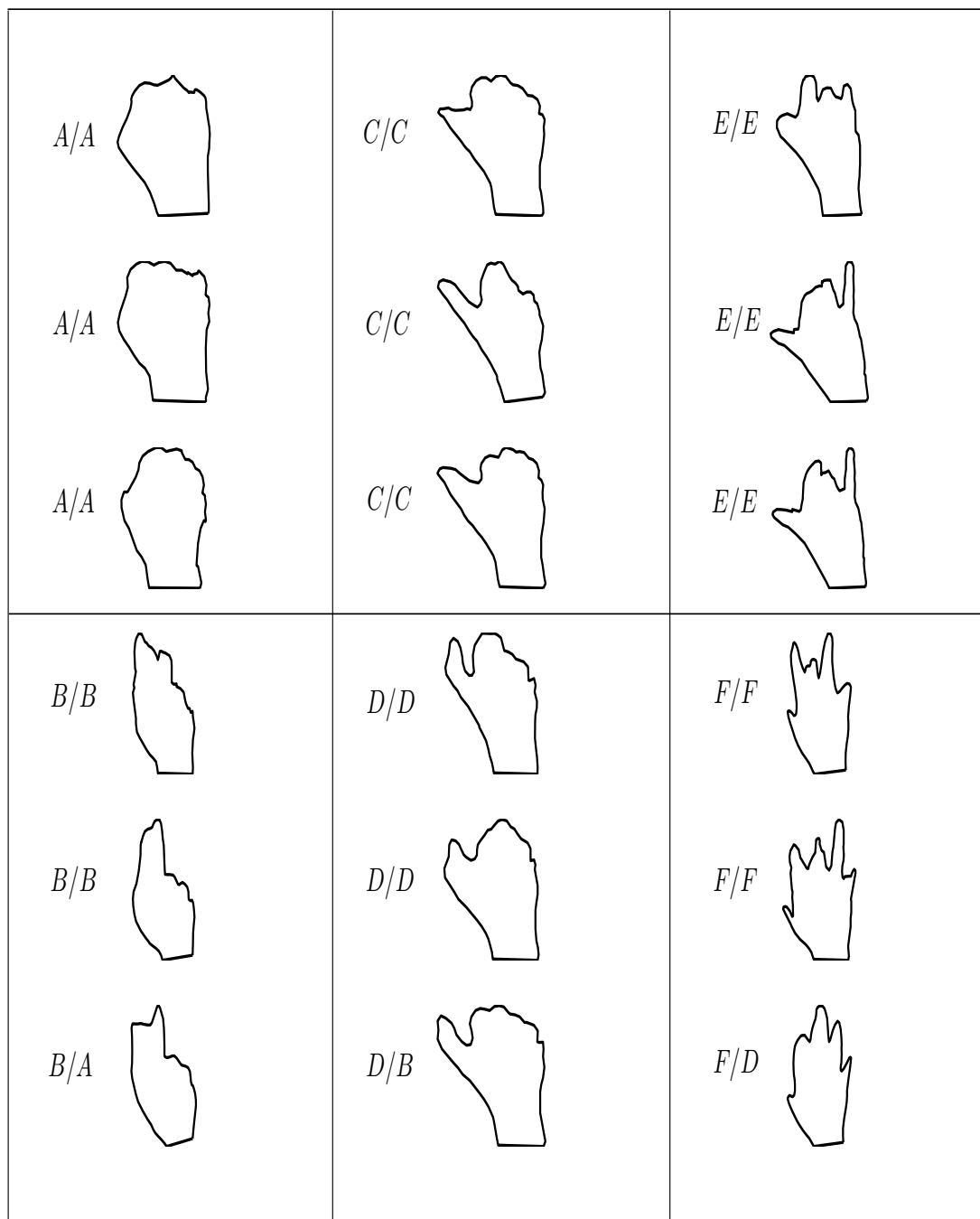


Figura 32: Alguns contornos de consultas não-deformados e as suas classificações.

## 5 CONCLUSÕES

Nesta tese, propusemos dois conjuntos de características invariantes sob transformação de similaridade e transformação afim, respectivamente, para o problema de reconhecimento de formas.

O nosso conjunto de características invariantes sob transformações de similaridade permite reconhecer objetos com uma baixa taxa de erro. Uma técnica baseada na aproximação poligonal do contorno do objeto também é proposta. Esta técnica faz uso da transformada de Fourier para calcular os pontos de alta curvatura. Um conjunto de características são extraídas desde estes pontos de alta curvatura e o classificador de distancia mínima é usado para o reconhecimento do objeto. A técnica proposta é aplicado para o reconhecimento supervisionado de gestos estáticos de mão e a alta taxa de reconhecimento obtido permite aplicações práticas da nossa técnica.

A nossas características invariantes sob transformações afins permitem reconhecer objetos planos, mesmo que eles sejam observados de diferentes pontos de vista. Estas características são baseados numa caracterização local, global e invariante sob transformação afim do contorno das formas. Parametrizamos o contorno das formas usando o parâmetro comprimento de arco afim, para tornar as nossas características robustos à parametrização da forma. Logo que o contorno das formas é descrito pelas características invariantes, o algoritmo do  $k$ -vizinho mais próximo é usado para a classificação de qualquer contorno de consulta. Utilizamos estas características para o reconhecimento supervisionado de gestos estáticos de mão, mas elas podem também ser usadas

no reconhecimento não-supervisionado de formas e na indexação e recuperação de imagens baseado no contorno do objeto. Os resultados experimentais mostraram que as nossas características são capazes de discriminar corretamente formas severamente deformadas sob transformações afins.

Um aspecto fundamental do problema de reconhecimento de formas invariante a mudança do ponto de vista reside na parametrização do contorno da forma. Esta parametrização não deve ser arbitrária, ela deve ser invariante sob transformação afim. As características são extraídas a partir do contorno dos objetos. Conseqüentemente, para este tipo de problema, o valor das características dependem fortemente da parametrização do contorno do objeto. Portanto, para obter um bom desempenho neste tipo de problema, a parametrização dos contornos e as características devem ser invariantes sob transformações afins.

## REFERÊNCIAS

- AHMADIAN, A.; FARAMARZI, E.; SAYADIAN. Image indexing and retrieval using Gabor wavelet and Legendre moments. In: 25TH ANNUAL INTERNATIONAL CONFERENCE ON ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY. *Proceedings*. Cancun, Mexico, 2003. v. 1, p. 560–563.
- ARBTER, K. Affine-invariant Fourier descriptors. In: FROM PIXELS TO FEATURES. *Proceedings*. Amsterdam, the Netherlands: Elsevier Science, 1989.
- ARBTER, K.; SNYDER, W.; BURKHARDT, G. H. H. Applications of affine-invariant Fourier descriptors to recognition of 3-d objects. *IEEE Trans. Pattern Analysis and Machine Intelligence*, v. 12, n. 7, p. 640–646, 1990.
- ARICA, N.; VURAL, F. T. Y. Bas: a perceptual shape descriptor based on the beam angle statistics. *Pattern Recognition Letters*, v. 24, n. 9-10, p. 1627–1639, June 2003.
- ATTNEAVE, F. Some informational aspects of visual perception. *Psychological Review*, v. 61, p. 183–193, 1954.
- BALLARD, D. H.; BROWN, C. M. *Computer Vision*. Englewood Cliffs, N.J.: Prentice-Hall, 1982.
- BEUS, H. L.; TIU, S. S. H. An improved corner detection algorithm based on chain-coded plane curves. *Pattern Recognition*, v. 20, p. 291–296, 1987.
- CHAKER, F.; BANNOUR, M. T.; GHORBEL, F. A complete and stable set of affine-invariant Fourier descriptors. In: INTERNATIONAL CONFERENCE ON IMAGE ANALYSIS AND PROCESSING. *Proceedings*. Mantova, Italy, 2003. p. 578–581.
- CHAKRABARTI, K. et al. Similar shape retrieval in mars. In: IEEE INTERNATIONAL CONFERENCE ON MULTIMEDIA AND EXPO. *Proceedings*. NY, USA, 2000. v. 2, p. 709–712.
- COSTA, L. F.; CESAR, R. M. *Shape Analysis and Classification: Theory and Practice*. Boca Raton, FL: CRC Press, 2001.
- DIONISIO, C. R. P.; KIM, H. Y. A supervised shape classification technique invariant under rotation and scaling. In: INT. TELECOMMUNICATIONS SYMPOSIUM. *Proceedings*. Natal, Brasil, 2002. p. 533–537.
- \_\_\_\_\_. New features for affine-invariant shape classification. In: IEEE INT. CONF. ON IMAGE PROCESSING. *Proceedings*. Singapore, 2004. p. 2135–2138.



- \_\_\_\_\_. Novas características para classificação de formas invariantes sob transformações afins. In: XV CONGRESSO BRASILEIRO DE AUTOMÁTICA (CBA2004). *Proceedings*. Gramados, RS, Brasil, 2004.
- FISCHLER, M. A.; WOLF, H. C. Locating perceptually salient points on planar curves. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 16, n. 2, p. 113–129, February 1994.
- FLUSSER, J.; SUK, T. Pattern recognition by affine moment invariants. *Pattern Recognition*, v. 26, n. 1, p. 167–174, 1993.
- FORTUNA, A. O. *Técnicas Computacionais para Dinâmica dos Fluidos - Conceitos Básicos e Aplicações*. São Paulo: Editora da Universidade de São Paulo, 2000.
- GAUTHIER, J. P.; BORNARD, G.; SILBERMANN, M. Motions and pattern analysis: harmonic analysis on motion groups and their homogeneous spaces. *IEEE Trans. on Systems, Man and Cybernetics*, v. 21, n. 1, Jan/Feb 1991.
- GHORBEL, F. Stability of invariant Fourier descriptors and its inference in the shape classification. In: IEEE 11TH INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION. *Proceedings*. Hague, NL, 1992.
- GOLD, S.; RANGARAJAN, A. A graduated assignment algorithm for graph matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 18, n. 4, p. 377–388, 1996.
- GONZALEZ, R. C.; WOODS, R. E. *Digital Image Processing*. Upper Saddle River, NJ: Prentice Hall, 2002.
- HADDADNIA, J.; AHMADI, M.; FAEZ, K. An efficient method for recognition of human faces using higher orders pseudo Zernike moment invariant. In: FIFTH IEEE INTERNATIONAL CONFERENCE ON AUTOMATIC FACE AND GESTURE RECOGNITION. *Proceedings*. Washington, USA, 2002. p. 315–320.
- HORN, B. K. P. *Robot Vision*. Cambridge MA: MIT Press, 1986.
- HUANG, Z.; COHEN, F. Affine-invariant b-spline moment for curve matching. In: IEEE COMPUTER SOCIETY CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION. *Proceedings*. Seattle, WA, 1994. p. 490–495.
- KHOTANZAD, A.; HONG, Y. H. Invariant image recognition by Zernike moments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 12, n. 5, p. 489–497, May 1990.
- KIM, H.-K. et al. A modified Zernike moment shape descriptor invariant to translation, rotation and scale for similarity-based image retrieval. In: IEEE INTERNATIONAL CONFERENCE ON MULTIMEDIA AND EXPO, 2000. *Proceedings*. NY, USA, 2000. v. 1, p. 307–310.
- KOPLOWITZ, J.; PLANTE, S. Corner detection for chain coded curves. *Pattern Recognition*, v. 28, n. 6, p. 843–852, 1995.

- KWOK, S. K. W.; POON, J. C. H. Viewpoint-invariant Fourier descriptors for 3d planar shape representation. *Electronics Letters*, v. 32, n. 19, p. 1775–1776, Sept. 1996.
- LATECKI, L. J.; LAKÄMPER, R. Shape similarity measure based on correspondence of visual parts. *IEEE Trans. PAMI*, v. 22, n. 10, p. 1185–1190, 2000.
- LATECKI, L. J.; LAKÄMPER, R.; ECKHARDT, U. Shape descriptors for non-rigid shapes with a single closed contour. In: IEEE CONFERENCE COMPUTER VISION AND PATTERN RECOGNITION. *Proceedings*. South Carolina, USA, 2000. v. 1, p. 424–429.
- LATECKI, L. J.; LAKÄMPER, R.; WOLTER, D. Shape similarity and visual parts. In: INT. CONF. ON DISCRETE GEOMETRY FOR COMPUTER IMAGERY (DGCI). *Proceedings*. Napoli, Italy, 2003. p. 34–51.
- LIN, I.-J.; KUNG, S. Y. Coding and comparison of dags as a novel neural structure with applications to on-line handwritten recognition. *IEEE Trans. Signal Processing*, 1996.
- LIU, H. C.; SRINATH, M. D. Corner detection from chain-code. *Pattern Recognition*, v. 23, n. 1-2, p. 51–68, 1990.
- MARTINEZ, J.; THOMAS, F. Efficient computation of local geometric moments. *IEEE Transactions on Image Processing*, v. 11, n. 9, p. 1102–1111, 2002.
- MILIOS, E.; PETRAKIS, E. G. M. Shape retrieval based on dynamic programming. *IEEE Transactions on Image Processing*, v. 9, n. 1, p. 141–146, 2000.
- MOKHTARIAN, F. Silhouette-based isolated object recognition through curvature-scale space. *IEEE Trans. PAMI*, v. 17, n. 5, p. 539–544, 1995.
- MOKHTARIAN, F.; ABBASI, S. Affine curvature scale space with affine length parametrisation. *Pattern Analysis & Applications*, v. 4, p. 1–8, 2001.
- \_\_\_\_\_. Shape similarity retrieval under affine transforms. *Pattern Recognition*, v. 35, n. 1, p. 31–41, 2002.
- MOKHTARIAN, F.; ABBASI, S.; KITTLER, J. Efficient and robust retrieval by shape content through curvature scale space. In: SMEULDERS, A. W. M.; JAIN, R. (Ed.). *Proceedings*. Singapore: World Scientific Publishing, 1997. p. 51–58.
- MOKHTARIAN, F.; MACKWORTH, A. K. A theory of multiscale, curvature-based shape representation for planar curves. *IEEE Pattern Analysis and Machine Intelligence*, v. 14, n. 8, p. 789–805, Aug. 1992.
- MOONS, T. et al. Foundations of semi-differential invariants. *International Journal of Computer Vision*, v. 14, n. 1, p. 25–47, January 1995.
- NALWA, V. S. *A Guided Tour of Computer Vision*. Reading, MA: Addison-Wesley, 1993.

- O'HIGGINS, P. *Methodological Issues in the Description of Forms*. P. E. Lestrel. [S.l.]: Cambridge University Press, 1997.
- PAUWELS, E. J. et al. Recognition of planar shapes under affine distortion. *International Journal of Computer Vision*, v. 14, n. 1, p. 49–65, January 1995.
- PINHEIRO, A. M. G.; IZQUIERDO, E.; GHANHARI, M. Shape matching using a curvature based polygonal approximation in scale-space. In: INTERNATIONAL CONFERENCE ON IMAGE PROCESSING. *Proceedings*. Vancouver, Canada, 2000. v. 2, p. 10–13.
- PREPARATA, F.; SHAMOS, M. *Computational Geometry: An Introduction*. New York: Springer-Verlag, 1985. (Texts and Monographs in Computer Science).
- RATTARANGSI, A.; CHIN, R. Scale-based detection of corners of planar curves. *IEEE Trans. PAMI*, v. 14, n. 4, p. 430–449, 1992.
- RIPLEY, B. D. *Pattern Recognition and Neural Networks*. [S.l.]: Cambridge University Press, 1996.
- RIVLIN, E.; WEISS, I. Local invariants for recognition. *IEEE Trans. PAMI*, v. 17, n. 3, p. 226–238, 1995.
- ROSENFELD, A.; WESZKA, J. S. An improved method of angle detection on digital curves. *IEEE Tran. Comput.*, v. 24, p. 940–941, 1975.
- RUSS, J. C. *The Image Processing Handbook*. 3rd. ed. Boca Raton, FL: CRC Press, 1999.
- SAPIRO, S.; TANNENBAUM, A. Affine invariant scale-space. *International Journal of Computer Vision*, v. 11, n. 1, p. 25–44, 1993.
- SHEN, D.; WONG, W.; IP, H. Affine-invariant image retrieval by correspondence matching of shapes. *Image and Vision Computing*, v. 17, p. 489–499, 1999.
- SIDDIQI, K. et al. Shock graphs and shape matching. *International Journal of Computer Vision*, v. 35, n. 1, p. 13–32, November 1999.
- STANISLAW, M. *Description invariante et locale des formes planes - Application à l'indexation d'une base d'images*. Tese (Doutorado) — UNIVERSITE DE VALENCIENNES ET DU HAINAUT CAMBRESIS - ENIC/INT, 1999.
- SUNDAR, H. et al. Skeleton based shape matching and retrieval. In: SHAPE MODELING INTERNATIONAL. *Proceedings*. Seoul, Korea, 2003. p. 130–139.
- TEH, C. H.; CHIN, R. T. On the detection of dominant points on digital curves. *IEEE Trans. PAMI*, v. 11, p. 859–872, 1989.
- WOOD, J. Invariant pattern recognition: A review. *Pattern Recognition*, v. 29, n. 1, p. 1–17, 1996.

YINAN, S.; WEIJUN, L.; YUECHAO, W. United moment invariants for shape discrimination. In: IEEE INTERNATIONAL CONFERENCE ON ROBOTICS, INTELLIGENT SYSTEMS AND SIGNAL PROCESSING. *Proceedings*. Changsha, China, 2003. v. 1, p. 88–93.

YU, M. P.; LO, K. C. Object recognition by combining viewpoint invariant Fourier descriptor and convex hull. In: INTERNATIONAL SYMPOSIUM ON INTELLIGENT MULTIMEDIA, VIDEO AND SPEECH PROCESSING. *Proceedings*. Hong Kong, 2001. p. 401–404.

ZHAO, A.; CHEN, J. Affine curve moment invariants for shape recognition. *Pattern Recognition*, v. 30, n. 6, p. 895–901, 1997.

# APÊNDICE A - ÁREA DE UM POLÍGONO

Seja  $\mathbf{u} : [0, T] \rightarrow \mathbb{R}^2$  uma curva paramétrica fechada, de período  $T$ , ou seja:

$$\mathbf{u}(t) = (x(t), y(t)) \quad \text{para } t \in [0, T].$$

## A.1 Área de um Triângulo

Seja  $\mathbf{u}(t_1) = (x_1, y_1)$ ,  $\mathbf{u}(t_2) = (x_2, y_2)$  e  $\mathbf{u}(t_3) = (x_3, y_3)$  três pontos pertencentes ao contorno  $\mathbf{u}$ . A área “sinalizada” do triângulo determinado pelos pontos  $\mathbf{u}(t_1)$ ,  $\mathbf{u}(t_2)$  e  $\mathbf{u}(t_3)$  é dada pela seguinte fórmula:

$$\text{Area}(\Delta(\mathbf{u}(t_1), \mathbf{u}(t_2), \mathbf{u}(t_3))) = \frac{1}{2} \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} \quad (\text{A.1})$$

onde  $|\cdot|$  denota a função determinante. Se  $\mathbf{u}(t_1)$ ,  $\mathbf{u}(t_2)$  e  $\mathbf{u}(t_3)$  são orientados no sentido anti-horário o valor da área é positivo, caso contrário é negativo.

## A.2 Área de um Polígono

O contorno fechado  $\mathbf{u}$  é um polígono simples. Então, a área de  $\mathbf{u}$  pode ser calculada pela seguinte fórmula:

$$\text{Area}(\mathbf{u}) = \sum_{i=0}^{T-1} \text{Area}(\Delta(\mathbf{u}(0), \mathbf{u}(i), \mathbf{u}(i+1))) \quad (\text{A.2})$$

## APÊNDICE B - PROGRAMAS E DADOS

Todos os programas usados neste trabalho e os dados usados para o problema de reconhecimento de formas invariante afim, se encontram no disquete incluído nesta tese. Cada programa contém a descrição sobre sua função. Todos estes programas foram implementadas em Matlab 6.5.

Os seguintes diretórios contém os dados usados para o reconhecimento de formas invariante sob transformação afim:

- **treinamento**: são os dados para treinamento do sistema. A partir deste conjunto é gerado o espaço de características.
- **BD?Hand\_Query\_Afim**, onde ?=1,2,3 denota a taxa de deformação. São as formas que foram geradas a partir do conjunto de formas **queries\_representatives**. Estes conjuntos de formas são usadas para testar a invariância e robustez das características sobre formas deformadas.
- **consulta**: conjunto de formas usado para testar as características sobre formas não-deformadas.
- **queries\_ruído**: Este conjunto contém formas ruidosas. Ele é usado para testar a robustez e invariância das características sob ruído.

Os principais programas para reconhecimento afim de formas são:

- **treina\_afim**: treina o sistema calculando o espaço de características.

- **erro\_afim**: este programa classifica qualquer conjunto de formas.
- **par\_affine\_length2**: parametriza uma curva através do parâmetro comprimento de arco afim.
- **features\_affine**: dado o contorno de um objeto, calcula as características invariante afim.
- **matriz\_area**: calcula a matriz de áreas de um contorno.
- **normaliza\_matriz\_area**: normaliza uma matriz de áreas.
- **simplif\_matriz\_area**: extraí os três vetores de informação a partir de uma matriz de áreas normalizada.
- **normaliza\_feat**: normaliza o espaço de características para terem a média zero e o desvio-padrão um.
- **classif\_k\_vizinho**: algoritmo para calcular os  $k$ -vizinhos mais próximos.
- **areacontorno**: calcula a área de um contorno fechado.
- **area2**: calcula a área de um triângulo.
- **curv**: lê uma curva contida num arquivo texto.
- **TAffine**: calcula a transformação afim de um contorno.

Os arquivos Fa, Fb, Fc, Fd, Fe e Ff contém o espaço de características para cada uma das seis classes. Os arquivos Wa, Wb, Wc, Wd, We e Wf são os vetores de pesos para cada uma das seis classes.