

UNIVERSIDADE FEDERAL FLUMINENSE

Raphael dos Santos Evangelista

**Detecção e Descrição de Características Visuais
Invariantes à Transformações de Similaridade, Afins e
Projetivas**

NITERÓI

2016

UNIVERSIDADE FEDERAL FLUMINENSE

Raphael dos Santos Evangelista

**Detecção e Descrição de Características Visuais
Invariantes à Transformações de Similaridade, Afins e
Projetivas**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Computação da Universidade Federal Fluminense como requisito parcial para a obtenção do Grau de Mestre em Computação. Área de concentração: Computação Visual.

Orientador:

Leandro Augusto Frata Fernandes

NITERÓI

2016

Raphael dos Santos Evangelista

Detecção e Descrição de Características Visuais Invariantes à Transformações de
Similaridade, Afins e Projetivas

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Computação da Universidade Federal Fluminense como requisito parcial para a obtenção do Grau de Mestre em Computação. Área de concentração: Computação Visual.

Aprovada em Abril de 2016.

BANCA EXAMINADORA

Prof. Leandro Augusto Frata Fernandes - Orientador, UFF

Prof. Luiz Carlos Pacheco Rodrigues Velho, IMPA

Prof. Cristina Nader Vasconcelos, UFF

Niterói

2016

"Audentis fortuna iuvat"

- Virgilius

Agradecimentos

Ofereço meus sinceros agradecimentos à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela bolsa de estudos à mim providenciada e à Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) pelo financiamento parcial de equipamentos para o projeto *Reconhecimento Automático de Objetos Tridimensionais a Partir de Características Salientes Invariantes*, processo: E-26/110.484/2014.

Sou grato ao meu orientador, Leandro Augusto Frata Fernandes, por seu auxílio e suporte excepcionais, que moldaram fortemente uma ainda jovem formação acadêmica estímulo para a pesquisa. Seus conhecimentos afiados tornaram-se meus guias em diversos momentos de contratempo, muitas das vezes, em reuniões de fim de semana. Espero que o resultado final tenha alcançado o mesmo nível de excelência inspirado por você.

À minha mãe, Ana Paula Rodrigues dos Santos, por seu apoio incansável diante de todos os momentos difíceis. Não há mais ninguém à quem eu me dedique tanto para dar orgulho, às vezes com um certo receio de falhar. A minha força para seguir, no entanto, advém dos muitos conselhos que recebi (e ainda recebo) da minha avó, Lúcia Rodrigues dos Santos, a quem devo o meu otimismo diante dos obstáculos que jamais deixarão de surgir. Desafios existem e são superados com preparo. Devo à meu pai, Júlio César Cardozo Evangelista, àquele que me apresentou desde cedo sua profissão e, principalmente, como se comportar diante de um problema ainda não solucionado, que a persistência é aliada ao êxito.

Agradeço, por fim, à minha amada, Luana Magalhães Uchôa, por seu apoio e carinho, além de sempre acreditar na minha capacidade. Muitas foram as madrugadas e os fins de semana dedicados à esta dissertação e você sempre esteve ao meu lado, paciente, sabendo que sua segurança seria fundamental para me prover serenidade para alcançar meus objetivos.

Agradeço a todos que me apoiaram, tenham certeza que cada um de vocês contribuiu fortemente para a conclusão deste trabalho.

Resumo

A habilidade de reconhecer padrões visuais é essencial para o desempenho de tarefas do dia-a-dia. Desde sua criação até os dias atuais, os computadores foram desenvolvidos e melhorados para nos auxiliar em atividades, incluindo-se desafios que exijam discernimento visual. Sistemas computacionais com habilidades de visão foram projetados e aplicados em diversas áreas, como na indústria para automação de processos e montagens, mapeamento de ambientes e posicionamento automatizado para a robótica, reconhecimento facial para controle de acesso, dentre outras.

Independentemente do desafio visual proposto, a tarefa mais elementar, ainda que complexa, envolve o ato de identificar e reconhecer objetos ou partes de um. Diversas técnicas foram propostas e estão disponíveis na literatura para resolver este problema. Porém, todas apresentam limitações quanto ao reconhecimento de uma mesma região sob diferentes perspectivas. As soluções existentes na literatura reconhecem o poder discriminativo de *junções* em imagens. No entanto, não incorporam em suas abordagens informações geométricas intrínsecas à morfologia de uma junção. A utilização da geometria das junções possui motivação biológica ao se basear no próprio sistema visual humano e, também, pois as junções de um objeto são bem localizadas e podem ser identificadas, ainda que tais objetos se encontrem em cenas completamente distintas.

A abordagem proposta neste trabalho se fundamenta na formalização matemática de junções em imagens contínuas e na formulação de monômios invariantes à transformações projetivas para serem aplicados sobre a geometria das junções extraídas de imagens digitais, assim como uma estratégia para representar uma imagem a partir de histogramas baseados nos valores invariantes calculados para suas junções. A palavra *invariante* deve ser reforçada, pois implica na não modificação de um valor independentemente da transformação projetiva aplicada, o que se diferencia da afirmação de que o *comportamento é invariante*, como ocorre com regiões e texturas. Tradicionalmente, as técnicas para descrever características e regiões interessantes se baseiam somente em informações de variações de intensidade e magnitude, como no uso de regiões e texturas, o que pode ser um fator limitante para reconhecer objetos sob transformações afins e projetivas.

Os experimentos realizados mostram a viabilidade da estratégia proposta. Eles abrangem tanto um conjunto de imagens transformadas artificialmente quanto transformações obtidas em situações reais, onde as imagens são capturadas em pontos de vista diferentes de uma mesma cena.

Palavras-chave: detecção de características visuais, descrição de características visuais, comparação de características visuais, junção, invariante, monômios invariantes, teoria de invariantes, teorema de Cayley-Hamilton, transformação de similaridade, transformação afim, transformação projetiva.

Abstract

The ability to recognize visual patterns is essential to complete our most common daily tasks. Since the beginning, computers have been developed and enhanced to support us in our activities, including those requiring visual capabilities. Computational vision systems were developed and applied to a countless number of problems including autonomous processes and assembly in industry, mapping and positioning automated by robots inside unknown places, face recognition to provide security and access control, among others.

Regardless of the visual challenge, the most basic task, but complex as well, is the identification, recognition and classification of objects or parts of them. A range of solutions were proposed and are available in the literature that solves this problem but, they all present limitations to recognize the same region under different perspectives. Some of them use junctions to determine whether a region is *distinct* from others, but their geometric properties are overlooked. The use of junction's geometry has motivation in biology, based on our own visual system, and because they are well localized and could be identified, even when they were in a different scene configuration.

The main goal of this dissertation is to present the mathematical formalization of continuous junctions into images and the formulation of monomials invariant under projective transformations which could be calculated from the geometry of junctions extracted from digital images and a strategy to describe images using histograms based on invariant values calculated from that junctions. The word *invariant* has a point. It means that a value does not change even when the junction is transformed by the application of a homography. As a matter of fact, it is not the same as to affirm that something has the *same behaviour*, as usually occurs with regions and textures. Even though regions are considered well localized, the lack of using geometry information of the inside objects, as the most of the solutions used to consider, could make the recognition, under affine and projective transformation, a harder task to achieve.

The experiments show the feasibility of the proposed approach. They consider a set of images transformed artificially and transformed by changing camera's position, using a set of images truly taken under different points of view of the same scene.

Keywords: visual features detection, visual features description, visual features matching, junction, invariant, invariant monomials, invariant theory, Cayley-Hamilton theorem, homography, similarity transformation, affine transformation, projective transformation.

Lista de Figuras

1.1	Destaque de junções e resumo do processo proposto	3
2.1	Importância das junções para o reconhecimento visual	7
2.2	Processo de descrição de características pelo SIFT.	9
2.3	Mapas de comparação entre pixels para descritores binários.	11
2.4	Padrões circulares para comparação de intensidades em descritores binários.	11
3.1	Importância das junções para o reconhecimento visual	16
3.2	Interpretação geométrica de vetores em \mathbb{P}^2 como pontos no plano homogêneo π	18
3.4	Diferentes tipos de seções cônicas	20
3.5	Representações primal e dual de uma mesma elipse	20
3.6	Hierarquia do grupo de transformações projetivas	23
3.7	Diferentes projeções de pontos p no subespaço \mathbb{P}^2	24
3.8	Produto cruzado entre os pontos \mathbf{a} , \mathbf{b} , \mathbf{c} e \mathbf{d} colineares	25
5.1	Diferentes classificações de junções por formato.	36
5.2	<i>Plot</i> de pontos $(\beta_{123}, \beta_{231}, \beta_{312})$ calculados a partir de n amostras aleatórias de junção utilizando o monômio invariante afim	40
6.1	Versões discretizadas de uma 2-junção, uma 3-junção e uma 4-junção	45
6.2	Resumo do processo de descrição de regiões de interesse	46
6.3	Procedimento para quebra de cadeias de pixel em segmentos retilíneos	49
6.4	Exemplo de conversão de uma 4-junção discreta em um 4-junção contínua.	50
6.5	Contribuição de cada pixel \mathbf{p} para com o segmento ao qual pertence	52
6.6	Discriminação dos segmentos digitais a partir do mapa de distâncias	53

6.7	Contribuição de cada pixel \mathbf{p} para com a orientação relativa ao eixo de referência \mathbf{ref}	54
6.8	Identificação de junções impróprias	55
6.9	<i>Plot</i> do monômio invariante projetivo	56
6.10	Processo de amostragem e aproximação da função acumulada para discretização uniforme do monômio invariante sob transformações projetivas. . .	57
6.11	Região do <i>plot</i> do monômio invariante projetivo que representa a função acumulada de \tilde{a}	58
6.12	<i>Plot</i> do monômio invariante projetivo e da função \hat{u}	58
6.13	Processo de amostragem e aproximação da função acumulada para discretização uniforme do monômio invariante sob transformações de similaridade	61
7.1	Experimento para ilustrar a robustez dos valores invariantes	64
7.2	Experimento para atestar a robustez dos valores invariantes sob transformações afim	65
7.3	Exemplos de imagens utilizadas para testar a técnica sobre situações reais com transformações induzidas	66
7.4	Exemplos de transformações induzidas nas imagens	68
7.5	Características mais semelhantes com transformações induzidas: caso <i>Kid</i> .	70
7.6	Características mais semelhantes com transformações induzidas: caso <i>Fourviere</i>	71
7.7	Características mais semelhantes com transformações induzidas: caso <i>Pueblo Bonito</i>	72
7.8	Características mais semelhantes com transformações livres: caso <i>Michelangelo</i>	73
7.9	Características mais semelhantes com transformações livres: caso <i>OpenCV</i>	74
7.10	Características mais semelhantes com transformações livres: caso <i>La Mexicana</i>	74

Lista de Tabelas

7.1	Matriz de confusão sobre cada um dos casos experimentados	66
7.2	Matriz de confusão sobre todas as imagens transformadas artificialmente .	69
7.3	Matriz de confusão sobre todas imagens capturadas em diferentes pontos de vista	73

Lista de Símbolos

SUBESPAÇOS

\mathbb{R}	Conjunto dos números reais.
\mathbb{R}^3	Espaço vetorial 3-dimensional sobre \mathbb{R} .
$\mathbb{R}^{3 \times 3}$	Espaço das matrizes quadradas 3×3 .
\mathbb{P}^2	Espaço projetivo 2-dimensional.

CONSTANTES E VETORES

a, b, \dots	Valores constantes em geral.
$\mathbf{a}, \mathbf{b}, \dots$	Vetores n -dimensionais.

MATRIZES

$\mathcal{A}, \mathcal{B}, \dots$	Matrizes em geral.
a_{ij}, b_{ij}, \dots	Elementos matriciais.
$\mathcal{A}^\top, \mathcal{B}^\top, \dots$	Transposição de matrizes.
$\mathcal{A}^{-1}, \mathcal{B}^{-1}, \dots$	Inversão de matrizes.

CONJUNTOS

$\mathcal{A}, \mathcal{B}, \dots$	Conjuntos em geral.
-----------------------------------	---------------------

MONÔMIOS INVARIANTES

α, β, \dots	Monômios invariantes propostos.
$\Gamma, \Upsilon, \Lambda, \dots$	Monômios invariantes existentes.

Sumário

1	Introdução	1
1.1	Proposta de Dissertação	2
1.2	Resultados e Contribuições	4
2	Trabalhos Relacionados	5
2.1	Detectores de Características	6
2.2	Descritores de Características	7
2.2.1	Descritores derivados do SIFT	8
2.2.2	Descritores derivados do SURF	9
2.2.3	Descritores derivados do BRIEF	10
2.2.4	Descritores baseados em contorno	11
2.3	Comparador de Características	12
2.4	Discussão	13
3	Conceitos sobre Características Visuais, Geometria Projetiva e Transformações Lineares	15
3.1	Características Visuais de Interesse	15
3.2	Estruturas Geométricas em Imagens	17
3.2.1	Pontos finitos e direções em \mathbb{P}^2	17
3.2.2	Retas em \mathbb{P}^2	18
3.2.3	Seções cônicas em \mathbb{P}^2	19
3.2.4	Feixe e rede de cônicas	21

3.2.5	Equivalência entre pontos e cônicas duais	21
3.3	Transformações Lineares Projetivas	22
3.3.1	Grupo das isometrias (transformações rígidas)	22
3.3.2	Grupo das similaridades (transformações de similaridade)	22
3.3.3	Grupo afim (transformações afins)	23
3.3.4	Grupo projetivo (transformações projetivas)	24
3.3.5	Aplicação de transformações	25
3.4	Discussão	25
4	Teoria de Invariantes	27
4.1	Teorema de Cayley-Hamilton	27
4.2	Adaptação para Múltiplas Matrizes	29
4.3	Coefficientes Característicos de Matrizes Simétricas	30
4.4	Composição de Coeficientes Característicos	31
4.5	Discussão	34
5	Monômios Invariantes sobre Junções	35
5.1	Formalização de Junções	35
5.2	Invariante sobre Similaridades para 2-Junções	37
5.3	Invariante Afim sobre 3-Junções	38
5.4	Invariante Projetivo sobre 4-Junções	40
5.5	Invariante Projetivo sobre 4-Junções Impróprias	42
5.6	Discussão	43
6	Descritor de Junções em Imagens	45
6.1	Identificação de Ramos Digitais Bem Definidos	48
6.2	Identificação de Ramos Próprios por Contribuições em Segmentos Digitais	51

6.3	Identificação de Ramos Impróprios por Contribuições em Histogramas de Orientação	52
6.4	Assinatura de Subregiões	54
6.5	Comparação de Características	59
6.6	Discussão	59
7	Experimentos e Resultados	62
7.1	Análise de Robustez dos Invariantes	62
7.2	Experimentos sobre Imagens Naturais com Transformações Induzidas . . .	65
7.3	Experimentos sobre Imagens Naturais de Cenas Capturadas sobre Diferentes Pontos de Vista	69
7.4	Discussão	73
8	Conclusão	76
	Referências	79
	Apêndice A - Função Acumulada Analítica para o Invariante Projetivo (script)	84

Capítulo 1

Introdução

A habilidade de reconhecer padrões visuais é essencial para o desempenho de tarefas do dia-a-dia. Por exemplo, a procura por objetos perdidos em meio à outros, a identificação do humor das pessoas a partir de seus gestos, a escolha de um livro em uma estante e até mesmo a própria leitura do livro exigem do cérebro humano uma resposta veloz e condizente com a atividade exercida em determinado momento. Do início do século XX até os dias de hoje, os computadores foram criados e desenvolvidos para auxiliar as pessoas em suas atividades e, mais recentemente, incluindo-se incontáveis desafios que exigem discernimento visual.

Tarefas visuais consideradas corriqueiras para o sistema visual humano exigem complexas pesquisas para que sejam realizadas por um sistema automatizado. Podem ser destacadas algumas aplicações, tais como atividades industriais para supervisão e controle de processos, execução de procedimentos de montagem automatizada que exigem atenção total e cálculos complexos [18, 57, 58], detecção automatizada de células não saudáveis em mamografias [11, 18], análise automatizada do ambiente por *rovers* ou *drones* em locais de difícil acesso ou inóspitos para seres humanos [58, 60], melhora na experiência do espectador amante de esportes nas transmissões dos jogos [30, 45, 51], reconhecimento facial para segurança e controle de acesso [26, 41], dentre muitas outras. Sejam quais forem os desafios visuais, a atividade mais elementar, ainda que complexa, envolve o ato de identificar e reconhecer objetos ou partes de um, dada uma cena com excessiva informação. Até mesmo o cérebro não é capaz de processar toda a informação visual existente em uma única cena, tendo desenvolvido, ao longo de toda a evolução, um complexo sistema de atenção com o objetivo de focar em informações úteis. Da mesma maneira, os sistemas computacionais, em suas atividades visuais, devem ser capazes de detectar características suficientes para representar as informações úteis da imagem.

Pesquisas recentes realizadas nas áreas de Psicologia, Neurociência e Atenção Visual demonstram que o contorno total dos objetos é mais do que suficiente para que o sistema visual humano reconheça o objeto. De fato, apenas os pontos de inflexão e seus contornos em direções diferentes são suficientes para inferir que tipo de objeto está em análise [12]. Esses pontos de encontro de bordas e suas ramificações são denominados *junções*.

Este trabalho apresenta técnicas para detectar e descrever junções de maneira invariante à diferentes transformações. Este desafio é de suma importância para a solução de problemas complexos por sistemas visuais automatizados, como, por exemplo, o processo de reconhecimento visual de objetos. A Figura 1.1 apresenta o processo seguido pela técnica desenvolvida nesta dissertação para caracterização de regiões de imagens por meio de descritores baseados em junções, compostos de histogramas de valores invariantes à diferentes transformações. Na Figura 1.1, os procedimentos em azul representam os passos efetivamente desenvolvidos e os procedimentos em cinza representam os passos cujas implementações utilizadas advêm de técnicas em seu estado-da-arte.

1.1 Proposta de Dissertação

Detectar e extrair características a partir de uma imagem representam desafios distintos da etapa de representá-la ou descrevê-la computacionalmente. Ambas possuem suas dificuldades. A extração deve lidar com os problemas causados pela rasterização da imagem e ruídos introduzidos no processo de captura por uma câmera. A descrição se responsabiliza por representar a característica de tal maneira que variações causadas pelos diferentes posicionamentos da câmera não interfiram na identificação de um mesmo objeto sob pontos de vista distintos. Além do mais, a propagação das incertezas inerentes à detecção podem interferir substancialmente na descrição final, o que torna este um problema à ser tratado nas duas etapas.

O grupo de transformações provocadas pela variação de pontos de vista de um espectador é denominado como o grupo das *transformações lineares projetivas* e representa um grande desafio para as soluções descritoras de características visuais. As soluções existentes na literatura baseiam-se pouco em informações geométricas presentes em suas estratégias para descrever características visuais, saturando suas representações em função de variações de intensidade e cores, o que provavelmente torna a invariância verdadeira à transformações afim e projetivas algo muito complexo de se alcançar. A proposição principal deste trabalho estabelece que:

É possível representar sistematicamente porções de imagens a partir de valores calculados somente por informações geométricas extraídas de suas junções, de tal maneira que esses valores sejam invariantes sob o grupo de transformações projetivas.

A utilização de junções tem motivação biológica ao simular parte do comportamento do sistema visual humano e também motivação qualitativa, pois as junções de um objeto são bem localizadas e podem ser utilizadas em pareamentos e correspondências, ainda que tais objetos se encontrem em cenas completamente distintas.

O termo *valores invariantes* possui denotação de impacto, pois indica que são inalterados independente da transformação aplicada. Perceba que é diferente afirmar que algo possui *comportamento invariante*, o que usualmente se associa à texturas e variações de

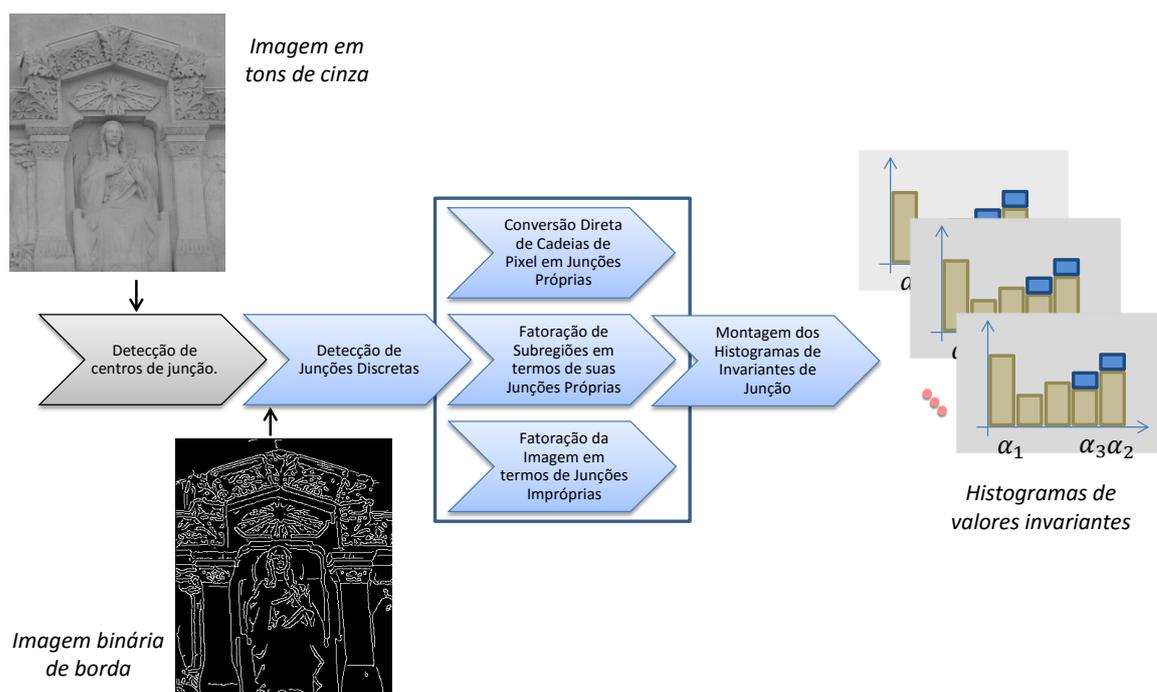


Figura 1.1: O procedimento para encontrar centros de junções utiliza técnicas para detecção de cantos cujos resultados representam localizações de pixels, interpretados neste trabalho como o vértice central de junções significativas. Em seguida, é realizada a detecção das junções discretas, *i.e.*, junções como conjuntos de segmentos digitais. O processo segue para conversão das junções discretas em junções contínuas, *i.e.*, cujas ramificações são segmentos de reta exatos e o vértice central é a intersecção de tais segmentos. As junções contínuas são utilizadas como entrada para cálculo de valores invariantes para cada uma, que contribuem para a construção de um histograma de valores invariantes, aqui denominado *assinatura de junção*. As assinaturas de junção são as características finais que devem ser comparadas quando se deseja verificar se duas imagens estão correlacionadas por uma transformação linear projetiva.

intensidade [65]. Com base na Teoria de Invariantes (Capítulo 4), neste trabalho foram formulados monômios invariantes para junções sob diferentes grupos de transformações, incluindo as de similaridade, afins e projetivas (Capítulo 5). Os monômios invariantes propostos foram incorporados como parte de um sistema descritor que acumula os valores calculados num histograma que se comporta como uma *assinatura de junção* da imagem em questão (Capítulo 6). Os fundamentos e trabalhos relacionados estão apresentados nos Capítulos de 2 à 4. O Capítulo 7 apresenta uma análise da robustez dos descritores propostos e dos resultados obtidos.

Junções extraídas diretamente de imagens digitais podem apresentar problemas quanto à discretização, o que pode levar a modificação de sua geometria, justificando a decisão de utilizar junções contínuas onde apenas o vértice central e os vértices finais de ramificação são preservados. Ao final do processo proposto, monômios invariantes são calculados sobre a geometria das junções e, por fim, contribuem para a assinatura da junção.

1.2 Resultados e Contribuições

Dentre as contribuições deste trabalho, destacam-se:

- Derivações de monômios invariantes à transformações de similaridade, afim e projetivas, sobre junções contínuas;
- Procedimentos para decomposição de regiões de interesse em termos de junções contínuas;
- Estratégias para composição de descritores invariantes para regiões de interesse em imagens.

Os experimentos realizados foram desenvolvidos para mostrar a viabilidade da estratégia proposta, a partir do processamento de um conjunto de imagens transformadas artificialmente e de um conjunto de imagens capturadas de pontos de vista diferentes em cenas reais (Capítulo 7). Os resultados obtidos são promissores, indicando que a aplicação prática dos descritores é plausível.

Capítulo 2

Trabalhos Relacionados

Este capítulo apresenta o estado-da-arte das técnicas atuantes no problema de detecção e descrição de características visuais invariantes, com foco especial em descritores de características locais.

O processo de reconhecimento de padrões visuais se divide em *detectar*, *descrever* e *comparar* características visuais. As abordagens relacionadas à detecção, apresentadas na Seção 2.1, devem partir da definição prévia de quais características visuais são interessantes para serem descritas posteriormente. Algumas técnicas se baseiam em informações de intensidade, enquanto que outras são baseadas em contornos. Cada abordagem com suas vantagens e desvantagens.

Uma vez detectadas e extraídas, é preciso transcrever computacionalmente cada uma das características visuais, para que essas possam ser comparadas automaticamente. Técnicas que se propõem à descrever características locais, apresentadas na Seção 2.2, não devem produzir diferentes resultados diante dos mesmos objetos ainda que sob diferentes transformações causadas pelas variações de pontos de vista. A definição de uma estratégia de detecção e descrição implica em estabelecer a robustez do descritor à determinadas transformações. O pareamento entre imagens sob diferentes transformações é uma necessidade para o funcionamento correto de diversas soluções em Visão Computacional. Por fim, conforme descrito na Seção 2.3, as características descritas para uma imagem devem ser comparadas com características advindas de outras imagens a fim de determinar o quão semelhantes elas são.

2.1 Detectores de Características

A detecção de características está diretamente relacionada com a estratégia de descrição. Porém, são procedimentos que podem ser apresentados de forma independente. As principais estratégias de descrição são baseadas na geometria de um contorno ou intensidades de uma região e, assim, os detectores também serão classificados. As características a serem detectadas ainda podem advir de elementos pontuais de uma cena específica (*i.e.*, características locais) ou da cena como um todo (*i.e.*, características globais).

Características globais são muito sensíveis à ruídos de qualquer tipo e à diversos outros fatores como alterações de escala, oclusão de objetos, mudanças de pontos de vista, variações de iluminação, dentre outras. No entanto, exigem menos esforço computacional. Este tipo de característica costuma ser usada em ambientes de automação industrial onde a iluminação é controlada e, comumente, o único objeto em cena é o que se deseja descrever de fato. Algumas soluções se baseiam em histogramação dos sinais de intensidade da imagem [5]. Algumas abordagens calculam momentos invariantes da imagem [44] e outras descrevem contornos inteiros de um objeto [7, 8, 47]. As transformadas sob bases ortonormais distintas das euclidianas (*e.g.*, transformada de Fourier [16], de Haar [27], Harmônica Esférica [43]) também geram vetores de coeficientes que podem ser consideradas características globais de uma imagem.

Os detectores de características locais focam em elementos pontuais presentes na imagem e também podem ser interpretados como uma representação esparsa da imagem. Em geral, buscam por pontos de inflexão nos contornos da imagem. Além da motivação biológica, esses pontos são mais discriminativos do que regiões lisas ou até mesmo bordas contínuas, conforme mostra a Figura 2.1. Dentre os detectores mais conhecidos destacam-se o detector de cantos e junções de borda de Moravec [46] e Harris [28], que correlacionam as informações da vizinhança de um ponto para concluir se o mesmo pertence à uma borda, canto ou uma região lisa. Outras técnicas se baseiam em convoluções da região de interesse, como a Laplaciana de uma Gaussiana [35], a Diferença entre Gaussianas utilizada por Lowe [40] e, também, a Diferença entre Hessianas aplicada por Bay [36, 37]. Essas técnicas são especializadas em identificar pontos interessantes que não se limitam à estruturas simples, tais como, bordas, cantos e junções.

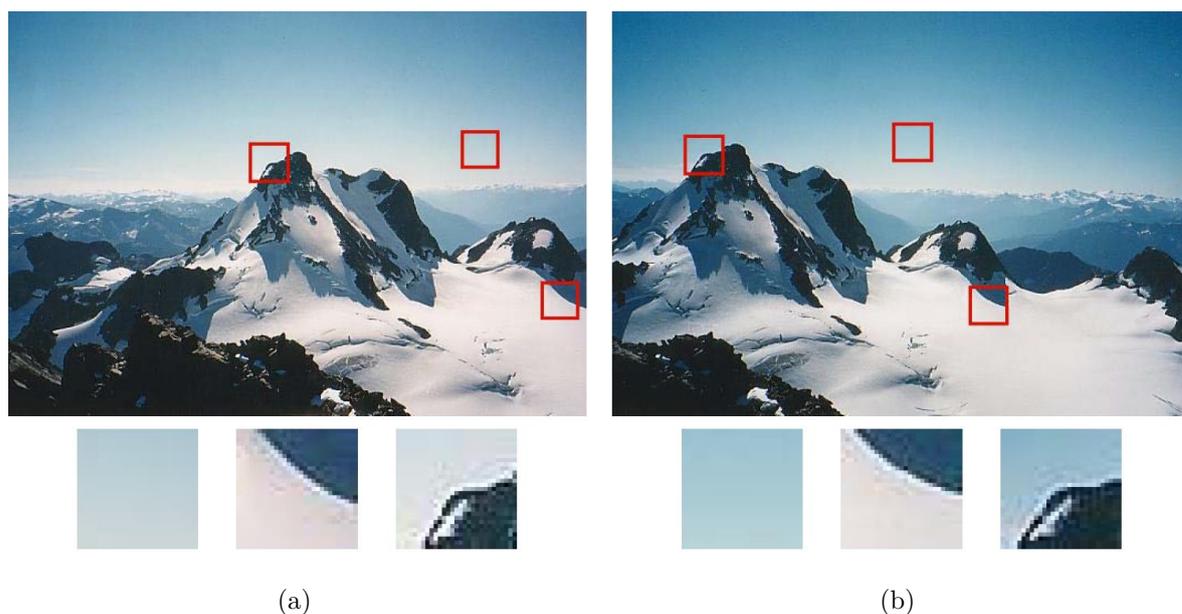


Figura 2.1: Duas fotografias de uma geleira sob um deslocamento de pontos de vista para demonstrar a distinguibilidade de cantos. Os quadrados vermelhos em ambas as imagens destacam as mesmas regiões de interesse, as quais os descritores deveriam correlacioná-las: uma parte do céu, uma borda de geleira e um pico de montanha. Perceba que a região de céu aberto se confunde com qualquer área do céu nas fotografias, o que a torna muito ruim para se considerar uma característica. A borda de geleira já possui uma distinguibilidade maior que a do céu aberto mas é fácil perceber que este pedaço de borda ainda se confundiria com outros pedaços de borda de geleira. As descontinuidades do pico de montanha é o que a torna uma região tão distinta.

2.2 Descritores de Características

A descrição invariante para um único ponto é inviável em aplicações práticas. Por isso, os descritores que utilizam detectores de pontos, em geral, descrevem correlações entre as intensidades de uma região em torno de um ponto, especialmente denominada como uma *região de interesse* [65]. A detecção de contornos de objetos inteiros, usualmente em locais onde há controle de iluminação e de ruídos externos, se destina à descritores de parâmetros intrínsecos da geometria do contorno, denominada como *contorno de contexto* ou *contorno de interesse* [7]. Descritores de contorno são apresentados na Subseção 2.2.4.

Descritores que se baseiam em regiões de interesse podem ser classificados de acordo com sua estratégia principal para descrição. Lowe e Bay *et al.* foram responsáveis pela criação, respectivamente, dos trabalhos, *Scale-Invariant Feature Descriptor (a.k.a., SIFT)* [39] e *Speed-Up Robust Features (a.k.a., SURF)* [6]. A partir daí, diversas melhorias foram propostas, tanto em eficiência quanto em representatividade, sobre essas técnicas. Recentemente, uma estratégia inovadora foi proposta por Calonder, denominada *Binary Robust*

Independent Elementary Features (*a.k.a.*, BRIEF) [13], que objetiva descrever regiões de interesse em sequências binárias. A vantagem do uso de sequências binárias está na etapa posterior, de comparação, que é executada de maneira mais eficiente que as utilizadas no SIFT e SURF.

Os descritores de contorno fundamentam-se na representação geométrica dos objetos e no cálculo de propriedades específicas, como, por exemplo, distância dos pontos de borda até seu centro de massa, excentricidade de seções cônicas que melhor se ajustam e proporções entre áreas, dentre outras.

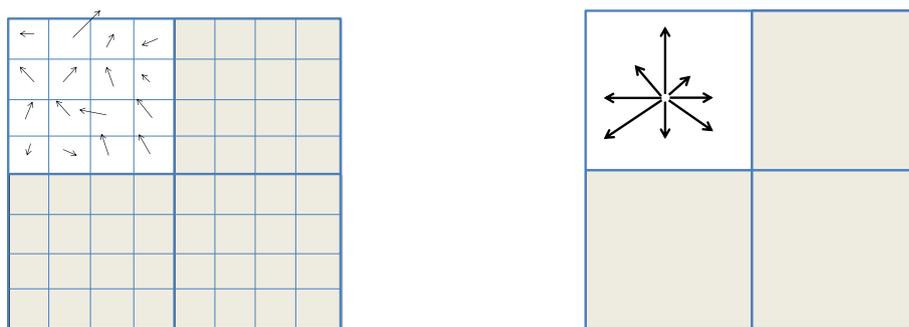
2.2.1 Descritores derivados do SIFT

O método para detectar e descrever características SIFT [39] foi projetado para proporcionar invariância às transformações euclidianas (*i.e.*, rotações, translação e escala uniforme), definidas na Seção 3.3. Posteriormente, diversos autores propuseram melhorias para a técnica, tanto para aumentar a eficiência computacional quanto para estender sua robustez para outras transformações [10, 33, 67].

SIFT opera sobre imagens em tons de cinza. Para obter invariância às modificações de escala, técnicas baseadas em SIFT utilizam como detector de características a Diferença entre Gaussianas numa pirâmide de aplicações sucessivas de filtro gaussiano sob diferentes desvios padrão, uma aproximação mais eficiente do filtro laplaciano, para detectar extremos no espaço de escalas.

Para descrever uma região de interesse na imagem, Lowe propôs uma janela de tamanho 16×16 pixels onde, para cada pixel, é calculada a magnitude e orientação do gradiente. Esses dados são aglomerados em quatro histogramas de orientação cujas entradas representam angulações múltiplas de 45° , como ilustrado na Figura 2.2. Perceba que, ao dividir a região em dezesseis subregiões de tamanho 4×4 pixels e num total de 8 angulações possíveis para os mapas de orientação, teremos uma característica composta por 128 coeficientes.

Trabalhos posteriores [24, 33] objetivam alterar o domínio das características, reduzindo sua dimensionalidade e otimizando, assim, a etapa seguinte de comparação. Ke *et al.* [33], por exemplo, utiliza a técnica de análise dos componentes principais para reduzir a dimensionalidade das características originais de 128 para apenas 20 coeficientes. Algumas variantes estendem a técnica original para imagens coloridas [1], enquanto que outras estendem o SIFT para prover invariância às diferentes transformações. Um exemplo



(a) Orientações do gradiente em cada pixel. (b) Soma das magnitudes em direções fixadas.

Figura 2.2: Processo de descrição de uma região de interesse 8×8 pelo SIFT. A imagem (a) mostra as direções dos gradientes em cada ponto numa subregião 4×4 , enquanto que (b) apresenta a soma das magnitudes da mesma subregião em 8 direções fixadas. Esta subregião poderá ser descrita por 32 valores numéricos. A região total proposta por Lowe, que utiliza regiões 16×16 , necessita de 128 valores para ser descrita.

interessante dessa variação é o ASIFT [67], que aplica SIFT diversas vezes em diferentes transformações afins simulada da imagem de entrada, levando à construção de descritores invariantes à transformações afins.

2.2.2 Descritores derivados do SURF

SURF [6] surgiu como uma alternativa mais computacionalmente eficiente que o SIFT. Ele provê invariância às transformações euclidianas e seu uso prático demonstra robustez à pequenas mudanças de pose da cena.

Bay *et al.* propõem uma detecção de regiões interessantes a partir da identificação de extremos locais, num espaço de escala, semelhante ao detector de cantos de Harris modificado para detectar pontos em regiões ainda que sob transformações afim. No entanto, o SURF utiliza o determinante hessiano da vizinhança para selecionar os pontos mais relevantes. Pesquisas apontam o detector hessiano afim de regiões interessantes como o mais estável [36, 37], quando comparados à Diferença de Gaussianas utilizado pelo SIFT.

Após a identificação das regiões, Bay *et al.* propõem a definição de um entorno circular, com raio parametrizado. Dos pontos pertencentes à esta região os autores utilizam uma transformada de Haar [4] sobre a segunda derivada da imagem de intensidades para identificar as direções preferenciais. Em cada direção possível, as orientações são somadas para formar vetores de orientação, onde o de maior comprimento é tomado como a orientação preferencial de toda a região.

A transformada de Haar, assim como a transformada de Fourier, permite a represen-

tação de uma função sobre bases de funções ortonormais, sendo caracterizada por obter sinais discretos como resultado.

2.2.3 Descritores derivados do BRIEF

Os descritores binários surgiram como uma alternativa compacta para a descrição de características em imagens. Além da simplicidade, BRIEF [13] apresenta características com alto poder discriminativo, descritas à partir de sequências binárias, podendo ser comparadas facilmente através de distâncias binárias, como a distância de Hamming. Segundo Calonder *et al.* [13], BRIEF apresenta uma alta sensibilidade à distorções e ruídos.

O método foi desenvolvido com base em trabalhos anteriores [50], que provam a possibilidade de se descrever as características de uma imagem à partir de um conjunto de comparações entre intensidades, dando origem à sequências binárias obtidas dos resultados dessas comparações.

Seja

$$\text{bin}_{n_d}(p) = \sum_{i=1}^{n_d} 2^{i-1} \text{patch}(p; \mathbf{a}_i, \mathbf{b}_i), \quad (2.1)$$

onde bin_{n_d} é uma função responsável por sequenciar valores, n_d é a cardinalidade desejada para a sequência gerada, p denota um padrão de comparação de intensidades pré-definido, \mathbf{a}_i e \mathbf{b}_i caracterizam o i -ésimo par do padrão p cujas intensidades serão comparadas e patch é a função responsável por comparar \mathbf{a}_i e \mathbf{b}_i , retornando um valor binário como resposta:

$$\text{patch}(p; \mathbf{a}, \mathbf{b}) = \begin{cases} 1 & \text{se } p(\mathbf{a}) < p(\mathbf{b}); \\ 0 & \text{c.c.} \end{cases} \quad (2.2)$$

O padrão de comparação p em uma dada região, à princípio, não precisa respeitar nenhum modelo específico.

Diferentes pares de pontos são sorteados pelos autores, respeitando distribuições de probabilidade, os pares sorteados podem ser vistos na Figura 2.3. Calonder *et al.* realizaram testes para verificar os mapas de comparação que apresentariam a maior taxa de acerto em um conjunto conhecido de imagens. Entretanto, nenhum obteve uma vantagem conclusiva para ser definido como o mais adequado.

Ao se estabelecer uma nova estratégia para descrição de características, que mostrou ser bastante simples e eficiente, diversos trabalhos foram desenvolvidos à partir do BRIEF na intenção de melhorar as características da técnica ou embutir uma nova invariância

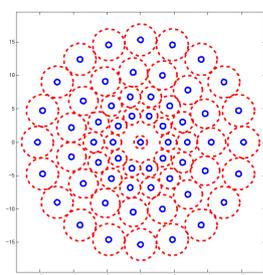


Figura 2.3: Mapas de comparação entre pixels. Cada mapa possui pares de pixels sorteados sob uma distribuição especificada e cada segmento de reta interliga dois pixels que formarão um par cujas intensidades serão comparadas. Da esquerda para a direita, as distribuições de probabilidade aplicadas foram: uniforme, gaussiana com $\sigma = 0.25$ para p_1 e p_2 , gaussiana em a_1 com $\sigma = 0.25$ e em a_2 com $\sigma = 0.01$, randômico e, o último, tomando um dos pontos como o centro da região. Imagem adaptada do artigo original [13].

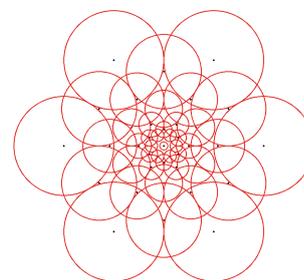
à transformações. Daí surgiram técnicas implementando, assim como o SIFT, a detecção de pontos interessantes dentro de um espaço de escala [49], adicionando invariância à rotações [34] e, o estabelecimento de um padrão circular de comparação entre pontos de uma região [55]. Os padrões circulares, chamados BRISK e FREAK, ilustrados pela Figura 2.4, auxiliam na determinação da direção preferencial do gradiente naquela região, tornando possível estabelecer uma posição canônica para a região, alcançando-se a invariância também à rotações sobre o plano.

2.2.4 Descritores baseados em contorno

Os descritores baseados em contorno existentes na literatura buscam informações de posicionamento geométrico dos pixels do contorno para gerar as chamadas assinaturas ou vetores de características do formato à ser analisado. Tais soluções são classificadas em: (i) descritores de silhuetas, (ii) descritores de aparência e (iii) descritores de formato.



(a) Padrão circular BRISK.



(b) Padrão circular FREAK.

Figura 2.4: Padrões circulares para comparação de intensidades: A imagem (a) apresenta o padrão BRISK, proposto como uma extensão do BRIEF, provendo invariância à rotações [34]. A imagem (b) apresenta o padrão circular FREAK, estratégia baseada em pesquisas sobre a biologia do olho humano [49]. Em ambos, os círculos em vermelho indicam a região de aplicação de um filtro passa-baixa.

Descritores de silhuetas tem por característica aproximar toda a região definida pelo contorno do objeto em questão, por geometrias conhecidas, tais como retângulos, círculos e elipses, a fim de obter parâmetros intrínsecos, como o eixo medial [23, 61], divisão da área total em triângulos formados por sucessivos pontos de contorno e seu centróide onde cada quantitativo de área seria o descritor da dupla de pontos [2, 3], dentre outras. Algumas soluções desta categoria propõem a descrição da silhueta em coeficientes da transformada de Fourier, garantidamente invariantes à isometrias [17, 52, 68].

As técnicas classificadas como descritoras de aparência foram propostas posteriormente, utilizando os mesmos conceitos das descritoras de silhuetas, porém, medindo diretamente valores de intensidade e cores da região interior do contorno do objeto. A identificação de semelhanças entre dois objetos pode ser realizada através de uma análise de componentes principais e medindo a distância entre seus centros de massa [48, 64];

A última classificação, descritores de formato, é referente à técnicas que se baseiam no conjunto de pixels que compõem o contorno. Diferentemente das anteriores, soluções dessa categoria não são sensíveis à diferenças de iluminação por fazerem uso apenas de informações geométricas dos pontos de borda. Esta abordagem se mostra bastante eficaz no reconhecimento de caracteres e de formatos isolados em condições onde a segmentação total dos objetos é possível [3, 23, 68].

2.3 Comparador de Características

Descritores derivados do SIFT, SURF e BRIEF constroem características visuais a partir de valores numéricos vetorizados, permitindo a aplicação de técnicas para cálculo de distâncias entre vetores numéricos (*e.g.*, distância euclidiana, geométrica, de Hamming, dentre outras).

Os descritores de características baseados em contornos, comumente, utilizam histogramas acumuladores de assinaturas dos pixels de contorno. Na etapa de comparação de características deve então se utilizar técnicas que servem à este fim, tais como, *Qui-Squared's Distance* [21], *Earth-Mover's Distance* [56], dentre outras.

Como será detalhado no Capítulo 6, este trabalho utiliza histogramas como objeto matemático para descrever junções em imagens raster e aplica o método *Earth-Mover's distance* para cálculo de distância entre histogramas.

2.4 Discussão

Os métodos para identificação e extração de características podem ser globais (*i.e.*, que consideram informações do contexto total da imagem) ou locais (*i.e.*, que consideram propriedades geométricas ou de intensidade de uma região da imagem). No preâmbulo deste capítulo foi explicitada a falta de representatividade e robustez das características globais, enquanto que no decorrer da Seção 2.1, o estado-da-arte de técnicas locais foi apresentado.

Dos detectores de características locais, os métodos ainda podem ser subclassificados entre os que utilizam informações geométricas e os que utilizam informações de intensidade. Ambas as abordagens possuem suas vantagens e desvantagens e a escolha entre uma delas dependerá da estratégia que se deseja utilizar para descrever as características encontradas. As técnicas baseadas em intensidade são comuns dentre os descritores de características mais conhecidos da literatura, como o SIFT e o SURF (Seção 2.2), pois os mesmos descrevem orientações em uma região de interesse.

As técnicas SIFT e SURF disputam entre si em termos de robustez das características que descrevem, enquanto que os descritores binários se apresentam como uma alternativa simples e eficiente, porém, bastante sensíveis à distorções e ruídos. É importante notar que os descritores em seu estado-da-arte, mesmo os binários, dependem de estratégias distintas para se obter invariância à diferentes transformações de regiões, como, por exemplo, a pirâmide de escalas para a escala uniforme e detecção de orientação preferencial para invariância à rotações, tornando o processo complexo como um todo.

Não há dúvidas sobre a importância das informações de intensidade e gradiente obtidas à partir de uma vizinhança de interesse para descrever adequadamente a mesma. Mas apesar das regiões de interesse serem consideradas características locais, os descritores não levam em consideração a estrutura dos objetos em seu interior, tornando a invariância à transformações afins e projetivas computacionalmente complicada de se obter. Até mesmo as técnicas estabelecidas para descrever contornos, também locais, consideram contornos inteiros como característica, o que limita a solução para objetos de formato específico que possa ser totalmente segmentado do restante da cena. Além do que, esta última ainda possui limitações quanto à garantir invariância sob transformações afins e projetivas.

Diferentemente das estratégias propostas para descrever regiões vistos nesse capítulo, informações geométricas sobre estruturas interiores podem ser utilizadas para alcançar invariância à transformações de similaridade, afim e projetivas, como será demonstrado

ao longo deste trabalho. A técnica defendida neste trabalho propõe a simplificação do processo, tomando junções como características principais para a obtenção de valores escalares invariantes por junção ajustada à região de interesse e, invariante à transformações de determinado grupo. Perceba que, de maneira distinta, a solução proposta fundamenta-se na representação de topologias localizadas e não de contornos inteiros.

Capítulo 3

Conceitos sobre Características Visuais, Geometria Projetiva e Transformações Lineares

Este capítulo tem por objetivo introduzir o que são características visuais e o que as faz serem consideradas interessantes (Seção 3.1). Na Seção 3.2 são apresentadas bases conceituais que permitem a representação geométrica de informações extraídas de uma imagem. A Seção 3.3 demonstra, brevemente, como as geometrias são afetadas diante das diferentes transformações existentes, não abstando-se de apresentar, ainda, as propriedades que tornem as geometrias invariantes.

3.1 Características Visuais de Interesse

Por definição, uma *característica visual* é uma propriedade, um detalhe ou uma qualidade capaz de distinguir um objeto sob análise, enquanto que uma imagem é uma representação visual de um aspecto particular de uma cena que, por sua vez, pode vir a possuir incontáveis características visuais.

Não seria viável, nem para os computadores nem para o cérebro humano, interpretar todo e qualquer tipo de característica visual presente numa cena e, por isso, são necessários alguns direcionamentos prévios. A evolução moldou um sistema de atenção visual complexo para que o cérebro fosse capaz de focar nas características interessantes, com o mínimo de desgaste neurológico. Especialistas acreditam que esse complexo sistema de filtragem se manifesta de duas formas [12, 19, 31]. A primeira, do inglês *bottom-up*, estabelece a atenção a partir dos sensores físicos que captam a cena, à exemplo da biologia do olho, que conduz o foco para a região central. A segunda, do inglês *top-down*, manifesta a

atenção como uma imposição dos objetivos da atividade à ser finalizada, por exemplo, ao se buscar um objeto vermelho numa caixa com outros objetos de cores variadas. Sistemas computacionais de visão baseados na biologia natural, em sua maioria, fundamentam suas soluções em filtragens *bottom-up*.

A retina, *i.e.*, região do globo ocular onde a cena visualizada projeta-se, é composta por células especializadas e sensíveis às propriedades luminosas específicas. Dentre tais células estão os bastonetes e os cones que são sensíveis, respectivamente, à intensidade luminosa e às cores e nitidez. As informações captadas que codificam formato, cores e noções de movimento são guiadas para o córtex visual central, presente na parte de trás do cérebro [29, 31, 66].

Existem muitas pesquisas e teorias nas áreas de Psicologia, Neurociência e Atenção Visual que têm por objetivo determinar a importância de cada informação captada para tarefas específicas. Para as atividades de reconhecimento e classificação de objetos, em especial, constatou-se que o formato, ou seja, o contorno dos objetos são mais do que suficientes para que o cérebro infira que tipo de objeto está em cena. Mais do que isso, as junções de borda, por si só, já seriam suficientes [63], justificando o uso de junções como características de alto interesse. A Figura 3.1 ilustra, objetivamente, o contorno de objetos conhecidos (à esquerda) e apenas com as junções de contorno (ao centro) e suas versões apenas com as bordas contínuas (à direita). Perceba como é mais natural identificar o objeto a partir de suas junções.



Figura 3.1: Na coluna mais à esquerda, seguem formatos de objetos conhecidos: uma taça e uma tesoura. Na coluna central, suas junções, ignorando-se as bordas contínuas, e na coluna mais à direita, seus contornos, ignorando-se as junções. Perceba como é mais natural inferir o objeto sob análise quando consideradas apenas as junções de borda. Imagens originais de Biedermen [9].

Uma vez definido o que é interessante para ser encontrado será necessário detectar e transcrever as características de tal maneira que sistemas automatizados sejam capazes de compará-las. Tradicionalmente, dada uma característica de interesse, a transcrição pode ser fundamentada sobre informações e propriedades presentes em sua região vizinha (*i.e.*, orientações e magnitude de gradientes, geometrias, intensidades e cores, etc.). Este processo de transcrição é usualmente conhecido como *descrição de características*. O *descriptor* é o representante da característica.

3.2 Estruturas Geométricas em Imagens

Seja

$$I(x, y) : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (3.1)$$

uma função que denota uma imagem contínua de intensidades avaliada em um ponto $\mathbf{a} = (x, y)$. Os contornos materializados em I são, de fato, variações bruscas de intensidade, dada uma direção em x ou em y . Seja, então,

$$\nabla I(x, y) = \left(\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y} \right), \quad (3.2)$$

onde ∇I denota o gradiente calculado no ponto \mathbf{a} .

A definição formal das funções de imagem estabelecem que um ponto $\mathbf{a} = (x, y)$ qualquer deve pertencer ao espaço planar \mathbb{R}^2 e são representações bidimensionais de uma cena com três dimensões. À este espaço bidimensional com uma dimensão homogênea extra w , denominamos espaço projetivo \mathbb{P}^2 . Por conveniência, o plano π paralelo aos eixos \mathbf{x} e \mathbf{y} , posicionado em $w = 1$, denomina-se plano homogêneo (Figura 3.2). As Seções 3.2.1 à 3.2.3 apresentam as representações algébricas das primitivas geométricas de interesse, definidos sob o espaço \mathbb{P}^2 , para esta dissertação.

A interpretação de informações visuais em termos de elementos geométricos conhecidos (*i.e.*, direções, pontos, retas, seções cônicas, etc.) e definidos sobre o espaço \mathbb{P}^2 é bastante útil, pois representa de maneira natural as diferentes transformações projetivas (ou homografias).

3.2.1 Pontos finitos e direções em \mathbb{P}^2

Seja o vetor $\mathbf{u} \in \mathbb{P}^2$, representado pelo trio ordenado $\mathbf{u} = (x, y, w)^\top$. Geometricamente, \mathbf{u} pode ser interpretado como um *ponto próprio* (ou finito) quando $w \neq 0$ ou *ponto impróprio*

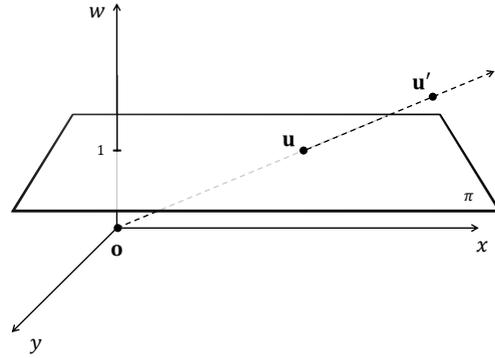


Figura 3.2: Interpretação geométrica de vetores em \mathbb{P}^2 como pontos no plano projetivo π para $w = 1$. Pontos e direções são escaláveis em \mathbb{P}^2 , o que significa que $\mathbf{u} = (x, y, w)^\top \equiv (kx, ky, kw)^\top$, para $k \neq 0$. Quando $k = \frac{1}{w}$, para $w \neq 0$, essa operação é denominada normalização pela coordenada homogênea.

(ou direção) quando $w = 0$. Pontos e direções são escaláveis em \mathbb{P}^2 , o que significa que $\mathbf{u} = (x, y, w)^\top \equiv (kx, ky, kw)^\top$, para $k \neq 0$.

O conceito de plano projetivo homogêneo é, comumente, aplicado para representar o plano de imagem onde os elementos se encontram originalmente. Os vetores que definem o plano π possuem coordenada homogênea $w = 1$. Ao se aplicar uma transformação projetiva, eventualmente, tais vetores são mapeados para elementos com coordenada homogênea w' distinta. Se $w' \neq 0$, então basta dividir todas as coordenadas por w' e o mesmo é reprojeto para o plano π . Se $w' = 0$ então o vetor resultante é interpretado como uma direção, ou como se o mesmo fosse um ponto infinitamente distante da origem.

3.2.2 Retas em \mathbb{P}^2

Retas respeitam à equação geral $ax + by + c = 0$ e também são representadas como vetores $\mathbf{r} = (a, b, c)^\top \in \mathbb{P}^2$. Retas são escaláveis em \mathbb{P}^2 , o que significa que $\mathbf{r} = (a, b, c)^\top \equiv (ka, kb, kc)^\top$. Ou seja, independentemente da escala $k \neq 0$, o vetor codificaria a mesma reta. Geometricamente, uma reta em \mathbb{P}^2 representa a intersecção entre o plano homogêneo e o plano de normal $(a, b, c)^\top$ e que passa pela origem do espaço vetorial tridimensional \mathbb{P}^2 , conforme ilustra a Figura 3.3. Na Figura 3.3a, dois pontos \mathbf{p} e \mathbf{q} são apresentados sobre o plano homogêneo. A reta \mathbf{r} que contém \mathbf{p} e \mathbf{q} é dada por $\mathbf{r} = \mathbf{p} \times \mathbf{q}$, onde \times indica o produto vetorial entre os pontos \mathbf{p} e \mathbf{q} . Como resultado, \mathbf{r} é o vetor perpendicular tanto à \mathbf{p} quanto à \mathbf{q} , ou seja, o vetor normal ao plano expandido pelos vetores \mathbf{p} e \mathbf{q} .

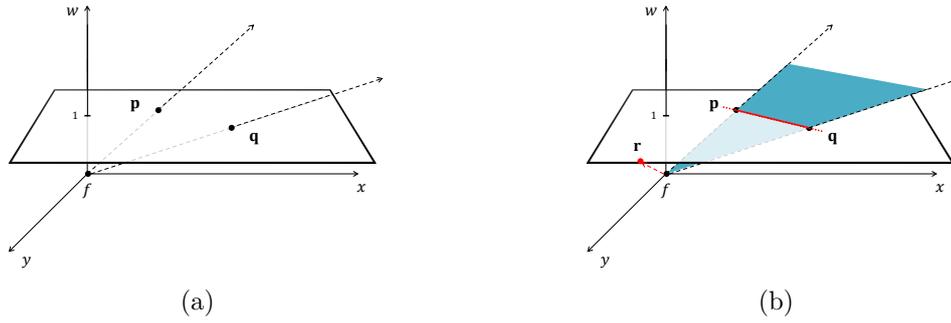


Figura 3.3: Interpretação geométrica de uma reta em \mathbb{P}^2 . Dados dois vetores \mathbf{p} e \mathbf{q} em \mathbb{P}^2 , a reta \mathbf{r} que contém ambos os pontos é dada pela intersecção do plano homogêneo e ao plano expandido pelos vetores \mathbf{p} e \mathbf{q} . Perceba que $\mathbf{r} = \mathbf{p} \times \mathbf{q}$, onde \times indica o produto vetorial, ou seja, \mathbf{r} é o vetor normal ao plano expandido.

3.2.3 Seções cônicas em \mathbb{P}^2

As *seções cônicas* representam um conjunto de curvas algébricas de segundo grau e que são classificadas em *próprias* ou *degeneradas*, caso essas sejam irredutíveis. As cônicas próprias se materializam em grupos de elipses, hipérboles, parábolas e círculos. As degeneradas também podem ser agrupadas em retas coincidentes, retas paralelas, retas concorrentes ou pontos [63]. A Figura 3.4 apresenta os diferentes tipos de seções cônicas, as cônicas da linha superior, da esquerda para a direita, um círculo, uma hipérbole, uma elipse e uma parábola, representam todos os tipos de cônicas próprias e a linha inferior, da esquerda para a direita, duas retas coincidentes, duas retas paralelas, duas retas concorrentes e um ponto, representam todos os grupos de cônicas degeneradas.

As seções cônicas podem ser definidas pela união de todos os pontos que as formam (*i.e.*, *representação primal*) e, de maneira análoga, também podem ser definidas pela união de todas as retas tangentes aos pontos que formam a cônica (*i.e.*, *representação dual*), o que caracteriza um envelope de cônicas. Cônicas primais e duais são duas maneiras distintas de se interpretar uma seção cônica. A Figura 3.5 ilustra uma mesma elipse em sua representação primal e, também, em sua representação dual.

As seções cônicas respeitam a equação

$$a_{11}x^2 + a_{12}xy + a_{22}y^2 + a_{13}xw + a_{23}yw + a_{33}w^2 = 0, \quad (3.3)$$

onde $\mathbf{p} = (x, y, w)^\top$ é um ponto que pertence à seções cônicas em \mathbb{P}^2 . Em notação

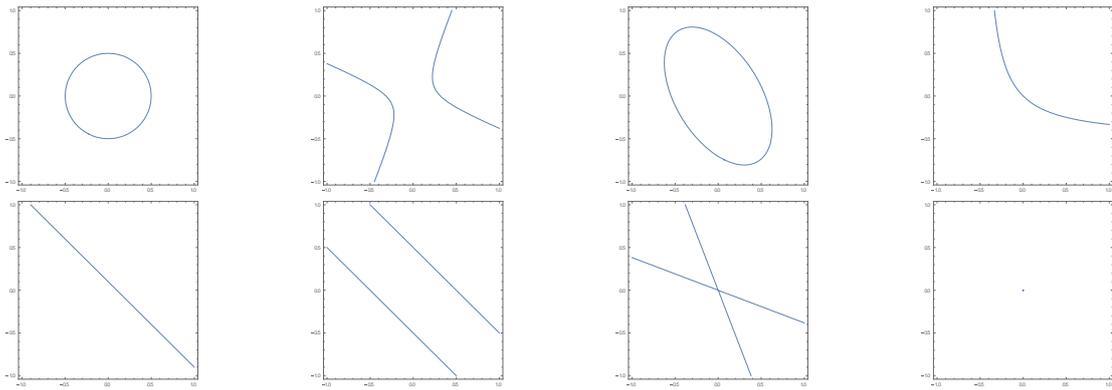


Figura 3.4: Na linha superior, da esquerda para a direita, seções cônicas próprias: um círculo, uma hipérbole, uma elipse e uma parábola. Na linha inferior, da esquerda para a direita, seções cônicas degeneradas: duas retas coincidentes, duas retas paralelas, duas retas concorrentes e um ponto.

matricial, a Equação 3.3 é escrita como

$$\mathbf{p}^\top \cdot \mathcal{A} \cdot \mathbf{p} = \begin{pmatrix} x & y & w \end{pmatrix} \cdot \begin{pmatrix} a_{11} & \frac{a_{12}}{2} & \frac{a_{31}}{2} \\ \frac{a_{12}}{2} & a_{22} & \frac{a_{32}}{2} \\ \frac{a_{31}}{2} & \frac{a_{32}}{2} & a_{33} \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ w \end{pmatrix} = 0, \quad (3.4)$$

onde a_{11} , a_{12} e a_{22} não são simultaneamente iguais à zero. As matrizes \mathcal{A} associadas à seções cônicas podem ser multiplicadas por um valor escalar, não alterando a interpretação geométrica das mesmas, o que significa que $\mathcal{A} \equiv k\mathcal{A}$, para $k \neq 0$. Logo, sem perda de representação, é possível determinar um $k = \frac{1}{a_{ij}}$, para $a_{ij} \neq 0$, diminuindo o número mínimo de coeficientes necessários para definir unicamente uma seção cônica para 5 coeficientes. A notação matricial facilita a interpretação de algumas propriedades em termos de operadores matriciais (*e.g.*, determinantes, traços, etc.). Matrizes associadas que possuem determinante nulo não são inversíveis e representam cônicas degeneradas.

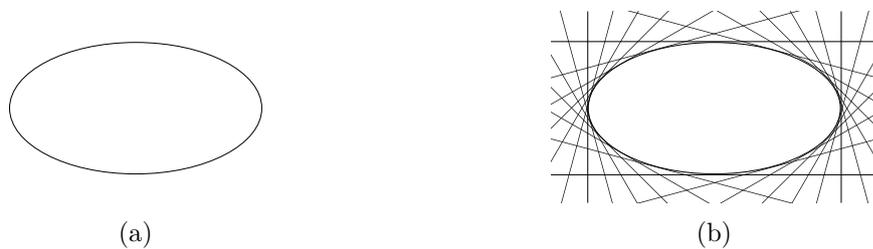


Figura 3.5: Uma única elipse representada em sua forma primal (a) e dual (b). A primal cujo contorno é dado pelos pontos de borda da cônica e a dual cujo contorno é dado por todas as tangentes dos pontos da primal.

Entende-se por discriminante de uma cônica o valor calculado por

$$\text{dsc}(\mathcal{A}) = \det \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}, \quad (3.5)$$

onde \mathcal{A} denota a matriz simétrica associada à cônica. O valor resultante do discriminante indica o tipo da cônica associada, como segue:

- Se $\text{dsc}(\mathcal{A}) < 0$, \mathcal{A} é uma elipse;
- Se $\text{dsc}(\mathcal{A}) > 0$, \mathcal{A} é uma hipérbole;
- Se $\text{dsc}(\mathcal{A}) = 0$, \mathcal{A} é uma parábola.

3.2.4 Feixe e rede de cônicas

Cônicas próprias ou degeneradas, podem ser combinadas linearmente para dar origem à uma nova matriz \mathcal{M} associada à uma seção cônica, formando assim um *feixe de cônicas* [59, 63]:

$$\mathcal{A} = k_1\mathcal{C}_1 + k_2\mathcal{C}_2 + \dots + k_n\mathcal{C}_n, \quad (3.6)$$

onde \mathcal{C}_i são matrizes associadas à cônicas e $k_i \in \mathbb{R}$, para $k_i \neq 0$.

De maneira análoga, cônicas duais também podem ser combinadas linearmente para dar origem à novas cônicas duais. A operação é similar à definida para cônicas primais na Equação 3.6. No entanto, são conceitualmente distintas e são nomeadas de forma distinta, também. A combinação de cônicas duais é denominada *rede de cônicas* [59] e é fundamental para entender parte da motivação que deu origem a um dos monômios invariantes propostos no Capítulo 5.

3.2.5 Equivalência entre pontos e cônicas duais

Uma equivalência entre pontos finitos no espaço projetivo \mathbb{P}^2 e matrizes simétricas associadas à cônicas duais no espaço das cônicas duais $\mathbb{R}^{3 \times 3}$ pode ser definida como segue:

$$T(\mathbf{u}) : \mathbf{u} \in \mathbb{P}^2 \rightarrow \begin{pmatrix} x^2 & xy & xw \\ xy & y^2 & yw \\ xw & yw & w^2 \end{pmatrix} \in \mathbb{R}^{3 \times 3}, \quad (3.7)$$

onde $\mathbf{u} = (x, y, w)^\top$.

Essa equivalência será utilizada no Capítulo 5 para mapear os pontos finitos que delimitam junções para o espaço dual das matrizes associadas à seções cônicas.

3.3 Transformações Lineares Projetivas

Transformações geométricas são operações aplicadas sobre primitivas geométricas que permitem mapear a disposição dos elementos pertencentes à um domínio para um outro, igual ou distinto. O grupo de transformações ditas lineares projetivas em \mathbb{P}^2 é composto por operações que mapeiam elementos pertencentes à um plano projetivo em outro através da aplicação de uma sequência finita de projeções.

Em álgebra matricial, uma *transformação projetiva* \mathcal{H} , em \mathbb{P}^2 , pode ser descrita por uma matriz 3×3

$$\mathcal{H} = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix}, \quad (3.8)$$

onde \mathcal{H} é uma matriz com determinante não nulo.

Transformações lineares projetivas são classificadas em diversos subgrupos considerando suas propriedades geométricas. Conforme ilustrado na Figura 3.6, o grupo mais geral de transformações lineares projetivas ainda pode ser especializado e subdividido em outros grupos com características bem definidas, como os grupos isométrico, de similaridade e afim, descritos nas seções que se sucedem.

3.3.1 Grupo das isometrias (transformações rígidas)

São transformações que preservam a forma geral e a escala dos objetos envolvidos. Os axiomas definidos pela geometria euclidiana são preservados e, por isso, também são conhecidas como transformações euclidianas. São elas, as rotações, translações e reflexões. Ao preservar o formato e tamanho originais dos objetos, pode-se garantir que distâncias entre pontos e angulações entre vetores ou retas são ditas grandezas invariantes sobre transformações rígidas, assim como a área de estruturas geométricas.

3.3.2 Grupo das similaridades (transformações de similaridade)

Este grupo inclui as transformações capazes de modificar o tamanho dos objetos envolvidos, no entanto, sem alterar o formato dos mesmos. Tais transformações, além de

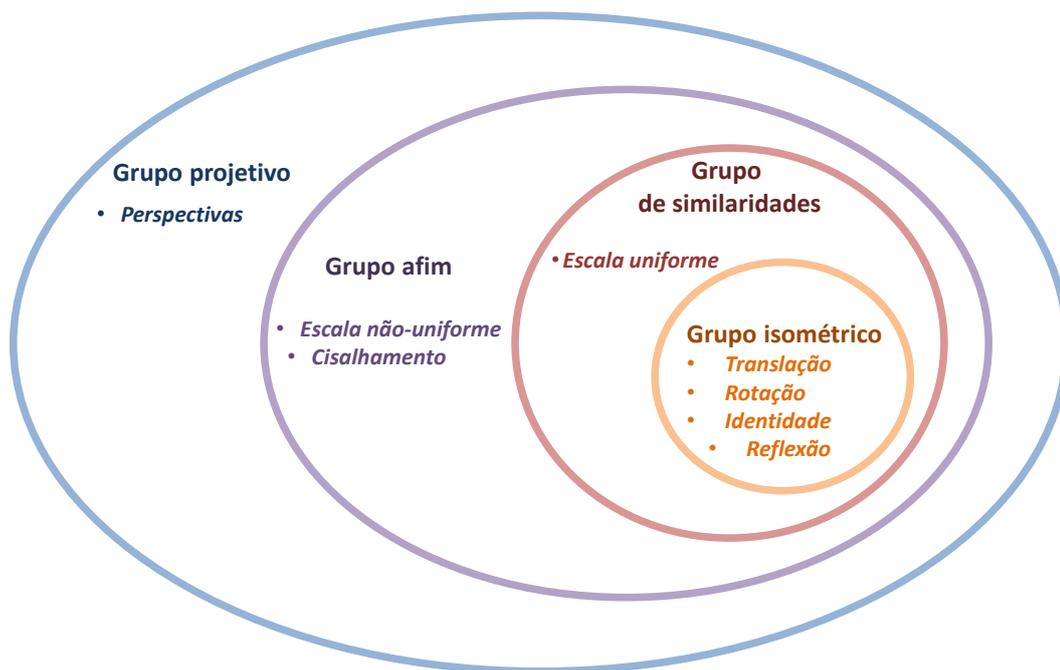


Figura 3.6: Hierarquia do grupo de transformações projetivas. O grupo mais interno representa o grupo isométrico. O grupo de similaridade, além de incluir o grupo isométrico, considera também a escala uniforme. O grupo afim engloba o grupo de similaridades e, ainda, as transformações de cisalhamento e escalas não-uniformes. O grupo projetivo considera a mudança perspectiva, além do grupo afim.

incluïrem as isométricas, se caracterizam pela inclusão da escala uniforme de um objeto.

Apesar de não preservarem a distância entre pontos, a razão entre distâncias, assim como as angulações, são grandezas invariantes sob este grupo de transformações.

3.3.3 Grupo afim (transformações afins)

As transformações afins formam um subgrupo especial das transformações projetivas que, além de permitir transformações de similaridade e euclidianas, permitem a distorção do formato dos objetos em cena, incorporando transformações de cisalhamento e escalas não-uniforme.

Transformações afins aplicadas sobre um plano projetivo qualquer mantém o paralelismo entre retas assim como a razão entre distâncias entre pontos pertencentes à retas paralelas, pontos de inflexão entre retas e razão entre áreas.

3.3.4 Grupo projetivo (transformações projetivas)

Além de englobar todos os outros grupos, o grupo projetivo inclui as transformações que mapeiam elementos geométricos em \mathbb{P}^2 para um plano projetivo diferente. Note que, como ilustrado na Figura 3.7, uma transformação projetiva completa pode ser visualizada como a aplicação de uma sequência finita de projeções de um plano sobre outro, dado pontos em comum de convergência.

Pode ser percebido, a partir da Figura 3.7, que ao se aplicar sequências de projeções de planos sobre planos a colinearidade dos pontos será mantida mesmo que as distâncias entre eles não o sejam. No entanto, a relação cruzada (do inglês *cross-ratio*) das distâncias entre pontos colineares se mantém invariante sobre qualquer transformação projetiva [59]. A relação cruzada, também denominada razão anarmônica, entre quatro pontos colineares **a**, **b**, **c** e **d** (Figura 3.7), é dada pela fórmula:

$$\{\mathbf{a}, \mathbf{b}; \mathbf{c}, \mathbf{d}\} = \frac{\|\mathbf{a} - \mathbf{c}\| \|\mathbf{b} - \mathbf{d}\|}{\|\mathbf{b} - \mathbf{d}\| \|\mathbf{a} - \mathbf{d}\|}, \quad (3.9)$$

onde $\|\mathbf{x} - \mathbf{y}\|$ denota a distância euclidiana orientada entre os pontos **x** e **y**.

O conceito de relação cruzada pode ser aplicado não somente sobre quatro pontos colineares, mas, também, sobre o conjunto de quatro retas concorrentes sobre o plano. Seja o conjunto de quatro retas **p**, **q**, **r** e **s** coplanares e concorrentes em um ponto **o**. A quinta reta **l** de referência deve ser fixada de tal forma que interceccione todas as outras em pontos distintos de **o** (Figura 3.7). Sobre os pontos de intersecção das retas com a

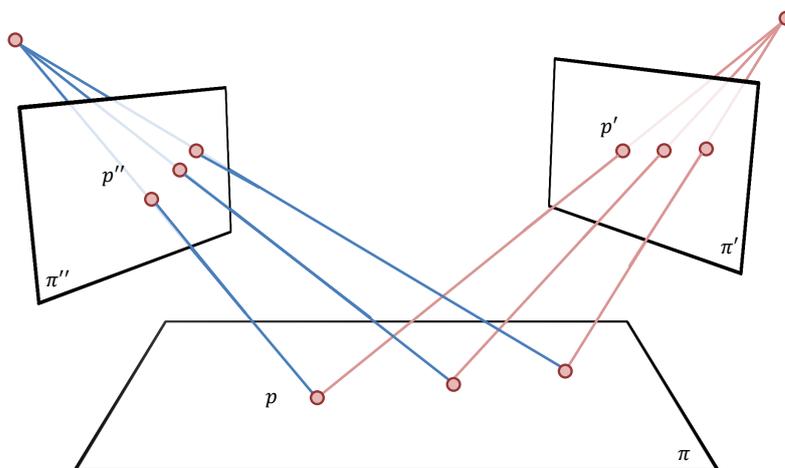


Figura 3.7: Diferentes posicionamentos de um plano projetivo π em \mathbb{P}^2 . Em cada plano projetivo, os mesmos pontos p assumem configurações distintas, perceba que a colinearidade do trio de pontos se mantém.

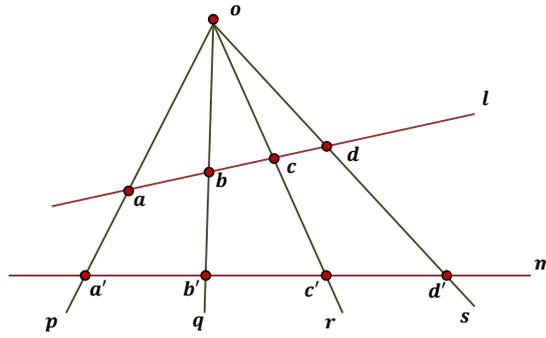


Figura 3.8: Dado um ponto de convergência \mathbf{o} e quatro outros pontos \mathbf{a} , \mathbf{b} , \mathbf{c} e \mathbf{d} , a relação cruzada $\{\mathbf{a}, \mathbf{b}; \mathbf{c}, \mathbf{d}\}$ das distâncias entre os pontos é invariante às transformações projetivas.

reta de referência, a relação cruzada segue como na Equação 3.9.

3.3.5 Aplicação de transformações

Transformações são aplicáveis sobre os elementos geométricos apresentados na Seção 3.2 através de multiplicações matriciais. Seja \mathcal{H} uma matriz de transformação projetiva, temos:

$$\mathbf{u}' = \mathcal{H}\mathbf{u}, \quad \text{onde } \mathbf{u} \text{ é um ponto.} \quad (3.10a)$$

$$\mathbf{l}' = \mathcal{H}^{-\top}\mathbf{l}, \quad \text{onde } \mathbf{l} \text{ é uma reta.} \quad (3.10b)$$

$$\mathcal{A}' = \mathcal{H}^{-\top} \cdot \mathcal{A} \cdot \mathcal{H}^{-1}, \quad \text{onde } \mathcal{A} \text{ é matriz associada à uma cônica primal.} \quad (3.10c)$$

$$\mathcal{A}' = \mathcal{H} \cdot \mathcal{A} \cdot \mathcal{H}^{\top}, \quad \text{onde } \mathcal{A} \text{ é matriz associada à uma cônica dual.} \quad (3.10d)$$

3.4 Discussão

Este capítulo apresentou alguns conceitos e princípios básicos para o entendimento da proposição desta dissertação. A motivação para utilizar somente estruturas de junção como características suficientes para descrever uma parte de interesse da imagem foi baseada em pesquisas anteriores sobre o funcionamento do sistema visual humano e nos monômios invariantes às transformações de similaridade, afim e projetivas propostos, cujas derivações estão formalizadas no Capítulo 5.

Além do conceito de características interessantes, foi introduzido o espaço projetivo \mathbb{P}^2 , onde elementos geométricos conhecidos, como pontos, direções, retas e seções cônicas podem ser utilizados para representar informações visuais extraídas da imagem. Por fim,

foi realizada uma breve apresentação sobre os diferentes grupos de transformações lineares projetivas passíveis de serem induzidas por uma câmera, e como estas afetam os elementos geométricos listados.

Capítulo 4

Teoria de Invariantes

Este capítulo apresenta conceitos específicos da teoria de invariantes, o suficiente para possibilitar o entendimento das formulações propostas neste trabalho. Em especial, será abordada uma das muitas linhas de pesquisa sobre derivação de invariantes, que tem por conceito básico a aplicação do Teorema de Cayley-Hamilton [14, 59] sobre matrizes simétricas.

A Seção 4.1 apresenta o Teorema de Cayley-Hamilton que estabelece que toda matriz possui coeficientes característicos, *i.e.*, quantificações, que compõem o polinômio característico de uma matriz. Na Seção 4.2 é demonstrada uma extensão do teorema para múltiplas matrizes, quando combinadas. Na Seção 4.3 passam a ser consideradas apenas matrizes simétricas, pois ao definir as transformações sobre matrizes simétricas como transformações sobre seções cônicas é possível combinar os coeficientes característicos para obter invariância à diversos grupos de transformações, combinação esta definida na Seção 4.4.

4.1 Teorema de Cayley-Hamilton

Seja

$$\mathcal{M} = \begin{pmatrix} m_{11} & m_{12} & \cdots & m_{1n} \\ m_{21} & m_{22} & \cdots & m_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ m_{n1} & m_{n2} & \cdots & m_{nn} \end{pmatrix} \quad (4.1)$$

uma matriz quadrada $n \times n$ e com determinante diferente de zero. O Teorema de Cayley-Hamilton estabelece que toda matriz \mathcal{M} possui suas próprias grandezas representativas,

derivadas a partir da equação

$$\text{cht}(k\mathcal{I}, \mathcal{M}) = \det(\mathcal{M} - k\mathcal{I}), \quad (4.2)$$

onde \mathcal{I} representa a matriz identidade $n \times n$, k é um coeficiente linear ainda desconhecido e \det denota a operação de determinante aplicada sobre matrizes.

O polinômio que se desenrola a partir da Equação 4.2 é tido como o *polinômio característico* da matriz \mathcal{M} . Ao se estender

$$\text{cht}(k\mathcal{I}, \mathcal{M}) = k^n - c_1 k^{n-1} + \dots + (-1)^n c_n, \quad (4.3)$$

onde c_1, c_2, \dots, c_n são coeficientes lineares, também denominados *coeficientes característicos* da matriz \mathcal{M} . Os coeficientes característicos são representações quantitativas da matriz \mathcal{M} . Se transformações em \mathcal{M} forem definidas assim como para matrizes associadas à seções cônicas, então os coeficientes característicos seriam ditos invariantes à transformações isométricas.

Exemplo 1 (Derivação do polinômio característico e dos coeficientes característicos associados à uma matriz 2×2).

Seja a matriz

$$\mathcal{M} = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}.$$

O polinômio característico associado é dado por:

$$\text{cht}(k\mathcal{I}, \mathcal{M}) = k^2 - (m_{11} + m_{22})k + (m_{11}m_{22} - m_{12}m_{21}).$$

Portanto, os três coeficientes característicos são

$$\begin{aligned} c_0 &= (m_{11}m_{22} - m_{12}m_{21})k^0 = \det(\mathcal{M})k^0, \\ c_1 &= (m_{11} + m_{22})k^1 = \text{tr}(\mathcal{M})k^1, \\ c_2 &= k^2, \end{aligned}$$

onde tr define a operação de traço aplicada sobre matrizes.

4.2 Adaptação para Múltiplas Matrizes

O polinômio característico também pode ser associado à combinação linear de mais de uma matriz, dada sua adaptação

$$\begin{aligned} \text{chtn}(k_1\mathcal{I}, \mathcal{M}_1, k_2\mathcal{I}, \mathcal{M}_2, \dots, k_n\mathcal{I}, \mathcal{M}_n) \\ = \det(k_1\mathcal{I} \cdot \mathcal{M}_1 + k_2\mathcal{I} \cdot \mathcal{M}_2 + \dots + k_n\mathcal{I} \cdot \mathcal{M}_n) \\ = f_1c_1 + \dots + \underbrace{k_1^{p_{i1}} k_2^{p_{i2}} \dots k_n^{p_{in}}}_{f_i} c_i + \dots + f_n c_n, \end{aligned} \quad (4.4)$$

onde f_i denota uma combinação entre os fatores combinatórios de entrada k_1, k_2, \dots, k_n e $p_{i1}, p_{i2}, \dots, p_{in}$ denotam escalares apenas para demonstração da fórmula geral.

Exemplo 2 (Derivação do polinômio característico e dos coeficientes característicos associados à duas matrizes 3×3).

Sejam as matrizes

$$\mathcal{M} = \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{pmatrix}, \quad \mathcal{N} = \begin{pmatrix} n_{11} & n_{12} & n_{13} \\ n_{21} & n_{22} & n_{23} \\ n_{31} & m_{32} & n_{33} \end{pmatrix}.$$

O polinômio característico associado é dado por $\text{chtn}(k_1\mathcal{I}, \mathcal{M}, k_2\mathcal{I}, \mathcal{N})$.

Portanto, os quatro coeficientes característicos são

$$\begin{aligned} c_0 f_0 &= \det(\mathcal{M}_1) \overbrace{k_1^3 k_2^0}^{f_0}, \\ c_1 f_1 &= \text{tr}(\mathcal{M}_1^{-1} \cdot \mathcal{M}_2) \det(\mathcal{M}_2) \overbrace{k_1^2 k_2^1}^{f_1}, \\ c_2 f_2 &= \text{tr}(\mathcal{M}_2^{-1} \cdot \mathcal{M}_1) \det(\mathcal{M}_1) \overbrace{k_1^1 k_2^2}^{f_2}, \\ c_3 f_3 &= \det(\mathcal{M}_2) \overbrace{k_1^0 k_2^3}^{f_3}. \end{aligned}$$

4.3 Coeficientes Característicos de Matrizes Simétricas

Ao se considerar apenas matrizes simétricas 3×3 , é possível definir a aplicação de uma homografia \mathcal{H} de maneira análoga à aplicação sobre seções cônicas (ver Equação 3.10c). Definidas como transformações serão aplicadas sobre as matrizes de entrada, é possível verificar a invariância dos coeficientes característicos sob rotações. A translação, como mudança de ponto referencial para a origem do plano projetivo, é assumida como uma operação realizada previamente, o que permite que a invariância ocorra sob o grupo de transformações rígidas.

Seja a transformação rígida

$$\mathcal{R} = \begin{pmatrix} r_{11} & r_{12} & 0 \\ r_{21} & r_{22} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{tal que } \det(\mathcal{R}) = 1, \quad (4.5)$$

aplicada sobre a matriz simétrica \mathcal{M} , a invariância dos coeficientes característicos segue:

$$\begin{aligned} c_0 &= \det(\mathcal{R}^{-\top} \cdot \mathcal{M} \cdot \mathcal{R}^{-1}) \\ &= \det(\mathcal{R}^{\top})^{-1} \det(\mathcal{M}) \det(\mathcal{R})^{-1} \\ &= \det(\mathcal{M}), \quad \text{pois } \det(\mathcal{R}) = 1. \end{aligned} \quad (4.6a)$$

$$\begin{aligned} c_1 &= \text{tr}(\mathcal{R}^{-\top} \cdot \mathcal{M} \cdot \mathcal{R}^{-1}) \\ &= \text{tr}(\mathcal{M} \cdot (\mathcal{R}^{\top} \cdot \mathcal{R})^{-1}) \\ &= \text{tr}(\mathcal{M}), \end{aligned}$$

A propriedade distributiva da operação de determinante sobre a multiplicação de matrizes e a invariância cíclica de multiplicações no traço associadas ao valor unitário do determinante das transformações rígidas permite provar que coeficientes característicos são valores invariantes sob o grupo de transformações rígidas.

Perceba que, quando o determinante de uma homografia que opere sobre as matrizes de entrada for diferente de 1, esses coeficientes deixam de ser invariantes. No entanto, é possível verificar que taxas entre os coeficientes podem ser suficientes para prover invariância à transformações de similaridade. Seja o invariante sob similaridades

$$\Lambda = \frac{\text{dsc}(\mathcal{M})}{\text{tr}(\mathcal{M})^2}, \quad (4.7)$$

onde dsc denota a operação de discriminante de uma matriz simétrica apresentado pela Equação 3.5 e tr representa a operação de traço matricial, ambas estão relacionadas aos coeficientes característicos.

Seja a transformação de similaridade

$$\mathcal{S} \cdot \mathcal{R} = \begin{pmatrix} s & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} r_{11} & r_{12} & 0 \\ r_{21} & r_{22} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (4.8)$$

onde $s \neq 0$ é um fator uniforme de escala e \mathcal{R} é uma matriz de rotação (Equação 4.5), aplicando a transformação definida na matriz simétrica \mathcal{M} na Equação 4.3, vemos que os coeficientes característicos são afetados como segue:

$$\begin{aligned} c_0 &= \text{dsc} \left((\mathcal{S} \cdot \mathcal{R})^{-\top} \cdot \mathcal{M} \cdot (\mathcal{S} \cdot \mathcal{R})^{-1} \right) \\ &= \text{dsc} \left((\mathcal{R}^{-1} \cdot \mathcal{S}^{-1})^{\top} \cdot \mathcal{M} \cdot (\mathcal{R}^{-1} \cdot \mathcal{S}^{-1}) \right) \\ &= \text{dsc} \left(\mathcal{R}^{-1} \cdot \mathcal{S}^{-1} \right) \text{dsc} (\mathcal{M}) \text{dsc} \left(\mathcal{R}^{-1} \cdot \mathcal{S}^{-1} \right) \\ &= \text{dsc} \left(\mathcal{S}^{-1} \right)^2 \text{dsc} (\mathcal{M}) \\ &= \text{dsc} \left(\mathcal{S}^{-1} \right)^2 \text{dsc} (\mathcal{M}) \\ &= \frac{1}{s^4} \text{dsc} (\mathcal{M}), \end{aligned} \quad (4.9)$$

$$\begin{aligned} c_1 &= \text{tr} \left((\mathcal{S} \cdot \mathcal{R})^{-\top} \cdot \mathcal{M} \cdot (\mathcal{S} \cdot \mathcal{R})^{-1} \right) \\ &= \text{tr} \left(\mathcal{M} \cdot (\mathcal{S} \cdot \mathcal{R})^{-1} \cdot (\mathcal{S} \cdot \mathcal{R})^{-\top} \right) \\ &= \text{tr} \left(\mathcal{M} \cdot (\mathcal{R}^{-1} \cdot \mathcal{S}^{-1} \cdot \mathcal{S}^{-\top} \cdot \mathcal{R}^{-\top}) \right) \\ &= \text{tr} \left(\mathcal{M} \cdot \mathcal{S}^{-1} \cdot \mathcal{S}^{-1} \right) \\ &= \frac{1}{s^2} \text{tr} (\mathcal{M}), \quad \text{pois } \mathcal{S} \text{ é diagonal.} \end{aligned}$$

Portanto, apenas uma razão entre c_0 e c_1 , como apresentada pela Equação 4.7 é capaz de anular os efeitos da escala, tornando-a invariante à transformações de similaridade.

4.4 Composição de Coeficientes Característicos

A composição de coeficientes característicos pode ser realizada por análise dos efeitos que grupos específicos de transformações causam sobre as operações de determinante e traço, assim como foi formulado o invariante sob similaridades da Equação 4.7 [59].

Portanto, os quatro coeficientes característicos são

$$\begin{aligned} c_0 f_0 &= \det(\mathcal{M}) \overbrace{k_1^3 k_2^0}^{f_0} \\ c_1 f_1 &= \operatorname{tr}(\mathcal{M}^{-1} \cdot \mathcal{N}) \det(\mathcal{M}) \overbrace{k_1^2 k_2^1}^{f_1} \\ c_2 f_2 &= \operatorname{tr}(\mathcal{N}^{-1} \cdot \mathcal{M}) \det(\mathcal{N}) \overbrace{k_1^1 k_2^2}^{f_2} \\ c_3 f_3 &= \det(\mathcal{N}) \overbrace{k_1^0 k_2^3}^{f_3}. \end{aligned}$$

Os fatores k_i associados aos coeficientes característicos são particularidades de uma combinação específica e, de fato, precisam ser eliminados completamente diante do monômio da Equação 4.10 para que sejam considerados verdadeiramente invariantes

$$\begin{aligned} \prod_{i=0}^n (f_i c_i)^{a_i} &= (k_1^3 k_2^0 c_0)^{a_0} (k_1^2 k_2^1 c_1)^{a_1} (k_1^1 k_2^2 c_2)^{a_2} (k_1^0 k_2^3 c_3)^{a_3} \\ &= (k_1^{3a_0} k_2^{0a_0} c_0^{a_0}) (k_1^{2a_1} k_2^{1a_1} c_1^{a_1}) (k_1^{1a_2} k_2^{2a_2} c_2^{a_2}) (k_1^{0a_3} k_2^{3a_3} c_3^{a_3}) \\ &= k_1^{3a_0} c_0^{a_0} k_1^{2a_1} k_2^{a_1} c_1^{a_1} k_1^{a_2} k_2^{2a_2} c_2^{a_2} k_2^{3a_3} c_3^{a_3}. \end{aligned}$$

Se a influência dos fatores particulares k_i deve ser anulada do produtório, diante de uma multiplicação com potências, perceba que a soma das potências deve ser nula, tanto para as potências de k_1 quanto para as potências de k_2 , simultaneamente. Portanto,

$$\begin{cases} 3a_0 + 2a_1 + a_2 = 0 \\ a_1 + 2a_2 + 3a_3 = 0 \end{cases}.$$

As duas soluções linearmente independentes para o sistema apresentado são: $a_0 = 1$, $a_1 = -2$, $a_2 = 1$, $a_3 = 0$ e $a_0 = 0$, $a_1 = 1$, $a_2 = -2$, $a_3 = 1$. Logo, dois monômios característicos linearmente independentes entre si são obtidos:

$$\Gamma = \frac{c_0 c_2}{c_1^2} \quad e \quad \Upsilon = \frac{c_1 c_3}{c_2^2}.$$

Para a combinação específica de duas matrizes simétricas, \mathcal{M} e \mathcal{N} estas podem ser interpretadas como seções cônicas e as transformações são aplicadas de maneira análoga.

Os monômios característicos

$$\Gamma = \frac{c_0 c_2}{c_1^2} = \frac{\text{tr}(\mathcal{N}^{-1} \cdot \mathcal{M}) \det(\mathcal{N})}{\text{tr}(\mathcal{M}^{-1} \cdot \mathcal{N})^2 \det(\mathcal{M})} \quad (4.13a)$$

e

$$\Upsilon = \frac{c_1 c_3}{c_2^2} = \frac{\text{tr}(\mathcal{M}^{-1} \cdot \mathcal{N}) \det(\mathcal{M})}{\text{tr}(\mathcal{N}^{-1} \cdot \mathcal{M})^2 \det(\mathcal{N})}, \quad (4.13b)$$

cujos processos de derivação foram apresentados pelo Exemplo 3, são geometricamente provados invariantes sob transformações projetivas em [25].

4.5 Discussão

Este capítulo introduziu uma das muitas linhas de pesquisa sobre a Teoria de Invariantes, apresentando definições iniciais e teoremas estabelecidos para quantizar matrizes. O Teorema de Cayley-Hamilton, aplicado sobre matrizes quadradas, introduziu os coeficientes característicos como sendo grandezas invariantes obtidas de operações de determinante e traço. Especificamente, para matrizes simétricas 3×3 , verificou-se sua invariância sob a aplicação de transformações rígidas.

Ao aprofundar um pouco mais sobre os efeitos da aplicação de transformações sobre matrizes simétricas, mostrou-se que é possível combinar os coeficientes característicos para alcançar invariância às transformações de similaridade. Também foi mostrado que uma única matriz não possui pistas suficientes para permitir a formulação de um invariante projetivo, sendo necessária a combinação de duas matrizes simétricas 3×3 para obter tal invariante.

Os invariantes sobre pares de seções cônicas, *i.e.*, matrizes simétricas, já são aplicados para resolver problemas de reconhecimento de padrões visuais [20]. No entanto, não foram especializados para estruturas interessantes em imagens, tais como as junções. No Capítulo 5 é proposta uma abordagem estável e de comportamento previsível sobre matrizes obtidas à partir da estrutura de junções.

Capítulo 5

Monômios Invariantes sobre Junções

A literatura apresenta soluções detectoras de características visuais em centros de junções como referências para características representativas (Capítulo 2). No entanto, as abordagens existentes para descrever o vértice central de uma junção desprezam as propriedades geométricas intrínsecas à morfologia da junção, limitando-se à assinaturas de orientação e magnitude de gradientes da região vizinha ao vértice.

Neste capítulo são formalizadas as junções como estruturas exatas, desconsiderando pormenores de discretização e ruídos inerentes ao processo de formação da imagem, a fim de viabilizar a análise de comportamento das junções sob diferentes transformações lineares. A fundamentação teórica sobre invariantes, apresentada no Capítulo 4, combinada à uma análise contundente dos efeitos de homografias permite a formulação de monômios invariantes especificamente para estruturas de junção.

A formalização matemática de junção e sua representação a partir de matrizes é descrita na Seção 5.1. As Seções 5.2 à 5.5 detalham a formulação dos monômios invariantes, respectivamente, à transformações de similaridade, à transformações afins e à transformações projetivas.

5.1 Formalização de Junções

Seja a junção contínua

$$J_n = \{\mathbf{p} | \mathbf{p} \in \overline{\mathbf{p}_0 \mathbf{p}_k}, \forall k \in \{1, 2, \dots, n\}\}, \quad (5.1)$$

como um conjunto formado por pontos $\mathbf{p} = (x, y)^\top$, pertencentes à ramificações $\overline{\mathbf{p}_0 \mathbf{p}_k}$, onde \mathbf{p}_0 denota o *vértice central* da junção e os pontos \mathbf{p}_k indicam os *vértices delimitadores*

de ramificação. Uma junção pode assumir $n \in \{2, 3, \dots\}$ ramificações. Porém, aqui serão limitadas à quatro por restrições impostas em capítulos futuros, pelo uso de imagens digitais em retículo quadrado.

Os vértices centrais sempre serão finitos. Entretanto, a definição mais elementar de junção estabelece que uma ramificação pode ser, simplesmente, uma direção (*i.e.*, ponto impróprio). Por isso, classifica-se como *junção própria* ou *junção imprópria* aquelas junções que possuem vértices delimitadores, respectivamente, próprios ou impróprios.

O total de ramificações, ou *grau* n , também é um fator característico para a estrutura de uma junção, já que junções de graus mais elevados são mais dessemelhantes quando especificado um grupo de transformações. Por exemplo, à partir de uma transformação afim, é possível fazer um mapeamento entre quaisquer junções com duas ramificações, o que restringe a dessemelhança para o grupo homotético de transformações. Estabelece-se, então, que uma junção com grau n será denotada como n -junção ou J_n .

Junções, ainda, podem ser classificadas de acordo com seu formato, de acordo com a disposição das ramificações. As 2-junções podem assumir o formato de um L. As 3-junções assumem formatos de Y, T ou \nearrow . As 4-junções se formatam em X ou K. A Figura 5.1 ilustra todos os tipos de junção e suas classificações.

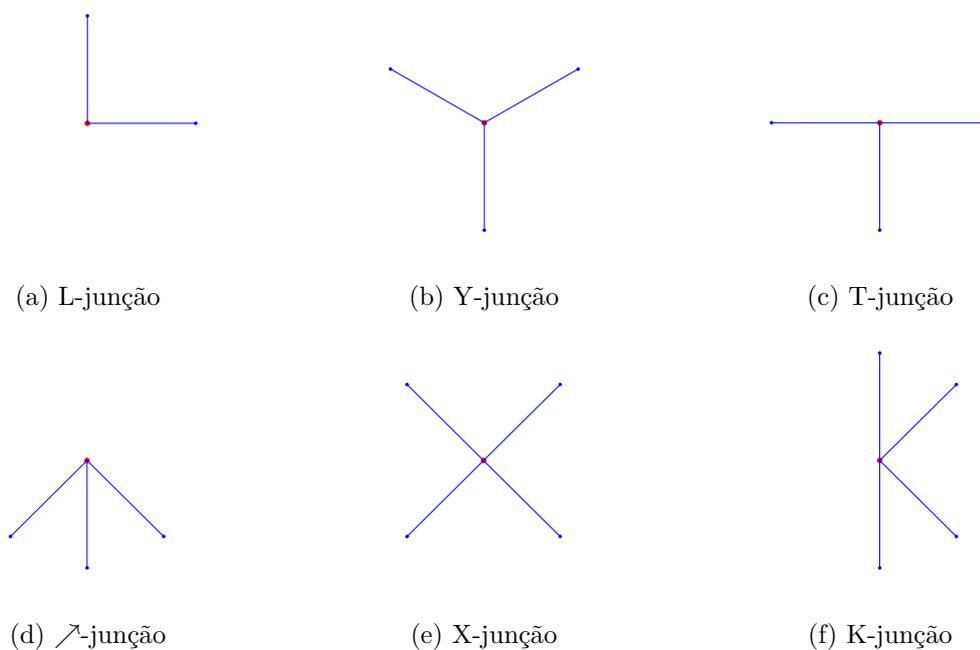


Figura 5.1: Diferentes classificações de junções por formato. As classificações por formato variam dentre (a) L-junção, (b) Y-junção, (c) T-junção, (d) \nearrow -junção, (e) X-junção e (f) K-junção.

Para a posterior aplicação dos monômios formulados, são assumidas como entrada

matrizes simétricas, propriedade esta essencial para descrever invariantes segundo os conceitos apresentados no Capítulo 4.

5.2 Invariante sobre Similaridades para 2-Junções

O grande trunfo das formulações expostas no Capítulo 4 é o aproveitamento das particularidades das matrizes simétricas de entrada. Alguns monômios invariantes necessitam de duas ou mais matrizes combinadas para preencherem os graus de liberdade em transformações afins e projetivas, por exemplo. Para o caso de transformações de similaridade, o monômio invariante Λ , apresentado pela Equação 4.7, necessita de apenas uma matriz simétrica como entrada.

Seja

$$\mathcal{Q}_{1,2} = \mathcal{D}_1 + \mathcal{D}_2 + \mathcal{D}_0, \quad (5.2)$$

onde $\mathcal{Q}_{1,2}$ é a rede de cônicas ou rede de matrizes simétricas associada aos vértices delimitadores de J_2 e \mathcal{D}_i denota a matriz associada à cônica dual equivalente ao vértice $\mathbf{p}_i \in \mathbb{P}^2$ (Equação 3.7).

Transformações aplicadas sobre os vértices delimitadores de J_2 afetam a matriz simétrica associada $\mathcal{Q}_{1,2}$ da mesma maneira que afetam cônicas duais (ver Equação 3.10d). Perceba que

$$\begin{aligned} \mathcal{Q}'_{1,2} &= \mathcal{H} \cdot \mathcal{D}_1 \cdot \mathcal{H}^\top + \mathcal{H} \cdot \mathcal{D}_2 \cdot \mathcal{H}^\top + \mathcal{H} \cdot \mathcal{D}_0 \cdot \mathcal{H}^\top \\ &= \mathcal{H} \cdot \mathcal{Q}_{1,2} \cdot \mathcal{H}^\top. \end{aligned} \quad (5.3)$$

Portanto, um monômio invariante às transformações de similaridade para junções segue pela aplicação direta da matriz $\mathcal{Q}_{1,2}$ (Equação 5.2) na Equação 4.7.

Tratando-se de transformações de similaridade, é sabido que a simples razão entre distâncias representa um quantitativo invariante, o que possibilita calcular razões entre os comprimentos de seus ramos. Isso bastaria para alcançar a invariância mas não seria suficiente para distinguir um par de 2-junção, pois o ângulo entre as ramificações é uma característica importante para a estrutura de uma junção. Formulemos, portanto, o invariante proposto sobre similaridades para junções,

$$\gamma = \cos(\angle \mathbf{p}_1 \mathbf{p}_0 \mathbf{p}_2) \frac{d_1}{d_2}, \quad (5.4)$$

onde \mathbf{p}_0 denota o vértice central da junção, \mathbf{p}_1 e \mathbf{p}_2 denotam os vértices delimitadores de

ramificação, d_1 e d_2 são as distâncias, respectivamente, entre \mathbf{p}_0 e \mathbf{p}_1 e entre \mathbf{p}_0 e \mathbf{p}_2 e por fim, $\angle_{\mathbf{p}_1\mathbf{p}_0\mathbf{p}_2}$ representa o ângulo interno entre $\overline{\mathbf{p}_0\mathbf{p}_1}$ e $\overline{\mathbf{p}_0\mathbf{p}_2}$.

Independentemente de qualquer rotação 2D aplicada sobre a junção, os efeitos serão aplicados igualmente sobre os vértices delimitadores, o que mantém o ângulo interno constante entre eles. Transformações de escala uniforme podem afastar os vértices delimitadores do vértice central. No entanto, não modificarão a angulação entre eles. Rotações não são suficientes para modificar a distância entre cada vértice delimitador e o vértice central, apenas uma escala uniforme, no âmbito das similaridades. Apesar disso, por ser uniforme, a escala vai afastar ambos os vértices na mesma proporção, o que torna a razão entre distâncias, uma grandeza constante. A translação é anulada ao tomar p_0 como a origem do sistema de coordenadas.

5.3 Invariante Afim sobre 3-Junções

O invariante afim proposto apresenta-se como segue:

$$\beta_{123} = \frac{\det(Q_{\{1,2\}})}{\text{tr}(Q_{\{2,3\}}^{-1} \cdot Q_{\{1,2\}})^2 \det(Q_{\{2,3\}})}, \quad (5.5)$$

onde $Q_{\{i,j\}}$ representa a matriz simétrica construída a partir Equação 5.2, utilizando apenas os vértices i e j . A invariância afim será alcançada pela combinação de duas matrizes simétricas, onde a fundamentação remete-se à intersecção de seções cônicas.

Considerar ao menos duas matrizes simétricas implica, segundo os conceitos definidos, na total impotência de se distinguir quaisquer 2-junções em âmbito afim. Em uma análise sucinta, de fato, uma 2-junção possui três vértices e um mapeamento afim é cabível entre quaisquer três pontos. Portanto, invariantes afim somente serão válidos para junções com, ao menos, três ramificações.

Os três ramos podem ser permutados e, efetivamente, gerar seis valores invariantes. As matrizes simétricas geradas, no entanto, também possuem simetria em relação à ordem de escolha dos vértices. Ou seja, $Q_{\{i,j\}} \equiv Q_{\{j,i\}}$. Por isso, as ramificações devem seguir ordenação combinatória, assim gerando apenas três valores invariantes distintos.

É importante ressaltar que as matrizes simétricas criadas são baseadas em seções cônicas, sejam elas primais ou duais, e como visto na Subseção 3.2.3, as equações de seções cônicas podem ser escaladas por um fator genérico k , diferente de zero, e continuam representando a mesma seção cônica. Devido à liberdade de se escalar as matrizes simétricas

associadas às cônicas, o termo $\text{tr} \left(\mathcal{Q}_{\{2,3\}}^{-1} \cdot \mathcal{Q}_{\{1,2\}} \right)$ por si só não é invariante, dependendo da razão de determinantes que o acompanha na Equação 5.5.

Seja a transformação afim \mathcal{H} aplicada sobre os vértices delimitadores de uma junção J_3 , o invariante β_{123} reage como segue:

$$\begin{aligned}
\beta_{123} &= \frac{\det \left(\mathcal{H} \cdot \mathcal{Q}_{\{1,2\}} \cdot \mathcal{H}^\top \right)}{\text{tr} \left(\mathcal{H}^{-\top} \cdot \mathcal{Q}_{\{2,3\}}^{-1} \cdot \mathcal{H}^{-1} \cdot \mathcal{H} \cdot \mathcal{Q}_{\{1,2\}} \cdot \mathcal{H}^\top \right)^2 \det \left(\mathcal{H} \cdot \mathcal{Q}_{\{2,3\}} \cdot \mathcal{H}^\top \right)} \\
&= \frac{\det \left(\mathcal{H} \right)^2 \det \left(\mathcal{Q}_{\{1,2\}} \right)}{\text{tr} \left(\mathcal{H}^{-\top} \cdot \mathcal{Q}_{\{2,3\}}^{-1} \cdot \mathcal{H}^{-1} \cdot \mathcal{H} \cdot \mathcal{Q}_{\{1,2\}} \cdot \mathcal{H}^\top \right)^2 \det \left(\mathcal{H} \right)^2 \det \left(\mathcal{Q}_{\{2,3\}} \right)} \\
&= \frac{\det \left(\mathcal{Q}_{\{1,2\}} \right)}{\text{tr} \left(\mathcal{H}^{-\top} \cdot \mathcal{Q}_{\{2,3\}}^{-1} \cdot \mathcal{Q}_{\{1,2\}} \cdot \mathcal{H}^\top \right)^2 \det \left(\mathcal{Q}_{\{2,3\}} \right)} \\
&= \frac{\det \left(\mathcal{Q}_{\{1,2\}} \right)}{\text{tr} \left(\mathcal{Q}_{\{2,3\}}^{-1} \cdot \mathcal{Q}_{\{1,2\}} \cdot \mathcal{H}^\top \cdot \mathcal{H}^{-\top} \right)^2 \det \left(\mathcal{Q}_{\{2,3\}} \right)} \\
&= \frac{\det \left(\mathcal{Q}_{\{1,2\}} \right)}{\text{tr} \left(\mathcal{Q}_{\{2,3\}}^{-1} \cdot \mathcal{Q}_{\{1,2\}} \right)^2 \det \left(\mathcal{Q}_{\{2,3\}} \right)}.
\end{aligned}$$

Perceba que o termo $\text{tr} \left(\mathcal{Q}_{\{2,3\}}^{-1} \cdot \mathcal{Q}_{\{1,2\}} \right)$ suprime a transformação por si só e a razão entre os determinantes também aparenta ser suficiente para prover invariância. Ao considerar que a matriz simétrica associada à uma seção cônica pode ser escalada por fatores quaisquer diferentes de zero e diferentes entre si, o invariante β_{123} , considerando $k_1 \mathcal{Q}_{\{1,2\}}$ e $k_2 \mathcal{Q}_{\{2,3\}}$, é afetado da seguinte forma:

$$\begin{aligned}
\beta_{123} &= \frac{\det \left(k_1 \mathcal{Q}_{\{1,2\}} \right)}{\text{tr} \left(\frac{1}{k_2} \mathcal{Q}_{\{2,3\}}^{-1} \cdot k_1 \mathcal{Q}_{\{1,2\}} \right)^2 \det \left(k_2 \mathcal{Q}_{\{2,3\}} \right)} \\
&= \frac{\det \left(k_1 \mathcal{Q}_{\{1,2\}} \right)}{\text{tr} \left(\frac{k_1}{k_2} \mathcal{Q}_{\{2,3\}}^{-1} \cdot \mathcal{Q}_{\{1,2\}} \right)^2 \det \left(k_2 \mathcal{Q}_{\{2,3\}} \right)} \\
&= \frac{k_1^3 \det \left(\mathcal{Q}_{\{1,2\}} \right)}{\frac{k_1^3}{k_2^3} \text{tr} \left(\mathcal{Q}_{\{2,3\}}^{-1} \cdot \mathcal{Q}_{\{1,2\}} \right) k_2^3 \det \left(\mathcal{Q}_{\{2,3\}} \right)} \\
&= \frac{k_1^3 \det \left(\mathcal{Q}_{\{1,2\}} \right)}{\frac{k_1^3}{k_2^3} \text{tr} \left(\mathcal{Q}_{\{2,3\}}^{-1} \cdot \mathcal{Q}_{\{1,2\}} \right)^2 k_2^3 \det \left(\mathcal{Q}_{\{2,3\}} \right)} \\
&= \frac{\det \left(\mathcal{Q}_{\{1,2\}} \right)}{\text{tr} \left(\mathcal{Q}_{\{2,3\}}^{-1} \cdot \mathcal{Q}_{\{1,2\}} \right)^2 \det \left(\mathcal{Q}_{\{2,3\}} \right)},
\end{aligned}$$

onde k_1 e k_2 são fatores de escala diferentes de zero e associados, respectivamente, às matrizes simétricas $\mathcal{Q}_{\{1,2\}}$ e $\mathcal{Q}_{\{2,3\}}$. Logo, β é um monômio verdadeiramente invariante às transformações afim. Até o momento, todos os monômios exigiram a permutação dos três vértices de entrada, o que se simplifica numa trinca β_{123} , β_{231} e β_{312} de invariantes distintos.

A trinca de invariantes afim obtida possui uma característica interessante: os três valores resultantes, se interpretados como coordenadas num espaço \mathbb{R}^3 , definem as coordenadas de pontos contidos em um triângulo. A Figura 5.2 ilustra um *plot* de pontos $(\beta_{123}, \beta_{231}, \beta_{312})$ calculados sobre amostras aleatórias de junção. Note que o domínio se resume à um triângulo contido em um cubo unitário. A trinca de invariantes, portanto, se comporta como coordenadas baricêntricas, o que implica em $\beta_{123} + \beta_{231} + \beta_{312} = 1$. Isto reduz a trinca para apenas um par de invariantes, pois $\beta_{123} = 1 - \beta_{231} - \beta_{312}$.

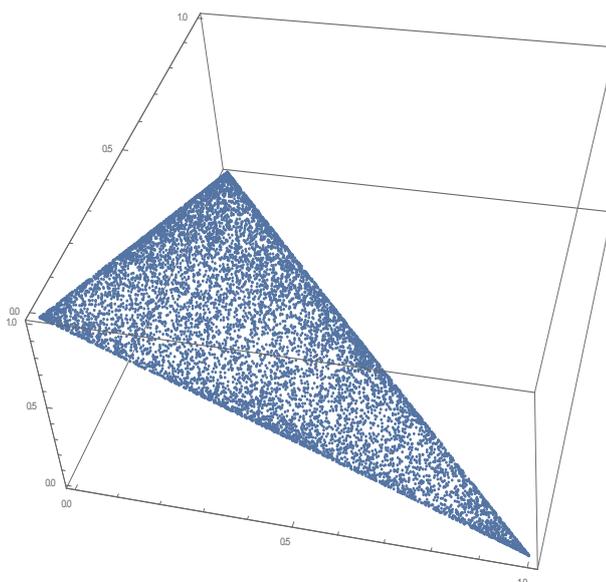


Figura 5.2: *Plot* de pontos $(\beta_{123}, \beta_{231}, \beta_{312})$ calculados a partir de 10000 amostras aleatórias de junção utilizando a Equação 5.5. O triângulo formado, em azul, representa o domínio dos invariantes β_{123} , β_{231} , e β_{312} quando interpretados como coordenadas num espaço \mathbb{R}^3 . Essa característica implica em $\beta_{123} = 1 - \beta_{231} - \beta_{312}$, reduzindo a trinca de invariantes para um par.

5.4 Invariante Projetivo sobre 4-Junções

As transformações de similaridade e afim possuem a característica de não afetar a reta imprópria no espaço \mathbb{P}^2 . Ou seja, pontos impróprios (direções), após transformados, são mapeados para pontos impróprios por esses grupos de transformações. Quando considera-

das transformações projetivas, a reta imprópria pode ser mapeada para uma reta própria e vice-versa. Logo, transformações projetivas afetam a coordenada homogênea de pontos, o que leva à necessidade de normalização de pontos próprios (*i.e.*, divisão das coordenadas $(x, y, w)^\top$ do vetor por w) para a obtenção de sua localização. Os invariantes apresentados nas Seções 5.2 e 5.3 não são robustos à normalizações de pontos próprios, sendo essa operação desnecessária naqueles contextos. O invariante apresentado nesta seção é robusto à transformações projetivas e à normalização pela coordenada homogênea.

Apresenta-se, por conseguinte, o invariante projetivo

$$\alpha = \frac{\det(\mathcal{P} \cdot \mathcal{P})}{\text{tr}(\mathcal{P} \cdot \mathcal{P})^3}, \quad (5.6)$$

onde \mathcal{P} é uma matriz gerada em função dos vértices de uma 4-junção, conforme a Equação 5.7. A multiplicação $\mathcal{P} \cdot \mathcal{P}$ é aplicada, pois observamos experimentalmente que o domínio de α aparentou ser computacionalmente mais estável devido à ordem de grandeza dos valores gerados. Defina então,

$$\mathcal{P} = \sum_{f \in S_4} \prod_{i=1}^4 (\mathcal{D}_{f_i} \cdot \mathcal{K}), \quad (5.7)$$

onde S_4 denota todas as permutações dos quatro vértices delimitadores; f_i é o i -ésimo vértice da combinação $f \in S_4$ corrente; \mathcal{K} é uma matriz associada à cônica primal ajustada sobre o vértice central e todos os quatro vértices f_i ; e \mathcal{D}_{f_i} denota a seção cônica dual equivalente ao vértice f_i (conforme Equação 3.7).

A normalização da coordenada homogênea implica, na prática, em dividir todas as coordenadas de um vetor por $w \neq 0$. Quanto à sua matriz simétrica dual, uma divisão análoga ocorre. A interferência da normalização pela coordenada homogênea w_i sobre os vértices delimitadores \mathbf{p}_i é refletida sobre a matriz simétrica \mathcal{D}_i , definida pela Equação 3.7, como $\mathcal{D}_i = \frac{1}{w_i^2} \mathcal{D}'_i$. De maneira análoga, a seção cônica primal \mathcal{K} também é afetada pela normalização, como $\mathcal{K} = \frac{1}{w_{\mathcal{K}}} \mathcal{K}'$. Assim, tome a multiplicação de todas as matrizes simétricas duais \mathcal{D}_i :

$$\prod_{i=1}^4 (\mathcal{D}_{f_i} \cdot \mathcal{K}) = \frac{1}{(w_1 w_2 w_3 w_4)^2 w_{\mathcal{K}}^4} (\mathcal{D}'_1 \cdot \mathcal{K}' \cdot \mathcal{D}'_2 \cdot \mathcal{K}' \cdot \mathcal{D}'_3 \cdot \mathcal{K}' \cdot \mathcal{D}'_4 \cdot \mathcal{K}'),$$

onde \mathcal{K} é a matriz associada à seção cônica primal ajustada sobre os quatro vértices delimitadores e o vértice central. O produtório de matrizes $\mathcal{D}_{f_i} \cdot \mathcal{K}$ foi introduzido na Equação 5.7 com dois objetivos: o primeiro é tornar α independente da ordenação dos vértices delimitadores da junção; e o segundo é permitir que o produtório de todas as

permutações de vértices sejam multiplicados pela mesma composição de produtórios de $\frac{1}{w_i}$, o que possibilita coletar esse fator escalar multiplicador para fora do somatório da Equação 5.7.

A homografia \mathcal{H} , quando aplicada sobre os vértices da 4-junção, reflete seus efeitos sobre a matriz \mathcal{P} como segue:

$$\begin{aligned}
\mathcal{P}' &= \frac{1}{(w_1 w_2 w_3 w_4)^2 w_{\mathcal{K}}^4} \sum_{f \in S_4} \prod_{i=1}^4 (\mathcal{H} \cdot \mathcal{D}_{f_i} \cdot \mathcal{H}^\top \cdot \mathcal{H}^{-\top} \cdot \mathcal{K} \cdot \mathcal{H}^{-1}) \\
&= \frac{1}{(w_1 w_2 w_3 w_4)^2 w_{\mathcal{K}}^4} \sum_{f \in S_4} \prod_{i=1}^4 (\mathcal{H} \cdot \mathcal{D}_{f_i} \cdot \mathcal{K} \cdot \mathcal{H}^{-1}) \\
&= \frac{1}{(w_1 w_2 w_3 w_4)^2 w_{\mathcal{K}}^4} \sum_{f \in S_4} \prod_{i=1}^4 \mathcal{H} \cdot (\mathcal{D}_{f_i} \cdot \mathcal{K}) \cdot \mathcal{H}^{-1} \\
&= \underbrace{\left(\frac{1}{(w_1 w_2 w_3 w_4)^2 w_{\mathcal{K}}^4} \right)}_{w_{all}} \mathcal{H} \cdot \mathcal{P} \cdot \mathcal{H}^{-1}. \tag{5.8}
\end{aligned}$$

Note como a homografia \mathcal{H} e o valor escalar w_{all} são coletados para fora do somatório. A coleta de \mathcal{H} é possível, pois a introdução da multiplicação de \mathcal{D}_{f_i} por \mathcal{K} faz com que \mathcal{H}^\top e $\mathcal{H}^{-\top}$ se cancelem.

Substituindo a Equação 5.8 na Equação 5.6, temos a prova da invariância de α :

$$\begin{aligned}
\alpha &= \frac{\det(\mathcal{P}' \cdot \mathcal{P}')}{\text{tr}(\mathcal{P}' \cdot \mathcal{P}')^3} \\
&= \frac{\det(w_{all} \mathcal{H} \cdot \mathcal{P} \cdot \mathcal{H}^{-1} \cdot w_{all} \mathcal{H} \cdot \mathcal{P} \cdot \mathcal{H}^{-1})}{\text{tr}(w_{all} \mathcal{H} \cdot \mathcal{P} \cdot \mathcal{H}^{-1} \cdot w_{all} \mathcal{H} \cdot \mathcal{P} \cdot \mathcal{H}^{-1})^3} \\
&= \frac{(w_{all})^6 \det(\mathcal{H} \cdot \mathcal{P} \cdot \mathcal{H}^{-1} \cdot \mathcal{H} \cdot \mathcal{P} \cdot \mathcal{H}^{-1})}{(w_{all})^6 \text{tr}(\mathcal{H} \cdot \mathcal{P} \cdot \mathcal{H}^{-1} \cdot \mathcal{H} \cdot \mathcal{P} \cdot \mathcal{H}^{-1})^3} \\
&= \frac{\det(\mathcal{H} \cdot \mathcal{P} \cdot \mathcal{P} \cdot \mathcal{H}^{-1})}{\text{tr}(\mathcal{H} \cdot \mathcal{P} \cdot \mathcal{P} \cdot \mathcal{H}^{-1})^3} \\
&= \frac{\det(\mathcal{P} \cdot \mathcal{P})}{\text{tr}(\mathcal{P} \cdot \mathcal{P})^3}.
\end{aligned}$$

5.5 Invariante Projetivo sobre 4-Junções Impróprias

O invariante projetivo sobre junções impróprias é, basicamente, a aplicação da taxa anarmônica entre quatro direções distintas. Essa taxa obtém como resultado um único valor λ diretamente relacionado à outros cinco, advindos da variação combinatória entre

as direções de entrada:

$$\lambda, \frac{1}{\lambda}, \frac{1}{1-\lambda}, 1-\lambda, \frac{\lambda}{\lambda-1}, \frac{\lambda-1}{\lambda}. \quad (5.9)$$

Seja a 4-junção imprópria J_4 cujos vértices delimitadores de ramificação são direções $\mathbf{d}_a, \mathbf{d}_b, \mathbf{d}_c$ e \mathbf{d}_d e seu vértice central é dado por \mathbf{o} , conforme a Figura 3.8. Temos, portanto, as quatro retas $\mathbf{r}_i = \mathbf{o} \times \mathbf{d}_i$. Definida uma reta de referência \mathbf{l} não paralela à nenhuma das retas \mathbf{r}_i , temos a aplicação do produto vetorial entre os quatro pontos de intersecção $\mathbf{r}_i \times \mathbf{l}$, exatamente como mostra

$$\{\mathbf{r}_1 \times \mathbf{l}, \mathbf{r}_2 \times \mathbf{l}; \mathbf{r}_3 \times \mathbf{l}, \mathbf{r}_4 \times \mathbf{l}\} = \frac{\|(\mathbf{r}_1 \times \mathbf{l} - \mathbf{r}_3 \times \mathbf{l})\| \|(\mathbf{r}_2 \times \mathbf{l} - \mathbf{r}_4 \times \mathbf{l})\|}{\|(\mathbf{r}_2 \times \mathbf{l} - \mathbf{r}_4 \times \mathbf{l})\| \|(\mathbf{r}_1 \times \mathbf{l} - \mathbf{r}_3 \times \mathbf{l})\|}. \quad (5.10)$$

Dentre todas as possibilidades existentes de λ (Equação 5.9), apenas um valor será negativo e contido no intervalo $[-1, 0)$. Portanto, ao identificar este resultado e ignorando o seu sinal, temos um único valor invariante projetivo para 4-junções impróprias. Em suma, por este procedimento, a ordem que tomamos os ramos de J_4 passa a ser arbitrada pelo valor de λ .

5.6 Discussão

Este capítulo propôs uma formalização matemática sobre junções contínuas como estruturas geométricas de interesse. A definição apresentada pela Equação 5.1 é independente de ruídos e incertezas de imagens nas quais são aplicados os monômios invariantes.

Os monômios invariantes foram baseados em teorias que assumem, como entrada, matrizes simétricas com determinante diferente de zero. A Equação 5.2 apresenta uma maneira de representar junções a partir de matrizes simétricas. Nesta abordagem, cada vértice \mathbf{p}_i delimitador de junção é mapeado para o espaço dual de cônicas \mathcal{D}_i . Esse mapeamento leva à seções cônicas degeneradas, *i.e.*, com determinante igual à zero. Entretanto, podem ser combinadas a partir da aplicação de um conceito denominado rede de cônicas, o que leva à matrizes simétrica não degenerada.

As Seções 5.2 à 5.4 apresentaram monômios invariantes, *i.e.*, fórmulas cujos resultados são invariantes à determinados grupos de transformações, respectivamente, transformações de similaridade, transformações afim e transformações projetivas.

A literatura apresenta outros trabalhos com aplicações em reconhecimento de padrões visuais que se baseiam na busca e no ajuste de seções cônicas diretamente sobre as silhuetas

presentes na imagem e aplicação de monômios invariantes. A proposição deste trabalho, entretanto, visa a transcrição de estruturas locais, como as junções, em valores invariantes sob diferentes grupos de transformações. Uma proposta inédita, até onde sabemos.

Capítulo 6

Descritor de Junções em Imagens

Junções extraídas diretamente de imagens digitais propagam consigo erros advindos de ruídos e rasterização. Essas serão denominadas *junções discretas*. Este capítulo tem por objetivo expor estratégias para identificar as junções contínuas que se ajustam às junções discretas, de modo que os monômios invariantes descritos no Capítulo 5 possam ser empregados na construção de descritores de características locais da imagem, na forma de histogramas.

Seja

$$\mathcal{J}_n = \{\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_n | n \in \{2, 3, 4\}\}, \quad (6.1)$$

onde \mathcal{J}_n denota uma junção discreta composta por um conjunto de n segmentos digitais \mathcal{B}_i . Cada segmento digital \mathcal{B}_i corresponde à uma cadeia de pixels aproximadamente colineares. A Figura 6.1 ilustra três junções discretas: uma 2-junção discreta em (a), uma 3-junção discreta em (b) e uma 4-junção discreta em (c), onde cada um dos seus ramos é definido por um segmento digital que converge para um pixel central.

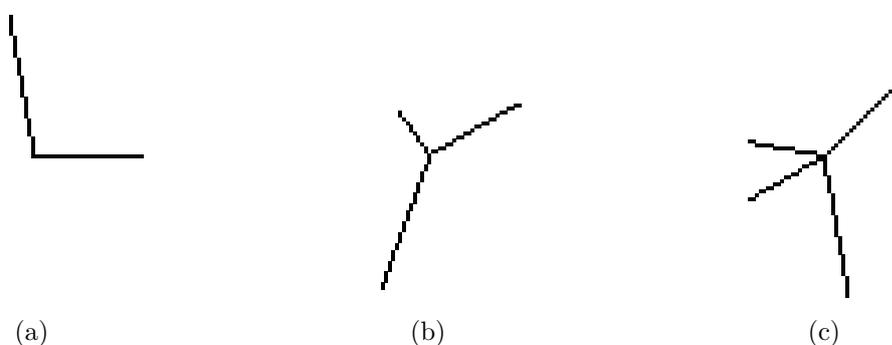


Figura 6.1: Versões discretizadas de uma 2-junção, em (a), uma 3-junção, em (b) e de uma 4-junção, em (c). Perceba que suas ramificações não são segmentos de reta, são conjuntos de pixels que formam segmentos digitais.

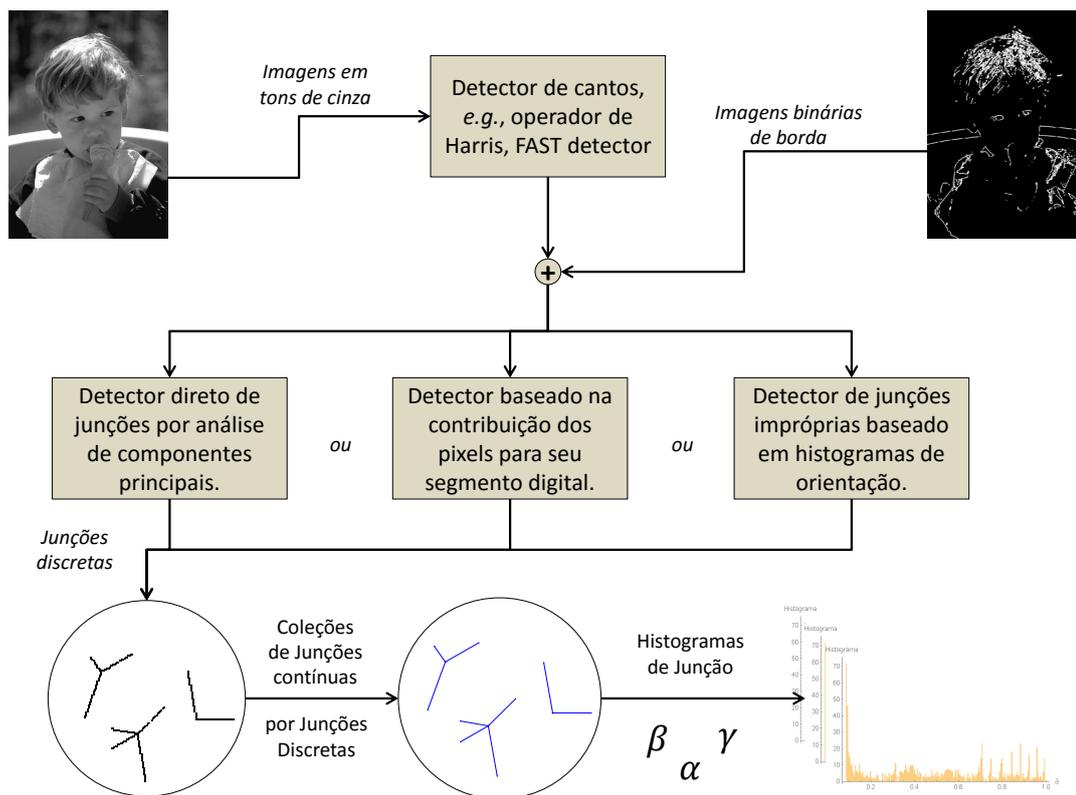


Figura 6.2: Resumo do processo de descrição de regiões de interesse. Como entrada, o algoritmo espera imagens em tons de cinza para que as localizações aproximadas dos vértices centrais de junções discretas possam ser identificadas. A partir daí, somente serão utilizadas imagens binárias de borda. Um dos três métodos descritos pelas Seções 6.1, 6.2 e 6.3 deverá ser executado para que um conjunto de junções discretas seja montado e suas versões contínuas devidamente identificadas para cada junção discreta. Por fim, para cada junção discreta, uma assinatura de junção é gerada de acordo com os monômios apresentados no Capítulo 5, calculados sobre as representações contínuas.

Junções que se encaixem perfeitamente na definição dada pela Equação 6.1 são difíceis de serem identificadas em imagens reais [42]. As junções utilizadas para calcular os valores invariantes dependem de detectores de junção com precisão elevada, um problema cuja solução está em aberto na literatura.

Este capítulo é dedicado às alternativas desenvolvidas neste trabalho para contornar as interferências e ruídos de discretização em imagens digitais. Técnicas para identificação de junções, *e.g.*, operador de Harris [28], detector FAST [50] e o detector de autovalor mínimo [62], identificam localizações aproximadas para os vértices centrais e podem ser utilizadas como ponto de partida na extração de junções discretas. Neste trabalho foi utilizado o detector de autovalor mínimo, pois seu resultado aparentou identificar mais centros de junção que os outros. O objetivo é analisar o máximo possível de junções válidas.

Todo o processo proposto para descrição de regiões de interesse em termos de suas junções segue conforme ilustrado pela Figura 6.2. Como entrada, o processo recebe uma imagem em tons de cinza, pois os detectores de vértices centrais utilizados necessitam de informações de orientação e magnitude do gradiente. Uma vez determinados os vértices centrais, o restante dos procedimentos se baseia em imagens binárias de borda.

Seja

$$B(u, v) : \mathbb{Z}^2 \rightarrow \{0, 1\}, \quad (6.2)$$

onde $B(u, v)$ denota uma imagem binária de bordas onde os valores de intensidade são restritos à 1, onde há bordas, e 0, onde não há, e o par $(u, v) \in \mathbb{Z}^2$ definem as coordenadas de um pixel da imagem.

A Seção 6.1 apresenta a extração de junções discretas para o caso em que junções possam ser perfeitamente identificadas, ou seja, os segmentos digitais que materializam as ramificações não podem ser facilmente extraídos. A Figura 6.4a ilustra uma junção extraída de uma imagem real, cujas ramificações podem ser curvas e a junção contínua que melhor se ajusta é sobreposta à primeira em 6.4b. Note que há uma quebra no segmento quando a curvatura excede um determinado limiar. As direções das ramificações são determinadas pela análise de componentes principais do conjunto de pixels que formam o segmento digital e o vértice central é ajustado pela localização que melhor descreve a intersecção das ramificações formadas.

Quando não é possível a identificação perfeita de junções, é proposta uma abordagem alternativa onde toda a subregião envolta do pixel central \mathbf{c} da junção é utilizada. Todo pixel de borda \mathbf{p} da subregião contribui, de certo modo, com o suposto segmento digital que interliga \mathbf{p} e \mathbf{c} . Este segmento digital é apenas candidato à ramificação. Ao utilizar este procedimento, podem surgir mais do que quatro ramificações candidatas. Essas, então, são combinadas, quatro à quatro, para formar uma coleção de junções possíveis. Os valores invariantes gerados para cada junção possível são acumulados em um histograma de junções da subregião de interesse. Esta abordagem está melhor descrita pela Seção 6.2.

Na Seção 6.3 é proposta uma terceira estratégia, não mais baseada em junções próprias e sim em junções impróprias. Neste caso, a subregião envolta do pixel central aproximado passa a ser a imagem como um todo. Todos os pixels de borda contribuem para com histogramas de orientação que codificam direções e acumulam a quantidade de pixels aproximadamente colineares naquela direção.

Ao final deste capítulo, a Seção 6.4 faz uma análise do domínio do monômio invari-

ante projetivo utilizando a função acumulada para realizar discretizações. A Seção 6.5 apresenta brevemente a técnica utilizada para comparar assinaturas de junção.

6.1 Identificação de Ramos Digitais Bem Definidos

Uma junção discreta bem definida é aquela em que os pixels que compõem as suas ramificações não se misturam e não são confundidos com pixels de borda de outros elementos que não a junção em questão. Este tipo de isolamento é raro de ser obtido em imagens naturais. No entanto, em informação textual, plantas baixas e imagens sintéticas, a estratégia apresentada nesta seção poderá ser utilizada diretamente para calcular descritores de junções discretas.

Identificado um pixel central \mathbf{c} na imagem binária de bordas que se aproxime do verdadeiro vértice central de uma junção, é iniciada uma varredura recursiva por cada uma de suas ramificações. Esse percorrimto tem por objetivo interligar os pixels vizinhos e aproximadamente colineares em n conjuntos que representem as ramificações. Ou seja, construir \mathcal{J}_n definido na Equação 6.1. A varredura de pixels deve parar quando a curvatura do segmento formado até o momento ultrapassar um determinado limiar. Para este trabalho, uma variante da técnica de detecção de segmentos de reta apresentado por Lowe foi desenvolvida.

A Figura 6.3 ilustra como cadeias de pixels são quebradas em segmentos de reta pela técnica original, onde dado um segmento curvilíneo completo, o algoritmo busca pelo pixel mais distante da reta suporte que interliga o primeiro e o último pixel do segmento. Se essa distância for maior que um determinado limiar, o segmento curvilíneo é dividido em dois e suas partes são analisadas recursivamente. Se a distância não for grande o suficiente, o algoritmo para. A técnica desenvolvida nesta dissertação não recebe o segmento curvilíneo completo. Durante a varredura recursiva de cada novo pixel, analisa-se se o segmento formado pode ser dividido ou não. Se puder ser dividido, então significa que o máximo segmento retilíneo foi encontrado e o algoritmo para, conforme mostra a Figura 6.4, onde uma 4-junção discreta com ramificações curvilíneas é ajustada para uma 4-junção contínua com um vértice central, em amarelo, e vértices delimitadores, em vermelho.

Para cada $\mathcal{B}_i \in \mathcal{J}_n$, é realizada uma análise dos componentes principais para determinar a direção preferencial que se ajusta aos pixels de \mathcal{B}_i . O vértice delimitador \mathbf{d}_i que

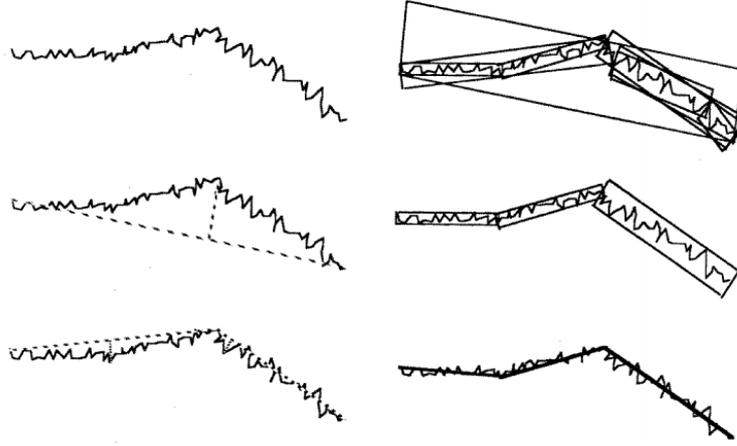


Figura 6.3: Procedimento para quebra de cadeias de pixels em segmentos retilíneos da técnica de Lowe [38]. Na imagem superior esquerda, uma cadeia qualquer extraída, onde logo abaixo é identificado o pixel mais distante da reta suporte formada pelo primeiro e último vértices da cadeia corrente e neste ponto ocorre uma quebra de continuidade, conforme mostra a imagem inferior esquerda. O mesmo ocorre na coluna direita da figura recursivamente para todas as partes quebradas até que, conforme mostra a imagem inferior direita, a cadeia seja convertida na união de segmentos retilíneos.

melhor se ajusta à ramificação \mathcal{B}_i é dado por

$$\mathbf{d}_i = \mathbf{m}_i \pm 3\sqrt{u_{i+}}\mathbf{v}_{i+}, \quad (6.3)$$

onde \mathbf{m}_i denota o ponto médio da ramificação \mathcal{B}_i , u_{i+} representa o maior autovalor de \mathcal{B}_i e \mathbf{v}_{i+} representa o autovetor associado ao autovalor u_{i+} .

O ponto médio da ramificação é dado por

$$\mathbf{m}_i = \frac{1}{n_i} \sum_{\mathbf{a} \in \mathcal{B}_i} \mathbf{a}, \quad (6.4)$$

onde n_i indica o total de pixels pertencentes à \mathcal{B}_i . Para determinar os autovalores e autovetores de \mathcal{B}_i , é preciso definir a matriz de covariância experimental dos pixels de \mathcal{B}_i ,

$$\mathcal{M} = \frac{1}{n-1} \begin{pmatrix} \sum_{\mathbf{a} \in \mathcal{B}_i} \bar{\mathbf{a}}_x^2 & \sum_{\mathbf{a} \in \mathcal{B}_i} \bar{\mathbf{a}}_x \bar{\mathbf{a}}_y \\ \sum_{\mathbf{a} \in \mathcal{B}_i} \bar{\mathbf{a}}_x \bar{\mathbf{a}}_y & \sum_{\mathbf{a} \in \mathcal{B}_i} \bar{\mathbf{a}}_y^2 \end{pmatrix}, \quad (6.5)$$

onde m indica o total de pixels pertencentes à \mathcal{B}_i e $\bar{\mathbf{a}} = \mathbf{a} - \mathbf{m}_i$.

Seja

$$\det(\mathcal{M} - k\mathcal{I}) = 0, \quad (6.6)$$

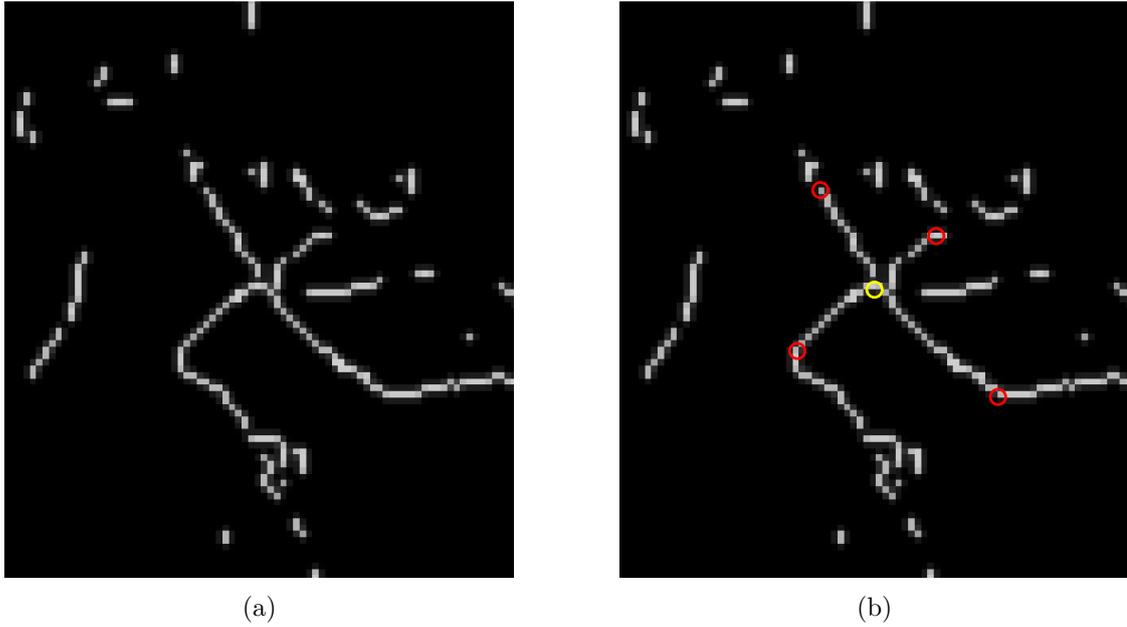


Figura 6.4: Exemplo de conversão de junção discreta em junção contínua a partir da análise de componentes principais sobre os conjuntos de pixels que formam as ramificações. Perceba que, quando o segmento torna-se curvo demais no momento da varredura, o algoritmo pára a varredura de novos pixels.

onde k é um coeficiente linear e \mathcal{I} é a matriz identidade. A Equação 6.6, quando utilizados pontos pertencentes à \mathbb{R}^2 , formam uma equação do segundo grau em função de k cujas soluções representam os autovalores associados à \mathcal{B}_i . Portanto,

$$u_{i\pm} = \frac{1}{2} \left(\mathcal{M}_{11} + \mathcal{M}_{22} \pm \sqrt{(\mathcal{M}_{11} + \mathcal{M}_{22})^2 + 4\mathcal{M}_{12}^2} \right), \quad (6.7)$$

onde u_{i+} denota o maior autovalor e u_{i-} denota o menor autovalor associados à \mathcal{B}_i . Os autovetores associados são definidos por

$$\mathbf{v}_{i+} = \begin{pmatrix} \mathcal{M}_{11} + \mathcal{M}_{12} - u_{i-} \\ \mathcal{M}_{22} + \mathcal{M}_{12} - u_{i-} \\ 0 \end{pmatrix} \quad (6.8)$$

O vértice central, inclusive, não está imune aos problemas de discretização e não deve ser confundido com o pixel \mathbf{c} . A localização do vértice central é o ponto que melhor se ajusta à intersecção de todas as restas suportes \mathbf{r}_i ajustadas para \mathcal{B}_i .

Seja

$$\mathbf{r}_i = \mathbf{m}_i \times \mathbf{d}_i, \quad (6.9)$$

a reta que passa pelo ponto \mathbf{m}_i e pelo ponto delimitador \mathbf{d}_i . Para cada $\mathcal{B}_i \in \mathcal{J}$, portanto, é ajustada uma reta \mathbf{r}_i . Os coeficientes de cada reta são dispostos por coluna para formar

a matriz \mathcal{N} utilizada para a decomposição em valores singulares da matriz resultante $\mathcal{N} \cdot \mathcal{N}^\top$. O autovetor associado ao menor autovalor é apontado como a solução que minimiza a distância algébrica. Este será o vértice central \mathbf{c} adotado na junção contínua.

A abordagem apresentada nesta seção faz um mapeamento $\mathcal{J}_n \rightarrow \mathcal{J}_n$ e a aplicação dos monômios invariantes ocorre uma única vez sobre \mathcal{J}_n , o que gera um único valor invariante.

6.2 Identificação de Ramos Próprios por Contribuições em Segmentos Digitais

O conceito de junção estabelece que são necessários, pelo menos, um vértice central e duas à quatro ramificações, sejam estas delimitadas por pontos próprios ou impróprios (Equação 5.1). Em teoria, uma ramificação digital pode ser composta inteiramente por um único ponto, distinto do vértice central, e sempre será possível traçar um segmento de reta entre o vértice central \mathbf{c} e qualquer pixel \mathbf{p} . Por isso, é possível definir que o pixel \mathbf{p} selecionado contribui explicitamente para que o segmento digital entre \mathbf{p} e \mathbf{c} seja considerado uma ramificação.

Especificado um pixel central \mathbf{c} , uma subregião quadrada da imagem binária de bordas B_{reg} é traçada no seu entorno para limitar o tamanho possível das ramificações. Qualquer pixel $\mathbf{p} \in B_{reg}$ pertence à um segmento candidato à ramificação. A contribuição de cada \mathbf{p} é realizada em duas etapas: determinação da sua orientação relativa à \mathbf{c} e, em seguida, determinação de sua distância em relação à \mathbf{c} . A discretização Δ utilizada para determinar direções possíveis está relacionada com o tamanho h do lado especificado para a subregião B_{reg} , sendo, portanto, $\Delta = \frac{\pi}{(2h+1)}$.

Essencialmente, são criados $m = (2h - 2)$ histogramas de distância, onde m é o total de compartimentos de uma discretização realizada sobre o intervalo de orientação $[0, 2\pi)$ e h denota a altura definida para a região B_{reg} . Ao identificar qual histograma \mathcal{H} de distância o pixel \mathbf{p} pertence, a contribuição ocorre de maneira acumulada em todos os compartimentos de \mathcal{H} até o compartimento que representa a distância entre \mathbf{c} e \mathbf{p} . A Figura 6.5 ilustra todo o processo de contribuição de um novo pixel $\mathbf{p} \in B_{reg}$ que se encontra à uma distância d e um ângulo θ ao eixo de referência \mathbf{ref} . O ângulo θ identifica qual histograma à ser considerado e a distância d permite que a contribuição seja acumulada para todos os pixels do segmento entre \mathbf{c} e \mathbf{p} .

Quando todos os pixels em B_{reg} tiverem contribuído para com o segmento ao qual pertencem, estará codificado um mapa de segmentos candidatos à ramificação relativos

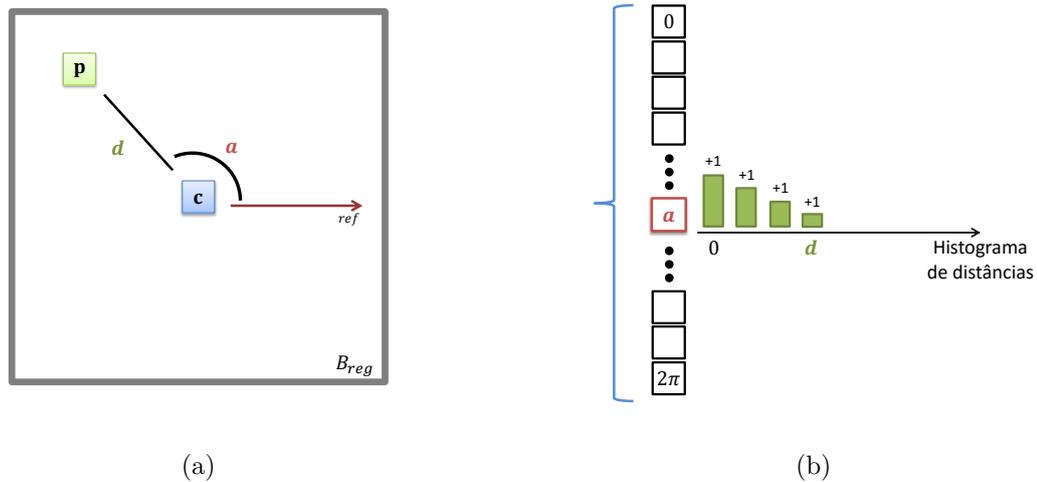


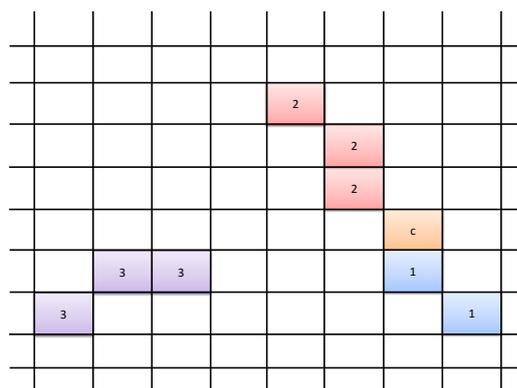
Figura 6.5: Contribuição de cada pixel \mathbf{p} para com o segmento ao qual pertence relativo ao pixel \mathbf{c} . Para cada pixel \mathbf{p} , identifica-se sua angulação θ e distância d em relação ao eixo de referência \mathbf{ref} , ilustrado em (a). Em (b), identifica-se em vermelho o histograma que receberá a contribuição de acordo com a angulação θ . A distância d permitirá que a contribuição seja acumulada para todo o segmento existente entre \mathbf{p} e \mathbf{c} .

ao pixel central \mathbf{c} . Entretanto, ainda é preciso identificar os segmentos que, por estarem muito afastados de \mathbf{c} , são descartados por não representarem ramificações de junção. Para determinar os segmentos que estão muito afastados, é possível avaliar cada histograma de distância criado.

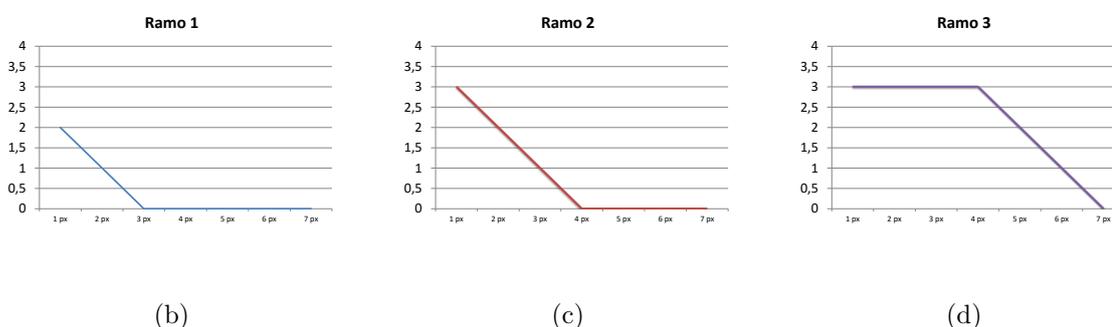
A contribuição acumulada de distâncias permite que sejam determinadas quebras na continuidade dos segmentos em uma dada direção θ . Quando a quebra de continuidade ocorre muito próximo ao pixel \mathbf{c} , significa que o segmento em questão não representa uma ramificação válida e, portanto, é descartado. Se a quebra de continuidade ocorrer à uma distância mínima determinada, tal distância representa a delimitação máxima do ramo, tornando-se o vértice delimitador da junção contínua. Para este trabalho, foi adotada uma distância máxima de 1, *i.e.*, uma tolerância de segmentos afastados em até um pixel a partir do pixel \mathbf{c} . A Figura 6.6 ilustra como os histogramas são interpretados visualmente na imagem binária de bordas e a identificação das ramificações válidas.

6.3 Identificação de Ramos Impróprios por Contribuições em Histogramas de Orientação

Para identificar junções impróprias, o procedimento é semelhante ao descrito na Seção 6.2, com a diferença de que a distância dos pixels \mathbf{p} ao pixel \mathbf{c} não serão necessárias. A



(a) Configuração específica de pixels



(b)

(c)

(d)

Figura 6.6: Discriminação dos segmentos digitais a partir do mapa de distâncias: (a) ilustra uma configuração exemplo de pixels onde existem três segmentos digitais, *i.e.*, possíveis ramificações. (b), (c) e (d) demonstram os mapas de distância para cada segmento. A rápida queda das contagens nos mapas (b) e (c) indicam que são segmentos muito próximo à \mathbf{c} e, portanto, ramificações. Quando o segmento se encontra à distância, a contribuição se caracteriza como ilustrado em (d), com um valor constante por alguns pixels de distância.

orientação de \mathbf{p} relativo ao \mathbf{c} , no entanto, será utilizada para montar um histograma de orientações. Outra diferença está relacionada à subregião \mathcal{B}_{reg} utilizada, que agora considerando a região total da imagem binária de bordas.

Para todo pixel \mathbf{c} identificado, será criado um histograma de tamanho m , de acordo com a discretização do intervalo de ângulações $[0, 2\pi)$. Na Seção 6.2, para cada compartimento, era traçado um histograma de distância relativa ao pixel \mathbf{p} . Para esta abordagem, os compartimentos angulares serão o próprio histograma final de orientações. A Figura 6.7 ilustra como ocorre o processo de votação do pixel \mathbf{p} no mapa de orientações relativas ao pixel central \mathbf{c} : para cada pixel \mathbf{p} , identifica-se sua angulação θ em relação ao eixo de referência \mathbf{ref} e, em vermelho, o compartimento que receberá a contribuição é determinado pelo ângulo θ .

Efetivamente, esse mapa de orientações acumula a quantidade de pixels colineares em

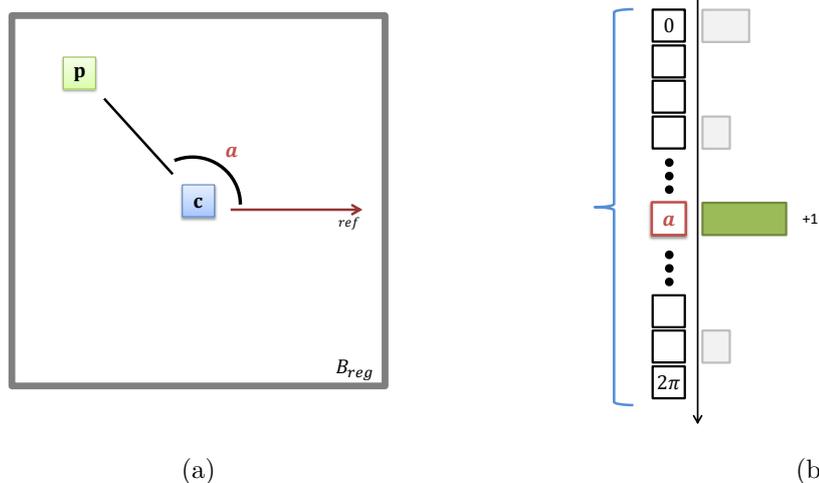


Figura 6.7: Contribuição de cada pixel \mathbf{p} para o mapa de orientações relativas ao eixo de referência \mathbf{ref} . Para cada pixel \mathbf{p} , identifica-se sua angulação θ em relação ao eixo de referência \mathbf{ref} , ilustrado em (a). Na Imagem (b), identifica-se em vermelho o compartimento que receberá a contribuição de acordo com a angulação θ . A distância não é uma grandeza útil neste processo. No histograma de orientações apresentado em (b), as barras em cinza são apenas para indicar que esse histograma já pode ter sido atualizado anteriormente por outros pixels, apenas a barra verde denota a contribuição do pixel \mathbf{p} .

determinada direção a partir do pixel \mathbf{c} . Após a montagem do histograma de orientações, uma estratégia para detecção de máximos locais é aplicada para selecionar as direções preferenciais. A Figura 6.8 ilustra, dado um pixel de referência, como as direções são selecionadas ao final do procedimento. Para cada pixel, em preto, contribui para o histograma de orientações, numa angulação θ relativa ao eixo de referência. Perceba que, pixels colineares contribuem para a mesma orientação, o que provê pesos diferentes para cada direção. Na etapa posterior de cálculo do invariante, ao combinarmos, quatro a quatro, as ramificações, direções com pesos menores podem ser filtradas, por questões de eficiência. O tamanho das setas indica a quantidade de pixels colineares naquela determinada direção.

6.4 Assinatura de Subregiões

A depender da estratégia utilizada para decompor subregiões em coleções de junções contínuas, eventualmente, uma série de valores invariantes pode ser calculada para uma única subregião da imagem. Todos esses valores são armazenados em histogramas cujo domínio varia de acordo com o monômio invariante desejado. À rigor, a única abordagem que não gera histogramas é a apresentada na Seção 6.1. No entanto, para considerar a incerteza

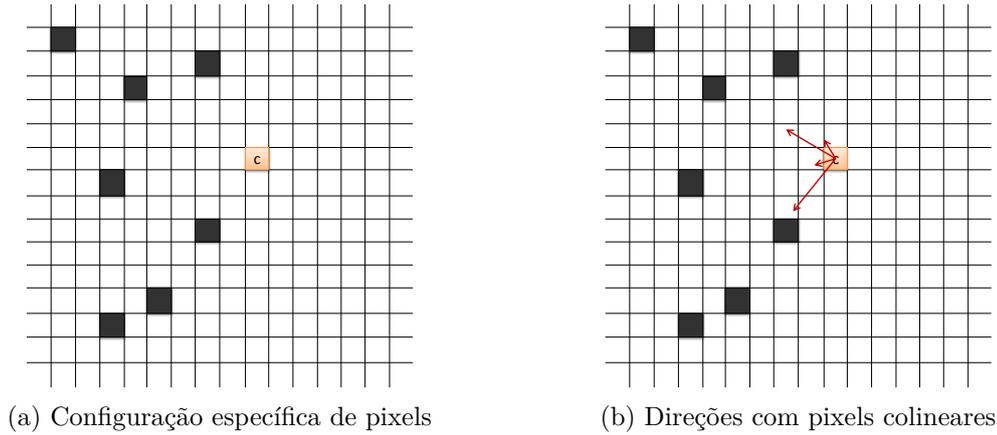


Figura 6.8: Identificação de junções impróprias: (a) ilustra uma configuração específica de pixels numa imagem e (b) demonstra as direções preferenciais dos pixels. O tamanho da seta é proporcional à quantidade de pixels naquela direção.

da localização dos pixels delimitadores, amostras podem ser geradas considerando seus vértices como variáveis aleatórias sob distribuição uniforme na região 1×1 envolta localização identificada. Amostras essas que geram valores invariantes e contribuem para um histograma.

O intervalo de valores assumidos pelos invariantes varia de acordo com o domínio do monômio invariante aplicado. O tamanho da assinatura de junção depende da discretização realizada sobre este domínio. A geração de m amostras aleatórias de junção é importante para uma análise do domínio do monômio aplicado e, quando os valores invariantes são dispostos num histograma, observa-se que a distribuição de valores possíveis não é uniforme ao longo do domínio. Especificamente para o monômio invariante projetivo apresentado pela Equação 5.6, considerando $m = 100000$ amostras aleatórias, a Figura 6.10a apresenta a configuração do histograma que acumula os m invariantes gerados. Perceba como o domínio não é uniforme nos extremos do intervalo.

A análise da função acumulada é importante para a equalização da distribuição total. A função acumulada equalizada das amostras varia no intervalo entre $[0, 1]$ e a discretização neste domínio é mais fiel à configuração uniforme de amostras. Desta forma, quando um invariante for calculado, este deverá ser equalizado para o seu valor respectivo na função acumulada, cuja fórmula analítica pode ser derivada.

Seja

$$\hat{\alpha} = \frac{\det(\mathcal{P} \cdot \mathcal{P})}{\text{tr}(\mathcal{P} \cdot \mathcal{P})^3} = \frac{u_1^2 u_2^2 u_3^2}{(u_1^2 + u_2^2 + u_3^2)^3}, \quad (6.10)$$

o monômio invariante projetivo da Equação 5.6 reescrito em função dos autovalores u_1 , u_2 , u_3 da matriz $\mathcal{P} \cdot \mathcal{P}$. Considerando que os autovalores serão normalizados pelo maior

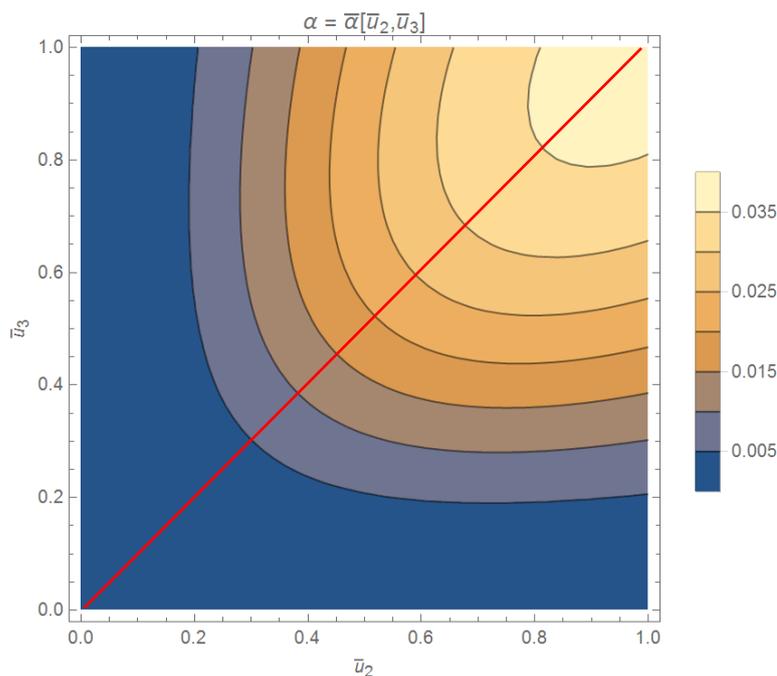


Figura 6.9: *Plot* do monômio invariante projetivo $\bar{\alpha}$ (Equação 6.11) ao variar \bar{u}_2 e \bar{u}_3 . A linha diagonal em vermelho representa os valores invariantes quando $\bar{u}_2 = \bar{u}_3$, explicitando a simetria da função.

dentre eles, digamos que seja u_1 ,

$$\bar{\alpha} = \frac{\bar{u}_2^2 \bar{u}_3^2}{(1 + \bar{u}_2^2 + \bar{u}_3^2)^3}, \quad (6.11)$$

onde $\bar{u}_i = \frac{u_i}{u_1}$ e portanto $-1 \leq \bar{u}_i \leq 1$. Os limites do monômio α estão definidos no intervalo entre $[0, \frac{1}{27})$. A Figura 6.9 apresenta um *plot* com todos os valores $\bar{\alpha}$ possíveis. A função $\bar{\alpha}$ é claramente crescente e simétrica quando $\bar{u}_2 = \bar{u}_3$, conforme ilustra a diagonal, traçada em vermelho.

A função acumulada da função $\bar{\alpha}$ para um valor invariante v específico é dado pela área em vermelho do *plot* da Figura 6.11a que representa o complementar da área em vermelho do *plot* da Figura 6.11b, que pode ser obtida de forma numérica a partir do cálculo aproximado da área da região, conforme apresentado no *script Mathematica*[©] presente no Apêndice A.

Por fim, com o auxílio da ferramenta *Mathematica*[©], conforme apresentado pelo *script* no Apêndice A, foi possível determinar uma fórmula analítica para a função acumulada, dada por:

$$\text{cumm}(\alpha) = 1 - 2 \int_{\bar{u}_2}^1 (\bar{u} - \hat{u}(\bar{u}, \alpha)) d\bar{u}, \quad (6.12)$$

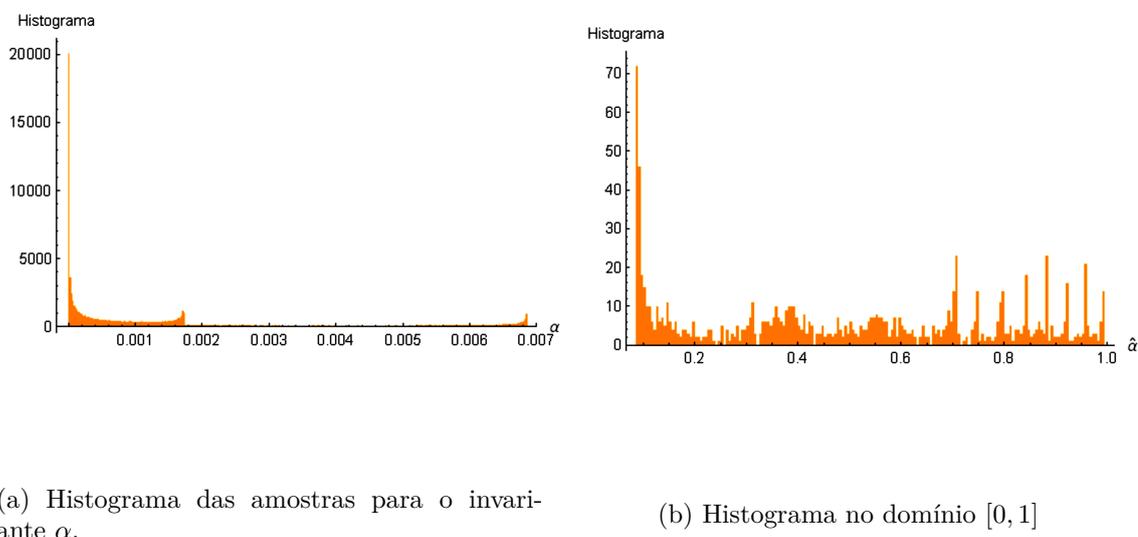


Figura 6.10: Processo de amostragem e aproximação da função acumulada para discretização uniforme sobre o intervalo $[0, 1]$. O gráfico (a) ilustra um histograma dos invariantes projetivos calculados sobre amostras de 100000 junções aleatórias. O gráfico (b) demonstra como a distribuição dos invariantes ficou mais uniforme sobre o intervalo $[0, 1]$ após a equalização da distribuição.

onde a integral denota a região em vermelho apresentada pela Figura 6.11b e $\hat{u}(\bar{u}, \alpha)$ representa uma função analítica pré-calculada, também disponível no Apêndice A, pois a equação é muito extensa. A Figura 6.12 demonstra, em vermelho, o domínio de $\hat{u}(\bar{u}, \alpha)$.

A integral apresentada pela Equação 6.12 é solucionada analiticamente à partir do cálculo aproximado da área da região em vermelho da Figura 6.11b cujo algoritmo *Recursive* está disponível no Apêndice A.

A Figura 6.10 ilustra um histograma e a frequência com que os invariantes são criados, todos no intervalo $[0, \frac{1}{27})$, por construção. A discretização neste domínio não é uniforme em todo o intervalo e pode ser mais fiel à configuração original ao analisar a função acumulada das amostras que, por definição, varia no intervalo $[0, 1]$. A discretização da assinatura de junção ocorre, portanto, no intervalo $[0, 1]$ e a função que mapeia, aproximadamente, o invariante calculado ao seu valor acumulado é obtida através da equalização pela função acumulada.

A mesma análise da função acumulada foi realizada para o monômio invariante à transformações de similaridades, conforme mostra a Figura 6.13. Para o monômio invariante afim, a mesma análise pode ser realizada para cada um dos valores pertencentes à trinca, no entanto, deixando de se comportar como coordenadas baricêntricas.

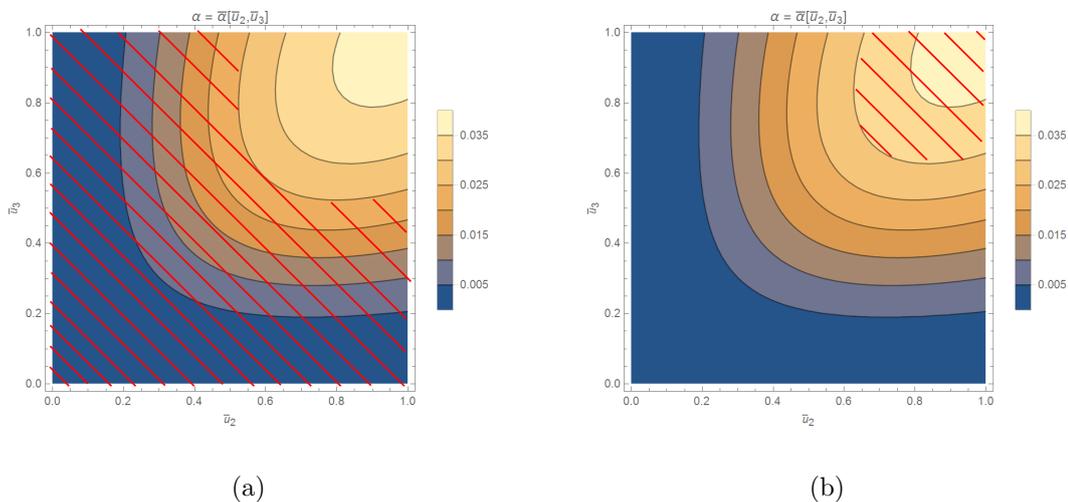


Figura 6.11: Região do *plot* da função do monômio invariante projetivo que representa a função acumulada de $\bar{\alpha}$ até determinado valor v . Em (a), a região em vermelho indica $\bar{\alpha} \leq v$. Em (b), a região complementar cuja área é calculada recursivamente, conforme *script Mathematica*[®] presente no Apêndice A.

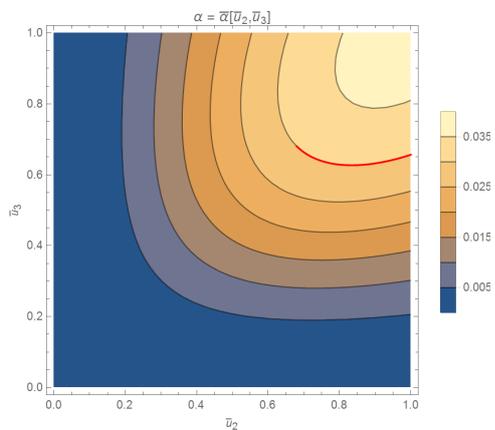


Figura 6.12: *plot* do monômio invariante projetivo e da função \hat{u} , em vermelho.

6.5 Comparação de Características

As características interessantes propostas se materializam em formato de histogramas de invariantes. Neste trabalho, foi utilizada na comparação de histogramas a medida de distância entre histogramas *Earth Mover's Distance* [56], cujo objetivo é indicar o custo necessário para transformar um histograma \mathcal{A} em outro \mathcal{B} .

Seja

$$\text{EMD}(\mathcal{A}, \mathcal{B}, \mathcal{F}) = \sum_{i=1}^m \sum_{j=1}^n f_{i,j} d_{i,j} \quad (6.13)$$

a função que codifica a distância entre os histogramas \mathcal{A} e \mathcal{B} , onde \mathcal{F} indica a função de fluxo, $d_{i,j}$ indica a *ground distance* entre os histogramas \mathcal{A} e \mathcal{B} e m, n representam, respectivamente, as dimensionalidade dos histogramas \mathcal{A} e \mathcal{B} . A implementação realizada assume que $m = n$, *i.e.*, as características possuem o mesmo número de elementos e as funções que representam \mathcal{A} e \mathcal{B} são discretas e unidimensionais. Desta forma, o cálculo da distância entre dois histogramas unidimensionais discretas \mathcal{A} e \mathcal{B} é dada pelo resultado da função $\text{EMDd}(\mathcal{A}, \mathcal{B})$, apresentada pelo Algoritmo 6.1.

O algoritmo de busca utilizado determina os k vizinhos mais próximos [22], onde $k = 3$ possíveis pareamentos. Além de calcular a medida de distância definida pela Equação 6.13, também foi utilizada a média de intensidades da vizinhança do vértice estabelecido como central. Quando houver mais de uma associação entre características cujas medidas de distância estiverem empatadas, com margem de erro em 0.1, a média de intensidades que mais se aproxima é utilizada como critério de desempate.

Algoritmo 6.1: Earth Mover's Distance

```

1 Função  $\text{EMDd}(\mathcal{A}, \mathcal{B})$ 
2    $emd \leftarrow 0;$ 
3   para todo  $i = 1, 2, \dots, m$  faça
4      $emd \leftarrow emd + |(\mathcal{A}_i + emd) - \mathcal{B}_i|;$ 
5   fim
6   retorna  $emd;$ 
7 fim

```

6.6 Discussão

Neste capítulo foram definidas e devidamente discriminadas as junções discretas, composta por conjuntos de pixels. Três são as estratégias propostas neste trabalho para

converter junções discretas ou subregiões em junções contínuas. A primeira fazendo uma análise de componentes principais para cada ramo, inferindo seus vértices delimitadores e, por fim, seu vértice central melhor ajustado. A segunda, ainda relacionada com junções próprias, traça um mapa de distâncias de todos os pixels da imagem, concluindo posteriormente quais os que efetivamente compõem uma ramificação. A terceira, para junções impróprias, inicializa um histograma circular no entorno de um vértice determinado como central e, para todos os pixels de uma região de interesse, verifica sua orientação e contribui para a montagem do histograma onde, por fim, um detector de máximos seleciona as orientações onde se encontram mais pontos colineares.

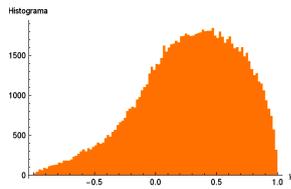
As três diferentes alternativas para conversão de junções discretas em contínuas, a partir de suas propriedades, possuem vantagens sobre as demais em certos tipos de imagem. A estratégia apresentada na Seção 6.1 é melhor aplicada sobre imagens sintéticas e bem comportadas, como em letras datilografadas e plantas baixa. A segunda alternativa, apresentada pela Seção 6.2 de decomposição de subregiões da imagem em termos de suas junções se aplica melhor em imagens naturais, onde as junções não são facilmente discriminadas. Por fim, a abordagem utilizada para extrair junções impróprias pode ser melhor aproveitada em imagens com baixa resolução onde nas quais haja a possibilidade de haver falhas na continuidade das bordas.

As abordagens propostas para interpretar junções não garantem que para cada junção haverá somente um valor invariante. Devido à possibilidade de múltiplas ramificações serem criadas durante o processo, é necessária a combinação, quatro à quatro, entre os ramos, formando múltiplas junções e, conseqüentemente, múltiplos invariantes. A composição dos invariantes pode ser um impeditivo à correta discriminância das junções, tornando o armazenamento desses valores num histograma um método mais conservador para descrição de uma região de interesse.

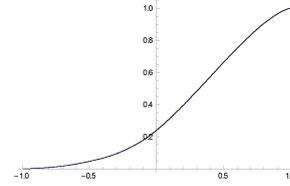
O intervalo de valores assumidos pelos invariantes varia de acordo com o domínio do monômio invariante aplicado. A alternativa proposta para equalizar as distribuições faz uso da função acumulada equalizada das amostras. A função acumulada varia no intervalo $[0, 1]$ e a discretização uniforme neste domínio é

Os histogramas criados devem variar seu domínio de acordo com cada monômio invariante. Isso implica na justa discretização do domínio, dando ênfase para os intervalos mais frequentes. A alternativa proposta parte da amostragem de junções e seus valores invariantes, revelando seu domínio e permitindo uma análise da sua função acumulada.

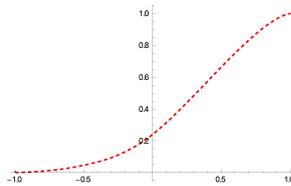
A função acumulada normalizada possui características importantes como o seu contra-



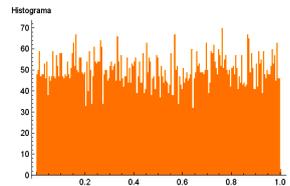
(a) Histograma das amostras para o invariante sob similaridade.



(b) Função acumulada do histograma (a).



(c) Função acumulada ajustada com o auxílio do *Mathematica*®.



(d) Histograma dos invariantes no domínio $[0, 1]$.

Figura 6.13: Processo de amostragem e aproximação da função acumulada para discretização uniforme sobre o intervalo $[0, 1]$. O gráfico (a) ilustra um histograma das de invariantes sob similaridade calculados sobre amostras de 100000 junções aleatórias. O gráfico (b) denota sua função acumulada e o gráfico (c), uma aproximação polinomial da função acumulada. O gráfico (d) demonstra como a distribuição dos invariantes ficou mais uniforme sobre o intervalo $[0, 1]$.

domínio variando entre $[0, 1]$ e sua configuração que reflete as variações de frequência nas amostras. Com o auxílio da ferramenta *Mathematica*®, foi possível determinar a função acumulada e determinar equações para mapear corretamente os invariantes calculados para o intervalo da acumulada. A discretização do domínio, por conseguinte, é realizada sobre $[0, 1]$.

Capítulo 7

Experimentos e Resultados

Este capítulo apresenta os experimentos e uma análise dos resultados obtidos. Os testes realizados têm por objetivo atestar a robustez dos valores invariantes diante de ruídos ou possíveis incertezas comuns em imagens digitais e verificar o comportamento dos descritores propostos em imagens naturais, inicialmente com transformações induzidas artificialmente e, posteriormente, diante de fotografias de cenas reais capturadas sob pontos de vista diferentes.

Na Seção 7.1 são definidos os experimentos sobre junções contínuas geradas artificialmente para demonstrar o comportamento dos valores invariantes diante de possíveis incertezas nas delimitações das junções e cuja implementação foi realizada inteiramente através do *Mathematica*[®]. A Seção 7.2 apresenta os resultados obtidos diante de imagens reais confrontadas com si mesmas transformadas artificialmente com o objetivo de mostrar a aplicabilidade dos descritores propostos em situações reais. Por fim, a Seção 7.3 apresenta resultados sobre fotografias de cenas reais capturadas sobre diferentes pontos de vista. Para os experimentos com imagens naturais, a técnica foi desenvolvida em *MATLAB*[®] e codificação em *C++*.

7.1 Análise de Robustez dos Invariantes

Os monômios formulados no Capítulo 5 são invariantes em condições exatas. Logo, podem sofrer variações por causa de ruídos ou incertezas na rasterização das imagens. Se há pouca ou nenhuma variação no resultado final, os monômios são classificados como robustos. Queremos avaliar um quão robusto pode ser um descritor construído com os invariantes propostos no Capítulo 5, conforme os processos de construção de descritores apresentados no Capítulo 6. Para realizar estes experimentos, foi preciso uma ferramenta para gerar e

manipular junções estatisticamente, baseadas na definição formal de junção contínua da Equação 5.1.

Gerar junções contínuas exige a definição de um vértice central e seus n vértices delimitadores, suficientes para calcular os valores invariantes. Os monômios são provadamente invariantes sobre junções contínuas e para analisar sua robustez é preciso induzir alguma incerteza no processo de criação da junção. Ao considerar cada vértice da junção, inclusive o central, como variável aleatória distribuída uniformemente em toda a região quadrangular associada ao pixel de referência, é possível gerar m amostras da mesma junção induzindo um possível erro de rasterização.

Para m amostras, são gerados m valores invariantes que contribuem para a assinatura da junção. A Figura 7.1a apresenta uma 4-junção utilizada como referência, onde sua assinatura indica que os $m = 100000$ valores são dispostos numa mesma região do histograma, conforme apresentado na Figura 7.1b. A Figura 7.1c apresenta uma 4-junção que diferencia-se da Figura 7.1a apenas pela rotação de um dos seus ramos, em vermelho, que por consequência, gera uma assinatura distinta da apresentada na Figura 7.1d. A Figura 7.1e apresenta uma 4-junção que possui um mapeamento projetivo com a 4-junção apresentada na Figura 7.1a, a região de acúmulo dos valores invariantes do histograma na Figura 7.1f é similar ao apresentado por na Figura 7.1b. Todos os valores calculados tomaram como base o monômio invariante projetivo da Equação 5.6.

No processo de captura de uma cena por uma câmera digital, a cena real passa por uma etapa de discretização. Os efeitos da discretização em imagens digitais podem ser simulados ao considerar uma etapa de discretização nos experimentos com a aplicação do algoritmo de Bresenham para rasterização das amostras. A escolha dos vértices delimitadores é realizada através da análise de componentes principais, conforme descreve a Seção 6.1.

Os monômios invariantes sob transformações afim, definido pela Equação 5.5, geram um par de invariantes para cada 3-junção. Neste caso, as assinaturas geradas serão construídas sobre histogramas bidimensionais. Para verificar a invariância sobre transformações afim, foi utilizado o mesmo procedimento para simulação de pontos de vista apresentado pela técnica ASIFT [67]. A Figura 7.2a apresenta uma 3-junção de referência submetida à transformações de ponto de vista, rasterizadas novamente pelo algoritmo de Bresenham e calculados seus pares de invariantes. A Figura 7.2b apresenta os valores calculados, eixo horizontal, para cada nova amostra, eixo vertical. Note que os valores invariantes se mantém estáveis mesmo com as modificações de pontos de vista.

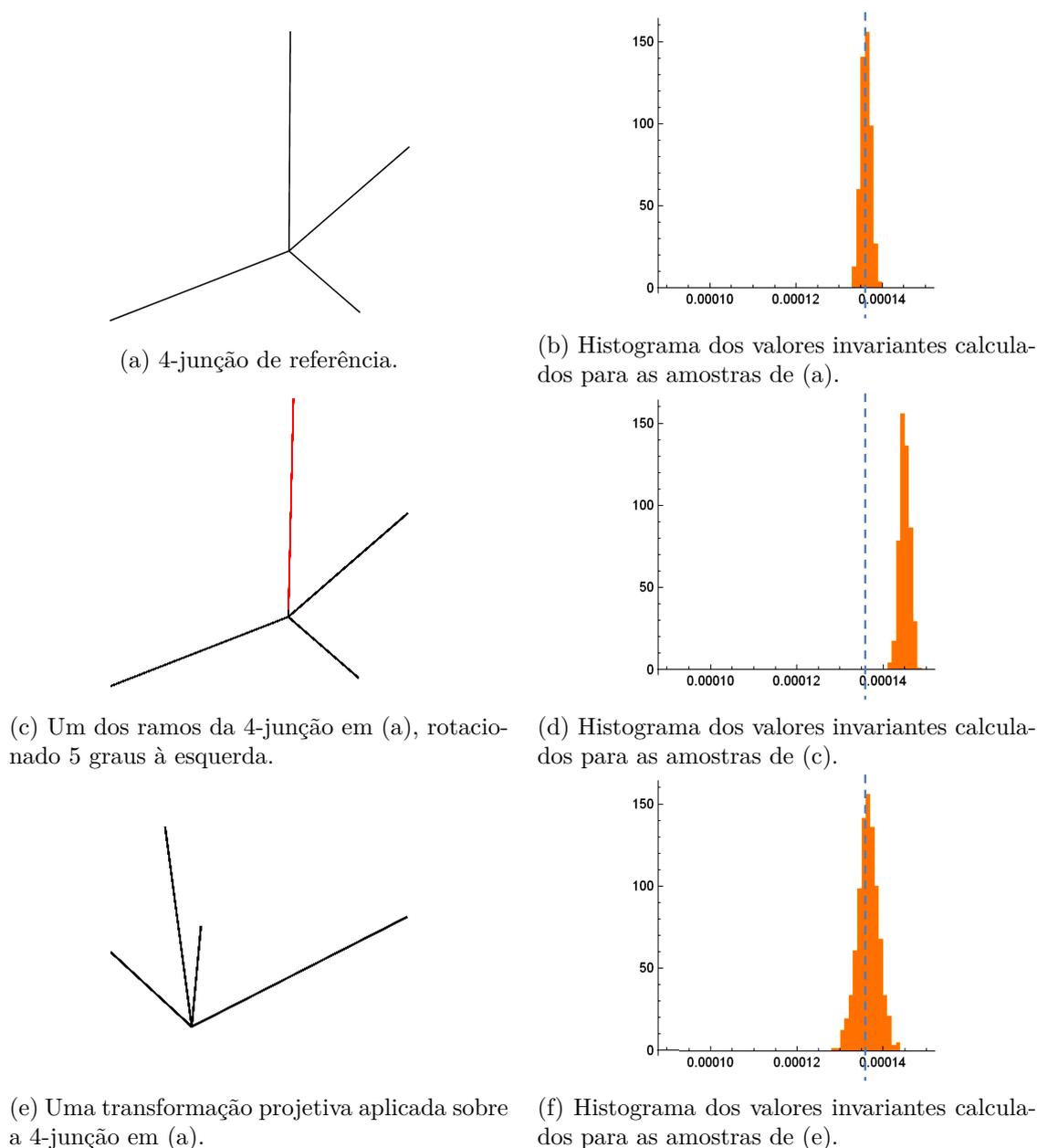


Figura 7.1: Experimento para ilustrar a robustez dos valores invariantes. Em (a), uma 4-junção tomada como referência cujas $m = 100000$ amostras geradas são utilizadas para calcular m valores invariantes, que contribuem para sua assinatura de junção em (b). Em (c), uma das ramificações, em vermelho, da 4-junção em (a) é rotacionada 5 graus à esquerda, o que faz com que os m valores invariantes de suas amostras contribuam numa região diferente do histograma, conforme apresentado em (d). Em (e), uma transformação projetiva foi aplicada sobre os vértices da 4-junção em (a), percebe-se que os m valores invariantes de suas amostras contribuam numa região semelhante ao histograma (b), conforme apresentado em (f). A linha tracejada, em azul, serve apenas como referência para facilitar a percepção de que as regiões se diferenciam.

Analogamente, o mesmo procedimento é executado para a 3-junção e os resultados são apresentados pelas Figuras 7.2c e 7.2d.

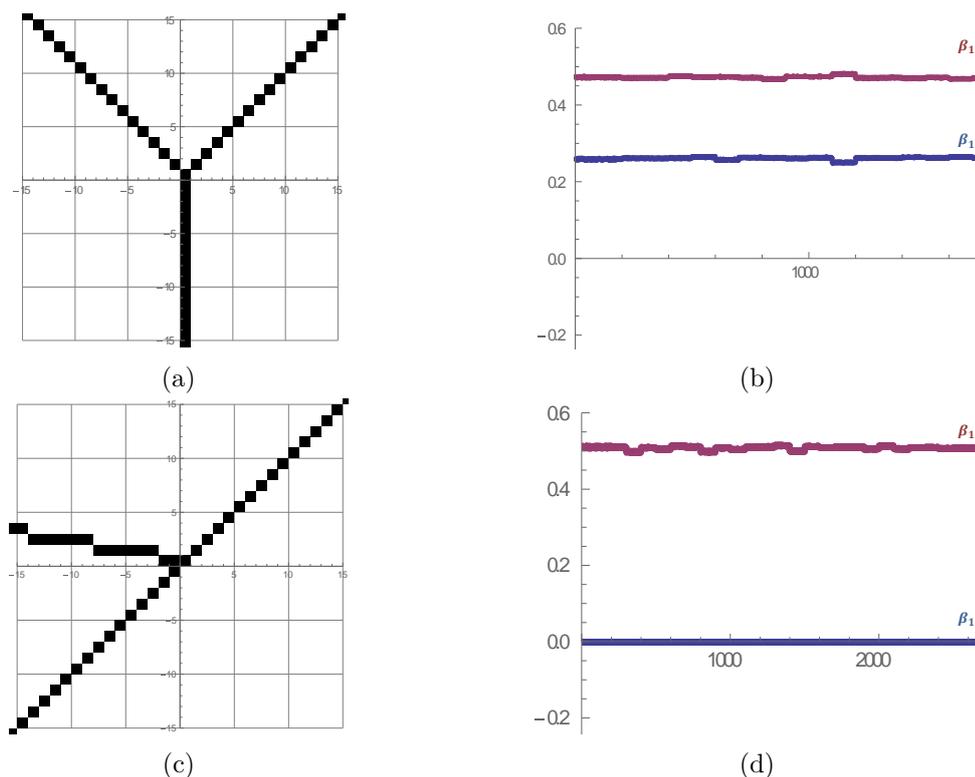


Figura 7.2: Experimento para atestar a robustez dos valores invariantes sob transformações afim. Em (a) é apresentada uma Y-junção de referência submetida à transformações de pontos de vista e em (b), para cada nova amostra, eixo x , pares de invariantes são calculados, eixo y . Perceba como os valores calculados se mantêm estáveis mesmo com as modificações de pontos de vista. As imagens (c) e (d) ilustra a execução da mesma estratégia de amostragem de (a) e (b), no entanto, sobre um T-junção de referência. Note como a sequência de valores são diferentes entre os invariantes da Y-junção e da T-junção em (b) e (c). As modificações de pontos de vista foram baseadas nas simulações de pontos de vista aplicados pelo ASIFT [67].

7.2 Experimentos sobre Imagens Naturais com Transformações Induzidas

Os testes de viabilidade visam a compreensão do comportamento do descritor desenvolvido sob fotografias reais, porém, ainda sob transformações pré-estabelecidas. As imagens utilizadas são encontradas na *Stanford Mobile Visual Search Dataset* [15] ou no banco de imagens próprio dos desenvolvedores do ASIFT [67]. Outras foram fotografias de portfólios independentes [53]. A Figura 7.3 apresenta algumas dessas imagens e que serão os alvos de análises no decorrer desta seção. A Tabela 7.1 apresenta todas as matrizes de confusão para cada caso específico da Figura 7.3.

Na condução dos testes, para cada imagem selecionada, foi aplicada uma transformação específica para gerar um banco de imagens transformadas e, posteriormente, tes-

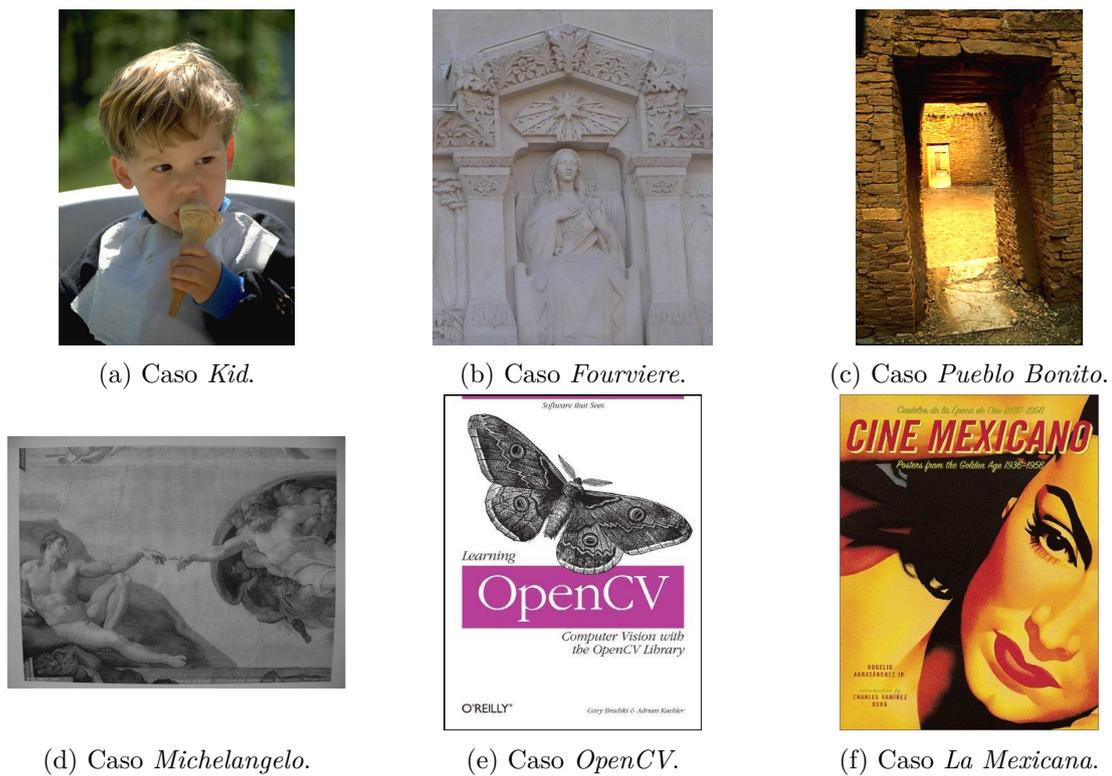


Figura 7.3: Exemplos de imagens utilizadas para testar a técnica sobre situações reais com transformações induzidas: (a) o caso *Kid*, (b) o caso *Fourviere*, (c) o caso *Pueblo Bonito*, (d) o caso *Michelangelo*, (e) o caso *OpenCV* e (f) o caso *La Mexicana*.

tar a aplicação. As transformações utilizadas foram: rotações, escala uniforme, escalas não-uniformes, cisalhamento e combinações entre elas. A Figura 7.4 ilustra o grupo de transformações que será exposto nessa dissertação.

O algoritmo de comparação entre assinaturas de junção, citado na Seção 6.5, mede o esforço necessário para transformar um descritor em outro. Basicamente, ele identifica o nível de dessemelhança entre histogramas, ou seja, quanto menor o nível de dessemelhança, mais próximos serão os descritores. Para a aplicação desenvolvida, não houve um estudo para identificar o valor de dessemelhança que pudesse representar um limiar de corte e, por isso, uma assinatura pode ser associada a outra muito diferente, caso não haja outra assinatura mais parecida.

A Figura 7.5 apresenta alguns resultados para o caso *Kid*, a Figura 7.6 demonstra os mesmos resultados para o caso *Fourvière* e a Figura 7.7 ilustra os resultados para o caso *Pueblo Bonito*. Para o caso *Kid* houveram algumas associações incorretas, que podem ser visualizadas na Figura 7.5a. Por exemplo, a associação entre uma junção do pano e uma junção do cabelo da criança. No entanto, boa parte das comparações se aglomeravam corretamente à mesma região da imagem, onde de fato, junções encontradas na região

	C (original)	I (original)
C (técnica)	1	85
I (técnica)	2	234

(a) Caso *Kid*

	C (original)	I (original)
C (técnica)	3	139
I (técnica)	2	376

(b) Caso *Fourvière*

	C (original)	I (original)
C (técnica)	20	291
I (técnica)	1	704

(c) Caso *Pueblo Bonito*

	C (original)	I (original)
C (técnica)	10	62
I (técnica)	1	145

(d) Caso *Michelangelo*

	C (original)	I (original)
C (técnica)	4	121
I (técnica)	19	338

(e) Caso *OpenCV*

	C (original)	I (original)
C (técnica)	9	160
I (técnica)	0	389

(f) Caso *La Mexicana*

Tabela 7.1: Matriz de confusão sobre cada um dos casos em separado: (a) representa os experimentos realizados sobre o caso *Kid*, (b) representa os resultados para o caso *Fourvière*, (c) representa o caso *Pueblo Bonito*, (d) representa o caso *Michelangelo*, (e) representa o caso *OpenCV* e (f) representa o caso *La Mexicana*. Em todos os casos, há um elevado número de casos incorretos (I) classificados como corretos (C) pela técnica proposta em decorrência das inconsistências na detecção. O caso *OpenCV* em (e) chama a atenção negativamente, pois classificou muitos pareamentos corretos como incorretos. Entretanto, ao analisarmos as imagens, são identificadas excessivas simetrias, o que pode ter confundido o comparador.

da cabeça eram mais parecidas. Para os casos *Fourvière* e *Pueblo Bonito*, boa parte das junções são repetidas por característica das próprias construções, a estrutura de tijolos de *Pueblo Bonito* cujas junções são confundidas entre si assim como as esculturas em formato de flor nas paredes de *Fourvière*, como exemplos.

A Tabela 7.2 apresenta uma matriz de confusão sobre os resultados obtidos para as imagens com transformações induzidas. O limiar de distância utilizado para identificar um pareamento incorreto é de 0,5 e o critério de desempate dentre vários possíveis pareamentos é a média de intensidade da vizinhança 3×3 do pixel central. Os pareamentos classificados como incorretos, *i.e.*, descartados pelo critério de desempate, mas estão relacionados pela homografia conhecida, o que indica que o critério precisa ser revisitado ou, ao menos, mais restritivo.

Alguns problemas foram identificados na implementação da técnica e que podem influenciar nos resultados incorretos. Junções discretas com ramificações muito pequenas não foram tratadas e seus invariantes (em todos os monômios apresentados) resultam em 0, o que representa um resultado válido para outras junções. Outro problema recai sobre

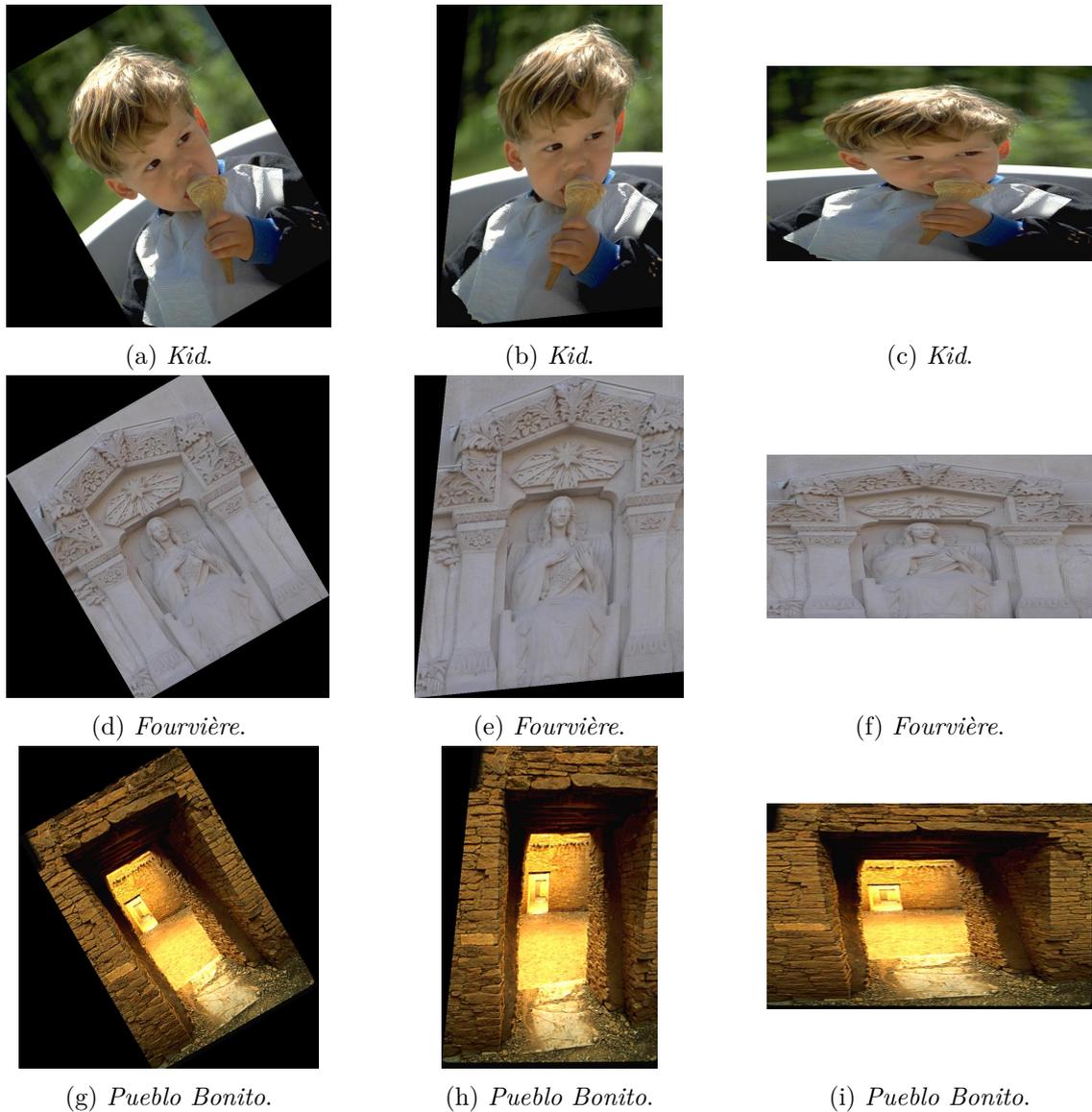


Figura 7.4: Exemplos de transformações induzidas nas imagens de teste. Diversas foram as transformações induzidas. Especificamente para esta dissertação, foram escolhidas, na coluna mais à esquerda, uma rotação de 45 graus para os casos *Kid*, *Fourviere* e *Pueblo Bonito*. Na coluna do meio, um cisalhamento de 0.175 graus para os casos *Kid*, *Fourviere* e *Pueblo Bonito* e na coluna mais à direita, uma escala não-uniforme de 1.2 sobre o eixo x e 0.5 sobre o eixo y também para os casos *Kid*, *Fourviere* e *Pueblo Bonito*.

	Correto (original)	Incorreto (original)
Correto (técnica)	24	515
Incorreto (técnica)	5	1314

Tabela 7.2: Matriz de confusão sobre todas as imagens transformadas artificialmente (Caso *Kid*, *Fourvière* e *Pueblo Bonito*, cujos resultados foram somados. A coluna indica como a técnica classificou os pareamentos identificados e, em linha, como originalmente deveriam ser classificadas de acordo com a homografia conhecida entre os pares de imagens. Os pareamentos classificados como incorretos, *i.e.*, descartados pelo critério de desempate, mas estão relacionados pela homografia conhecida indicam que o critério precisa ser revisitado ou, ao menos, mais restritivo.

a comparação realizada onde uma junção é correlacionada com outra quando o nível de semelhança entre seus histogramas é o menor dentre todos os outros e não foi avaliado um limiar que limitasse esse nível de semelhança. É preciso, também, enfatizar que os detectores de junção investigados nem sempre são consistentes em encontrar as mesmas junções em imagens relacionadas por uma transformação de similaridade, afim ou projetiva e, além disso, que os detectores de borda também são sensíveis a tais transformações.

7.3 Experimentos sobre Imagens Naturais de Cenas Capturadas sobre Diferentes Pontos de Vista

Por fim, para atestar a aplicabilidade da técnica em casos reais, um conjunto de imagens capturadas de pontos de vista diferentes foram descritas e comparadas. Os resultados, apesar de animadores, ainda carregam consigo inconsistências advindas da detecção das junções. As Figuras 7.8, 7.9 e 7.10 apresentam resultados de algumas cenas reais capturadas de pontos de vista distintos, respectivamente, da pintura *A Criação de Adão*, de *Michelangelo*, um livro do *OpenCV* e uma revista *La Mexicana*.

Os testes realizados mostraram que os descritores apresentam robustez até certo ponto à transformações de ponto de vista. Para o caso da pintura de *Michelangelo*, os resultados apresentam falhas nas comparações conforme o ponto de vista se torna mais oblíquo, devido à falhas na detecção de junções discretas. O mesmo ocorre para os outros casos. Conforme o ângulo de visualização diminui, as junções discretas tornam-se mais difíceis de serem identificadas e extraídas corretamente. A Tabela 7.3 apresenta uma matriz de confusão sobre os resultados obtidos para as imagens capturadas sob diferentes pontos de vista. O limiar de distância utilizado e o critério de desempate são os mesmos utilizados nos experimentos sob transformações induzidas. Os pareamentos classificados como incorretos, *i.e.*, descartados pelo critério de desempate, mas estão relacionados pela homografia

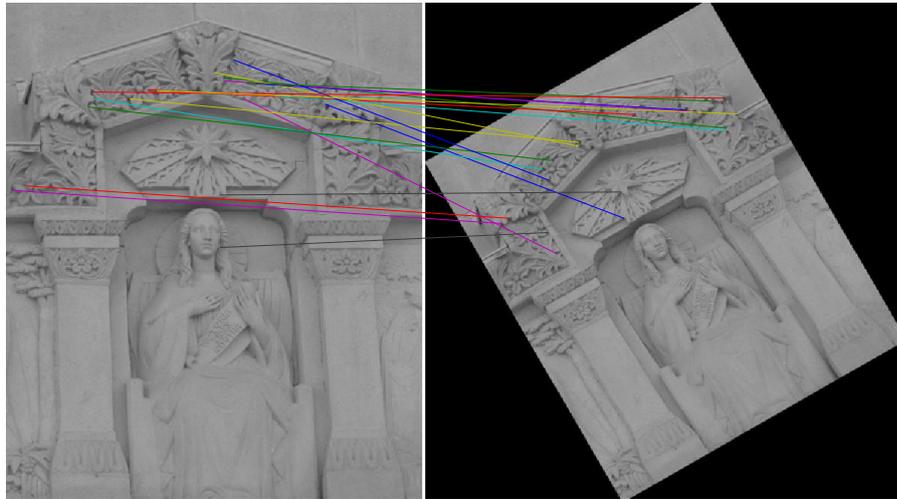


(a) Rotação

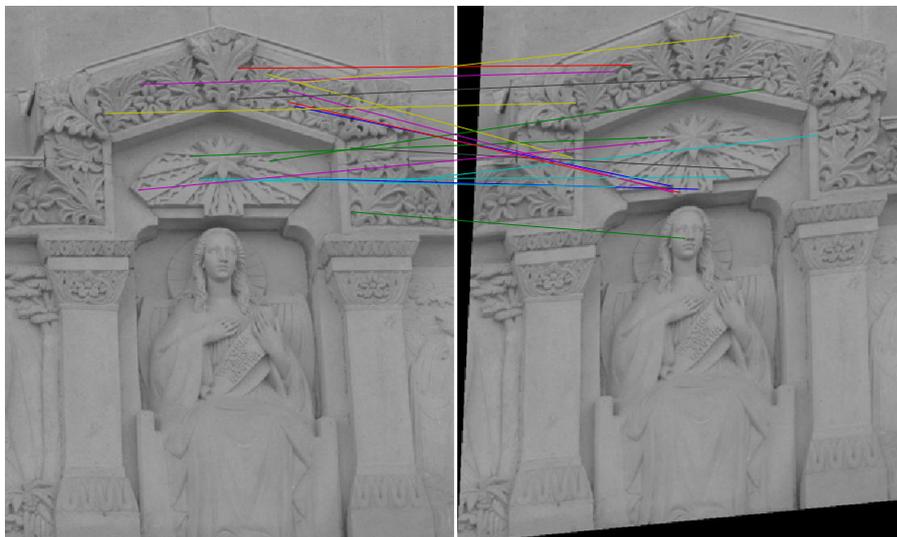


(b) Cisalhamento

Figura 7.5: Descritores mais semelhantes com transformações induzidas: caso *Kid*.

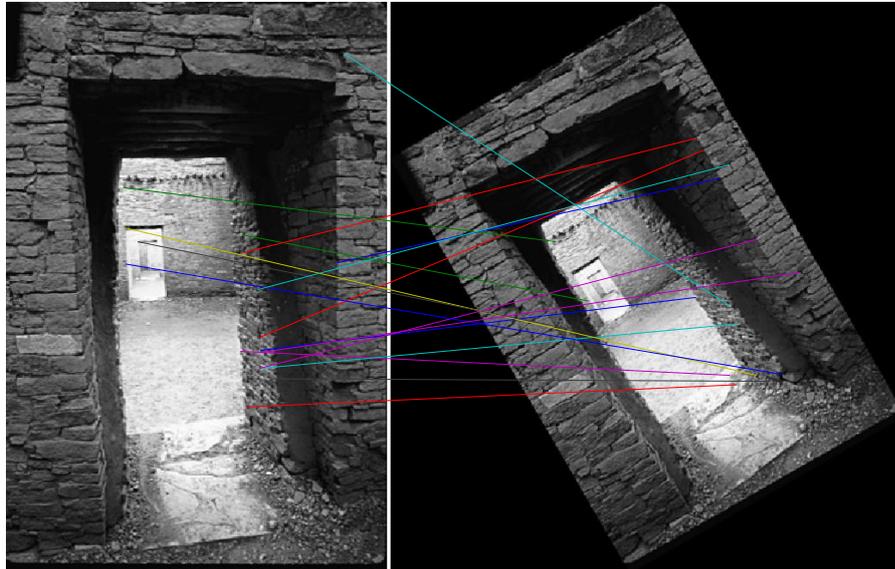


(a) Rotação

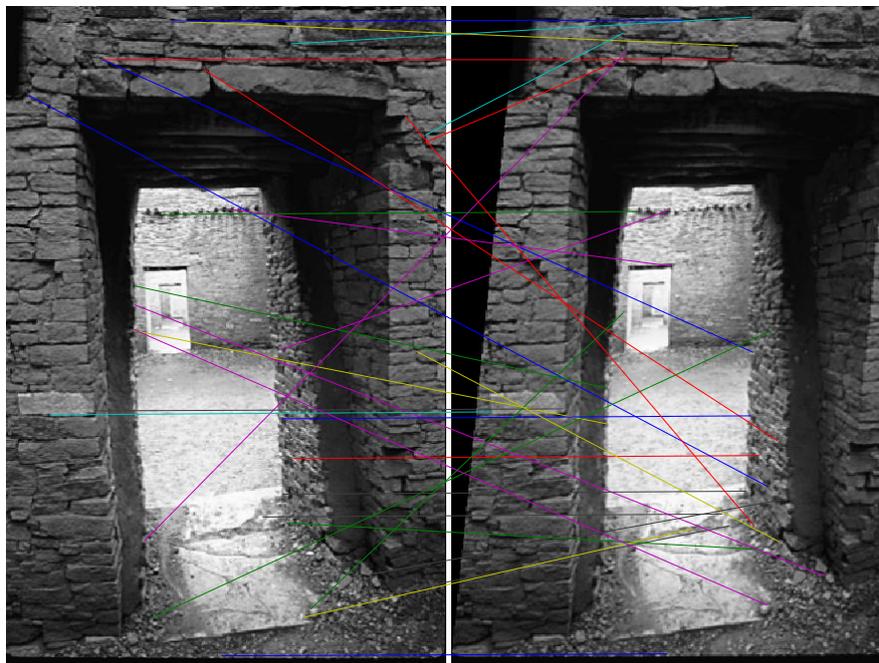


(b) Cisalhamento

Figura 7.6: Descritores mais semelhantes com transformações induzidas: caso *Fourviere*.



(a) Rotação



(b) Cisalhamento

Figura 7.7: Descritores mais semelhantes com transformações induzidas: caso *Pueblo Bonito*.

	Correto (original)	Incorreto (original)
Correto (técnica)	23	343
Incorreto (técnica)	20	872

Tabela 7.3: Matriz de confusão sobre todas as imagens capturadas em diferentes pontos de vista (Caso *Michelangelo*, *La Mexicana* e *OpenCV*, cujos resultados foram somados. A coluna indica como a técnica classificou os pareamentos identificados e, em linha, como originalmente deveriam ser classificadas de acordo com a homografia conhecida entre os pares de imagens. Os pareamentos classificados como incorretos, *i.e.*, descartados pelo critério de desempate, mas estão relacionados pela homografia conhecida indicam que o critério precisa ser revisitado ou, ao menos, mais restritivo.

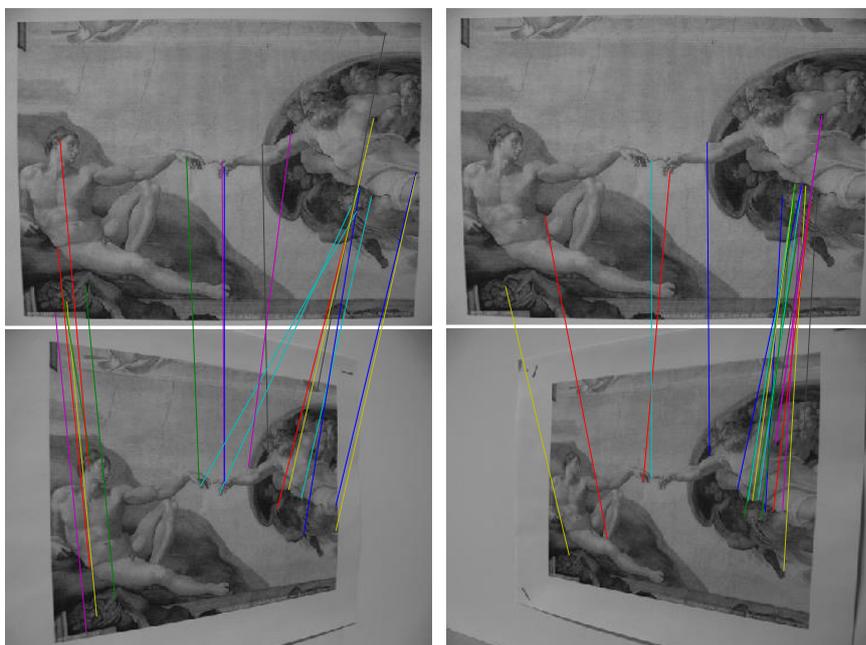


Figura 7.8: Descritores mais semelhantes com transformações livres: caso *Michelangelo*.

conhecida indicam que o critério precisa ser revisitado ou, ao menos, mais restritivo.

7.4 Discussão

Neste capítulo foram apresentados os experimentos realizados para atestar a robustez dos valores invariantes formulados e a viabilidade dos descritores propostos em situações mais próximas da realidade.

Na Seção 7.1, foi realizada uma análise sobre a robustez dos valores invariantes propostos e como as possíveis manipulações da junção surtem efeito nos valores finais. Na Seção 7.2, foram mostradas algumas fotografias transformadas artificialmente e confrontadas entre si através do descritor proposto com resultados interessantes. Por fim, na

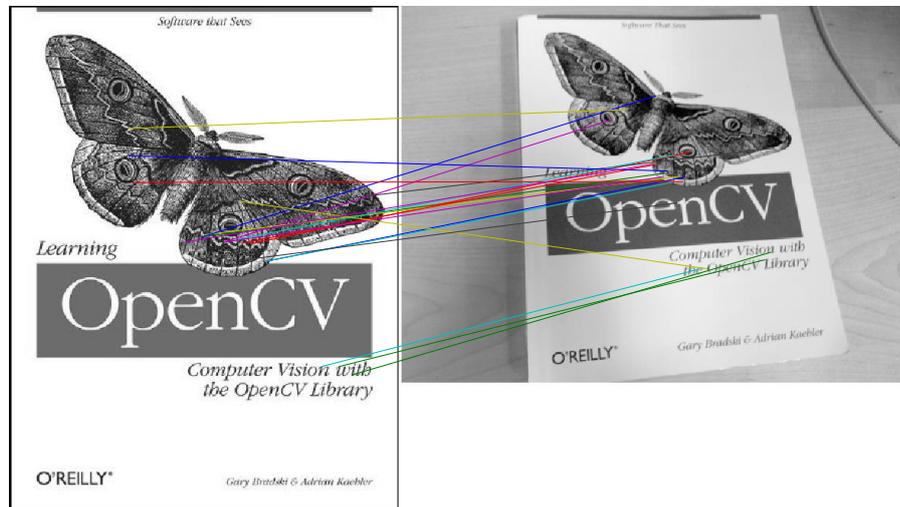


Figura 7.9: Descritores mais semelhantes com transformações livres: caso *OpenCV*.



Figura 7.10: Descritores mais semelhantes com transformações livres: caso *La Mexicana*.

Seção 7.3, foram realizados testes com imagens reais de cenas capturadas em pontos de vista diferentes para atestar a viabilidade da técnica em casos reais.

Os experimentos que envolvem manipulação de junções contínuas são consistentes e, de fato, atestam a estabilidade dos valores invariantes gerados. Os testes posteriores em imagens reais são animadores, porém, os resultados apresentaram falhas que advém, principalmente, de problemas na detecção das junções.

Os resultados obtidos não foram sistematicamente confrontados com os resultados de outras técnicas do estado-da-arte que se propõem à solucionar o mesmo problema, pois a aplicação prática da teoria proposta apresentou inconsistências que ainda precisam ser investigadas. Para a técnica proposta, não foi realizado um estudo para se determinar um limiar mínimo de semelhança para a comparação entre características, o que deixou os resultados finais susceptíveis à falhas pela falta de um filtro qualitativo. Apesar disso, os experimentos realizados apresentam bom desempenho computacional, o que caracteriza a técnica proposta como uma potencial alternativa de pesquisa para a área de detecção e descrição de características locais.

Capítulo 8

Conclusão

A detecção e descrição de características visuais em imagens são operações fundamentais para diversas aplicações da área de Computação Visual e Processamento de Imagens como a classificação de objetos, mapeamento automatizado do ambiente, dentre outras. As técnicas mais comuns da literatura se baseiam na detecção de regiões de interesse e análise de intensidade, sem levar em consideração a geometria dos objetos contidos nelas. Esse tipo de abordagem, apesar de comum, impossibilita ou, ao menos, dificulta o desenvolvimento de uma operação realmente invariante às transformações de graus maiores, como as afins e projetivas.

Diversas pesquisas nas áreas de Psicologia e Neurociência demonstram a importância que as inflexões de borda possuem para o cérebro humano na identificação e classificação de objetos, o que sugere o desenvolvimento de um descritor simplificado, completamente baseado sobre junções. Dentre as contribuições propostas nesta dissertação, está a definição formal de uma série de monômios invariantes às transformações homotéticas, afins e projetivas, calculadas sobre os pontos delimitadores de junções. No mais, também foi proposto como descritor um histograma acumulador de todos os valores invariantes encontrados. Esse descritor age como uma assinatura invariante de uma região de interesse da imagem. Os experimentos mostraram que este tipo de estratégia é promissora.

É importante notar que os descritores, em seu estado-da-arte, dependem de estratégias distintas para se obter invariância às diferentes transformações de regiões, como por exemplo, a pirâmide de escala para obter invariância à escala uniforme e detecção de orientação preferencial para invariância às rotações, tornando o processo complexo como um todo. Informações de orientação e magnitude do gradiente podem ser utilizadas para auxiliar possivelmente o descritor, porém, deixam de ser as informações mais relevantes.

A proposta deste trabalho é simplificar o processo de detecção, descrição e provimento de invariância às transformações, de tal maneira que esses procedimentos sejam uniformes. Para cada junção identificada, valores invariantes são calculados e utilizados para alimentar um histograma de junção. A invariância está embutida no monômio utilizado para gerar tais valores.

A principal contribuição deste trabalho é a derivação de monômios invariantes sobre estruturas visuais locais de junção. Os procedimentos descritos no Capítulo 5 mostram que junções podem ser representadas por um valor escalar que é invariante às transformações projetivas.

As junções são matematicamente formalizadas, porém, não se manifestam de maneira tão comportada em imagens digitais. As junções discretas são sensíveis à discretização e ruídos e muitas vezes impossíveis de serem isoladas localmente. O Capítulo 6 apresentou maneiras de se identificar e converter junções discretas em junções contínuas, utilizadas para cálculo dos monômios invariantes. Abordagens são propostas para a detecção e descrição de junções próprias e impróprias, descritas a partir de um histograma de valores invariantes para cada junção identificada.

Os monômios formulados são inegavelmente invariantes para junções contínuas, no entanto, apresentam distúrbios quando calculados diretamente sobre as junções extraídas da imagem, devido às dificuldades encontradas para se detectar e delimitar junções. O descritor proposto como assinatura de junções para uma imagem foi suficiente para mostrar resultados animadores, entretanto, ainda deve ser melhorado.

A derivação de novos monômios para junções é possível, inclusive com o conhecimento contido neste documento, por exemplo, combinando mais matrizes simétricas ou derivando monômios invariantes às múltiplas junções que permitam caracterizar inteiramente um objeto.

Especificamente neste trabalho, os objetos de interesse são as junções identificadas sobre o plano de imagem e um eventual mapeamento projetivo entre as mesmas. Entretanto, a aplicação dos monômios pode ser estendida para descrever junções em três dimensões, como em casos imagens *RGB-d* capturadas a partir dispositivos especializados.

A caracterização independente de informações de textura pode ser útil para os casos específicos de objetos abertos, *i.e.*, onde os objetos ao fundo se confundem com o objeto em questão, como é o caso da estrutura de uma bicicleta. Técnicas em seu estado-da-arte, baseadas em texturas de regiões, falhariam nestes casos, o que sugere uma solução híbrida.

A Teoria de Invariantes, no mais, é bastante ampla e possui aplicações práticas em diversas áreas. Em teoria dos grafos, por exemplo, existem métodos para determinação de um único valor invariante que represente a estrutura do grafo, porém, falham em condições específicas. Testes informais indicam que a representação de um grafo à partir de um histograma invariante, como apresentado neste trabalho, pode ser aplicada para determinar isomorfismo (inclusive em grafos com pesos em sua arestas ou em dígrafos), uma abordagem alternativa á busca por um único valor mas que permite uma comparação eficiente.

A detecção das junções é um desafio que merece atenção especial, pois sem um ajuste minimamente próximo ao correto das junções, muitas incertezas são propagadas para o histograma descritor final, o que pode confundir o algoritmo comparador. Além disso, existem problemas inerentes à esta etapa em específico que não foram consideradas neste trabalho, como por exemplo, possíveis falhas na detecção das bordas sob transformações elevadas de escala. A adaptação para o uso dos chamados *superpixels* pode ser uma linha interessante de pesquisa a ser trabalhada.

Referências

- [1] ABDEL-HAKIM, A. E., FARAG, A. A. Csfift: A sift descriptor with color invariant characteristics. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on* (2006), vol. 2, p. 1978–1983.
- [2] ALAJLAN, N., EL RUBE, I., KAMEL, M. S., FREEMAN, G. Shape retrieval using triangle-area representation and dynamic space warping. *Pattern Recogn.* 40, 7 (julho de 2007), 1911–1920.
- [3] ALAJLAN, N., KAMEL, M. S., FREEMAN, G. Multi-object image retrieval based on shape and topology. *Image Commun.* 21, 10 (novembro de 2006), 904–918.
- [4] ASHOK, V., BALAKUMARAN, T., GOWRISHANKAR, C., VENNILA, I. L. A., KUMAR, A. N. The fast haar wavelet transform for signal & image processing. *CoRR abs/1002.2184* (2010).
- [5] BANNOUR, H., HLAOUA, L., EL AYEYB, B. Survey of the adequate descriptor for content-based image retrieval on the web: Global versus local features. In *CORIA* (2009), LSIS-USTV, p. 445–456.
- [6] BAY, H., ESS, A., TUYTELAARS, T., VAN GOOL, L. Speeded-up robust features (surf). *Comput. Vis. Image Underst.* 110, 3 (junho 2008), 346–359.
- [7] BELONGIE, S., MALIK, J. Matching with shape contexts. In *Content-based Access of Image and Video Libraries, 2000. Proceedings. IEEE Workshop on* (2000), p. 20–26.
- [8] BELONGIE, S., MALIK, J., PUZICHA, J. Shape matching and object recognition using shape contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24, 4 (abril 2002), 509–522.
- [9] BIEDERMAN, I. Recognition-by-components: A theory of human image understanding. *Psychological Review* 94 (1987), 115–147.
- [10] BOSCH, A., ZISSERMAN, A., MUÑOZ, X. Scene classification using a hybrid generative/discriminative approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30, 4 (abril 2008), 712–727.
- [11] BOVIS, K., SINGH, S., FIELDSEND, J., PINDER, C. Identification of masses in digital mammograms with mlp and rbf nets. In *Neural Networks, 2000. IJCNN 2000, Proceedings of the IEEE-INNS-ENNS International Joint Conference on* (2000), vol. 1, p. 342–347.
- [12] BUNDESEN, C. A theory of visual attention. *Psychological Review* 97, 4 (outubro 1990), 523–547.

- [13] CALONDER, M., LEPETIT, V., OZUYSAL, M., TRZCINSKI, T., STRECHA, C., FUA, P. Brief: Computing a local binary descriptor very fast. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 34, 7 (2012), 1281–1298.
- [14] CAYLEY, A. A memoir on the theory of matrices. *The Collected Mathematical Papers* 24, 7 (1858), 475–496.
- [15] CHANDRASEKHAR, V. R., CHEN, D. M., TSAI, S. S., CHEUNG, N., CHEN, H., TAKACS, G., REZNIK, Y., VEDANTHAM, R., GRZESZCZUK, R., BACH, J., GIROD, B. The stanford mobile visual search data set. In *Proceedings of the Second Annual ACM Conference on Multimedia Systems* (New York, NY, USA, 2011), MMSys '11, ACM, p. 117–122.
- [16] COOLEY, J. W., TUKEY, J. W. An algorithm for the machine calculation of complex fourier series. In *Mathematics of Computation* (1965), p. 297–301.
- [17] EL-GHAZAL, A., BASIR, O., BELKASIM, S. Farthest point distance: A new shape signature for fourier descriptors. *Image Commun.* 24, 7 (agosto 2009), 572–586.
- [18] ELFARRA, B., ABUHAIBA, I. New feature extraction method for mammogram computer aided diagnosis. In *International Journal of Signal Processing, Image Processing and Pattern Recognition* (2013), vol. 1, p. 13–.
- [19] FILIPE, S., ALEXANDRE, L. A. From the human visual system to the computational models of visual attention: a survey. *Artificial Intelligence Review* 43, 4 (2015), 601.
- [20] FORSYTH, D., MUNDY, J. L., ZISSERMAN, A. Invariant descriptors for 3d object recognition and pose. In *In Applications of Invariance in Computer Vision* (Reykjavik, Iceland, março 1991), p. 171–208.
- [21] FREUND, R. J., WILSON, W. J., MOHR, D. L. *Statistical Methods and Applications*, vol. 3. Academic Press; ISBN 978-0123749703., 2010.
- [22] FRIEDMAN, J. H., BENTLEY, J. L., FINKEL, R. A. An algorithm for finding best matches in logarithmic expected time. *ACM Trans. Math. Softw.* 3, 3 (setembro de 1977), 209–226.
- [23] GDALYAHU, Y., WEINSHALL, D. Flexible syntactic matching of curves and its application to automatic hierarchical classification of silhouettes. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 21, 12 (dezembro 1999), 1312–1328.
- [24] GIONIS, A., INDYK, P., MOTWANI, R. Similarity search in high dimensions via hashing. In *Proceedings of the 25th International Conference on Very Large Data Bases* (San Francisco, CA, USA, 1999), VLDB '99, Morgan Kaufmann Publishers Inc., p. 518–529.
- [25] GROS, P., QUAN, L. Projective invariants for vision, 1992. LIFIA - Institut IMAG.
- [26] GUPTA, S. K., AGRWAL, S., MEENA, Y., NAIN, N. A hybrid method of feature extraction for facial expression recognition. In *Signal-Image Technology and Internet-Based Systems (SITIS), 2011 Seventh International Conference on* (novembro 2011), p. 422–425.

- [27] HAAR, A. Zur theorie der orthogonalen funktionensysteme. In *Mathematische Annalen* (1911), vol. 69, p. 331–371.
- [28] HARRIS, C., STEPHENS, M. A combined corner and edge detector. In *In Proceedings of Fourth Alvey Vision Conference* (1988), p. 147–151.
- [29] HARTLEY, R. I., ZISSERMAN, A. *Multiple View Geometry in Computer Vision*, second ed. Cambridge University Press, ISBN: 0521540518, 2004.
- [30] INTILLE, S. S., DAVIS, J. W., BOBICK, A. F. Real-time closed-world tracking. In *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on* (junho 1997), p. 697–703.
- [31] JEHEE, J. F. M., BRADY, D. K., TONG, F. Attention improves encoding of task-relevant features in the human visual cortex. *The Journal of Neuroscience* 31, 22 (2011), 8210–8219.
- [32] KAPUR, D., MUNDY, J. L. Fitting affine invariant conics to curves. In *In Applications of Invariance in Computer Vision* (Reykjavik, Iceland, março 1991), p. 209–233.
- [33] KE, Y., SUKTHANKAR, R. Pca-sift: A more distinctive representation for local image descriptors. In *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (Washington, DC, USA, 2004), CVPR’04, IEEE Computer Society, p. 506–513.
- [34] LEUTENEGGER, S., CHLI, M., SIEGWART, R. Y. Brisk: Binary robust invariant scalable keypoints. In *Proceedings of the 2011 International Conference on Computer Vision* (Washington, DC, USA, 2011), ICCV ’11, IEEE Computer Society, p. 2548–2555.
- [35] LINDBERG, T. *Scale-Space Theory in Computer Vision*, vol. 256. Springer, ISBN 0-7923-9418-6., 1994.
- [36] LINDBERG, T. Scale selection properties of generalized scale-space interest point detectors. *Journal of Mathematical Imaging and Vision* 46, 2 (2012), 177–210.
- [37] LINDBERG, T. Image matching using generalized scale-space interest points. *Journal of Mathematical Imaging and Vision* 52, 1 (2014), 3–36.
- [38] LOWE, D. G. Three-dimensional object recognition from single two-dimensional images. *Artificial Intelligence* 31, 3 (março 1987), 355–395.
- [39] LOWE, D. G. Object recognition from local scale-invariant features. In *Proceedings of the International Conference on Computer Vision* (Washington, DC, USA, 1999), ICCV ’99, IEEE Computer Society, p. 1150–.
- [40] LOWE, D. G. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* 60, 2 (2004), 91–110.
- [41] LU, J., PLATANIOTIS, K. N., VENETSANOPOULOS, A. N. Face recognition using kernel direct discriminant analysis algorithms. *Neural Networks, IEEE Transactions on* 14, 1 (janeiro 2003), 117–126.

- [42] MAIRE, M., ARBELAEZ, P., FOWLKES, C., MALIK, J. Using contours to detect and localize junctions in natural images. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on* (junho 2008), p. 1–8.
- [43] MARTIN, J. *A Fast Transform for Spherical Harmonics*. Tese de Doutorado, Yale University, New Haven CT, Estados Unidos, 1997.
- [44] MIKOLAJCZYK, K., TUYTELAARS, T., SCHMID, C., ZISSERMAN, A., MATAS, J., SCHAFFALITZKY, F., KADIR, T., GOOL, L. V. A comparison of affine region detectors. *International Journal of Computer Vision* 65, 1 (2005), 43–72.
- [45] MONIER, E., WILHELM, P., RÜCKERT, U. A Computer Vision Based Tracking System for Indoor Team Sports. In *The fourth International Conference on Intelligent Computing and Information Systems* (2009).
- [46] MORAVEC, H. Obstacle avoidance and navigation in the real world by a seeing robot rover. In *tech. report CMU-RI-TR-80-03, Robotics Institute, Carnegie Mellon University and doctoral dissertation, Stanford University*. Stanford University Press, setembro 1980.
- [47] MORI, G., MALIK, J. Estimating human body configurations using shape context matching. In *Proceedings of the 7th European Conference on Computer Vision-Part III* (London, UK, UK, 2002), ECCV '02, Springer-Verlag, p. 666–680.
- [48] MURASE, H., NAYAR, S. K. Visual learning and recognition of 3-d objects from appearance. *Int. J. Comput. Vision* 14, 1 (janeiro 1995), 5–24.
- [49] ORTIZ, R. Freak: Fast retina keypoint. In *Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (Washington, DC, USA, 2012), CVPR '12, IEEE Computer Society, p. 510–517.
- [50] OZUYSAL, M., CALONDER, M., LEPETIT, V., FUA, P. Fast keypoint recognition using random ferns. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32, 3 (março 2010), 448–461.
- [51] PERS, J., KOVACIC, S. Computer vision system for tracking players in sports games. In *Image and Signal Processing and Analysis, 2000. IWISPA 2000. Proceedings of the First International Workshop on* (2000), p. 177–182.
- [52] PERSON, E., FU, K. S. Shape discrimination using fourier descriptors. *IEEE Transactions on Systems, Man, and Cybernetics* 7, 3 (março 1977), 170–179.
- [53] PETITCOLAS, F. Digital watermarking and steganography. http://www.petitcolas.net/watermarking/image_database/, 1997. [ativo; acessado 28-mar-2016].
- [54] QUAN, L., GROS, P., MOHR, R. Invariants of a pair of conics revisited. In *British Machine Vision Conference* (Glasgow, Reino Unido, setembro 1991), p. 71–77.
- [55] RUBLEE, E., RABAUD, V., KONOLIGE, K., BRADSKI, G. Orb: An efficient alternative to sift or surf. In *Proceedings of the 2011 International Conference on Computer Vision* (Washington, DC, USA, 2011), ICCV '11, IEEE Computer Society, p. 2564–2571.

- [56] RUBNER, Y., TOMASI, C., GUIBAS, L. J. A metric for distributions with applications to image databases. In *Proceedings of the Sixth International Conference on Computer Vision* (Washington, DC, USA, 1998), ICCV '98, IEEE Computer Society, p. 59–.
- [57] SANZ, J., PETKOVIC, D. Machine vision algorithms for automated inspection thin-film disk heads. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 10, 6 (novembro 1988), 830–848.
- [58] SARIPALLI, S., MONTGOMERY, J. F., SUKHATME, G. Vision-based autonomous landing of an unmanned aerial vehicle. In *Robotics and Automation, 2002. Proceedings. ICRA '02. IEEE International Conference on* (2002), vol. 3, p. 2799–2804.
- [59] SEMPLE, J. G., KNEEBONE, G. T. *Algebraic Projective Geometry*. Oxford Science Publication, 1952.
- [60] SHARP, C. S., SHAKERNIA, O., SASTRY, S. S. A vision system for landing an unmanned aerial vehicle. In *Robotics and Automation, 2001. Proceedings 2001 ICRA. IEEE International Conference on* (2001), vol. 2, p. 1720–1727.
- [61] SHARVIT, D., CHAN, J., TEK, H., KIMIA, B. B. Symmetry-based indexing of image databases. In *Content-Based Access of Image and Video Libraries, 1998. Proceedings. IEEE Workshop on* (junho 1998), p. 56–62.
- [62] SHI, J., TOMASI, C. Good features to track. In *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR '94., 1994 IEEE Computer Society Conference on* (junho 1994), p. 593–600.
- [63] SOMMERVILLE, D. M. Y. *Analytical Conics*, third ed. London G. Bell and Sons, Ltd., 1961.
- [64] TURK, M., PENTLAND, A. Eigenfaces for recognition. *J. Cognitive Neuroscience* 3, 1 (janeiro 1991), 71–86.
- [65] TUYTELAARS, T., MIKOLAJCZYK, K. Local invariant feature detectors: A survey. *Found. Trends. Comput. Graph. Vis.* 3, 3 (julho 2008), 177–280.
- [66] WISCHNEWSKI, M., BELARDINELLI, A., SCHNEIDER, W. X., STEIL, J. J. Where to look next? combining static and dynamic proto-objects in a tva-based model of visual attention. In *Cogn Comput* (Reykjavik, Iceland, março 2010), vol. 2, p. 326–343.
- [67] YU, G., MOREL, J. Asift: An algorithm for fully affine invariant comparison. *Image Processing On Line* 1 (2011).
- [68] ZHANG, D., LU, G. A comparative study of fourier descriptors for shape representation and retrieval. In *Proceedings of 5th Asian Conference on Computer Vision (ACCV* (2002), Springer, p. 646–651.

APÊNDICE A - Função Acumulada Analítica para o Invariante Projetivo (script)

O objetivo deste *script* em *Mathematica* é mostrar a dedução de uma expressão fechada ou uma forma numérica para calcular a função acumulada utilizada na equalização no intervalo $[0, 1]$ dos valores assumidos pela invariante γ dada por:

$$\Gamma[\lambda1_ , \lambda2_ , \lambda3_] := \frac{\lambda1^2 \lambda2^2 \lambda3^2}{(\lambda1^2 + \lambda2^2 + \lambda3^2)^3};$$

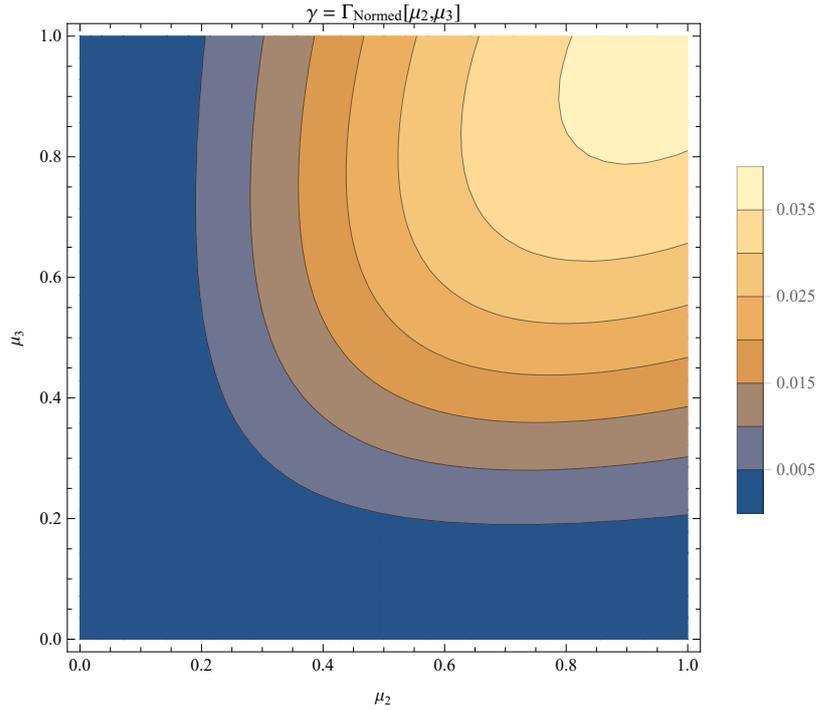
onde λ_1, λ_2 e λ_3 são autovalores de uma matriz quadrada. A função que calcula γ pode ser reescrita de forma normalizada como:

$$\Gamma\text{Normed}[\mu2_ , \mu3_] := \frac{\mu2^2 \mu3^2}{(1 + \mu2^2 + \mu3^2)^3};$$

onde $\mu_1 = \frac{\lambda_1}{\lambda_1} = 1, \mu_2 = \frac{\lambda_2}{\lambda_1}$ e $\mu_3 = \frac{\lambda_3}{\lambda_1}$. Nesse caso, é assumido que $\lambda_1 \neq 0$ e $|\lambda_1| \geq |\lambda_2| \geq |\lambda_3|$. Observada a simetria da expressão em termos de μ_1 e μ_2 , e visto que ambos os autovalores normalizados serão elevados a quadrado, podemos, sem perda de generalizada, definir seus interlados como $0 \leq \mu_2, \mu_3 \leq 1$. Os limites de γ são $\left[0, \frac{1}{27}\right]$.

```
{γMin, γMax} = {MinValue[Γ[λ1, λ2, λ3], {λ1, λ2, λ3}, Reals],
  MaxValue[Γ[λ1, λ2, λ3], {λ1, λ2, λ3}, Reals]};
```

```
ContourPlot[ΓNormed[μ2, μ3], {μ2, 0, 1}, {μ3, 0, 1}, FrameLabel → {"μ2", "μ3"},
  PlotLabel → "γ = ΓNormed[μ2, μ3]", PlotLegends → Automatic]
```



Dado um valor de γ , teremos $\mu_2 = \mu_1 = \mu$ no segmento de reta diagonal definidos pelos pontos (0, 0) a (1, 1). O valor de μ pode ser calculado pelo *Mathematica*® a partir de:

$$\text{Solve}[\Gamma\text{Normed}[\mu, \mu] == \gamma, \mu]$$

Após descartar todas as soluções inválidas (i.e., relacionadas aos limites de μ_2 e μ_3), temos:

$$\mu[\gamma_-] := \frac{1}{4 \sqrt{3}} \sqrt{\left(-\frac{1}{\gamma} \left(-2 + 24 \gamma + \frac{(1 - i \sqrt{3})(-1 + 24 \gamma)}{(-1 + 36(1 - 6 \gamma) \gamma + 24 \sqrt{3}(\gamma^3(-1 + 27 \gamma)))^{1/3}} - i(-i + \sqrt{3}) \left(-1 + 36(1 - 6 \gamma) \gamma + 24 \sqrt{3}(\gamma^3(-1 + 27 \gamma))^{1/3} \right) \right) \right)}$$

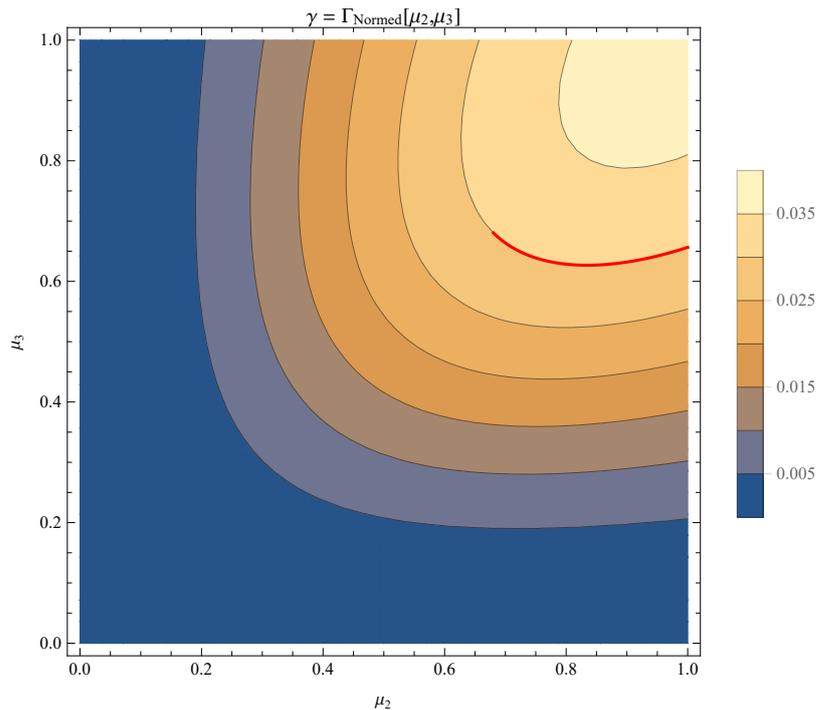
Fixando $\mu_2 = 1$, para um dado valor de γ , o valor assumido por μ_3 (i.e., μ_3) pode ser calculado pelo *Mathematica*® a partir de:

$$\text{Solve}[\Gamma\text{Normed}[1, \mu_3] == \gamma, \mu_3]$$

Após descartar todas as soluções inválidas (i.e., soluções complexas e fora dos limites de μ_3), temos:

$$\mu_3\text{Dot}[\gamma_-] := \frac{1}{\sqrt{3}} \left(\sqrt{\left(-6 + \frac{(-3)^{2/3}}{(-9 \gamma^2 + \sqrt{3} \sqrt{\gamma^3(-1 + 27 \gamma)})^{1/3}} - \frac{(-3)^{1/3}(-9 \gamma^2 + \sqrt{3} \sqrt{\gamma^3(-1 + 27 \gamma)})^{1/3}}{\gamma} \right) \right)}$$

A integral de todos os segmentos de reta verticais no intervalo no intervalo $\mu_2 \in [\mu[\gamma], 1]$ e que compreendem o intervalo $\mu_3 \in [\mu_3[\mu_2, \gamma], \mu_2]$ correspondem à metade da área da região do plano de autovalores normalizados que contém pares de autovalores que levam a invariantes maiores do que γ . O complemento de duas vezes essa área corresponde à função acumulada desejada, escrita no *Mathematica*® como:



Uma maneira de avaliar $\text{Cumulative}\Gamma[\gamma]$ numericamente é pela função recursiva descrita abaixo, com critério de parada relacionado à precisão numérica desejada:

```

NumericCumulativeGamma[ $\gamma$ _, Precision_] :=
  1 - 2 Recursive[{{ $\mu$ [ $\gamma$ ],  $\mu$ [ $\gamma$ ]}, {1,  $\mu$ 3Dot[ $\gamma$ ]}, {1, 1}},  $\gamma$ , Precision];

```

onde:

```

Recursive[{p1_, p2_, p3_},  $\gamma$ _, Precision_] := Module[{m, q, s},
  s = Abs[ $\frac{1}{2}$  (p1[[1]] (p2[[2]] - p3[[2]]) + p2[[1]] (p3[[2]] - p1[[2]]) + p3[[1]] (p1[[2]] - p2[[2]]) )];
  (*Área do triângulo {p1,p2,p3}*)
  s + If[s > 10-Precision,
    m =  $\frac{p1 + p2}{2}$ ;
    q = {m[[1]],  $\mu$ 3Hat[m[[1]],  $\gamma$ ]};
    Recursive[{p1, q, m},  $\gamma$ , Precision] + Recursive[{q, p2, m},  $\gamma$ , Precision], 0]
];

```