



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA DE TELEINFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE TELEINFORMÁTICA
MESTRADO ACADÊMICO EM ENGENHARIA DE TELEINFORMÁTICA

ARTUR RODRIGUES ROCHA NETO

**ATRIBUTOS GEOMÉTRICOS INVARIANTES ÀS TRANSFORMAÇÕES AFINS
PARA REPRESENTAÇÃO DE NUVENS DE PONTOS E UM ARCABOUÇO PARA
EXTRAÇÃO DE MOMENTOS ALGÉBRICOS**

FORTALEZA

2020

ARTUR RODRIGUES ROCHA NETO

ATRIBUTOS GEOMÉTRICOS INVARIANTES ÀS TRANSFORMAÇÕES AFINS PARA
REPRESENTAÇÃO DE NUVENS DE PONTOS E UM ARCABOUÇO PARA EXTRAÇÃO
DE MOMENTOS ALGÉBRICOS

Dissertação apresentada ao Curso de Mestrado Acadêmico em Engenharia de Teleinformática do Programa de Pós-Graduação em Engenharia de Teleinformática do Centro de Tecnologia da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Engenharia de Teleinformática. Área de Concentração: Engenharia de Teleinformática

Orientador: Prof. Dr. José Marques Soares

FORTALEZA

2020

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

R571a Rocha Neto, Artur Rodrigues.

ATRIBUTOS GEOMÉTRICOS INVARIANTES ÀS TRANSFORMAÇÕES AFINS PARA REPRESENTAÇÃO DE NUVENS DE PONTOS E UM ARCABOUÇO PARA EXTRAÇÃO DE MOMENTOS ALGÉBRICOS / Artur Rodrigues Rocha Neto. – 2020.

74 f. : il. color.

Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Tecnologia, Programa de Pós-Graduação em Engenharia de Teleinformática, Fortaleza, 2020.

Orientação: Prof. Dr. José Marques Soares.

1. Nuvens de Pontos. 2. Momentos. 3. Decomposição em Valores Singulares. 4. Polinômios Ortogonais. I. Título.

CDD 621.38

ARTUR RODRIGUES ROCHA NETO

ATRIBUTOS GEOMÉTRICOS INVARIANTES ÀS TRANSFORMAÇÕES AFINS PARA
REPRESENTAÇÃO DE NUVENS DE PONTOS E UM ARCABOUÇO PARA EXTRAÇÃO
DE MOMENTOS ALGÉBRICOS

Dissertação apresentada ao Curso de Mestrado Acadêmico em Engenharia de Teleinformática do Programa de Pós-Graduação em Engenharia de Teleinformática do Centro de Tecnologia da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Engenharia de Teleinformática. Área de Concentração: Engenharia de Teleinformática

Aprovada em:

BANCA EXAMINADORA

Prof. Dr. José Marques Soares (Orientador)

Prof. Dr. João Paulo Papa
Universidade Estadual Paulista Júlio de Mesquita
Filho (UNESP)

Profa. Dra. Michela Mulas
Universidade Federal do Ceará (UFC)

Prof. Dr. George André Pereira Thé
Universidade Federal do Ceará (UFC)

Para Artemísia e Antônio (*in memoriam*).

AGRADECIMENTOS

Ao Prof. Dr. José Marques Soares, por me orientar mais uma vez na vida acadêmica.

Ao Prof. Dr. George André Pereira Thé, por co-orientar esta dissertação de mestrado e por me ensinar a duvidar de tudo (mas sem perder a ternura jamais).

À Prof^a. Dr^a. Natália Maria Cordeiro Barroso, pela ajuda na revisão deste trabalho e por me ajudar a mostrar como a matemática pode ser mais rica do que aquela que eu aprendi na graduação.

A todos os membros do grupo interFACES, por serem os melhores colegas de pesquisa que eu poderia pedir.

Aos demais professores do Programa de Pós-Graduação em Engenharia de Teleinformática (PPGETI), com os quais eu tive o privilégio de aprender, debater e refletir sobre diversos temas e assuntos, não somente ciência e tecnologia.

À minha família (Ageisa, Samuel, Armízia, vovô Artur e Seu Antônio) e aos meus grandes amigos (Glícia, Hamlet e Jy), por serem verdadeiros pilares emocionais em minha vida.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), em nome dos seus membros sérios e verdadeiramente interessados pelo desenvolvimento da ciência e tecnologia no Brasil, pelo financiamento da pesquisa de mestrado via bolsa de estudos.

Por último, registro meu mais profundo agradecimento e imensurável respeito a todos os profissionais de saúde e de serviços essenciais que trabalharam para manter a sociedade cearense firme durante a primeira onda da pandemia de COVID-19, cuja ascendência e pico de número de casos ocorreram durante a formalização do texto desta dissertação.

“Para que serve a utopia? Serve para isso: para que eu não deixe de caminhar.”

(Eduardo Galeano)

RESUMO

As imagens tridimensionais (imagens 3D) permitem um estudo mais preciso de formas e superfícies pois estas guardam informação de profundidade. Em comparação às imagens bidimensionais (imagens 2D), as imagens 3D superam algumas limitações daquelas em alguns cenários práticos. Entretanto, muitas áreas de aplicação ainda dependem de etapas de pré-processamento bastante invasivas nas imagens 3D, frequentemente demandando parametrizações *a priori*. Tais dependências das imagens 3D impactam na etapa de extração de atributos, seja limitando as opções de métodos que possam ser usados, seja prejudicando os atributos extraídos e, conseqüentemente, nos resultados obtidos. A contribuição deste trabalho divide-se em duas. A primeira consiste de uma nova representação para imagens 3D calculada a partir de uma nuvem de pontos. A segunda é materializada pela proposta de um arcabouço para a extração de atributos baseado em Momentos Algébricos, que tem como entrada de dados única a nova representação para imagens 3D. Montada a partir de informações locais de superfície e sem a necessidade de quaisquer técnicas parametrizadas, o formato foi chamado de Matriz RABG e se propõe a resolver o problema das três invariâncias oriundas das transformações afins: translação, escala uniforme e rotação. O arcabouço de extração de atributos herda as propriedades de invariância e é definido em termos modulares, permitindo a construção de qualquer sistema de extração de Momentos Invariantes. Avalia-se o aspecto invariante através de demonstrações e de experimentos empíricos. O poder de descrição do arcabouço de extração proposto é testado a partir de três Funções Momentos que foram aplicadas em um cenário de classificação de indivíduos a partir de faces representadas por nuvens de pontos.

Palavras-chave: Nuvens de pontos. Momentos. Decomposição em Valores Singulares. Polinômios Ortogonais.

ABSTRACT

Three-dimensional images (3D images) allow a more precise study of shapes and surfaces by means of depth information. Compared to two-dimensional images (2D images), 3D images overcome some limitations imposed in some practical scenarios. However, many application areas still depend on very invasive pre-processing steps in 3D images, often requiring a priori parameterization. Such dependencies of 3D images reflect on the feature extraction step, either limiting the options of methods that can be used or by impairing the extracted attributes and, consequently, the obtained results. The contribution of this work is two-fold: a new representation for 3D images calculated from a point cloud and a feature extraction system framework based on Moments that has the aforementioned format as the only input. Assembled from local surface information and without the need for any parameterized techniques, the format was called RABG Matrix and proposes to solve the problem of the three invariances arising from the transformations of translation, uniform scaling and rotation. The feature extraction framework inherits the invariance properties and is defined in modular terms, allowing the construction of any Moment Invariants. The invariant aspect is evaluated through demonstrations and empirical experiments. The description capacity of the feature extraction framework is tested by designing three Moment Functions which were used in a scenario of classification of individuals from faces represented by point clouds.

Keywords: Point clouds. Moments. Singular Value Decomposition. Orthogonal Polynomials.

LISTA DE FIGURAS

Figura 1 – Algumas nuvens de pontos. I: <i>Armadillo</i> , II: <i>Happy Buddha</i> , III: <i>Bunny</i> , IV: <i>Dragon</i>	20
Figura 2 – Visualização do polinômio $R_{nm}(\rho)$: $n \in \{0, 1, 2, 3, 4, 5\}$, $m \in \{0, 1, 2\}$ e $\rho \in [0, 1]$	26
Figura 3 – Visualização das partes reais dos Harmônicos Esféricos $Y_{nm}(\theta, \phi)$ com $n \in [0, 3]$ e $m \in [-3, 3]$	29
Figura 4 – Efeito da Decomposição em Valores Singulares visualizado em uma nuvem de pontos.	33
Figura 5 – Visualização dos ângulos singulares.	39
Figura 6 – Visualização do polinômio $L_k(x)$	54
Figura 7 – Visualização do polinômio $C_k(x)$	56
Figura 8 – Nuvens utilizadas para o teste numérico de invariância.	58
Figura 9 – Exemplos de faces da Base Bosphorus. I: face neutra. II: expressão de raiva. III: expressão de alegria. IV: oclusão de boca.	63
Figura 10 – Visualização das etapas de pré-processamento submetidas às nuvens de faces.	64
Figura 11 – Resultados: Momentos de Legendre.	65
Figura 12 – Resultados: Momentos de Chebyshev.	66
Figura 13 – Resultados: Momentos de Zernike.	67

LISTA DE TABELAS

Tabela 1 – Invariância dos Momentos de Legendre RABG	60
Tabela 2 – Invariância dos Momentos de Chebyshev RABG	61
Tabela 3 – Invariância dos Momentos de Zernike RABG	62

LISTA DE SÍMBOLOS

X	Nuvem de pontos em representação cartesiana
\bar{c}	Centro geométrico (ou centróide) de uma nuvem de pontos
μ	Momento
g	Função de base (ou aproximadora) de um Momento
f	Função alvo (ou de descrição) de um Momento
ρ	Distância radial (sistema de coordenadas polares/esféricas)
θ	Ângulo de abertura (sistema de coordenadas polares/esféricas)
n	Ordem do Polinômio de Zernike e dos Harmônicos Esféricos
m	Repetição do Polinômio de Zernike e dos Harmônicos Esféricos
$R_{nm}(\rho)$	Polinômio Radial de Zernike
$V_{nm}(\rho, \theta)$	Polinômio Complexo de Zernike
$\mathcal{Z}_{nm}(f)$	Momento de Zernike 2D
$Z_{nm}(f)$	Momento Discreto de Zernike 2D
ϕ	Ângulo de elevação (sistema de coordenadas esféricas)
$Y_{nm}(\theta, \phi)$	Harmônicos Esféricos
$\Xi_{nm}(\rho, \theta, \phi)$	Polinômios Esféricos
$\Omega_{nm}(f)$	Momentos Harmônicos Esféricos
U	Matriz geral da Decomposição em Valores Singulares
Σ	Matriz de valores singulares da Decomposição em Valores Singulares
V	Matriz de vetores singulares da Decomposição em Valores Singulares
σ	Valor singular
v	Vetor singular
M	Nuvem de pontos representada em Matriz RABG
\bar{r}	Distância radial normalizada da Matriz RABG
α	Primeiro ângulo singular da Matriz RABG
β	Segundo ângulo singular da Matriz RABG

γ	Terceiro ângulo singular da Matriz RABG
$g(\alpha, \beta, \gamma)$	Função de base do arcabouço de extração de Momentos
$f(\bar{r})$	Função alvo do arcabouço de extração de Momentos
$\mu(M)^{(g,f)}$	Função Momento proposta
Υ	Conjunto de parâmetros da Função Momento Geral proposta
$\mu_{\Upsilon}(M)^{(g,f)}$	Função Momento Geral proposta
$\Lambda(M)_{\Upsilon}^{(g,f)}$	Vetor de Momentos extraídos a partir de uma Função Momento Geral
L_{k+1}	Polinômio recursivo de Legendre
C_{k+1}	Polinômio recursivo de Chebyshev

SUMÁRIO

1	INTRODUÇÃO	15
1.1	Objetivos	17
1.2	Organização deste Trabalho	18
2	FUNDAMENTAÇÃO TEÓRICA	19
2.1	Nuvens de Pontos	19
2.2	Momentos	22
2.2.1	<i>Propriedade da Ortogonalidade</i>	24
2.2.2	<i>Propriedade da Invariância</i>	25
2.2.3	<i>Momentos de Zernike</i>	25
2.2.4	<i>Momentos Harmônicos Esféricos</i>	28
2.3	Decomposição em Valores Singulares	30
2.4	Considerações finais sobre Fundamentação Teórica	33
3	METODOLOGIA	34
3.1	Motivação	34
3.2	Um Novo Formato para Nuvens de Pontos	36
3.2.1	<i>Etapas de geração da Matriz RABG</i>	36
3.3	Invariância da Matriz RABG	38
3.3.1	<i>Invariância à Translação</i>	40
3.3.2	<i>Invariância à Escala</i>	41
3.3.3	<i>Invariância à Rotação</i>	43
3.4	Invariância à rotação no eixo \hat{x}	44
3.4.0.1	<i>Invariância à rotação no eixo \hat{y}</i>	45
3.4.0.2	<i>Invariância à rotação no eixo \hat{z}</i>	46
3.5	Vantagens da Matriz RABG	47
3.6	Um Arcabouço para Extração de Atributos Baseado em Momentos	48
3.6.1	<i>Invariância do arcabouço de extração de atributos</i>	49
3.6.2	<i>Montagem de um vetor de atributos</i>	50
3.7	Vantagens do Extrator de Momentos Proposto	50
3.8	Considerações Finais sobre Metodologia	50
4	RESULTADOS	52

4.1	Funções de Base Escolhidas	53
4.1.1	<i>Sugestão de Momentos de Legendre</i>	53
4.1.2	<i>Sugestão de Momentos de Chebyshev</i>	55
4.1.3	<i>Sugestão de Momentos de Zernike</i>	56
4.2	Teste Numérico de Invariância	57
4.3	Resultados em Tarefas de Classificação	59
4.3.1	<i>Cenário e base de dados</i>	60
4.3.2	<i>Classificador</i>	63
4.3.3	<i>Taxas de classificação</i>	64
4.4	Considerações Finais sobre Resultados	68
5	CONCLUSÕES	69
	REFERÊNCIAS	71

1 INTRODUÇÃO

O ser humano percebe o ambiente ao seu redor através de seus cinco sentidos principais: olfato, paladar, audição, tato e visão. Desses, a visão desempenha um papel majoritário na experiência sensorial. O cérebro humano instintivamente procura por padrões dentro de estímulos visuais, que servem de informação para um grande conjunto de tomadas de decisão.

A curiosidade por modelar a visão em termos de sistemas matemáticos deu início a um série de estudos e pesquisas a partir da década de 1960, formalizando na década de 1970 uma área do conhecimento chamada Visão Computacional (SZELISKI, 2010). O objeto de estudo principal da Visão Computacional é a imagem, a estrutura que armazena a percepção de forma e conteúdo de uma cena (i.e partição) do mundo real.

Além de atenta com as maneiras de adquirir as imagens e os dispositivos eletrônicos utilizados, a Visão Computacional se alia com a estatística e a álgebra a partir do Reconhecimento de Padrões e, juntos, passam a compartilhar técnicas, modelos matemáticos e algoritmos computacionais para a análise de imagens, extraíndo delas significância.

Cada uma dessas três etapas (aquisição, análise e significância) apresenta um conjunto de desafios e limitações próprias. Nenhum sistema de aquisição de imagens é capaz de extrair de uma cena todos os seus detalhes; condições de iluminação, ângulo e posição do dispositivo de captura, além de falhas de natureza aleatória, influenciam na imagem gerada, tornando-a assim uma mera aproximação do real. A etapa de análise deve compensar a diferença entre a imagem aproximada e a real; como medidas de mitigação aplicáveis, pode-se delimitar características de mais importância, filtrar distorções e realçar detalhes. A significância de todo o processo deve estar alinhada com as etapas anteriores, tornando os resultados e as conclusões obtidas consistentes e alinhadas com as perguntas de pesquisa que se deseja responder.

Já é possível aplicar técnicas e modelos bastante robustos no trato de imagens bidimensionais (ou imagens 2D), como segmentação (AL-KOFAHI *et al.*, 2018), classificação (BENOUMINI *et al.*, 2019) e reconstrução (YAMNI *et al.*, 2019). É natural que, com o advento de novas técnicas de aquisição de imagens em novos formatos, também novas técnicas de análise e significância sejam propostas. É o caso da popularização de dispositivos que capturam imagens tridimensionais (ou imagens 3D), dando margem a um novo campo de pesquisa em análise e extração de atributos de cenas do mundo real. Além das vantagens inerentes das imagens bidimensionais em relação às tridimensionais (robustez à baixa iluminação e armazenamento de atributos geométricos, para citar apenas algumas), tarefas como alinhamento (também

conhecido como correção de pose ou registro, (BESL; MCKAY, 1992)), reconhecimento de objetos e odometria são desafios de engenharia que exploram as imagens 3D.

Vários estudos exploraram o uso de Funções Momento (ou apenas Momentos) como extratores de atributo em imagens tridimensionais; classificação de faces humanas, reconstrução de objetos e mapeamento de ambientes estão entre suas aplicações mais comuns (SIQUEIRA *et al.*, 2018; MESBAH *et al.*, 2016; WANG *et al.*, 2019). Entretanto, problemas como ruídos e independência de sistema de aquisição ainda são problemas que abrem espaço para mais pesquisas. Tais problemas são comumente resolvidos com a adição de etapas de pré-processamento nas imagens, que podem ou não serem invasivos (i.e modificam consideravelmente a superfície analisada).

Podemos delimitar um sistema completo de análise de imagens tridimensionais tal como o de qualquer outro experimento de análise exploratória e extração de atributos: definição dos dados de entrada, etapas de pré-processamento, extração de atributos, geração de modelos e avaliação de resultados. O formato dos dados de entrada influencia diretamente os passos de análise seguintes, sobretudo na técnica de extração de atributos e, portanto, nos resultados obtidos.

Entre os formatos de representação de imagens 3D, podemos citar as nuvens de pontos, as malhas triangulares, o cuboide (ou *voxels*) e as imagens de profundidade. Dessas, a nuvem de pontos é considerada a mais simples. Tal simplicidade de aquisição e armazenamento trazem limitações a todas as etapas de análise, sobretudo no pré-processamento.

O processo de análise teórica e prática e permeado por experimentações empíricas, realizado durante esta pesquisa, apontou que os formatos de representação de imagens 3D mais comuns apresentam grande dependência de etapas de pré-processamento para que, só então, possam ser usados como entrada de um modelo de extração de atributos. Neste contexto, podemos apontar como o maior desafio a busca por uma representação normalizada, mas que não imponha outras dependências externas (e.g sistema de aquisição bem controlado, parametrizações externas conhecidas *a priori*, filtragens invasivas) ao sistema de análise de imagens 3D.

Estudos preliminares que serviriam de base para esta dissertação foram inicialmente discutidos no artigo *Revisitando Momentos de Zernike para Reconhecimento de Padrões em Nuvens de Pontos*, apresentado no **XIV Congresso Brasileiro de Inteligência Computacional - CBIC**, em 04 de Novembro de 2019. Até então, pretendia-se resolver o problema das invariâncias na etapa de extração de atributos. O uso de Momentos, especificamente os Momentos Ortogonais

de Zernike, mostrou grande capacidade descritiva em problemas de classificação, mas para alcançar a invariância à rotação era necessário uma formulação com capacidade descritiva reduzida.

Iniciou-se então uma investigação em busca de conceber uma nova representação de imagens 3D, especificamente a partir de uma nuvem de pontos, que resolve os problemas de normalização espacial (translação, escala e rotação) e que seja montada a partir apenas de informações presentes na superfície (i.e natureza local), sem quaisquer dependências com sistema de aquisição, e que pudesse servir de entrada para um sistema de extração de atributos capaz de gerar vetores de atributos também invariantes.

Partindo das nuvens de pontos, este trabalho propõe um novo formato para análise de imagens 3D. Essa nova representação pretende resolver três problemas comuns em análise de imagens tridimensionais, todos relacionados ao pré-processamento espacial das superfícies: normalização da localização espacial, normalização do tamanho/escala e orientação espacial. Cada um desses desafios está relacionado às transformações geométricas de translação, escalamento e rotação, respectivamente, chamadas de transformações afins (STRANG, 2016).

De posse do novo formato, um sistema de extração de atributos se faz necessário para uso em vários cenários de Reconhecimento de Padrões, como, por exemplo, classificação. O sistema proposto é baseado em Momentos, entes algébricos que relacionam o grau de semelhança entre duas funções. O sistema de extração de atributos é invariante às transformações afins (por usar o formato invariante como entrada) e pretende ser modular, de forma que a proposta de extrator de atributos corresponde a um arcabouço de onde se pode montar facilmente qualquer função de extração de atributos.

1.1 Objetivos

O primeiro objetivo deste trabalho é propor um novo formato de representação para nuvens de pontos que agrupa atributos geométricos invariantes às transformações de translação, escalamento uniforme e rotação. Chamado de Matriz RABG, esse formato pode ser visto como uma representação geometricamente normalizada a partir de atributos de natureza radial e angular.

O segundo objetivo é apresentar um arcabouço de sistema de extração de atributos e montagem de vetor de atributos baseado em Momentos que tenha como entrada uma nuvem de pontos no formato proposto pelo primeiro objetivo. A redução de dimensionalidade, ou

seja, a diminuição do espaço do problema e retenção das características mais descritivas é um desafio presente em problemas de Reconhecimento de Padrões; mostramos que as características invariantes do formato são herdadas pelo sistema de extração de atributos.

São objetivos adicionais: avaliar numericamente a invariância do formato e dos Momentos calculados a partir do sistema de extração de atributos e experimentar o poder de descrição do método com um estudo de caso ilustrativo em um cenário de classificação de faces humanas.

1.2 Organização deste Trabalho

Este trabalho está dividido da seguinte maneira. No Capítulo 2 são apresentados todas as ferramentas matemáticas necessárias para a compreensão da proposta: definição formal de nuvem de pontos, conceituação de Momentos e Decomposição em Valores Singulares. No Capítulo 3, o método é apresentado em termos de definições, demonstrações e um algoritmo. O Capítulo 4 traz os resultados de experimentação: análise de invariância e taxas de classificação do cenário de teste. Por fim, propomos uma discussão final do método, suas limitações e sugestões de novas pesquisas no Capítulo 5.

2 FUNDAMENTAÇÃO TEÓRICA

Para alcançarmos o formato de representação proposto e o sistema de extração de atributos, iniciaremos por expor neste capítulo a teoria necessária para o entedimento, avaliação e crítica do método. Os conceitos que servem de base para a contribuição deste trabalho são: Nuvens de Pontos e Decomposição em Valores Singulares (formato de imagem tridimensional) e Momentos (arcabouço de sistema de extração de atributos).

2.1 Nuvens de Pontos

Imagens tridimensionais (ou imagens 3D) são imagens que armazenam informação de profundidade, também chamada espacial ou geométrica. Elas podem ser representadas em diversos formatos, como mapas de profundidade, malhas triangulares, cubóides (ou *voxels*) e nuvens de pontos. Os sensores capazes de capturar imagens 3D, como os de infravermelho ou os de tempo-de-vôo, têm se popularizado em ritmo acelerado graças ao barateamento de sua produção, facilitando a aquisição e análise desse tipo de informação. Independente da estrutura de dados escolhida, o uso de imagens 3D apresenta vantagens sobre as imagens bidimensionais (2D), como maior robustez a cenários de baixa iluminação e armazenamento de informações geométricas de superfície.

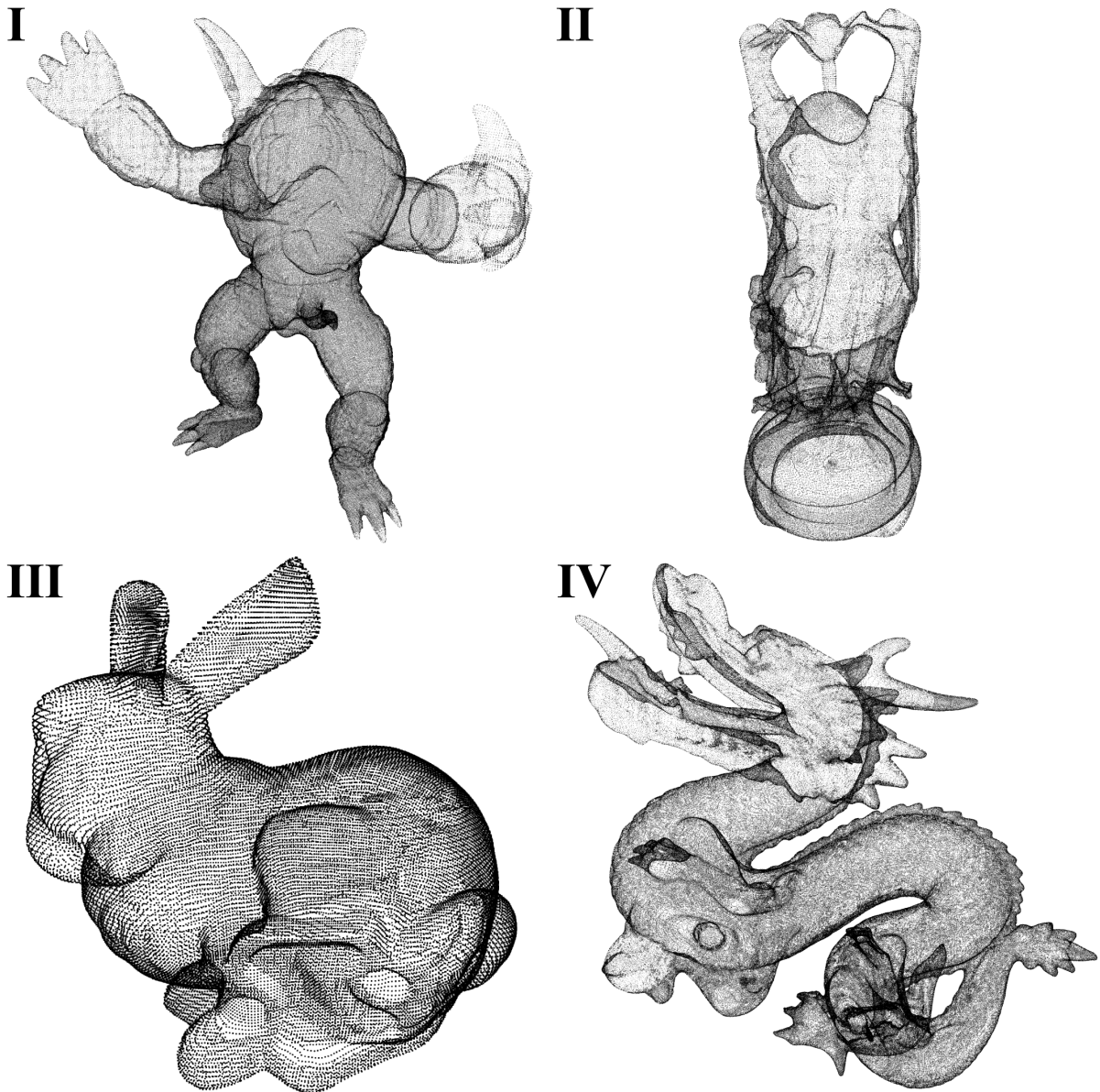
Podemos citar o uso de imagens tridimensionais em duas áreas de grande interesse humano. Na medicina se destaca o uso de sensores de ultra-som e de aquisição bidimensional para construção de modelos 3D baseado em triangularização com o objetivo de visualizar órgãos e outros elementos constitutivos do corpo humano (BELLE *et al.*, 2017). Na produção de alimentos, o uso de veículos aérios não-tripulados equipados com sensores de profundidade ajudam a captar imagens de plantações para ajudar no controle do crescimento de gêneros agrícolas e na previsão de colheitas (SONG *et al.*, 2020).

A escolha do formato de representação dependerá da aplicação em questão e trará vantagens e limitações no decorrer da análise e investigação que se pretende realizar com imagens 3D. O objeto de estudo deste trabalho é formato em nuvens de pontos.

A Nuvem de Pontos, um dos tipos mais comuns de formato de imagens tridimensionais, possui vantagens em relação a outras estruturas de representação 3D. Por ser definida como um lista não ordenada de pontos no \mathbb{R}^3 , não há necessidade de armazenamento de conexões entre vértices (como no caso das malhas triangulares), ou dos centróides e limites de um estrutura

volumétrica (como nos *voxels*). Dizemos, portanto, que as nuvens de pontos são estruturas que não guardam informação topológica (KOBBELT; BOTSCH, 2004). As nuvens de pontos são as estruturas mais simples que se pode escolher para representar uma imagem 3D. A Figura 1 traz exemplos de nuvens de pontos retiradas da base de dados da Universidade de Stanford (LEVOY *et al.*, 2005).

Figura 1 – Algumas nuvens de pontos. I: *Armadillo*, II: *Happy Buddha*, III: *Bunny*, IV: *Dragon*



Fonte: Levoy *et al.* (2005), adaptado pelo autor.

Matematicamente, podemos definir uma nuvem de pontos como uma matriz

$$X = \begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ \vdots & \vdots & \vdots \\ x_N & y_N & z_N \end{bmatrix}, \quad (2.1)$$

onde X é uma nuvem com N pontos, cada qual disposto em uma das linhas, também chamada de representação cartesiana. A representação de pontos por coluna, ou X^T , também é possível.

Uma entidade de grande interesse na análise de nuvens de pontos é o centro geométrico (ou centróide). O centro geométrico representa um lugar no espaço que descreve localmente a posição de um conjunto de pontos e é facilmente calculado como

$$\bar{c} = [\bar{x} \quad \bar{y} \quad \bar{z}] = \frac{1}{N} \sum_{i=1}^N [x_i \quad y_i \quad z_i]. \quad (2.2)$$

A resolução do sistema de aquisição indica a quantidade de pontos usados para representar a cena/superfície. Chamamos de densa uma nuvem com muitos pontos e esparsa aquela com poucos pontos. Essa categorização é bastante subjetiva e normalmente usada como ferramenta comparativa entre duas representações de uma mesma cena.

As etapas de pré-processamento mais comuns realizadas em nuvens de pontos são a de subamostragem (remoção de pontos, tornando a nuvem mais esparsa), preenchimento (adição de pontos, tornando a nuvem mais densa) remoção de *outliers* (pontos distantes da superfície principal ou pontos ruidosos) e alinhamento (correção da orientação em função de um sistema de direções de referência). Cada etapa de pré-processamento é responsável, em maior ou menor grau, por alterar a forma e posição da superfície, tentando descartar redundâncias e preservando características geométricas de maior significância.

Após ser pré-processada, a nuvem de pontos pode, finalmente, passar por uma etapa de extração de atributos. O objetivo aqui é promover uma transformação baseada em redução da dimensionalidade (i.e quantidade de dados) que preserve características de maior relevância. Chamamos de vetor de atributos o conjunto de valores calculados a partir de um conjunto de dados inicial; o vetor de atributos possui um tamanho consideravelmente reduzido em comparação à quantidade de valores correspondente aos dados iniciais. O método de extração de atributos aplicado neste trabalho é baseado no cálculo dos chamados Momentos.

2.2 Momentos

A extração de atributos é um processo de redução de dimensionalidade sobre um conjunto de dados que gera um novo conjunto de menor tamanho e mais facilmente processável. Espera-se que tal transformação conserve características que continuem a descrever bem o conjunto de dados original, mas simplificando as demais etapas de tratamento e armazenamento. A fase de extração de atributos está presente em vários modelos de classificação aplicados em Visão Computacional. São exemplos de extração de atributos nesse último: cálculo de curvaturas, detecção de bordas, análise de frequência e cálculo de Momentos.

Momentos não surgem como ferramenta de extração de atributos, mas como elemento constitutivo da análise de vários corpos algébricos, principalmente na área da Topologia, onde se destacam os estudos do matemático alemão Felix Hausdorff (HAUSDORFF, 1921). Já para o Reconhecimento de Padrões, a extração de atributos avalia um conjunto de dados (expressos por uma função) e dele extraem valores que o descreve. Existem duas perspectivas matemáticas que aquela área enxerga e aplica Momentos: a perspectiva estatística e a perspectiva algébrica.

Dada uma variável aleatória descrita por uma certa função densidade de probabilidade, dizemos que um Momento dessa variável aleatória representa uma característica da forma tomada por sua distribuição ao redor de um valor referencial; esse é o aspecto estatístico dos Momentos. Expresso em termos de ordens, podemos caracterizar os Momentos Estatísticos como sendo a probabilidade total (Momento de ordem 0, centrado em zero), a média (Momento de ordem 1, centrado em zero), a variância (Momento de ordem 2, centrado na média) a assimetria (Momento de ordem 3, centrado na média) e curtose (Momento de ordem 4, centrado na média). O Momento Estatístico de ordem n de uma variável aleatória x descrita pela função densidade de probabilidade $f(x)$ em torno de um valor de referência c é expresso por

$$\mu_n(x) = \int_{-\infty}^{\infty} (x - c)^n f(x) dx, \quad (2.3)$$

tal que c é comumente escolhido como a média da distribuição de dados (BUSSAB; MORETTIN, 2010).

O caráter estatístico dos Momentos, portanto, simboliza a extração de características de um conjunto de dados a partir do próprio conjunto de dados. Todavia em outros casos deseja-se obter características de natureza comparativa, ou seja, dada uma função, pretende-se encontrar valores de semelhança desta com uma outra; esse é o aspecto algébrico dos Momentos.

Algebricamente, os Momentos são escalares que exprimem a semelhança entre duas funções: uma função de base g (também chamada função aproximadora) e uma função alvo f (também chamada função de descrição). Essa semelhança pode ser vista como a "projeção" ou "sobreposição" de f sobre g (FLUSSER *et al.*, 2009).

A semelhança entre duas funções pode ser calculada através da integral de sobreposição ou produto interno funcional. Dada duas funções $\{(f(x), g(x)) \mid x \in K\}$ definidas em um corpo qualquer K , expressamos o Momento Algébrico μ como

$$\mu = \langle g, f \rangle = \int_a^b g(x)f(x)dx. \quad (2.4)$$

A Equação 2.4 expressa uma formulação geral, ou Função Momento. As funções g e f podem ser definidas em qualquer dimensão; por exemplo, a função que representa uma imagem pode ser escrita como um $f(x, y)$, enquanto a função que representa uma superfície tridimensional como um $f(x, y, z)$. Em outras palavras, um Momento, do ponto de vista algébrico, é uma função que retorna um escalar que representa a semelhança entre duas funções definidas em qualquer corpo de qualquer dimensão. Este trabalho se atém aos Momentos Algébricos que serão chamados daqui por diante apenas de Momentos.

Por exemplo, podemos associar a distribuição de intensidades de *pixels* de uma imagem bidimensional em tons de cinza $f(x, y) \in [0, 255]$ de dimensões $W \times H$ com uma função de potência $g_{pq}(x, y) = x^p y^q$, tal que

$$\mu_{pq} = \langle g, f \rangle = \int_W \int_H g_{pq}(x, y)f(x, y)dxdy = \int_W \int_H x^p y^q f(x, y)dxdy. \quad (2.5)$$

Da Equação 2.5, vemos que, ao varrermos a imagem, acumulamos, *pixel a pixel*, uma contribuição de posição (coordenadas do *pixel*, representada pela função $g(x, y)$) e uma contribuição de intensidade (o tom do *pixel* naquela posição, representado pela função $f(x, y)$). Por levar em consideração informações de posição e intensidade, dizemos que Momentos representam informação de forma (GONZALEZ; WOODS, 2009). A Equação 2.5 é chamada de Momento Geral ou Momento de Imagem e foi introduzida por Ming-Kuei Hu (1962), cujo trabalho pode ser apontado como a gênese do uso prático de Momentos em Visão Computacional.

Momentos podem ser divididos em Geométricos e Ortogonais. Essa categorização está associada à natureza das funções de base usadas na construção da Função Momento. Momentos Geométricos são aqueles cuja função de base é função de potência $g_{pq} = x^p y^q$,

apresentada anteriormente. O conceito de Momento Geométrico foi estendido posteriormente por Xu e Li (2008) para descrever Momentos escritos a partir de primitivas geométricas tais como distância euclidiana, área, produto interno e volume. Um Momento é dito Ortogonal quando é construído a partir de um conjunto de funções de base que gozam da propriedade da ortogonalidade (FLUSSER *et al.*, 2009). Além da ortogonalidade, outra característica bastante desejável e que é almejada pelos Momentos Ortogonais é a invariância.

A discussão que segue neste trabalho focará nos Momentos Ortogonais e nas características que podem levá-los a serem invariantes. A seguir, definiremos matematicamente as propriedades de ortogonalidade e invariância. Depois, mostraremos dois exemplos de Momentos, um bidimensional (Momentos de Zernike) e outro tridimensional (Momentos Harmônicos Esféricos), que serviram de inspiração tanto para o formato de nuvem de pontos quanto para a definição do extrator de atributos propostos.

2.2.1 Propriedade da Ortogonalidade

Seja $\Psi = \{\psi_i \mid i \in \mathbb{N}\}$ um conjunto de funções de base de uma certa Função Momento μ . As funções desse conjunto são ortogonais entre si caso

$$\langle \psi_i, \psi_j \rangle = \omega \delta_{i,j}, \quad (2.6)$$

onde $\psi_i, \psi_j \in \Psi$, ω é um escalar definido no mesmo corpo que $\langle \psi_i, \psi_j \rangle$ e chamado de constante de normalização, enquanto que $\delta_{i,j}$ é o operador Delta de Kronecker definido como

$$\delta_{i,j} = \begin{cases} 0, & \text{se } i \neq j, \\ 1, & \text{se } i = j. \end{cases} \quad (2.7)$$

A ortogonalidade está associada ao conceito de independência, ou baixa sobreposição dos dados. Note que, quando $i = j$ na Equação 2.6, temos o caso especial em que uma função é projetada sobre ela mesma; isto implica que a integral de produto interno dada pela Equação 2.4 leva a magnitude ao quadrado da função ψ_i em questão. É por esta razão que, quando $\omega = 1$, fala-se em ortonormalidade.

2.2.2 Propriedade da Invariância

Seja $\mathcal{F}(f)$ um atributo calculado a partir de uma função f e seja H um grupo de transformações. Dizemos que \mathcal{F} é um descritor invariante no espaço de H se

$$\mathcal{F}(hf) = \mathcal{F}(f), \quad (2.8)$$

para $\forall h \in H$. As transformações mais relevantes no contexto de extração de atributos em imagens, sejam bidimensionais ou tridimensionais, são translação, escalamento uniforme e rotação, chamadas transformações afins. Semanticamente, uma imagem que passa por algumas dessas transformações não se altera sobremaneira, portanto os atributos dela calculados devem refletir essa invariância.

2.2.3 Momentos de Zernike

Diversas funções de base podem ser escolhidas para gerar Momentos Ortogonais. Entretanto, alguns se destacam por sua influência teórico-prática nos campos de Visão Computacional. Entre os Momentos Ortogonais, destacaremos os chamados Momentos de Zernike, introduzidos por Teague (1980) que definiu um Momento cujas funções de base são construídas a partir de um conjunto de polinômios ortogonais descritos por Frits Zernike (1934) para análise e modelagem de aberrações ópticas e fenômenos de difração.

O estudo desses Momentos ainda encontra aplicações práticas na literatura recente. O reconhecimento de impressões digitais (KAUR; PANNU, 2019) e a classificação robusta de formas geométricas (ABBAS *et al.*, 2019) são exemplos de uso dos Momentos de Zernike no domínio de imagens bidimensionais.

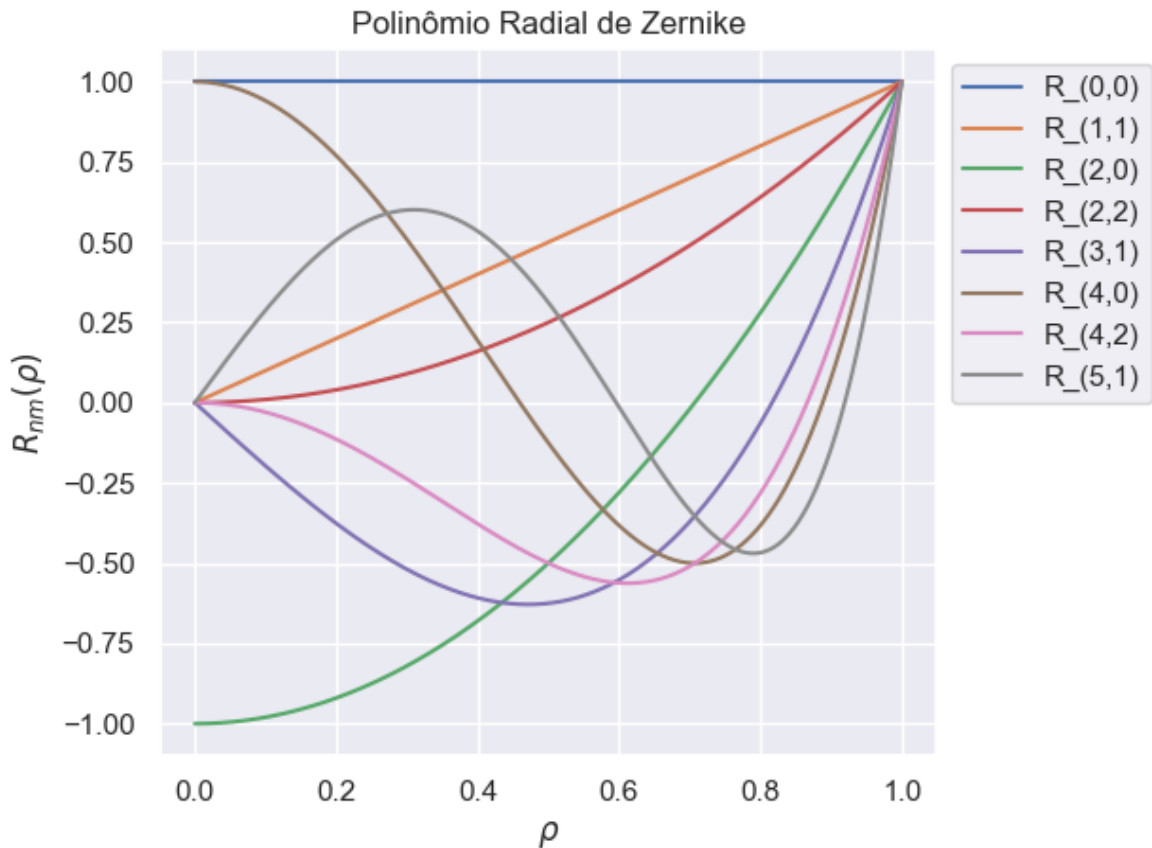
Os Momentos de Zernike são definidos dentro do disco de raio unitário, mais precisamente no corpo complexo \mathbb{C} . Para imagens bidimensionais, assume-se uma região circular, centrada no centro geométrico das intensidades dos *pixels* da imagem, da qual se extraem as distâncias radiais ρ (entre um certo *pixel* e o centróide) e os ângulos θ que o vetor direção dessas distâncias faz com o eixo horizontal do sistema de coordenadas; em outras palavras, uma conversão para um sistema de coordenadas polares. A função de base é construída a partir de uma parte radial (o Polinômio Radial de Zernike, de argumento ρ) e uma parte angular (exponencial complexa, de argumento θ).

A parte radial da função de base, o Polinômio Radial de Zernike $R_{nm}(\rho)$, é escrito como

$$R_{nm}(\rho) = \sum_{s=0}^{\frac{(n-|m|)}{2}} (-1)^s \frac{(n-s)!}{s! \left(\frac{n+|m|}{2} - s\right)! \left(\frac{n-|m|}{2} - s\right)!} \rho^{n-2s}, \quad (2.9)$$

onde ρ é a distância entre o *pixel* (i, j) e o centróide da imagem, n é a ordem do polinômio e m é a repetição do polinômio. n e m são parâmetros inteiros não-negativos que satisfazem as condições $n \geq |m| \geq 0$ e $n - |m|$ sempre par. A Figura 2 mostra a forma tomada por $R_{nm}(\rho)$ para diferentes n , m e ρ .

Figura 2 – Visualização do polinômio $R_{nm}(\rho)$: $n \in \{0, 1, 2, 3, 4, 5\}$, $m \in \{0, 1, 2\}$ e $\rho \in [0, 1]$



Fonte: O autor.

Como comentado anteriormente, a função de base é formada a partir de uma parte radial e uma parte angular. Definimos a função de base $V_{nm}(\rho, \theta)$, chamado Polinômio Complexo

de Zernike (ou apenas Polinômio de Zernike), como o produto

$$V_{nm}(\rho, \theta) = R_{nm}(\rho)e^{jm\theta}, \quad (2.10)$$

tal que a parte angular é $e^{jm\theta}$, a base ortogonal da Expansão por Série de Fourier.

Usando a notação da Seção 2.2.1, dizemos que o conjunto de funções de base dos Momentos de Zernike bidimensionais é definido como

$$\Psi_{zernike} = \{V_{nm}(\rho, \theta) \mid \rho \in [0, 1], \theta \in [0, 2\pi), n \geq |m| \geq 0, (n - |m|)_{par}\}. \quad (2.11)$$

Finalmente, a Função Momento de Zernike 2D $\mathcal{Z}_{nm}(f)$ é definida como o produto interno (Equação 2.4) entre $V_{nm}(\rho, \theta)$ e a imagem $f(x, y)$ (intensidade de cor do pixel numa dada coordenada (x, y)) de dimensões $k \times l$:

$$\mathcal{Z}_{nm}(f) = \langle V_{nm}(\rho, \theta), f(x, y) \rangle = \iint_{k,l} V_{nm}(\rho_{k,l}, \theta_{k,l}) f(x_k, y_l) dx dy. \quad (2.12)$$

Teague (1980) mostrou que as funções $\mathcal{Z}_{nm}(f)$ são ortogonais com constante $\omega = \frac{\pi}{n+1}$. A partir dessa análise, esse autor derivou a forma discreta dos Momentos de Zernike 2D para uma dada imagem de dimensões (W, H) como

$$Z_{nm}(f) = \frac{n+1}{\pi} \sum_{k=0}^{W-1} \sum_{l=0}^{H-1} V_{nm}^*(\rho_{k,l}, \theta_{k,l}) f(x_k, y_l), \quad (2.13)$$

tal que $*$ é o operador complexo conjugado.

Como $Z_{nm}(f)$ é definido no conjunto dos número complexos \mathbb{C} , temos liberdade para usar diferentes combinações de valores como atributos: apenas a parte real $\text{Re}(Z_{nm}(f))$, apenas a parte imaginária $\text{Im}(Z_{nm}(f))$, a sua magnitude $|Z_{nm}(f)|$ ou mesmo o próprio valor complexo e quaisquer outras combinações de adição ou multiplicação. Os Momentos de Zernike 2D Discretos são invariantes às transformações de translação e escalamento uniforme. É possível alcançar a invariância à rotação apenas com o uso da magnitude $|Z_{nm}(f)|$ (NOVOTNI; KLEIN, 2004).

O primeiro trabalho a descrever uma generalização dos Momentos de Zernike para o domínio 3D foi proposto por Canterakis (1999). Entretanto, esse autor abordou apenas aspectos

teóricos e não experimentou, por exemplo, a formulação de uma expressão discreta. A seguir, é apresentado o estudo de funções de base radiais e angulares para uso em Reconhecimento de Padrões em imagens tridimensionais através do uso de um conjunto de funções especiais chamadas Harmônicos Esféricos.

2.2.4 Momentos Harmônicos Esféricos

Chamamos de Harmônicos Esféricos um conjunto de funções especiais definidas na superfície da esfera unitária e introduzidas por Pierre Simon de Laplace em 1782. Sejam $Y_{nm}(\theta, \phi)$ os harmônicos de ordem n e repetição m , temos

$$Y_{nm}(\theta, \phi) = \sqrt{\frac{2n+1}{4\pi} \frac{(n-m)!}{(n+m)!}} e^{jm\theta} P_{nm}(\cos \phi), \quad (2.14)$$

onde P_{nm} são os Polinômios Associados de Legendre (DAVE; ARMSTRONG, 1970).

A Figura 3 mostra alguns harmônicos ao variarmos os parâmetros (n, m) e os argumentos (θ, ϕ) .

Novotni e Klein (2004) expandiram o estudo de Canterakis (1999) e propuseram um Momento cuja função de base é, assim como os Momentos de Zernike, composta de uma parte radial (o Polinômio Radial de Zernike, Equação 2.9) e uma parte angular (os Harmônicos Esféricos, Equação 2.14). Chamada de Polinômio Esférico, essa função de base é representada neste trabalho como

$$\Xi_{nm}(\rho, \theta, \phi) = R_{nm}(\rho) Y_{nm}(\theta, \phi). \quad (2.15)$$

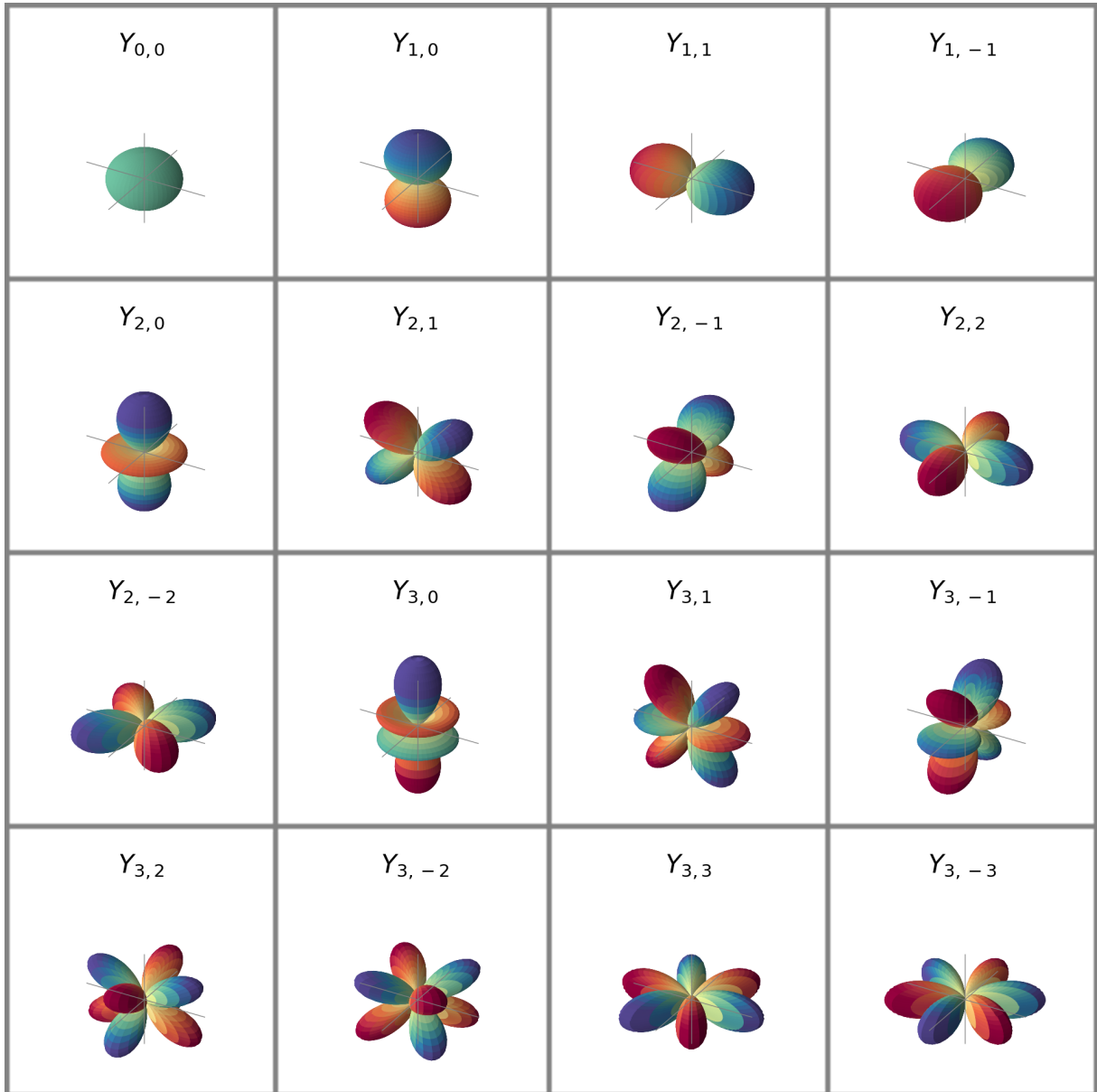
O conjunto de funções de base dos Momentos Harmônicos Esféricos são expressas por

$$\Psi_{harm} = \{\Xi_{nm}(\rho, \theta, \phi) \mid \rho \in [0, 1], \theta \in [0, 2\pi), \phi \in [0, \pi], n \geq |m| \geq 0, (n - |m|)_{par}\}. \quad (2.16)$$

Aplicando-se o produto interno da Equação 2.4 entre Ξ_{nm} e um função característica de um imagem tridimensional $f(x, y, z)$, chegamos à fórmula dos Momentos Harmônicos Esféricos

$$\Omega_{nm}(f) = \langle \Xi_{nm}(\rho, \theta, \phi), f(x, y, z) \rangle = \iiint_{k,l,q} \Xi_{nm}^*(\rho_{k,l,q}, \theta_{k,l,q}, \phi_{k,l,q}) f(x_k, y_l, z_q) dx dy dz. \quad (2.17)$$

Figura 3 – Visualização das partes reais dos Harmônicos Esféricos $Y_{nm}(\theta, \phi)$ com $n \in [0, 3]$ e $m \in [-3, 3]$



Fonte: O autor.

É fácil ver que, assim como os Momentos de Zernike, os Momentos Harmônicos Esféricos também são definidos no conjunto dos número complexos \mathbb{C} . Aqui também cabe a escolha de quais valores usar como atributos extraídos: parte real $\text{Re}(\Omega_{nm}(f))$, imaginária $\text{Im}(\Omega_{nm}(f))$, magnitude $|\Omega_{nm}(f)|$ ou mesmo o próprio valor complexo e quaisquer outras combinações de adição ou multiplicação. Todas as partes dos Momentos Harmônicos Esféricos são invariantes às transformações de translação, escalamento uniforme e rotação.

Entretanto, por serem definidos em um sistema de coordenadas esféricas, alcançar a propriedade de invariância completa dos Momentos Harmônicos Esféricos depende da definição *a priori* do sistema referencial 3D. Em termos práticos, supõe-se que o referencial adotado pelo

sistema de aquisição é fixo e padronizado ao longo de toda a cena. A maioria dos Momentos tridimensionais, não apenas os Harmônicos Esféricos, supõe a base canônica do \mathbb{R}^3 , tal que o eixo $\hat{z} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T$ deva ser o eixo de profundidade (direção oposta aonde aponta a câmera).

Entretanto, para casos mais gerais e, em especial, para situações nas quais o método de extração de atributos deva ser independente do sistema de aquisição, opta-se pela escolha de um sistema local de referência. Assim, os pontos da superfície podem ser reescritos em função de um sistema local de coordenadas; esse é o chamado problema do alinhamento ou correção de orientação.

Dado que a nuvem de pontos deve ter sua orientação normalizada e que os métodos de extração de atributos baseado em Momentos não conseguem preservar caráter descritivo e invariância completa ao mesmo tempo, propomos uma representação de nuvens de pontos que se traduz em uma normalização espacial afim. A contribuição do trabalho como um todo será discutida com mais detalhes no Capítulo 3; por hora apresentamos o conceito chave usado para alcançar a invariância à rotação: Decomposição em Valores Singulares.

2.3 Decomposição em Valores Singulares

Chamamos de decomposição (ou fatorização) de matrizes todo método que representa uma matriz como o produto de duas ou mais outras matrizes. Métodos de decomposição são valiosos pois as matrizes do produto resultante costumam carregar características importantes da matriz original.

Além disso, sendo K um corpo e A uma matriz de dimensão $m \times n$, com entradas em K , A é considerada uma transformação linear T de K^n em K^m dada

$$T(\vec{x}) = A\vec{x}, \tag{2.18}$$

em que \vec{x} é um vetor coluna com n linhas e os elementos de A são os coeficientes da transformação T . Neste caso, as propriedades de T podem ser separadas e analisadas individualmente a partir de uma decomposição de A .

Os autovalores e autovetores de uma matriz são duas propriedades bastante utilizadas em análise de grandes volumes de dados. A motivação por encontrar esses dois atributos parte da intuição de que toda transformação linear pode ser vista como uma composição entre um escalamento e uma rotação. Os autovalores são escalares que representam a parcela de

escalamento da transformação, já os autovetores são as direções de rotação. Seja uma matriz A de dimensão $n \times n$, um autovalor λ_i de A é um escalar cuja multiplicação pelo autovetor correspondente u_i representa o mesmo que calcular Au_i , ou seja:

$$Au_i = \lambda_i u_i. \quad (2.19)$$

Um exemplo de decomposição envolvendo autovalores e autovetores ocorre se A for uma matriz $n \times n$ simétrica. Nessas condições, A pode ser decomposta no produto

$$A = PDP^T = \begin{bmatrix} - & p_1 & - \\ - & p_2 & - \\ & \vdots & \\ - & p_n & - \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} \begin{bmatrix} | & & & | \\ p_1 & p_2 & \dots & p_n \\ | & & & | \end{bmatrix}, \quad (2.20)$$

chamado *autodecomposição de A*. A matrix P é ortogonal, de dimensão $n \times n$, e suas colunas são os autovetores de A . A matrix D , também $n \times n$, é diagonal e composta pelos autovalores de A . Portanto, a autodecomposição permite separar o caráter de escala e o de rotação de uma matriz, contanto que ela seja simétrica.

A autodecomposição é uma ferramenta de análise muito útil. Entretanto, a maioria dos conjuntos de dados não apresenta as condições necessárias para sua aplicação. Uma solução possível é a Decomposição em Valores Singulares, ou DVS, que pode ser descrita como uma generalização da fatorização em autovalores e autovetores e que pode ser aplicada em matrizes de qualquer dimensão, não somente matrizes quadradas.

A chave para a DVS está no fato de o produto da transposta de uma matriz por ela mesma resulta em uma matriz simétrica. Se aplicarmos a autodecomposição na matriz resultante, podemos extrair escalares e vetores que se relacionam aos autovalores e autovetores da matriz original.

Seja uma matriz B de dimensão $m \times n$. B pode ser representada como o produto entre três matrizes $U_{m,m}$, $\Sigma_{m,n}$ e $V_{n,n}$, tal que

$$B^T B = U \Sigma V^T = \begin{bmatrix} b_{1,1} & b_{1,2} & \dots & b_{1,m} \\ b_{2,1} & b_{2,2} & \dots & b_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m,1} & b_{m,2} & \dots & b_{m,m} \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_{m,n} \end{bmatrix} \begin{bmatrix} | & & & | \\ v_1 & v_2 & \dots & v_n \\ | & & & | \end{bmatrix}. \quad (2.21)$$

A matriz U é quadrada (e unitária, se for definida no corpo dos complexos) e representa uma rotação. Σ é uma matriz diagonal de valores reais e não negativos $\{\sigma_1, \sigma_2, \dots, \sigma_m \mid \sigma_i \geq \sigma_{i+1}\}$ chamados *valores singulares* e representa uma transformação de escalamento. Já V^T é uma matriz ortornormal que guarda em suas linhas $\{v_1, v_2, \dots, v_n\}$ os eixos (ou direções) de dispersão da matriz original B , denominados *vetores singulares*. Cada valor singular σ_i representa o quão disperso estão os pontos da nuvem ao longo do vetor singular correspondente v_i .

Dependendo da dimensão da matriz $B^T B$, sua DVS poderá ser bastante custosa em tempo de processamento computacional e armazenamento em memória, além de propícia a erros de precisão numérica. É muito comum aplicar uma representação compacta da Decomposição em Valores Singulares que reduz o espaço de dados para um valor igual ao posto r da matrix B . É possível mostrar que os valores singulares σ_1 a σ_r são não nulos, e os demais iguais a zero. A versão compacta da DVS pode ser escrita tal como o produto da Equação 2.21, mas agora com matrizes de dimensão $U_{m,r}$, $\Sigma_{r,r}$ e $V_{r,r}$, tal que

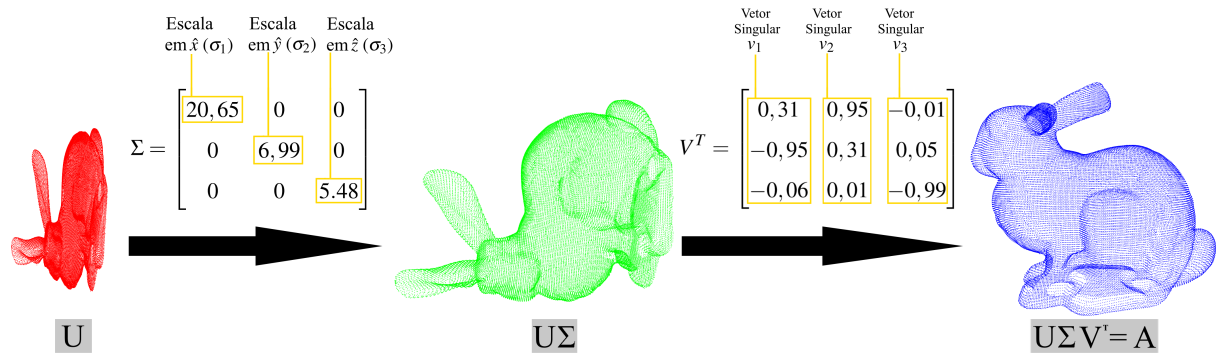
$$M = U\Sigma V^T = \begin{bmatrix} b_{1,1} & b_{1,2} & \dots & b_{1,r} \\ b_{2,1} & b_{2,2} & \dots & b_{2,r} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m,1} & b_{m,2} & \dots & b_{m,r} \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_r \end{bmatrix} \begin{bmatrix} | & | & & | \\ v_1 & v_2 & \dots & v_r \\ | & | & & | \end{bmatrix}. \quad (2.22)$$

O significado de cada uma das matrizes da DVS muda com a versão compacta. Para ilustrar essas ressignificações, ilustraremos o efeito da Decomposição em Valores Singulares em uma nuvem de pontos para se ter uma interpretação mais intuitiva do significado das matrizes U , Σ e V (Figura 4). Vamos tomar a nuvem de pontos de um coelho, tal como o mostrado na Figura 1, e representado por uma matriz A . Ao decompor A segundo a Equação 2.22, a matriz U está representada pelo coelho em vermelho. De mesma dimensão que a nuvem original ($N \times 3$), a matriz U guarda semelhanças com a superfície original, a menos de um efeito de escala e um de rotação. Uma vez que promovemos o produto US , estamos escalando esta superfície mais simples ao longo dos eixos de dispersão dos dados. No exemplo, vemos que a contribuição da direção \hat{x} é a maior que em \hat{y} e \hat{z} ; o coelho em verde possui as mesmas dimensões que o coelho original. Por fim, a correção de orientação é feita pelo produto final USV^T . Como V^T é uma matriz ortornormal, esta representa uma transformação de rotação, resultando no coelho original, em azul.

Este trabalho aborda o uso da versão compacta da Decomposição em Valores Singulares. Daqui por diante, toda menção a essa técnica ou a sua sigla farão referência à versão compacta.

O método proposto neste trabalho faz uso dos vetores singulares da matriz V da DVS como base do referencial local à nuvem de pontos. Como as vetores singulares formam uma base ortonormal e local à nuvem (i.e uma base do \mathbb{R}^3 com origem no centro geométrico da nuvem), podemos calcular ângulos em relação a estes que se preservam mesmo depois de operações de rotação. Este é o *rationale* para a representação aqui trazida como contribuição.

Figura 4 – Efeito da Decomposição em Valores Singulares visualizado em uma nuvem de pontos.



Fonte: O autor.

2.4 Considerações finais sobre Fundamentação Teórica

Neste capítulo é apresentado o referencial teórico que serviu de base para a contribuição deste trabalho. Discutimos a teoria de Momentos como extratores de atributos e alguns de seus derivados, os Momentos de Zernike e os Momentos Harmônicos Esféricos. Por fim, a Decomposição em Valores Singulares foi apresentada como solução para a definição de um referencial local à nuvem de pontos que depende apenas da informação contida em sua superfície, etapa chave para alcançarmos um formato de representação (e, por consequência, um extrator de atributos) invariante.

3 METODOLOGIA

Além da exposição da teoria, foram discutidos no Capítulo 2 alguns aspectos que alimentaram as reflexões durante a pesquisa na busca por solucionar o problema das invariâncias sem comprometer a etapa de extração de atributos.

O primeiro objetivo deste trabalho é um formato de representação de nuvens de pontos invariante às transformações afins (translação, escala e rotação). Alcançaremos tal representação a partir de um algoritmo que recebe como entrada uma nuvem de pontos e que gera como resultado uma matriz que agrupa atributos geométricos invariantes.

O segundo objetivo deste trabalho explora o formato de representação (primeiro objetivo) no contexto de redução de dimensionalidade: construir um sistema de extração de atributos baseado em Momentos, o qual necessita apenas a formulação geral da integral de sobreposição (Equação 2.4) e que herda invariância afim completa por ter como única entrada o formato de nuvem proposto.

Inspirados pelos Momentos apresentados nas Seções 2.2.3 e 2.2.4, Momentos de Zernike e Momentos Harmônicos Esféricos, encontramos a chave para a contribuição deste trabalho como um todo: a conjugação de uma informação de natureza radial e uma informação de natureza angular, seja no formato, seja no sistema de extração. Iniciamos com a motivação que levou à concepção das propostas aqui apresentadas. Depois, apresentaremos o algoritmo de geração do formato de matriz proposto. Em seguida, delimitaremos uma formulação geral de extração de Momentos que faz uso daquele formato e apresentaremos como definir um vetor de atributos a partir do cálculo de diversos Momentos.

3.1 Motivação

A motivação inicial traçada durante esta pesquisa de Mestrado foi a de solucionar o problema da classificação de faces humanas representadas por nuvens de pontos. Boa parte da investigação foi norteadada pelo uso de Momentos como extratores de atributo. Vimos que algumas Funções Momento geravam taxas de acerto promissoras, mas o processo de obtenção destas demandava etapas de pré-processamento profundas, principalmente a de correção da orientação (que no contexto de faces é conhecido como correção de pose (SIQUEIRA *et al.*, 2018)). Para resolver a correção de orientação, emprega-se um algoritmo de alinhamento (também chamado de algoritmo de registro), como o ICP (*Iterative Closest Point*, (BESL; MCKAY, 1992)). Essa e

outras técnicas semelhantes são muito dependentes de parâmetros que devem ser definidos *a priori*, o que influencia na generalização do método.

O passo a seguir na pesquisa foi buscar um extrator de atributos que não necessitasse que as nuvens tenham sido alinhadas previamente, ou seja, um extrator invariante à rotação. Era desejável também que este fosse invariante à escala, uma vez que vários cenários práticos resultam em nuvens adquiridas a diferentes distâncias do sensor, portanto gerando nuvens de dimensões distintas. O problema da translação é simples de se resolver e até mesmo as mais simples Funções Momento bastavam.

Entretanto, a formulação de Momentos com invariância afim completa não é trivial. Alguns estudos propõe modelos de definição de Momentos Invariantes, mas com fortes limitações em questão de formulação e sensibilidade a altas ordens (TAHRI, 2018; XU; LIAO, 2019). A invariância à escala é normalmente resolvida através de relações recursivas da equação geral do Momento, logo o algoritmo de extração final tende a ser muito custoso em tempo de processamento e recursos (MALLAHI *et al.*, 2015; PANDEY *et al.*, 2016). A invariância à rotação muitas vezes é resolvida por partes ou para uma transformação de rotação fixa (MALLAHI *et al.*, 2018), ou através do estudo dos chamados Invariantes Algébricos, valores calculados através de relações entre matrizes de Momentos. Os Invariantes Algébricos são poucos, normalmente 3 ou 6 atributos apenas, resultando em vetores de atributos pequenos e, portanto, com baixo poder de descrição, mas seu estudo foi de grande importância para a definição dos primeiros arcabouços de Momentos tridimensionais (GUO, 1993). Em outros casos, como discutido nos Momentos de Zernike e nos Momentos Harmônicos Esféricos, a invariância à rotação é alcançada apenas com uma formulação particular (no caso, a magnitude dos Momentos complexos), o que também implica em vetores de atributo pouco descritivos.

Mesmo com essas limitações, uma adaptação inicial dos Momentos de Zernike continuaram se mostrando os mais promissores dentre todos os outros estudados, o que nos levou a analisar mais profundamente a formulação deles. Concluímos que a chave para uma boa extração de atributos está no formato em nuvens de pontos, ou seja, numa descrição simples de três coordenadas, mas não no sistema cartesiano, e sim na tradução para um sistema de coordenadas esféricas, etapa base para o uso dos Momentos de Zernike e Momentos Harmônicos Esféricos. A extração de atributos em um sistema de coordenadas que conjuga informação de natureza radial e de natureza angular conseguia exprimir melhor a superfície que um extrator calculado no domínio de coordenadas cartesianas.

Assim, definiu-se a investigação do uso de outras Funções Momento, agora no domínio de ρ , θ e ϕ , e, com isso, encontramos melhoras nas taxas de classificação. Porém, o problema da orientação se mantinha; a tradução de um sistema de coordenadas para outro não resolvia por si só a invariância à rotação. A solução veio na concepção de um novo sistema hipersférico, um que fosse calculado em função de um sistema de coordenadas local e invariante.

Primeiro, uma normalização da nuvem de entrada X (em coordenadas x , y e z) em termos do centro geométrico soluciona a invariância à translação. A informação radial (distância de um ponto ao centro geométrico), agora chamada r , continuava sendo calculada em função do centro geométrico, mas agora passaria por um etapa de normalização em função da maior distância encontrada, gerando um \bar{r} . Da nuvem normalizada, aplicamos a Decomposição em Valores Singulares e da matriz V extraímos um sistema de coordenadas local que reflete a dispersão dos pontos ao longo da superfície. Imaginado que cada vetor singular v_j representa a normal de um plano π_j , calculamos, para cada ponto da superfície da nuvem, o ângulo que este faz com o plano. Assim, três ângulos, um para cada vetor singular, são encontrados: α , β e γ . Dada a nuvem de entrada X , a representação final foi armazenada numa matriz, chamada Matriz RABG e denotada por M , onde cada linha i de M guarda os valores $\left[\bar{r}_i \quad \alpha_i \quad \beta_i \quad \gamma_i \right]$ do respectivo ponto $\left[x_i \quad y_i \quad z_i \right]$ da entrada X .

3.2 Um Novo Formato para Nuvens de Pontos

Definição 3.2.1 (Matriz RABG). Dada uma nuvem de pontos X com N linhas em formato cartesiano (Equação 2.1) a Matriz RABG M resultante armazena quatro atributos geométricos invariantes para cada ponto correspondente i da matriz de entrada: distância radial normalizada \bar{r}_i e três ângulos singulares, α_i , β_i e γ_i , que cada um dos vetores-direção dos pontos da superfície faz com cada um dos planos π_j definidos pelas normais v_j (vetores singulares da matriz V da Decomposição em Valores Singulares) e que contêm o centro geométrico, respectivamente.

3.2.1 Etapas de geração da Matriz RABG

Dada uma nuvem de pontos X (tal como a Equação 2.1) de coordenadas do centro geométrico $\bar{c} = \left[\bar{x} \quad \bar{y} \quad \bar{z} \right]$ (calculadas a partir da Equação 2.2), montamos uma nuvem deslocada

\bar{X} , tal que

$$\bar{X} = \begin{bmatrix} - & \bar{p}_1 & - \\ - & \bar{p}_2 & - \\ & \vdots & \\ - & \bar{p}_N & - \end{bmatrix} = \begin{bmatrix} x'_1 & y'_1 & z'_1 \\ x'_2 & y'_2 & z'_2 \\ & \vdots & \\ x'_N & y'_N & z'_N \end{bmatrix} = \begin{bmatrix} x_1 - \bar{x} & y_1 - \bar{y} & z_1 - \bar{z} \\ x_2 - \bar{x} & y_2 - \bar{y} & z_2 - \bar{z} \\ & \vdots & \\ x_N - \bar{x} & y_N - \bar{y} & z_N - \bar{z} \end{bmatrix}. \quad (3.1)$$

A subtração entre cada ponto da nuvem pelas coordenadas do centro geométrico equivale à transformação de translação do centro geométrico para a origem do sistema de referência, ou seja, uma normalização espacial. Tal etapa garante a invariância à transformação de translação.

Calculamos um vetor coluna temporário Q com os tamanhos dos vetores \bar{p}_i :

$$Q = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_N \end{bmatrix} = \begin{bmatrix} \|\bar{p}_1\| \\ \|\bar{p}_2\| \\ \vdots \\ \|\bar{p}_N\| \end{bmatrix} = \begin{bmatrix} \sqrt{x_1'^2 + y_1'^2 + z_1'^2} \\ \sqrt{x_2'^2 + y_2'^2 + z_2'^2} \\ \vdots \\ \sqrt{x_N'^2 + y_N'^2 + z_N'^2} \end{bmatrix}. \quad (3.2)$$

Na representação prospectada, para perseguir a invariância à escala, encontramos o valor máximo dentro do vetor Q tal que $r_{max} = \max(r_i)$ para então normalizarmos as distâncias r_i no intervalo $[0, 1]$, gerando um vetor de distâncias normalizadas \bar{Q} :

$$\bar{Q} = \begin{bmatrix} \bar{r}_1 & \bar{r}_2 & \dots & \bar{r}_N \end{bmatrix}^T = \begin{bmatrix} \frac{r_1}{r_{max}} & \frac{r_2}{r_{max}} & \dots & \frac{r_N}{r_{max}} \end{bmatrix}^T. \quad (3.3)$$

Chegamos, assim, ao conjunto de valores que formam a primeira coluna da Matriz RABG. As medidas de distância que compõem o vetor \bar{Q} foram calculadas a partir da nuvem deslocada, as quais são invariantes à translação, e normalizadas no intervalo $[0, 1]$, o que torna a representação invariante à escala.

A partir da DVS (Seção 2.3) da nuvem deslocada \bar{X} , encontra-se a matriz V de vetores singulares v_j , a qual, por sua vez, pode ser usada para montar uma matriz temporária S com os ângulos que cada ponto da nuvem faz com o respectivo plano π_j cuja normal é o vetor

singular correspondente v_j :

$$S = \begin{bmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 \\ \vdots & \vdots & \vdots \\ \alpha_N & \beta_N & \gamma_N \end{bmatrix} = \begin{bmatrix} \angle(\bar{p}_1, v_0) & \angle(\bar{p}_1, v_1) & \angle(\bar{p}_1, v_2) \\ \angle(\bar{p}_2, v_0) & \angle(\bar{p}_2, v_1) & \angle(\bar{p}_2, v_2) \\ \vdots & \vdots & \vdots \\ \angle(\bar{p}_N, v_0) & \angle(\bar{p}_N, v_1) & \angle(\bar{p}_N, v_2) \end{bmatrix}, \quad (3.4)$$

em que

$$\angle(a, b) = \sin^{-1} \left(\frac{|\langle a, b \rangle|}{\|a\| \|b\|} \right), \quad (3.5)$$

onde $\langle a, b \rangle = a_x b_x + a_y b_y + a_z b_z$ expressa o produto interno entre dois vetores. Por serem ângulos de um vetor direção em relação a planos, podemos concluir que $\alpha, \beta, \gamma \in [0, \pi]$. A Figura 5 ilustra o eixo de referência local criado pelas DVS e os planos gerados por cada vetor singular, além do respectivo ângulo singular resultante.

Como mostraremos a seguir, uma transformação de rotação R aplicada à nuvem corresponde a essa mesma rotação na matriz V . Portanto, o ângulo entre um ponto qualquer \bar{p}_i e o plano π_j permanece o mesmo independente de transformações de rotação. Logo, os ângulos singulares são atributos invariantes à rotação.

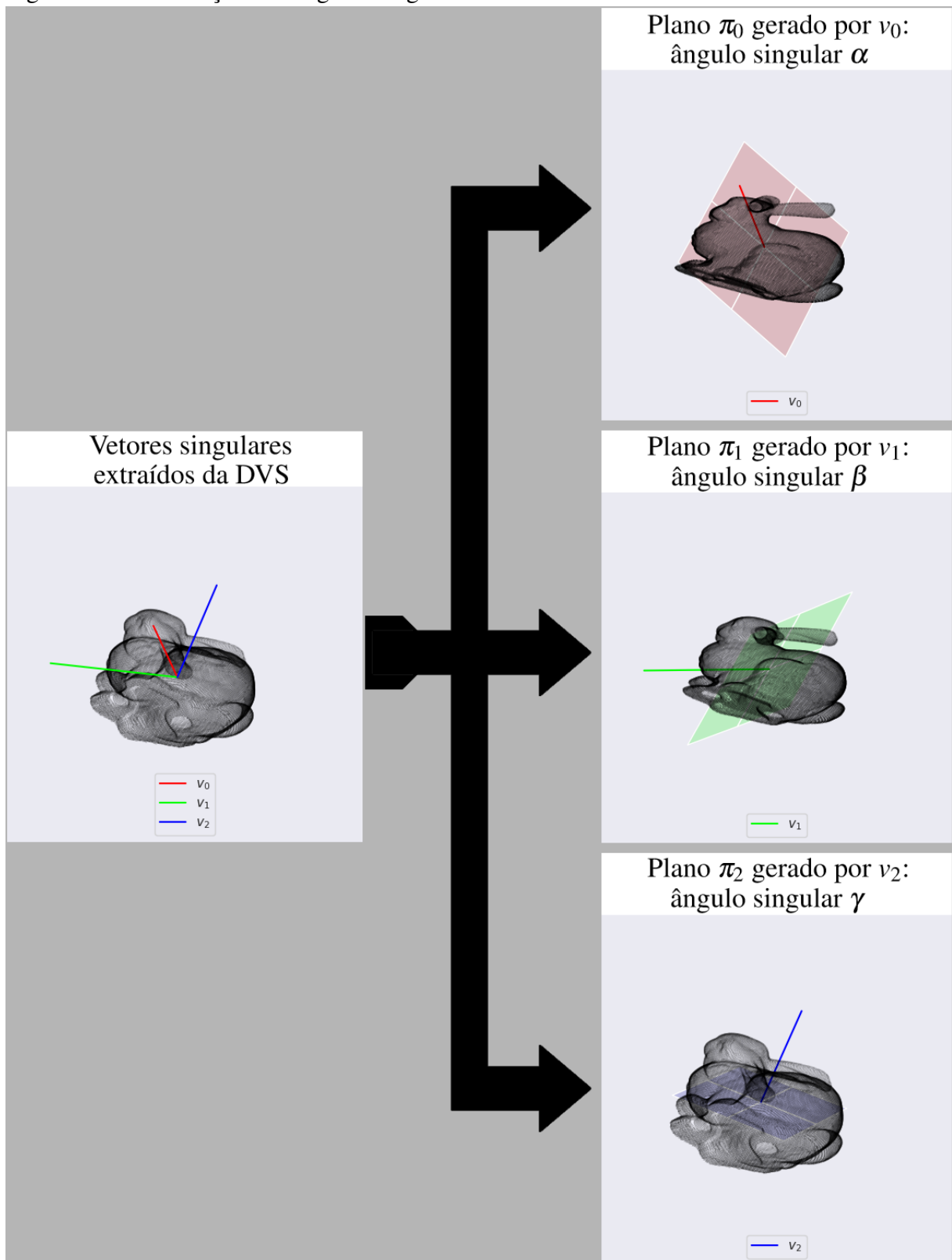
Por fim, montamos a Matriz RABG M a partir do vetor coluna \bar{Q} e da matriz S :

$$M = \begin{bmatrix} \bar{Q} & S \end{bmatrix} = \begin{bmatrix} \bar{r}_1 & \alpha_1 & \beta_1 & \gamma_1 \\ \bar{r}_2 & \alpha_2 & \beta_2 & \gamma_2 \\ \vdots & \vdots & \vdots & \vdots \\ \bar{r}_N & \alpha_N & \beta_N & \gamma_N \end{bmatrix}. \quad (3.6)$$

3.3 Invariância da Matriz RABG

Cada etapa de construção do formato de nuvem de pontos proposto adiciona uma característica invariante. Mostraremos a seguir, por meio de demonstrações e testes numéricos, que a Matriz RABG goza de invariância completa às transformações de translação, escala e rotação.

Figura 5 – Visualização dos ângulos singulares.



Fonte: O autor.

3.3.1 Invariância à Translação

Seja uma nuvem de pontos X_1 com centro geométrico $\bar{c}_1 = [\bar{x}_1 \quad \bar{y}_1 \quad \bar{z}_1]$ e a nuvem deslocada correspondente \bar{X}_1 (ver Equações 2.1, 2.2 e 3.1). A nuvem X_2 resultante da translação de X_1 ao longo da direção $D = [d_x \quad d_y \quad d_z]$ é

$$X_2 = \begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ \vdots & \vdots & \vdots \\ x_N & y_N & z_N \end{bmatrix} + \begin{bmatrix} d_x & d_y & d_z \\ d_x & d_y & d_z \\ \vdots & \vdots & \vdots \\ d_x & d_y & d_z \end{bmatrix} = \begin{bmatrix} x_1 + d_x & y_1 + d_y & z_1 + d_z \\ x_2 + d_x & y_2 + d_y & z_2 + d_z \\ \vdots & \vdots & \vdots \\ x_N + d_x & y_N + d_y & z_N + d_z \end{bmatrix}. \quad (3.7)$$

As coordenadas do centro geométrico $\bar{c}_2 = [\bar{x}_2 \quad \bar{y}_2 \quad \bar{z}_2]$ da nuvem X_2 podem ser escritas como, respectivamente:

$$\begin{aligned} \bar{x}_2 &= \frac{1}{N} \sum_{i=1}^N (x_i + d_x) = \frac{1}{N} \left(\sum_{i=1}^N x_i + \sum_{i=1}^N d_x \right) = \frac{1}{N} \left(\sum_{i=1}^N x_i + Nd_x \right) = \frac{1}{N} \sum_{i=1}^N x_i + d_x = \bar{x}_1 + d_x \\ \bar{y}_2 &= \frac{1}{N} \sum_{i=1}^N (y_i + d_y) = \frac{1}{N} \left(\sum_{i=1}^N y_i + \sum_{i=1}^N d_y \right) = \frac{1}{N} \left(\sum_{i=1}^N y_i + Nd_y \right) = \frac{1}{N} \sum_{i=1}^N y_i + d_y = \bar{y}_1 + d_y \\ \bar{z}_2 &= \frac{1}{N} \sum_{i=1}^N (z_i + d_z) = \frac{1}{N} \left(\sum_{i=1}^N z_i + \sum_{i=1}^N d_z \right) = \frac{1}{N} \left(\sum_{i=1}^N z_i + Nd_z \right) = \frac{1}{N} \sum_{i=1}^N z_i + d_z = \bar{z}_1 + d_z \end{aligned} \quad (3.8)$$

Se escrevermos a nuvem deslocada \bar{X}_2 em termos das coordenadas encontradas na

Equação 3.8, temos:

$$\begin{aligned}
 \bar{X}_2 &= \begin{bmatrix} x_1 + d_x - \bar{x}_2 & y_1 + d_y - \bar{y}_2 & z_1 + d_z - \bar{z}_2 \\ x_2 + d_x - \bar{x}_2 & y_2 + d_y - \bar{y}_2 & z_2 + d_z - \bar{z}_2 \\ & \vdots & \\ x_N + d_x - \bar{x}_2 & y_N + d_y - \bar{y}_2 & z_N + d_z - \bar{z}_2 \end{bmatrix} \\
 &= \begin{bmatrix} x_1 + d_x - (\bar{x}_1 + d_x) & y_1 + d_y - (\bar{y}_1 + d_y) & z_1 + d_z - (\bar{z}_1 + d_z) \\ x_2 + d_x - (\bar{x}_1 + d_x) & y_2 + d_y - (\bar{y}_1 + d_y) & z_2 + d_z - (\bar{z}_1 + d_z) \\ & \vdots & \\ x_N + d_x - (\bar{x}_1 + d_x) & y_N + d_y - (\bar{y}_1 + d_y) & z_N + d_z - (\bar{z}_1 + d_z) \end{bmatrix} \\
 &= \begin{bmatrix} x_1 - \bar{x}_1 & y_1 - \bar{y}_1 & z_1 - \bar{z}_1 \\ x_2 - \bar{x}_1 & y_2 - \bar{y}_1 & z_2 - \bar{z}_1 \\ & \vdots & \\ x_N - \bar{x}_1 & y_N - \bar{y}_1 & z_N - \bar{z}_1 \end{bmatrix} = \bar{X}_1
 \end{aligned} \tag{3.9}$$

Logo, a representação deslocada é invariante à translação \square .

3.3.2 Invariância à Escala

Seja uma nuvem de pontos X_1 com centro geométrico \bar{c}_1 e a nuvem deslocada correspondente \bar{X}_1 (ver Equações 2.1, 2.2 e 3.1). Seja a nuvem X_2 resultante do escalamento uniforme de X_1 por $a > 0$ e representada por

$$X_2 = a \begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ & \vdots & \\ x_N & y_N & z_N \end{bmatrix} = \begin{bmatrix} ax_1 & ay_1 & az_1 \\ ax_2 & ay_2 & az_2 \\ & \vdots & \\ ax_N & ay_N & az_N \end{bmatrix}, \tag{3.10}$$

e com coordenadas do centro geométrico \bar{c}_2 escritas como

$$\begin{aligned}\bar{x}_2 &= \frac{1}{N} \sum_{i=1}^N (ax_i) = a \frac{1}{N} \sum_{i=1}^N x_i = a\bar{x}_1 \\ \bar{y}_2 &= \frac{1}{N} \sum_{i=1}^N (ay_i) = a \frac{1}{N} \sum_{i=1}^N y_i = a\bar{y}_1 \\ \bar{z}_2 &= \frac{1}{N} \sum_{i=1}^N (az_i) = a \frac{1}{N} \sum_{i=1}^N z_i = a\bar{z}_1\end{aligned}\tag{3.11}$$

Logo, a representação deslocada \bar{X}_2 é

$$\bar{X}_2 = \begin{bmatrix} ax_1 - \bar{x}_2 & ay_1 - \bar{y}_2 & az_1 - \bar{z}_2 \\ ax_2 - \bar{x}_2 & ay_2 - \bar{y}_2 & az_2 - \bar{z}_2 \\ \vdots & \vdots & \vdots \\ ax_N - \bar{x}_2 & ay_N - \bar{y}_2 & az_N - \bar{z}_2 \end{bmatrix} = \begin{bmatrix} ax_1 - a\bar{x}_1 & ay_1 - a\bar{y}_1 & az_1 - a\bar{z}_1 \\ ax_2 - a\bar{x}_1 & ay_2 - a\bar{y}_1 & az_2 - a\bar{z}_1 \\ \vdots & \vdots & \vdots \\ ax_N - a\bar{x}_1 & ay_N - a\bar{y}_1 & az_N - a\bar{z}_1 \end{bmatrix} = a\bar{X}_1\tag{3.12}$$

A primeira coluna de atributos da Matriz RABG é um vetor \bar{Q} com os tamanhos dos vetores da nuvem de pontos, normalizados no intervalo $[0, 1]$ (Equação 3.3). Como visto na Equação 3.12, o efeito de escala se reflete na nuvem deslocada \bar{X}_2 . Portanto, o vetor de distâncias correspondente Q_2 também será igualmente escalado pelo fator a :

$$\begin{aligned}Q_2 &= \begin{bmatrix} \sqrt{(ax'_1)^2 + (ay'_1)^2 + (az'_1)^2} \\ \sqrt{(ax'_2)^2 + (ay'_2)^2 + (az'_2)^2} \\ \vdots \\ \sqrt{(ax'_N)^2 + (ay'_N)^2 + (az'_N)^2} \end{bmatrix} \\ &= \begin{bmatrix} \sqrt{a^2(x_1'^2 + y_1'^2 + z_1'^2)} \\ \sqrt{a^2(x_2'^2 + y_2'^2 + z_2'^2)} \\ \vdots \\ \sqrt{a^2(x_N'^2 + y_N'^2 + z_N'^2)} \end{bmatrix} \\ &= \begin{bmatrix} a\sqrt{x_1'^2 + y_1'^2 + z_1'^2} \\ a\sqrt{x_2'^2 + y_2'^2 + z_2'^2} \\ \vdots \\ a\sqrt{x_N'^2 + y_N'^2 + z_N'^2} \end{bmatrix} \\ &= aQ_1\end{aligned}\tag{3.13}$$

Por fim, se formos escrever o vetor de distâncias normalizadas \bar{Q}_2 , a distância máxima correspondente também terá sofrido o efeito de multiplicação por a , ou seja:

$$\bar{Q}_2 = \left[\frac{ar_1}{ar_{max}} \quad \frac{ar_2}{ar_{max}} \quad \cdots \quad \frac{ar_3}{ar_{max}} \right]^T = \left[\frac{r_1}{r_{max}} \quad \frac{r_2}{r_{max}} \quad \cdots \quad \frac{r_3}{r_{max}} \right]^T = \bar{Q}_1. \quad (3.14)$$

Logo, a representação deslocada é invariante à escala \square .

3.3.3 Invariância à Rotação

Relembrando a Seção 2.3, a DVS de uma matriz A resulta em $A = U\Sigma V^T$. Como escolhermos representar uma nuvem como sendo uma matriz onde os pontos estão dispostos nas linhas (Seção 2.1), precisamos escrever uma transformação de rotação R como a multiplicação matricial à direita. Seja A^T uma nuvem cujos pontos estão guardados ao longo das colunas, da propriedade da matriz transposta podemos escrever

$$(RA^T)^T = AR^T. \quad (3.15)$$

Portanto, a DVS da rotação de A por R pode ser escrita como

$$AR^T = U\Sigma V^T R^T. \quad (3.16)$$

A Equação 3.16 mostra que o efeito de rotação sobre a nuvem reflete na matriz de vetores singulares V . Se cada ponto da nuvem, além dos vetores singulares, são igualmente rotacionados, basta mostrar que os ângulos entre ponto e vetor singular se mantêm invariantes à rotação. Faremos uma demonstração por rotação sobre cada eixo canônico do \mathbb{R}^3 .

Seja $p = \begin{bmatrix} p_x & p_y & p_z \end{bmatrix}$ um ponto da nuvem A e $v = \begin{bmatrix} v_x & v_y & v_z \end{bmatrix}$ um dos vetores singulares extraídos da DVS de A . Seja $R_{\hat{x}}(\theta)$, $R_{\hat{y}}(\theta)$ e $R_{\hat{z}}(\theta)$ as matrizes que definem as rotações ao longos dos eixos canônicos $\hat{x} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T$, $\hat{y} = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^T$ e $\hat{z} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T$, respectivamente.

Queremos mostrar que a função ângulo $\angle(p, v)$ (Equação 3.5) é invariante à rotação,

ou seja

$$\begin{aligned}
 \angle(pR_j(\boldsymbol{\theta})^T, vR_j(\boldsymbol{\theta})^T) &= \angle(p, v) \\
 &= \sin^{-1} \left(\frac{|\langle p, v \rangle|}{\|p\| \|v\|} \right), \\
 &= \sin^{-1} \left(\frac{|p_x v_x + p_y v_y + p_z v_z|}{\|p\|} \right)
 \end{aligned} \tag{3.17}$$

em que $j \in \{\hat{x}, \hat{y}, \hat{z}\}$ e $\|v\| = 1$ pois os vetores singulares da matriz V são unitários.

3.4 Invariância à rotação no eixo \hat{x}

Seja $R_{\hat{x}}(\boldsymbol{\theta})$ a transformação de rotação ao longo do eixo \hat{x} definida por

$$R_{\hat{x}}(\boldsymbol{\theta}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \boldsymbol{\theta} & -\sin \boldsymbol{\theta} \\ 0 & \sin \boldsymbol{\theta} & \cos \boldsymbol{\theta} \end{bmatrix}. \tag{3.18}$$

Vamos desenvolver o produto $pR_{\hat{x}}(\boldsymbol{\theta})^T$:

$$\begin{aligned}
 pR_{\hat{x}}(\boldsymbol{\theta})^T &= \begin{bmatrix} p_x & p_y & p_z \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \boldsymbol{\theta} & \sin \boldsymbol{\theta} \\ 0 & -\sin \boldsymbol{\theta} & \cos \boldsymbol{\theta} \end{bmatrix} \\
 &= \begin{bmatrix} p_x & p_y \cos \boldsymbol{\theta} - p_z \sin \boldsymbol{\theta} & p_y \sin \boldsymbol{\theta} + p_z \cos \boldsymbol{\theta} \end{bmatrix}
 \end{aligned} \tag{3.19}$$

Analogamente, $vR_{\hat{x}}(\boldsymbol{\theta})^T = \begin{bmatrix} v_x & v_y \cos \boldsymbol{\theta} - v_z \sin \boldsymbol{\theta} & v_y \sin \boldsymbol{\theta} + v_z \cos \boldsymbol{\theta} \end{bmatrix}$.

A seguir, desenvolvendo o produto interno $\langle pR_{\hat{x}}(\theta)^T, vR_{\hat{x}}(\theta)^T \rangle$, temos

$$\begin{aligned}
 \langle pR_{\hat{x}}(\theta)^T, vR_{\hat{x}}(\theta)^T \rangle &= p_x v_x + p_y v_y \cos^2 \theta - p_y v_z \cos \theta \sin \theta - p_z v_y \cos \theta \sin \theta \\
 &\quad + p_z v_z \sin^2 \theta + p_y v_y \sin^2 \theta + p_y v_z \cos \theta \sin \theta \\
 &\quad + p_z v_y \cos \theta \sin \theta + p_z v_z \cos^2 \theta \\
 &= p_x v_x + p_y v_y (\cos^2 \theta + \sin^2 \theta) + p_z v_z (\cos^2 \theta + \sin^2 \theta) \\
 &= p_x v_x + p_y v_y + p_z v_z
 \end{aligned} \tag{3.20}$$

\therefore

$$\begin{aligned}
 \angle(pR_{\hat{x}}(\theta)^T, vR_{\hat{x}}(\theta)^T) &= \sin^{-1} \left(\frac{|\langle pR_{\hat{x}}(\theta)^T, vR_{\hat{x}}(\theta)^T \rangle|}{\|p\| \|v\|} \right) \\
 &= \sin^{-1} \left(\frac{|p_x v_x + p_y v_y + p_z v_z|}{\|p\|} \right) \\
 &= \angle(p, v)
 \end{aligned} \tag{3.21}$$

Portanto, a função ângulo é invariante à rotação ao longo do eixo $x = (1, 0, 0)$ \square .

3.4.0.1 Invariância à rotação no eixo \hat{y}

Seja $R_{\hat{y}}(\theta)$ a transformação de rotação ao longo do eixo \hat{y} definida por

$$R_{\hat{y}}(\theta) = \begin{bmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{bmatrix}. \tag{3.22}$$

Vamos desenvolver o produto $pR_{\hat{y}}(\theta)^T$:

$$\begin{aligned}
 pR_{\hat{y}}(\theta)^T &= \begin{bmatrix} p_x & p_y & p_z \end{bmatrix} \begin{bmatrix} \cos \theta & 0 & -\sin \theta \\ 0 & 1 & 0 \\ \sin \theta & 0 & \cos \theta \end{bmatrix} \\
 &= \begin{bmatrix} p_x \cos \theta + p_z \sin \theta & p_y & p_z \cos \theta - p_x \sin \theta \end{bmatrix}
 \end{aligned} \tag{3.23}$$

Analogamente, $vR_{\hat{y}}(\theta)^T = \begin{bmatrix} v_x \cos \theta + v_z \sin \theta & v_y & v_z \cos \theta - v_x \sin \theta \end{bmatrix}$.

A seguir, desenvolvendo o produto interno $\langle pR_{\hat{y}}(\theta)^T, vR_{\hat{y}}(\theta)^T \rangle$, temos

$$\begin{aligned}
 \langle pR_{\hat{y}}(\theta)^T, vR_{\hat{y}}(\theta)^T \rangle &= p_x v_x \cos^2 \theta + p_x v_z \cos \theta \sin \theta + p_z v_x \cos \theta \sin \theta + p_z v_z \sin^2 \theta \\
 &\quad + p_y v_y + p_z v_z \cos \theta - p_z v_z \cos^2 \theta - p_z v_x \cos \theta \sin \theta \\
 &\quad - p_z v_x \cos \theta \sin \theta - p_x v_z \cos \theta \sin \theta + p_x v_x \sin^2 \theta \\
 &= p_x v_x (\cos^2 \theta + \sin^2 \theta) + p_y v_y + p_z v_z (\cos^2 \theta + \sin^2 \theta) \\
 &= p_x v_x + p_y v_y + p_z v_z
 \end{aligned} \tag{3.24}$$

\therefore

$$\begin{aligned}
 \angle(pR_{\hat{y}}(\theta)^T, vR_{\hat{y}}(\theta)^T) &= \sin^{-1} \left(\frac{|\langle pR_{\hat{y}}(\theta)^T, vR_{\hat{y}}(\theta)^T \rangle|}{\|p\| \|v\|} \right) \\
 &= \sin^{-1} \left(\frac{|p_x v_x + p_y v_y + p_z v_z|}{\|p\|} \right) \\
 &= \angle(p, v)
 \end{aligned} \tag{3.25}$$

Portanto, a função ângulo é invariante à rotação ao longo do eixo $y = (0, 1, 0)$ \square .

3.4.0.2 Invariância à rotação no eixo \hat{z}

Seja $R_{\hat{z}}(\theta)$ a transformação de rotação ao longo do eixo \hat{z} definida por

$$R_{\hat{z}}(\theta) = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3.26}$$

Vamos desenvolver o produto $pR_{\hat{z}}(\theta)^T$:

$$\begin{aligned} pR_{\hat{z}}(\theta)^T &= \begin{bmatrix} p_x & p_y & p_z \end{bmatrix} \begin{bmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} p_x \cos \theta - p_y \sin \theta & p_x \sin \theta + p_y \cos \theta & p_z \end{bmatrix} \end{aligned} \quad (3.27)$$

Analogamente, $vR_{\hat{z}}(\theta)^T = \begin{bmatrix} p_x \cos \theta - p_y \sin \theta & p_x \sin \theta + p_y \cos \theta & p_z \end{bmatrix}$.

A seguir, desenvolvendo o produto interno $\langle pR_{\hat{z}}(\theta)^T, vR_{\hat{z}}(\theta)^T \rangle$, temos

$$\begin{aligned} \langle pR_{\hat{z}}(\theta)^T, vR_{\hat{z}}(\theta)^T \rangle &= p_x v_x \cos^2 \theta - p_x v_y \cos \theta \sin \theta - p_y v_x \cos \theta \sin \theta \\ &\quad + p_y v_y \sin^2 \theta + p_x v_x \sin^2 \theta + p_x v_y \cos \theta \sin \theta \\ &\quad + p_y v_x \cos \theta \sin \theta + p_y v_y \cos^2 \theta + p_z v_z \\ &= p_x v_x (\cos^2 \theta + \sin^2 \theta) + p_y v_y (\cos^2 \theta + \sin^2 \theta) + p_z v_z \\ &= p_x v_x + p_y v_y + p_z v_z \end{aligned} \quad (3.28)$$

\therefore

$$\begin{aligned} \angle(pR_{\hat{z}}(\theta)^T, vR_{\hat{z}}(\theta)^T) &= \sin^{-1} \left(\frac{|\langle pR_{\hat{z}}(\theta)^T, vR_{\hat{z}}(\theta)^T \rangle|}{\|p\| \|v\|} \right) \\ &= \sin^{-1} \left(\frac{|p_x v_x + p_y v_y + p_z v_z|}{\|p\|} \right) \\ &= \angle(p, v) \end{aligned} \quad (3.29)$$

Portanto, a função ângulo é invariante à rotação ao longo do eixo $z = (0, 0, 1)$ \square .

3.5 Vantagens da Matriz RABG

A proposta de concepção de um novo formato para representar nuvens de pontos nasce da necessidade de superar desafios comuns em análise de formas tridimensionais: posição relativa no espaço, escala da distribuição de pontos e alinhamento. Alguns destes são de fácil

solução, mas um método que agrega a todos em uma única representação numérica não havia sido explorado anteriormente.

Ao resolver o problema da posição relativa e o de escala através da geometria do grupo euclidiano, estamos promovendo uma normalização das coordenadas dos pontos da superfície. Essa forma intermediária representa a superfície de maneira mais simples no \mathbb{R}^3 . Dela, podemos ainda extrair, a partir de uma decomposição, as direções nas quais a superfície se dispersa no espaço; com os vetores singulares, podemos resolver o último problema, o do alinhamento.

3.6 Um Arcabouço para Extração de Atributos Baseado em Momentos

O processo de montagem da Matriz RABG consiste no cálculo de atributos geométricos invariantes (distância radial e ângulos singulares) de cada ponto de uma nuvem. Entretanto, é fácil notar que este processo resulta em um formato com maior dimensionalidade do que aquele de entrada; dada uma matriz $N \times 3$, agora temos uma matriz $N \times 4$.

Mesmo com uma representação completamente invariante, ainda é possível se valer de uma etapa de redução de dimensionalidade, pois é desejável reduzir o espaço de valores que serão usados por modelos de predição sem que haja uma perda na capacidade descritiva do conjunto de dados completo.

Para tanto, propomos um sistema de extração de atributos baseado em Momentos. Como visto na Seção 2.2, um Momento μ é definido como a aproximação de uma função alvo f sob uma função aproximadora g . Em nossa proposta, f e g serão funções cujas entradas são elementos da Matriz RABG M calculada a partir de uma nuvem de pontos X .

Definição 3.6.1 (Função alvo). A função alvo f deve ser escolhida tal que $f : \mathbb{R} \rightarrow \mathbb{R}$ cuja entrada seja a distância normalizada \bar{r} correspondente à primeira coluna da Matriz RABG. Essa escolha foi feita com base no estudo de Xu e Li (2008) que sugere que primitivas geométricas (tais como distâncias, áreas e produtos internos) são boas funções alvo, visto que gozam de invariâncias no grupo euclidiano. Dessa forma, a função alvo corresponde à natureza radial da superfície da nuvem de pontos. Dada uma Matriz RABG M , a função alvo deve ser escrita, portanto, como um $f(\bar{r})$.

Definição 3.6.2 (Função aproximadora). A função de base g deve ser escolhida tal que $g : \mathbb{R}^3 \rightarrow \mathbb{R}$, cujas entradas são os ângulos singulares guardados nas demais colunas de M : α , β e γ .

Por levar em consideração as direções singulares de cada ponto, a função de aproximação aqui proposta corresponde ao caráter angular da superfície da nuvem. Dada uma Matriz RABG M , a função de base deve ser escrita, portanto, como um $g(\alpha, \beta, \gamma)$.

Como o problema de redução de dimensionalidade ocorre no contexto de uma estrutura de dados discreta, a formulação do extrator de atributos será a interpretação discreta da integral de sobreposição definida na Equação 2.4. Aproximar $f(\bar{r})$ por $g(\alpha, \beta, \gamma)$ significa acumular o produto destas para cada ponto da Matriz RABG M . Logo, a partir das Definições 3.6.1 e 3.6.2, a formulação do extrator de atributos proposto é:

Definição 3.6.3 (Função Momento Proposto). A Função Momento $\mu(M)^{(g,f)}$ é o somatório dos produtos de g por f ao longo de todas as linhas da Matriz RABG de entrada M :

$$\mu(M)^{(g,f)} = \sum_{i=1}^N g(\alpha_i, \beta_i, \gamma_i) f(\bar{r}_i). \quad (3.30)$$

Tanto $f(\bar{r})$ quanto $g(\alpha, \beta, \gamma)$ podem ser funções que dependem de parâmetros adicionais. Por exemplo, a escolha de polinômios como funções de base é bastante comum, sendo o valor da ordem do polinômio um parâmetro de configuração deste. A partir dessa reflexão, podemos sintetizar uma definição mais versátil e que representa o arcabouço geral de extração de Momentos proposto:

Definição 3.6.4 (Função Momento Proposto Geral). Seja $\Upsilon = \{v_1, v_2, \dots, v_l\}$ o conjunto de todos os parâmetros adicionais das funções de base g e alvo f , definimos a Função Momento Geral $\mu_\Upsilon(M)^{(g,f)}$ como

$$\mu_\Upsilon(M)^{(g,f)} = \sum_{i=1}^N g_\Upsilon(\alpha_i, \beta_i, \gamma_i) f_\Upsilon(\bar{r}_i). \quad (3.31)$$

A invariância do arcabouço de extração de atributos é trivial e apresentado na Seção a seguir.

3.6.1 Invariância do arcabouço de extração de atributos

Seja X uma nuvem de pontos em formato cartesiano tal como na Equação 2.1 e hX o resultado da transformação de X por h , $\forall h \in H$ tal que H é o conjunto das transformações afins: translação, escala e rotação. Seja M a representação em Matriz RABG de X e M' a de hX .

Seja $\mu_{\Upsilon}(M)^{(g,f)}$ o momento extraído de M para uma dada função de base g e função alvo f , parametrizadas por Υ . Analogamente, $\mu_{\Upsilon}(M')^{(g,f)}$ é o momento igualmente definido, mas extraído a partir de M' .

Como visto nas Seções 3.3.1, 3.3.2 e 3.3.3, a representação em Matriz RABG é invariante a todas as transformações h , logo:

$$M = M'. \quad (3.32)$$

Da Equação 3.32, temos

$$\mu_{\Upsilon}(M)^{(g,f)} = \mu_{\Upsilon}(M')^{(g,f)}. \quad (3.33)$$

Portanto, o arcabouço de extração de atributos baseado em Momentos é completamente invariante às transformações afins \square .

3.6.2 Montagem de um vetor de atributos

Podemos montar um vetor de atributos $\Lambda(M)_{\Upsilon}^{(g,f)}$, correspondente à extração de Momentos da nuvem X com Matriz RABG M , na base g , com função alvo f e parametrizada por Υ . Os elementos deste vetor correspondem a todas as combinações dos parâmetros v_l variando a partir de zero. Chamamos de cardinalidade o número total de elementos do vetor de atributos.

3.7 Vantagens do Extrator de Momentos Proposto

As Definições 3.6.1 e 3.6.2 são escritas em termos gerais, sem quaisquer limitações de como escrever as funções escolhidas, de maneira proposital. Uma vez que a entrada do sistema de extração de atributos é o formato invariante Matriz RABG, os Momentos dela extraídos também se valerão de invariância completa. A Definição 3.6.4 condensa como escrever qualquer Função Momento. Aqui está a essência da contribuição quanto à extração de atributos: uma formulação geral de Momentos Invariantes Tridimensionais.

3.8 Considerações Finais sobre Metodologia

Apresentamos neste capítulo as contribuições deste trabalho: uma nova representação para nuvens de pontos e um sistema de extração de atributos baseado em Momentos que faz

uso das propriedades dessa nova representação. Por meio de demonstrações, vimos que a Matriz RABG é completamente invariante às transformações afins. Além disso, delimitamos o arcabouço necessário para montar um sistema de extração de atributos invariantes tridimensionais baseado em Momentos. A demonstração de uso do arcabouço proposto é feita no Capítulo 4, onde se avalia a natureza invariante dos Momentos extraídos e seu uso em um cenário de classificação de teste.

4 RESULTADOS

Podemos sintetizar a proposta deste trabalho como sendo o novo formato de representação de nuvens de pontos, chamado de Matriz RABG, e o sistema de extração de atributos baseado em Momentos que faz uso daquele formato. Afim de avaliar a contribuição da proposta em termos práticos, definiu-se dois tipos de experimentos: experimentos de invariância e experimentos de classificação.

Apresentaremos neste capítulo um conjunto de estudos de caso que servirão como ilustração do método proposto. Montaremos três Momentos distintos e os avaliaremos em termos de invariância e capacidade descritiva para classificação. O objetivo desta parte do trabalho é exemplificar o uso do método em cenário real.

Em outras palavras, queremos mostrar que o modelo de extração de atributos é versátil, ou seja, que dada qualquer função de base e qualquer função alvo, e tendo como entrada a Matriz RABG, os Momentos extraídos da formulação resultante a partir do arcabouço da Definição 3.6.4 terão invariância afim completa. Para isso, montamos três Funções Momento a partir de polinômios ortogonais como funções de base e duas formas de se escrever funções alvo.

Esses três Momentos sugeridos foram aplicados em um cenário de reconhecimento de faces humanas na popular base de dados Bosphorus (SAVRAN *et al.*, 2008). O desafio de classificação de indivíduos a partir de faces representadas em três dimensões é um desafio considerado resolvido, mas que ainda necessita de etapas de pré-processamento das superfícies consideradas custosas e com alto grau de dependência de parâmetros; dessas, a etapa de alinhamento é considerada a mais crítica (ZHOU; XIAO, 2018; ALEXANDRE *et al.*, 2020). A escolha por usar esse cenário como *benchmark* permite compararmos a capacidade descritiva do modelo de extração de atributos proposto e as vantagens que ele carrega frente ao estado da arte em reconhecimento facial 3D.

A implementação dos experimentos foi realizada em linguagem Python 3.6 com uso das bibliotecas *numpy*, *scipy*, *pandas*, *matplotlib* e *scikit-learn* (HARRIS *et al.*, 2020; VIRTANEN *et al.*, 2020; MCKINNEY, 2010; HUNTER, 2007; PEDREGOSA *et al.*, 2011). O código-fonte completo está disponível em repositório na *internet*¹.

¹ <https://github.com/keizerzilla/rocha-neto-dissertacao>

4.1 Funções de Base Escolhidas

O ponto de partida de todos os experimentos apresentados neste Capítulo é a montagem das funções de extração de Momentos, as quais se darão a partir da definição das respectivas funções de base g e funções alvo f de acordo com a Definição 3.6.4. Os Momentos sugeridos aqui servem de ilustração para a contribuição deste trabalho: a capacidade de montar qualquer Momento Invariante Tridimensional seguindo o arcabouço de extração de atributos e usando o formato Matriz RABG. As funções de base e funções alvo foram montadas a partir de três polinômios bastante explorados na literatura de Momentos: Polinômios de Legendre, Polinômios de Chebyshev e Polinômios de Zernike. Esse último já foi bastante discutido na Seção 2.2.3, portanto nos ateremos apenas à sugestão de Momentos baseado neles. Apresentaremos os dois primeiros com mais detalhes a seguir, além das respectivas sugestões de Momentos.

4.1.1 Sugestão de Momentos de Legendre

Os polinômios de Legendre foram inicialmente explorados por Teague (1980) como função de base para um extrator de Momentos de imagens bidimensionais, mas o interesse por eles em estudos de aproximação de funções é anterior (DAVE; ARMSTRONG, 1970). Vários trabalhos já avaliaram o uso dos Polinômios de Legendre em diferentes desafios de classificação, adaptação para imagens tridimensionais e avaliação de aspectos de implementação e custo computacional (HOSNY, 2011; HOSNY *et al.*, 2019; HJOUJI *et al.*, 2020).

Os polinômios de Legendre de ordem k podem ser escritos de forma recursiva como

$$L_{k+1}(x) = \frac{(2k+1)xL_k(x) - kL_{k-1}(x)}{k+1}, \quad (4.1)$$

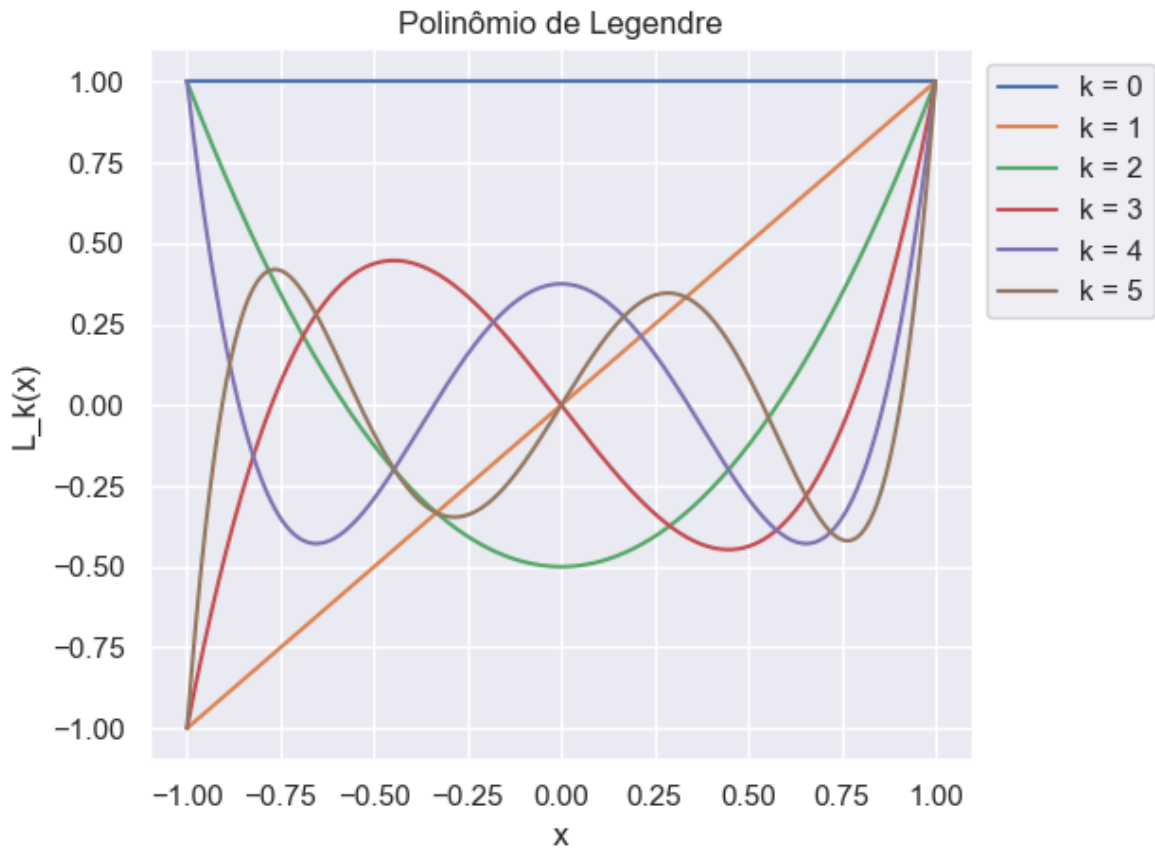
tal que os polinômios de ordem 0 e 1 são, respectivamente

$$\begin{aligned} L_0(x) &= 1 \\ L_1(x) &= x \end{aligned} \quad (4.2)$$

A Figura 6 ilustra os polinômios de Legendre de ordens $\{0, 1, 2, 3, 4, 5\}$ e $x \in [-1, 1]$.

Da Definição 3.6.2, a função de base deve ser escrita como um $g(\alpha, \beta, \gamma)$. Para a sugestão de Momentos de Legendre, optamos pela escolha do produto de três polinômios

Figura 6 – Visualização do polinômio $L_k(x)$.



Fonte: O autor.

$L_{k+1}(x)$, alternando-se as entradas entre os ângulos singulares. Portanto, a função de base da sugestão de Momentos de Legendre foi escrita como

$$g_{pqr}(\alpha, \beta, \gamma)^{(Legendre)} = L_p(\alpha)L_q(\beta)L_r(\gamma). \quad (4.3)$$

Note que poderíamos ter escrito a função $g_{pqr}(\alpha, \beta, \gamma)^{(Legendre)}$ como a soma de três polinômios ao invés do produto, dado o caráter modular do arcabouço de extração de atributos.

Da Definição 3.6.1, a função alvo deve ser escrita como um $f(\bar{r})$. Para a sugestão de Momentos de Legendre, a função alvo será a própria distância normalizada \bar{r} , ou seja

$$f(\bar{r})^{(Legendre)} = \bar{r}. \quad (4.4)$$

A partir das Equações 4.3 e 4.4, e aplicando a Definição 3.6.4 com $\Upsilon = \{p, q, r\}$,

temos a expressão para a sugestão de Momentos de Legendre

$$\mu_{pqr}(M)^{(Legendre)} = \sum_{i=1}^N L_p(\alpha_i) L_q(\beta_i) L_r(\gamma_i) \bar{r}_i. \quad (4.5)$$

4.1.2 Sugestão de Momentos de Chebyshev

Os polinômios de Chebyshev (também grafados como *Tchebichef*) foram introduzidos no contexto de análise de imagens por Mukundan *et al.* (2001). A primeira avaliação destes polinômios na construção de um extrator de Momentos Tridimensional foi realizada por Mallahi *et al.* (2015). Experimentos no contexto de classificação de objetos e reconstrução de imagens bidimensionais são áreas de interesse recente que exploram os polinômios de Chebyshev (XU; LIAO, 2019; WU; XU, 2019).

Os polinômios de Chebyshev de ordem k podem ser escritos de forma recursiva como

$$C_{k+1}(x) = 2xC_k(x) - C_{k-1}(x), \quad (4.6)$$

tal que os polinômios de ordem 0 e 1 são, respectivamente

$$\begin{aligned} C_0(x) &= 1 \\ C_1(x) &= x \end{aligned} \quad (4.7)$$

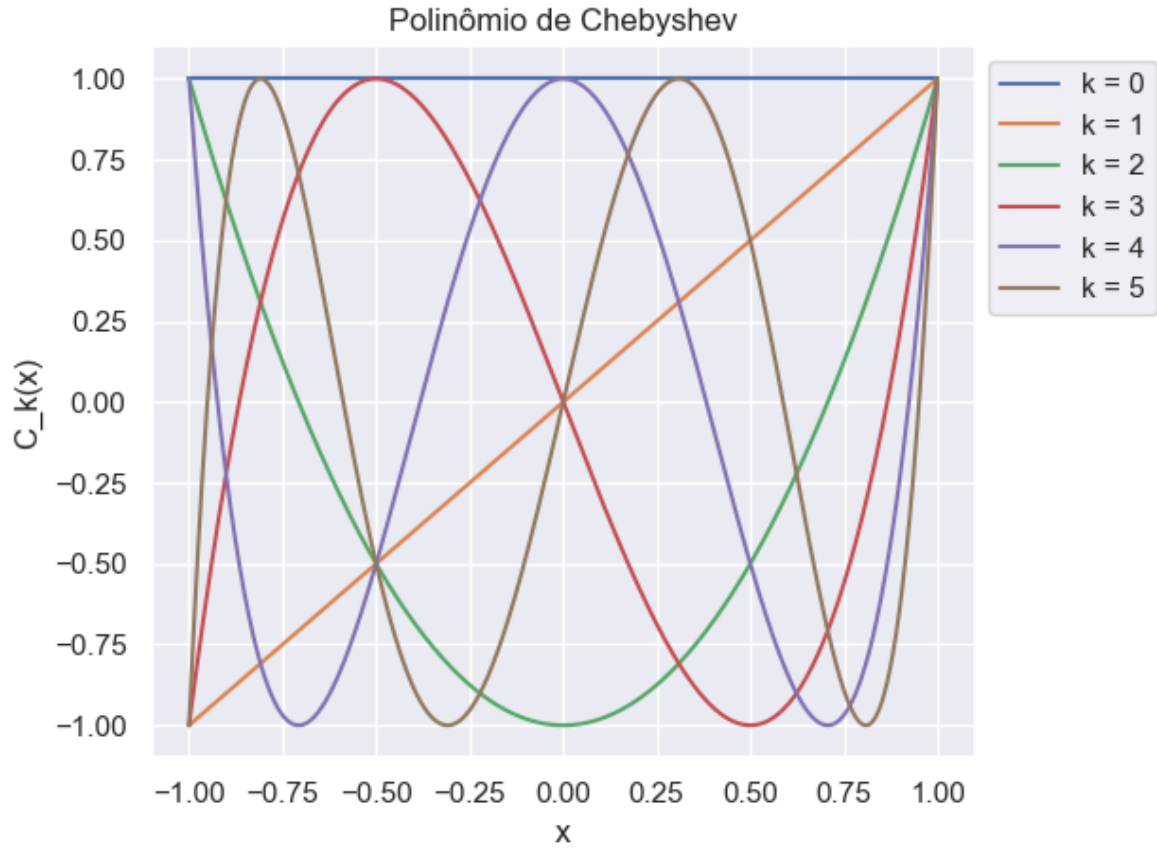
A Figura 7 ilustra os polinômios de Chebyshev de ordens $\{0, 1, 2, 3, 4, 5\}$ e $x \in [-1, 1]$.

Da Definição 3.6.2, a função de base deve ser escrita como um $g(\alpha, \beta, \gamma)$. Para a sugestão de Momentos de Chebyshev, optamos pela escolha do produto de três polinômios $C_{k+1}(x)$, alternando-se as entradas entre os ângulos singulares. Portanto, a função de base da sugestão de Momentos de Chebyshev foi escrita como

$$g_{pqr}(\alpha, \beta, \gamma)^{(Chebyshev)} = C_p(\alpha) C_q(\beta) C_r(\gamma). \quad (4.8)$$

Assim como pontuado para a função de base da sugestão de Momentos de Legendre (Equação 4.3), aqui também poderíamos ter escrito a função $g_{pqr}(\alpha, \beta, \gamma)^{(Chebyshev)}$ como a soma de três polinômios ao invés do produto.

Figura 7 – Visualização do polinômio $C_k(x)$.



Fonte: O autor.

Da Definição 3.6.1, a função alvo deve ser escrita como um $f(\bar{r})$. Para a sugestão de Momentos de Chebyshev, a função alvo será a própria distância normalizada \bar{r} , ou seja

$$f(\bar{r})^{(Chebyshev)} = \bar{r}. \quad (4.9)$$

A partir das Equações 4.8 e 4.9, e aplicando a Definição 3.6.4 com $\Upsilon = \{p, q, r\}$, temos a expressão para a sugestão de Momentos de Chebyshev

$$\mu_{pqr}(M)^{(Chebyshev)} = \sum_{i=1}^N C_p(\alpha_i) C_q(\beta_i) C_r(\gamma_i) \bar{r}_i. \quad (4.10)$$

4.1.3 Sugestão de Momentos de Zernike

As construções das sugestões de Momentos de Legendre e de Chebyshev, abordadas nas Seções anteriores, tiveram um processo criativo idêntico: dado um polinômio, a função de base foi definida como o produto de três polinômios, onde cada um recebe de entrada um dos três

ângulos singulares. Já a função alvo foi escrita como a própria distância normalizada. A seguir, veremos uma outra maneira de construir uma Função Momento com base nas peculiaridades dos Polinômios Complexos de Zernike e avaliando mais uma vez o caráter modular da proposta de arcabouço de extração de atributos.

Os Polinômios Complexos de Zernike, tal como discutidos na Seção 2.2.3 e apresentados na Equação 2.10, são definidos como o produto de uma função radial e uma função angular, cujas entradas são calculadas a partir das coordenadas dos *pixels* de uma imagem bidimensional mapeados dentro do círculo de raio unitário (i.e coordenadas polares). A função angular dos Polinômios Complexos de Zernike é a exponencial complexa, cujo ângulo é escalado pelo valor de repetição m . Por sua vez, a função radial é o chamado Polinômio Radial (Equação 2.9), cuja entrada é a distância radial e parametrizada pelos valores de ordem n e repetição m .

Para a sugestão de Momentos de Zernike, propomos uma função angular que serve de função de base, um $g(\alpha, \beta, \gamma)$ como dita a Definição 3.6.2. Resultado da soma entre o produtório dos senos e o produtório dos cossenos dos ângulos singulares escalados pela parâmetro m , temos

$$g_m(\alpha, \beta, \gamma)^{(Zernike)} = \prod_{\theta \in \{\alpha, \beta, \gamma\}} \cos(m\theta) + \prod_{\phi \in \{\alpha, \beta, \gamma\}} \sin(m\phi). \quad (4.11)$$

Usaremos o Polinômio Radial como função alvo para a sugestão de Momentos de Zernike. Essa função atende ao requisito único da Definição 3.6.1: uma função cujo único parâmetro é a distância normalizada \bar{r} . Temos então

$$f_{nm}(\bar{r})^{(Zernike)} = R_{nm}(\bar{r}). \quad (4.12)$$

Das Equações 4.11 e 4.12, e aplicando a Definição 3.6.4, chegamos à formulação da sugestão de Momentos de Zernike que usa o arcabouço de extração de atributos proposto, escritos como

$$\mu_{nm}(M)^{(Zernike)} = \sum_{i=1}^N \left[\left(\prod_{\theta \in \{\alpha, \beta, \gamma\}} \cos(m\theta) + \prod_{\phi \in \{\alpha, \beta, \gamma\}} \sin(m\phi) \right) R_{nm}(\bar{r}_i) \right]. \quad (4.13)$$

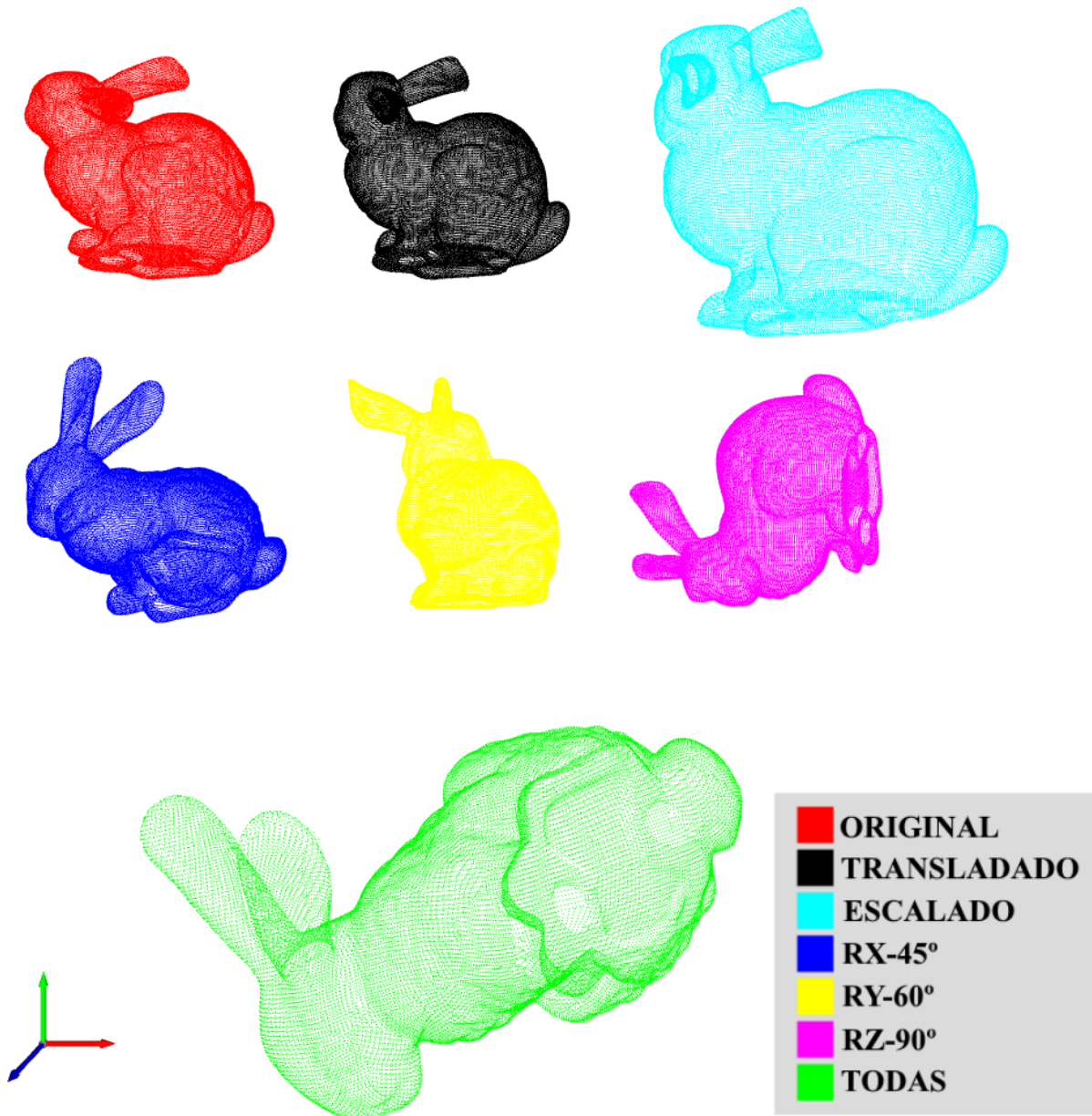
4.2 Teste Numérico de Invariância

Afim de mostrar empiricamente que os Momentos construídos carregam a invariância afim completa provida pelo formato Matriz RABG (como demonstrado na Seção 3.6.1),

construímos um cenário de extração com um conjunto de nuvens de controle. Tomamos a nuvem do coelho da base de *scans* da Stanford (LEVOY *et al.*, 2005) e aplicamos um conjunto de transformações de translação, escala e rotação.

A Figura 8 traz uma visualização de cada uma das nuvens do conjunto de controle, composto por: nuvem original (**ORIGINAL**), nuvem original transladada ao longo das três direções canônicas (**TRANSLADADO**), nuvem original escala uniformemente com o dobro de tamanho (**ESCALADO**), nuvem original rotacionada 45° em torno do eixo \hat{x} (**RX-45°**), nuvem original rotacionada 60° em torno do eixo \hat{y} (**RY-60°**), nuvem original rotacionada 90° em torno do eixo \hat{z} (**RZ-90°**) e nuvem original com todas as transformações afins aplicadas (**TODAS**).

Figura 8 – Nuvens utilizadas para o teste numérico de invariância.



Fonte: Levoy *et al.* (2005), adaptado pelo autor.

Para cada uma das nuvens do conjunto de controle, foram extraídos os Momentos sugeridos de Legendre, Chebyshev e Zernike. A seguir, verificou-se a variância estatística dos Momentos para cada parametrização ao longo de cada nuvem transformada. Os parâmetros escolhidos para cada Função Momento foram:

$$\begin{aligned}\Upsilon^{(Legendre)} &= \{p \in \{0, 1, 2\}, q \in \{0, 1, 2\}, r \in \{0, 1, 2\}\} \\ \Upsilon^{(Chebyshev)} &= \{p \in \{0, 1, 2\}, q \in \{0, 1, 2\}, r \in \{0, 1, 2\}\} \\ \Upsilon^{(Zernike)} &= \{n \in \{0, 1, 2, 3, 4, 5, 6\}, m \in \{0, 1, 2, 3, 4, 5, 6\}\}\end{aligned}\tag{4.14}$$

As Tabelas 1, 2 e 3 trazem os resultados numéricos de invariância para os Momentos sugeridos de Legendre, Chebyshev e Zernike, respectivamente. Cada linha representa um dos Momentos μ parametrizados por ν . Os valores ao longo das colunas estão associados a cada uma das nuvens transformadas do conjunto de controle.

Para os Momentos de Legendre e Chebyshev, vemos que a variância dos valores está sempre acomodada no intervalo entre a quarta e oitava casa decimais; a menos da precisão numérica de uma variável ponto flutuante, podemos concluir que estes Momentos são invariantes para todas as transformações afins. Ainda com relação às Tabelas 1 e 2, vemos que os atributos extraídos com ordens não maiores que 1 são iguais entre Legendre e Chebyshev, já que os polinômios em ordens mais baixas de ambos são iguais (Equações 4.2 e 4.7).

Em consideração aos resultados obtidos para a sugestão de Momentos de Zernike (Tabela 3), temos, mais uma vez, valores de variância dentro da margem de precisão numérica, mas um pouco maiores que para as sugestões de Momentos de Legendre e Chebyshev, com exceção do momento $\mu_{00}^{(Zernike)}$; para $n = 0$ e $m = 0$, a função angular retorna 1 e a função radial também retorna 1, portanto a Equação 4.13 calcula a acumulação de pontos da nuvem de entrada, ou seja, sempre o valor constante N , resultando na menor variância possível dentro da precisão discreta.

4.3 Resultados em Tarefas de Classificação

Após mostrar o caráter invariante dos Momentos calculados a partir do arcabouço de extração de atributos baseado em Momentos e na Matriz RABG, avaliaremos a seguir a capacidade descritiva dos vetores de atributo montados a partir da extração de Momentos de faces humanas representadas por nuvens de pontos para classificação de indivíduos.

Tabela 1 – Invariância dos Momentos de Legendre RABG

Momento	Original	Transladado	Escalado	RX-45°	RY-60°	RZ-90°	Todas	Variância
$\mu_{000}^{(Legendre)}$	19335.90	19335.90	19335.90	19335.94	19335.89	19335.92	19335.87	4.65e-04
$\mu_{001}^{(Legendre)}$	7645.90	7645.90	7645.90	7645.92	7645.90	7645.91	7645.89	6.84e-05
$\mu_{001}^{(Legendre)}$	-2349.16	-2349.16	-2349.16	-2349.17	-2349.16	-2349.17	-2349.16	8.78e-06
$\mu_{010}^{(Legendre)}$	10213.83	10213.83	10213.83	10213.86	10213.83	10213.85	10213.82	1.32e-04
$\mu_{011}^{(Legendre)}$	3444.31	3444.31	3444.31	3444.32	3444.31	3444.32	3444.31	1.38e-05
$\mu_{012}^{(Legendre)}$	-2240.35	-2240.35	-2240.35	-2240.35	-2240.35	-2240.35	-2240.34	7.02e-06
$\mu_{020}^{(Legendre)}$	2219.17	2219.17	2219.17	2219.17	2219.17	2219.17	2219.16	1.08e-05
$\mu_{021}^{(Legendre)}$	-347.77	-347.77	-347.77	-347.77	-347.77	-347.77	-347.77	1.40e-07
$\mu_{022}^{(Legendre)}$	-2220.43	-2220.43	-2220.43	-2220.43	-2220.43	-2220.43	-2220.43	7.89e-06
$\mu_{100}^{(Legendre)}$	15128.62	15128.62	15128.62	15128.66	15128.62	15128.64	15128.60	2.84e-04
$\mu_{101}^{(Legendre)}$	4897.43	4897.43	4897.43	4897.44	4897.43	4897.43	4897.42	2.83e-05
$\mu_{102}^{(Legendre)}$	-3731.33	-3731.33	-3731.33	-3731.33	-3731.32	-3731.33	-3731.32	1.77e-05
$\mu_{110}^{(Legendre)}$	6094.18	6094.18	6094.18	6094.19	6094.18	6094.19	6094.17	4.30e-05
$\mu_{111}^{(Legendre)}$	1785.31	1785.31	1785.31	1785.31	1785.31	1785.31	1785.31	3.37e-06
$\mu_{112}^{(Legendre)}$	-1751.10	-1751.10	-1751.10	-1751.10	-1751.10	-1751.10	-1751.10	3.71e-06
$\mu_{120}^{(Legendre)}$	-1956.32	-1956.32	-1956.32	-1956.32	-1956.32	-1956.32	-1956.31	4.89e-06
$\mu_{121}^{(Legendre)}$	-972.11	-972.11	-972.11	-972.11	-972.11	-972.11	-972.11	1.33e-06
$\mu_{122}^{(Legendre)}$	51.44	51.44	51.44	51.44	51.44	51.44	51.44	1.26e-08
$\mu_{200}^{(Legendre)}$	12409.93	12409.93	12409.93	12409.96	12409.93	12409.94	12409.91	2.01e-04
$\mu_{201}^{(Legendre)}$	2380.74	2380.74	2380.74	2380.74	2380.73	2380.74	2380.73	6.85e-06
$\mu_{202}^{(Legendre)}$	-5678.32	-5678.32	-5678.32	-5678.33	-5678.32	-5678.32	-5678.31	4.10e-05
$\mu_{210}^{(Legendre)}$	2276.71	2276.71	2276.71	2276.72	2276.71	2276.71	2276.71	4.74e-06
$\mu_{211}^{(Legendre)}$	218.40	218.40	218.40	218.40	218.40	218.40	218.40	4.94e-08
$\mu_{212}^{(Legendre)}$	-1321.48	-1321.48	-1321.48	-1321.48	-1321.48	-1321.48	-1321.48	1.73e-06
$\mu_{220}^{(Legendre)}$	-6457.59	-6457.59	-6457.59	-6457.60	-6457.58	-6457.59	-6457.58	6.26e-05
$\mu_{221}^{(Legendre)}$	-1554.59	-1554.59	-1554.59	-1554.59	-1554.59	-1554.59	-1554.59	3.41e-06
$\mu_{222}^{(Legendre)}$	2531.71	2531.71	2531.71	2531.72	2531.71	2531.71	2531.71	9.41e-06

Fonte: o autor.

4.3.1 Cenário e base de dados

Usaremos como cenário de referência a tarefa de classificação de indivíduos a partir do reconhecimento de faces. Aqui, o desafio é treinar um modelo de classificação a partir de uma única amostra por cada classe envolvida no problema; para cada indivíduo, uma única nuvem de pontos da sua face será usada para extração de Momentos. Como queremos montar um cenário de referência, usaremos a tarefa de classificar as faces a partir de feições neutras (i.e sem nenhuma expressão facial). O problema de classificação dessa natureza já foi considerado como solucionado, podendo assim ser usada como *benchmark* para novos modelos de predição (ZHOU; XIAO, 2018).

Usaremos a base de dados Bosphorus (SAVRAN *et al.*, 2008), uma das mais explo-

Tabela 2 – Invariância dos Momentos de Chebyshev RABG

Momento	Original	Transladado	Escalado	RX-45°	RY-60°	RZ-90°	Todas	Variância
$\mu_{000}^{(Chebyshev)}$	19335.90	19335.90	19335.90	19335.94	19335.89	19335.92	19335.87	4.65e-04
$\mu_{001}^{(Chebyshev)}$	7645.90	7645.90	7645.90	7645.92	7645.90	7645.91	7645.89	6.84e-05
$\mu_{002}^{(Chebyshev)}$	-9577.52	-9577.52	-9577.52	-9577.54	-9577.51	-9577.53	-9577.50	1.23e-04
$\mu_{010}^{(Chebyshev)}$	10213.83	10213.83	10213.83	10213.86	10213.83	10213.85	10213.82	1.32e-04
$\mu_{011}^{(Chebyshev)}$	3444.31	3444.31	3444.31	3444.32	3444.31	3444.32	3444.31	1.38e-05
$\mu_{012}^{(Chebyshev)}$	-6391.74	-6391.74	-6391.74	-6391.76	-6391.74	-6391.75	-6391.73	5.41e-05
$\mu_{020}^{(Chebyshev)}$	-3486.41	-3486.41	-3486.41	-3486.42	-3486.41	-3486.41	-3486.41	1.07e-05
$\mu_{021}^{(Chebyshev)}$	-3012.33	-3012.33	-3012.33	-3012.34	-3012.33	-3012.33	-3012.33	1.01e-05
$\mu_{022}^{(Chebyshev)}$	-1741.22	-1741.22	-1741.22	-1741.22	-1741.22	-1741.23	-1741.22	8.74e-06
$\mu_{100}^{(Chebyshev)}$	15128.62	15128.62	15128.62	15128.66	15128.62	15128.64	15128.60	2.84e-04
$\mu_{101}^{(Chebyshev)}$	4897.43	4897.43	4897.43	4897.44	4897.43	4897.43	4897.42	2.83e-05
$\mu_{102}^{(Chebyshev)}$	-10017.97	-10017.97	-10017.97	-10018.00	-10017.97	-10017.98	-10017.96	1.26e-04
$\mu_{110}^{(Chebyshev)}$	6094.18	6094.18	6094.18	6094.19	6094.18	6094.19	6094.17	4.30e-05
$\mu_{111}^{(Chebyshev)}$	1785.31	1785.31	1785.31	1785.31	1785.31	1785.31	1785.31	3.37e-06
$\mu_{112}^{(Chebyshev)}$	-4366.19	-4366.19	-4366.19	-4366.20	-4366.19	-4366.19	-4366.18	2.26e-05
$\mu_{120}^{(Chebyshev)}$	-7651.30	-7651.30	-7651.30	-7651.32	-7651.30	-7651.31	-7651.29	7.33e-05
$\mu_{121}^{(Chebyshev)}$	-2928.62	-2928.62	-2928.62	-2928.63	-2928.62	-2928.62	-2928.61	1.09e-05
$\mu_{122}^{(Chebyshev)}$	4300.26	4300.26	4300.26	4300.27	4300.25	4300.26	4300.25	2.19e-05
$\mu_{200}^{(Chebyshev)}$	10101.27	10101.27	10101.27	10101.30	10101.27	10101.28	10101.26	1.38e-04
$\mu_{201}^{(Chebyshev)}$	625.68	625.68	625.68	625.68	625.68	625.68	625.68	6.22e-07
$\mu_{202}^{(Chebyshev)}$	-12417.81	-12417.81	-12417.81	-12417.84	-12417.80	-12417.82	-12417.79	1.96e-04
$\mu_{210}^{(Chebyshev)}$	-369.00	-369.00	-369.00	-369.00	-369.00	-369.00	-369.00	2.42e-06
$\mu_{211}^{(Chebyshev)}$	-856.91	-856.91	-856.91	-856.91	-856.91	-856.91	-856.91	1.34e-06
$\mu_{212}^{(Chebyshev)}$	-1230.59	-1230.59	-1230.59	-1230.59	-1230.59	-1230.59	-1230.59	1.01e-06
$\mu_{220}^{(Chebyshev)}$	-15833.54	-15833.54	-15833.54	-15833.58	-15833.53	-15833.56	-15833.52	3.75e-04
$\mu_{221}^{(Chebyshev)}$	-2817.71	-2817.71	-2817.71	-2817.71	-2817.71	-2817.71	-2817.70	1.15e-05
$\mu_{222}^{(Chebyshev)}$	15611.66	15611.66	15611.66	15611.70	15611.65	15611.68	15611.63	3.50e-04

Fonte: o autor.

radas nesse contexto, para definição do cenário de classificação. Montada a partir do *scanner* Inspeck Mega Capturor II e totalizando 4666 amostras de 105 indivíduos, esse conjunto de dados contém imagens 3D de faces em formato de nuvens de pontos com variação de expressões (e.g feliz, neutro, raiva), poses (i.e rotações ao longo do eixo do pescoço) e diferentes tipos de oclusão (e.g mão no rosto, óculos, mão na boca, cachecol). A Figura 9 traz exemplos de rostos encontrados na base de dados Bosphorus.

Não usaremos o conjunto de dados completo, mas apenas o subconjunto das faces com expressão neutra. No total, 299 nuvens são usadas nos testes de classificação, sendo 105 para o subconjunto de treinamento e as 194 nuvens restantes responsáveis pelo subconjunto de teste. Como explicado anteriormente, montaremos um classificador que aprenderá com uma única amostra por classe.

Tabela 3 – Invariância dos Momentos de Zernike RABG

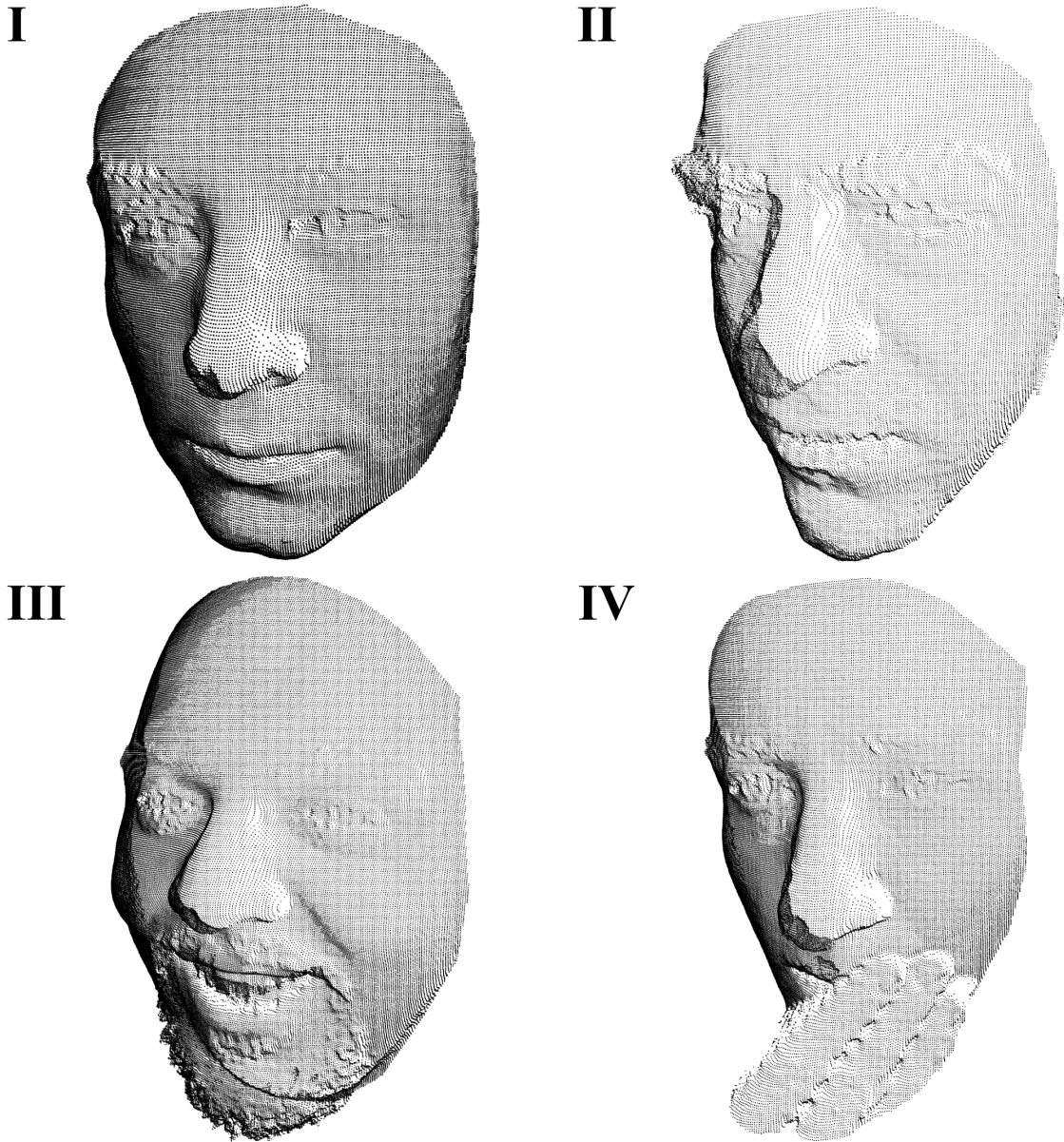
Momento	Original	Transladado	Escalado	RX-45°	RY-60°	RZ-90°	Todas	Variância
$\mu_{00}^{(Zernike)}$	11442.29	11442.29	11442.29	11442.29	11442.29	11442.29	11442.29	3.30e-24
$\mu_{11}^{(Zernike)}$	12384.89	12384.89	12384.89	12384.92	12384.89	12384.91	12384.87	1.79e-04
$\mu_{20}^{(Zernike)}$	-31884.65	-31884.65	-31884.65	-31884.41	-31884.66	-31884.52	-31884.80	1.31e-02
$\mu_{22}^{(Zernike)}$	-1210.51	-1210.51	-1210.51	-1210.52	-1210.50	-1210.51	-1210.50	1.76e-05
$\mu_{31}^{(Zernike)}$	-30426.03	-30426.03	-30426.03	-30425.95	-30426.04	-30425.99	-30426.08	1.44e-03
$\mu_{33}^{(Zernike)}$	-803.60	-803.60	-803.60	-803.60	-803.60	-803.60	-803.59	9.79e-06
$\mu_{40}^{(Zernike)}$	-20653.73	-20653.73	-20653.73	-20654.03	-20653.71	-20653.89	-20653.54	2.06e-02
$\mu_{42}^{(Zernike)}$	475.41	475.41	475.41	475.40	475.41	475.41	475.43	1.10e-04
$\mu_{44}^{(Zernike)}$	-2368.22	-2368.22	-2368.22	-2368.24	-2368.23	-2368.23	-2368.21	8.04e-05
$\mu_{51}^{(Zernike)}$	17036.03	17036.03	17036.03	17035.78	17036.05	17035.90	17036.19	1.38e-02
$\mu_{53}^{(Zernike)}$	-2289.09	-2289.09	-2289.09	-2289.14	-2289.09	-2289.12	-2289.06	5.31e-04
$\mu_{55}^{(Zernike)}$	-3323.07	-3323.07	-3323.07	-3323.10	-3323.07	-3323.09	-3323.05	3.00e-04
$\mu_{60}^{(Zernike)}$	51828.82	51828.82	51828.82	51828.67	51828.82	51828.74	51828.90	4.44e-03
$\mu_{62}^{(Zernike)}$	5734.00	5734.00	5734.00	5734.04	5734.01	5734.03	5733.97	4.91e-04
$\mu_{64}^{(Zernike)}$	1485.92	1485.92	1485.92	1485.86	1485.93	1485.88	1485.97	1.06e-03
$\mu_{66}^{(Zernike)}$	1983.84	1983.84	1983.84	1983.86	1983.84	1983.84	1983.82	1.66e-04

Fonte: o autor.

É corriqueiro aplicarmos etapas de pré-processamento às nuvens de pontos antes do processo de extração de atributos. São passos de pré-processamento mais comuns em sistema de classificação de face: remoção de *outliers* (remoção de pontos não constituintes da superfície do rosto), subamostragem (diminuição da quantidade de pontos da superfície para processamento mais rápido), correção de pose (alinhamento das faces para corrigir pequenas rotações durante aquisição) e definição da região de interesse (retenção apenas de parte característica da superfície). Para mostrar a capacidade de descrição do método de extração proposto, apenas duas etapas de pré-processamento foram realizadas: remoção de *outliers* e definição da região de interesse (também conhecido pela sigla em inglês ROI, ou *Region of Interest*).

Primeiro, cada uma das 299 faces foi submetida a um algoritmo estatístico de remoção de *outliers*, onde a distância de cada ponto para o centro geométrico foi comparada com um desvio padrão $\sigma = 1$ da massa de pontos da superfície. Depois, a etapa de definição da região de interesse foi escolhida como a definição de uma região central da face a partir do corte radial ao redor da ponta do nariz dada uma distância em milímetros. Cinco configurações diferentes de cortes foram testadas: 100mm, 90mm, 80mm, 70mm e 60mm. A localização da ponta do nariz usada para o corte foi provida pela própria base de dados; o problema de detecção deste ponto é tratado em outros trabalhos, como o de Liu *et al.* (2016). A Figura 10 ilustra os passos de pré-processamento das faces.

Figura 9 – Exemplos de faces da Base Bosphorus. I: face neutra. II: expressão de raiva. III: expressão de alegria. IV: oclusão de boca.



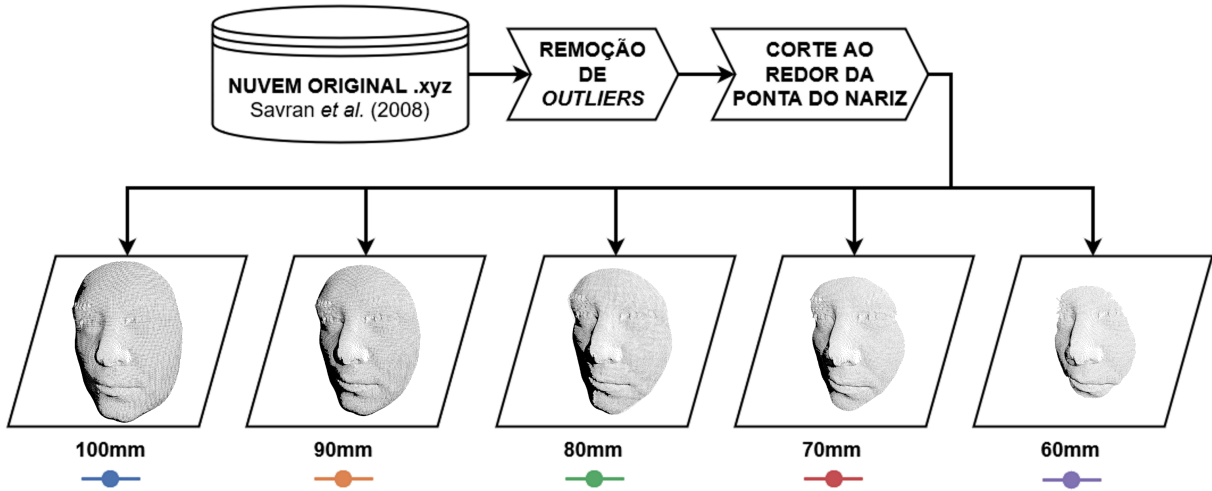
Fonte: Savran *et al.* (2008), adaptado pelo autor.

A etapa de subamostragem não foi executada para que a Matriz RABG pudesse guardar a maior quantidade de atributos geométricos invariantes possível, representando mais fielmente a superfície do rosto. A etapa de correção de pose não é necessária vista a propriedade de invariância à rotação da Matriz RABG.

4.3.2 *Classificador*

Usaremos o classificador Vizinho mais Próximo como modelo de predição. Seu funcionamento é descrito pelos seguintes passos:

Figura 10 – Visualização das etapas de pré-processamento submetidas às nuvens de faces.



Fonte: O autor.

Passo 1: Armazenar em memória o conjunto de treinamento X

Passo 2: Para cada nova observação x , buscar em X pelo índice do vetor de atributos mais próximo de x :

$$i^* = \arg \min_{i=1, \dots, N} \|x - x_i\|^2 \quad (4.15)$$

onde $\|v\|$ denota a norma euclidiana do vetor v .

Passo 3: Atribuir à observação x o índice de x_{i^*}

Sugerimos a análise do comportamento dos Momentos extraídos através do arcabouço proposto com outros modelos de classificador como pesquisa futura (Capítulo 5).

4.3.3 Taxas de classificação

Para cada um dos Momentos sugeridos, foram montados vetores de atributos de diferentes tamanhos definidos a partir dos parâmetros Υ de cada um deles. Os parâmetros dos Momentos sugeridos para o teste de classificação foram diferentes daqueles usados para o teste de invariância:

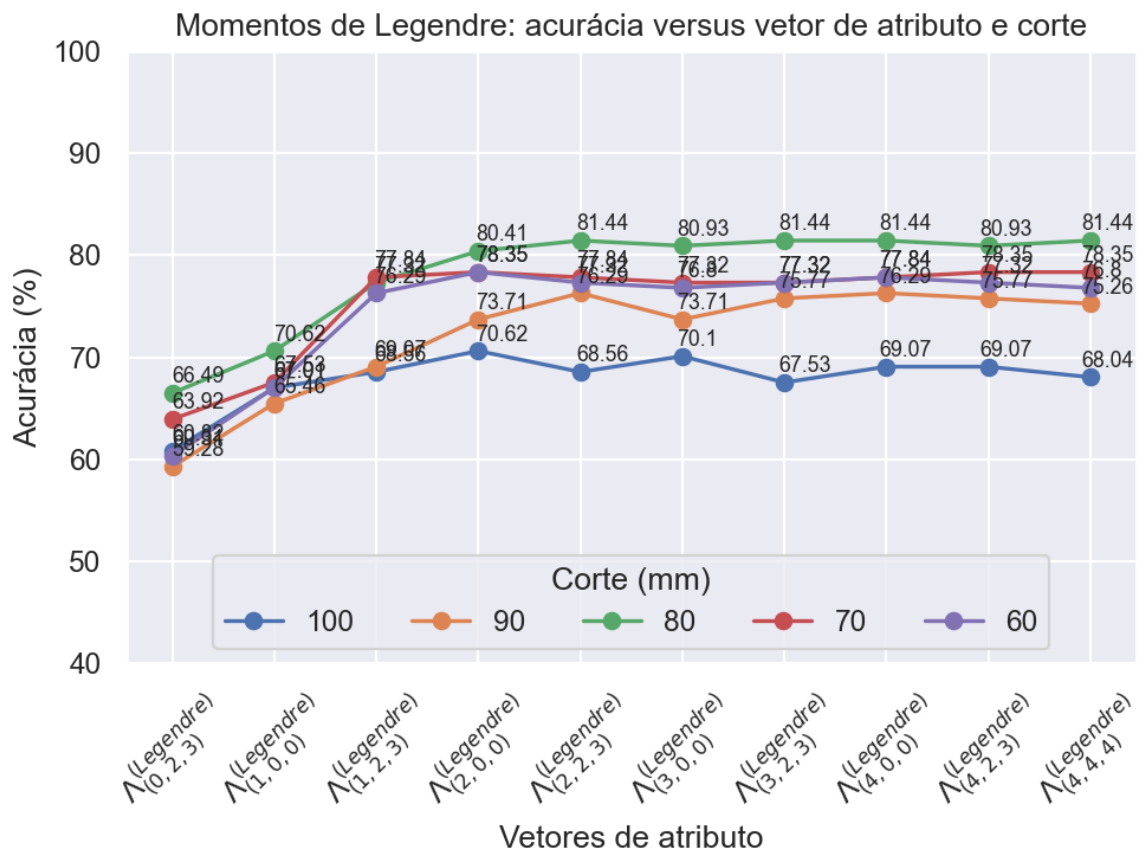
$$\begin{aligned} \Upsilon^{(Legendre)} &= \{p \in \{0, 1, 2, 3, 4\}, q \in \{0, 1, 2, 3, 4\}, r \in \{0, 1, 2, 3, 4\}\} \\ \Upsilon^{(Chebyshev)} &= \{p \in \{0, 1, 2, 3, 4\}, q \in \{0, 1, 2, 3, 4\}, r \in \{0, 1, 2, 3, 4\}\} \\ \Upsilon^{(Zernike)} &= \{n \in \{0, 1, 2, \dots, 20\}, m \in \{0, 1, 2, \dots, 20\}\} \end{aligned} \quad (4.16)$$

O processo de extração-classificação foi repetido cinco vezes, um para cada configuração de corte radial. Os resultados foram condensados nas Figuras 11, 12 e 13. Estão ordenados,

no eixo horizontal, a referência do vetor de atributos $\Lambda(M)_{\Upsilon}^{(g,f)}$ que gerou a acurácia respectiva (eixo vertical). Cada linha do gráfico representa um dos cortes radiais.

A Figura 11 traz os resultados para o vetor de atributos dos Momentos de Legendre sugeridos $\mu_{pqr}(M)^{(Legendre)}$. Vemos que existe uma tendência inicial de aumento na taxa de classificação quanto maior for o vetor de atributos extraído, mas que se estabiliza a partir de ordens maiores que 3. As faces cortadas 80mm radialmente ao redor da ponta do nariz correspondem aos cenários onde o classificador melhor identifica os indivíduos, com performance marginalmente superior em relação aos cortes de 60mm, 70mm e 90mm. A acurácia máxima obtida com o corte de 80mm foi de 81,44%, compartilhada por quatro diferentes vetores de atributo.

Figura 11 – Resultados: Momentos de Legendre.

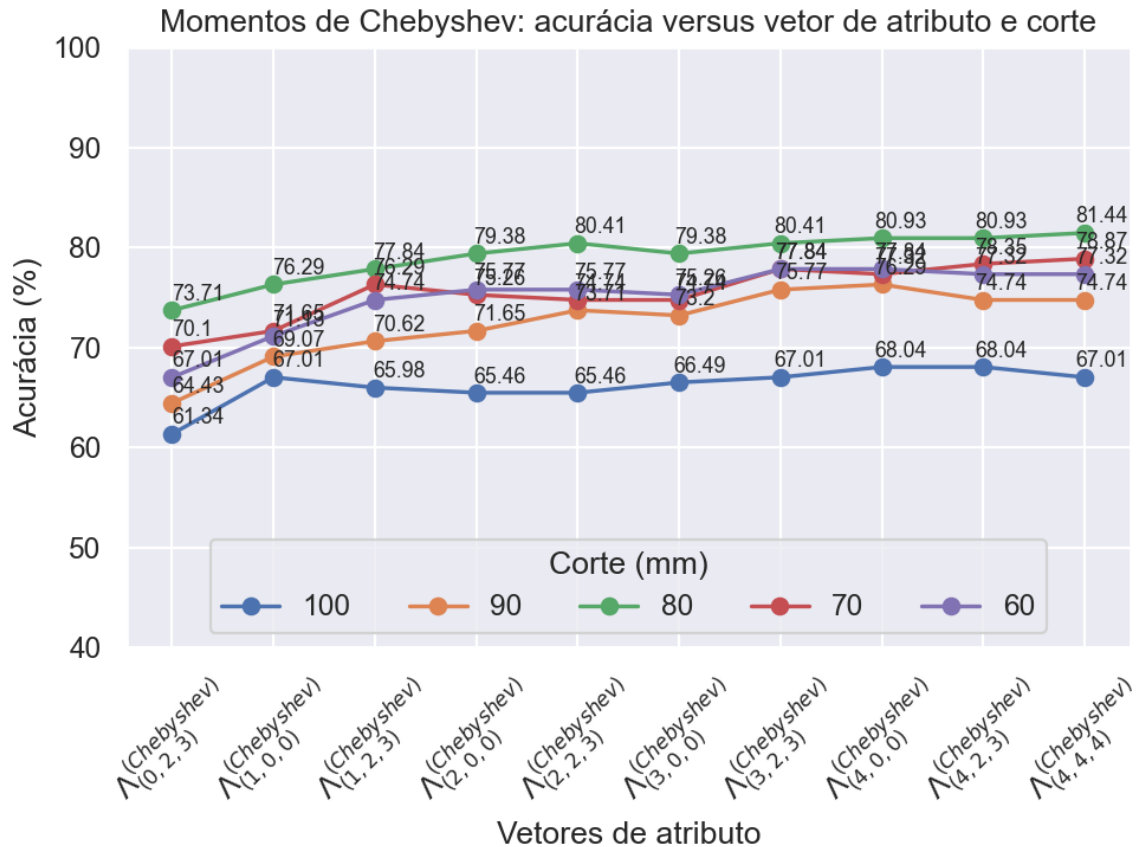


Fonte: O autor.

A Figura 12 traz os resultados para o vetor de atributos dos Momentos de Chebyshev sugeridos $\mu_{pqr}(M)^{(Chebyshev)}$. O padrão de evolução das taxas de acurácia versus vetor de atributos encontradas nos Momentos de Legendre também estão presente nesta outra formulação; a semelhança das funções de base de ambos, principalmente para baixas ordens, explica a

semelhança nos resultados. O corte de 80mm mais uma vez resultou em melhores taxas de classificação, mas apenas para o vetor de atributos de maior ordem $\Lambda(M)_{4,4,4}^{(Chebyshev)}$.

Figura 12 – Resultados: Momentos de Chebyshev.



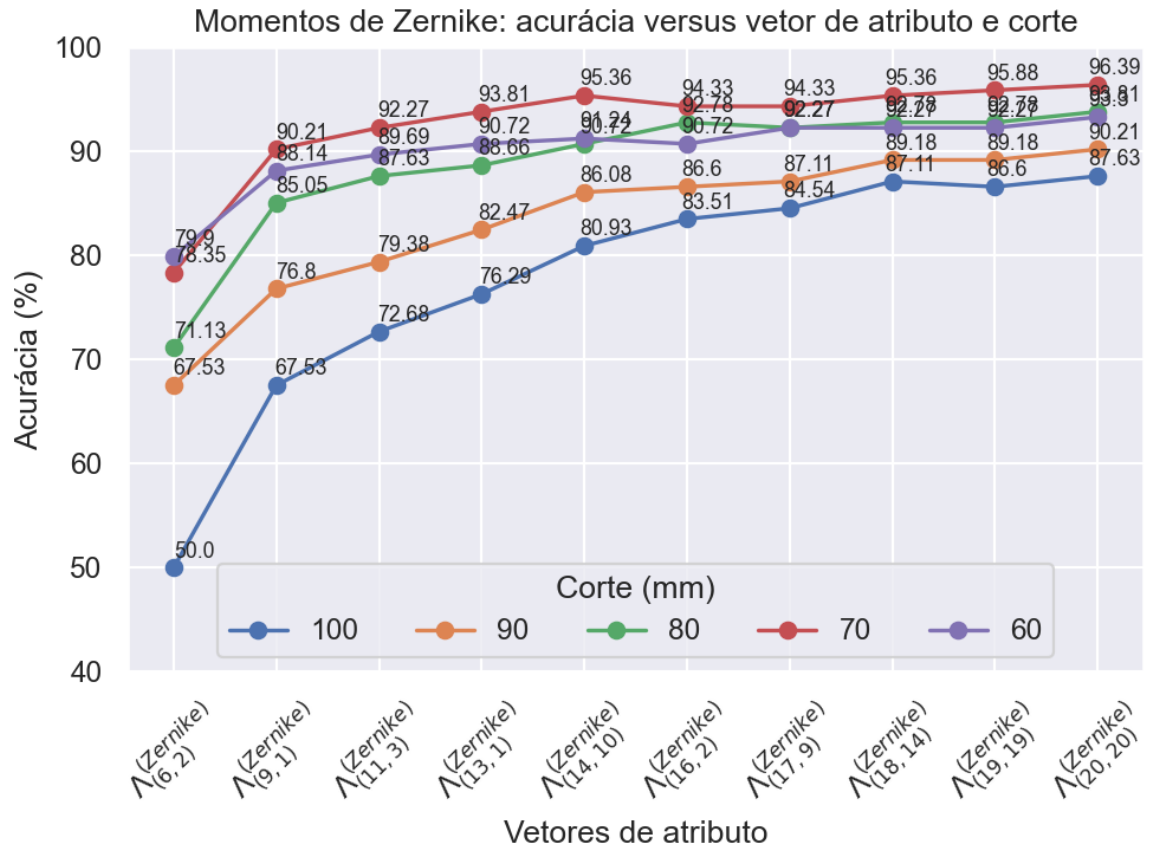
Fonte: O autor.

Vemos que a sugestão de Momentos de Legendre e Momentos de Chebyshev alcançaram resultados bastante semelhantes em termos de taxa de acurácia. A formulação semelhante dessas duas sugestões, tanto em termos de baixas ordens dos polinômios de base, quanto na escolha de função alvo (igual para ambos), explica os resultados de classificação equivalentes.

A Figura 13 traz os resultados para o vetor de atributos dos Momentos de Zernike $\mu_{nm}(M)^{(Zernike)}$. Essa formulação de extrator de atributos resultou nas melhores taxas de classificação, alcançando acerto de 96,39% para o vetor $\Lambda(M)_{20,20}^{(Zernike)}$. Além de uma performance de classificação superior aos dois primeiros Momentos analisados, a sugestão de Momentos de Zernike se diferenciaram com relação ao corte radial preferível, sendo aqui a distância de 70mm ao redor da ponta do nariz a melhor configuração para definição da região de interesse. Em comparação a outros métodos de classificação de indivíduos a partir de nuvens de pontos de faces, podemos concluir que a sugestão de Momentos de Zernike mostra como promissora a

capacidade descritiva do arcabouço de extração de atributos proposto.

Figura 13 – Resultados: Momentos de Zernike.



Fonte: O autor.

A pesquisa de Li *et al.* (2015) propõe um modelo de extração de atributos livre da etapa de registro a partir do cálculo de pontos de interesse na superfície do rosto que não variam com a rotação. Entretanto, essa etapa do algoritmo proposto por ele tem forte dependência com a superfície da face humana, não representando um método realmente invariante e, portanto, dificultando a generalização de seu uso em outros cenários.

O trabalho de Emambakhsh e Evans (2017) apontou a área próxima da ponta do nariz como a melhor região de interesse para classificação, resultado que converge com os obtidos pelos Momentos sugeridos. Entretanto, a região de interesse definida nesse trabalho exige, além da ponta do nariz, mais outros 6 pontos de interesse (*landmarkings*), além de um etapa de extração de atributos que apenas mitiga as possíveis invariâncias a partir de estimações de orientação e escala.

4.4 Considerações Finais sobre Resultados

Apresentamos neste capítulo resultados práticos de experimentos no contexto da invariância da Matriz RABG, montagem de um função de extração baseada em Momentos e uso do sistema de extração em tarefas de classificação, como por exemplo identificação de indivíduos. Os testes numéricos de invariância para os três Momentos sugeridos provam que um sistema de extração de atributos baseado em Momentos e na Matriz RABG goza de invariância afim completa. Os resultados de classificação na Base Bosphorus evidenciam a capacidade de descrição do método e as invariâncias carregadas pelos vetores de atributo ajudam a contornar exigências comuns ao sistema de extração-classificação, como, por exemplo, a etapa de correção de pose, desnecessária visto a invariância à rotação da Matriz RABG.

5 CONCLUSÕES

Os dois objetivos deste trabalho, um novo formato para nuvens de pontos e um arcabouço de sistema de extração de atributos baseado em Momentos, foram apresentados em termos de definições das partes que compõem o modelo proposto, de demonstrações matemáticas das características que se queria alcançar e por meio de experimentos numéricos com ilustrações em estudos de caso controlados. Espera-se que este trabalho contribua no debate sobre a análise de imagens tridimensionais e suas aplicações em Visão Computacional.

As demonstrações no Capítulo 3 e os estudos de caso apresentados no Capítulo 4 mostraram a invariância afim do formato de nuvem proposto, além de exibir o poder descritivo do método de extração frente um cenário de *benchmark* controlado. As taxas de classificação obtidas, com máximo em 96,39% a partir da sugestão de Momentos de Zernike que usa o arcabouço de extração proposto, se mostraram bastante competitivos frente a métodos de referência sem a grande dependência de etapas de pré-processamento destes. Alcançada a invariância completa e com resultados promissores no reconhecimento de indivíduos a partir da classificação de faces, mostramos que a contribuição é viável do ponto de vista prático.

É importante salientar algumas limitações. O primeiro atributo da Matriz RABG, a coluna de distância normalizadas, é bastante sensível a *outliers*. Poucos pontos dispersos não alteram a localização do centro geométrico, mas podem implicar numa distância máxima r_{max} discrepante em relação ao restante da distribuição da superfície, resultando uma normalização em escala comprometida.

Além disso, como os atributos da Matriz RABG são calculados a partir do centro geométrico e de uma primeira normalização na posição espacial (i.e invariância à translação), podemos dizer que o formato em si agrupa atributos que levam em consideração a superfície por completo da nuvem de pontos analisada. Algumas tarefas específicas podem exigir um particionamento da superfície para análises mais granulares e, portanto, o método de particionamento influenciará diretamente no formato das Matrizes RABG de saída. Sugere-se um estudo de como adaptar o algoritmo de geração do formato proposto para casos onde se quer analisar nuvens localmente, e não globalmente. Um possível ponto de partida para essa investigação é a reformulação da etapa de normalização na translação em função de um outro ponto de referência, contanto que este se mantenha fixo.

O estudo mais profundo de diferentes funções de base e funções alvo podem ajudar a esclarecer os limites do método de extração, suas vantagens e limitações, de maneira mais

objetiva. No contexto de classificação, sugere-se uma investigação cruzada entre Momentos e cenários de experimentação distintos, como reconhecimento facial, classificação de objetos e aprendizado por transferência.

O uso do formato Matriz RABG como entrada para um sistema de extração de atributos não se limita a uma formulação de Momentos, como proposto neste trabalho. De fato, podem-se aplicar outros modelos de extração, como análise de histogramas ou mesmo redes neurais, contanto que a investigação aponte como usar corretamente cada uma das colunas do formato de representação aqui apresentado no contexto do modelo de predição em questão.

Também não foi contemplado neste trabalho um estudo da capacidade de reconstrução de nuvens de pontos a partir dos atributos extraídos da Matriz RABG e do extrator de Momento propostos. Além da própria nuvem de pontos, incentiva-se as investigações em outros domínios, como as malhas triangulares e a representação em *voxels*, a fim de explorar a capacidade de reconstrução em formatos onde essa tarefa é mais explorada.

A implementação do método proposto e suas possíveis otimizações em arquiteturas do tipo GPU (*Graphics Processing Unit*, ou Unidade de Processamento Gráfico) poderão se favorecer de processamento paralelo, permitindo a extração de mais atributos em menos tempo e, portanto, abrindo espaço para uso em sistemas de tempo real.

Por fim, é possível avaliar uma generalização total do modelo de representação proposto. O formato Matriz RABG agrega atributos invariantes de natureza radial e angular. De fato, seu processo de montagem não faz limitação quanto à dimensão do espaço de entrada. Este trabalho abordou o contexto tridimensional com nuvens de pontos, mas é perfeitamente possível adaptar o formato para análise de imagens bidimensionais, ou quaisquer formas hiperdimensionadas. Um estudo da generalização dimensional da Matriz RABG pode revelar outros domínios de uso.

REFERÊNCIAS

- ABBAS, S.; FARHAN, S.; FAHIEM, M. A.; TAUSEEF, H. Efficient shape classification using zernike moments and geometrical features on mpeg-7 dataset. **Advances in Electrical and Computer Engineering**, Stefan cel Mare University of Suceava, v. 19, n. 1, p. 45–51, 2019.
- AL-KOFAHI, Y.; ZALTSMAN, A.; GRAVES, R.; MARSHALL, W.; RUSU, M. A deep learning-based algorithm for 2-d cell segmentation in microscopy images. **BMC bioinformatics**, BioMed Central, v. 19, n. 1, p. 1–11, 2018.
- ALEXANDRE, G. R.; SOARES, J. M.; THÉ, G. A. P. Systematic review of 3d facial expression recognition methods. **Pattern Recognition**, Elsevier, v. 100, p. 107108, 2020.
- BELLE, M.; GODEFROY, D.; COULY, G.; MALONE, S. A.; COLLIER, F.; GIACOBINI, P.; CHÉDOTAL, A. Tridimensional visualization and analysis of early human development. **Cell**, Elsevier, v. 169, n. 1, p. 161–173, 2017.
- BENOUINI, R.; BATIOUA, I.; ZENKOUAR, K.; ZAHI, A.; FADILI, H. E.; QJIDAA, H. Fast and accurate computation of racah moment invariants for image classification. **Pattern Recognition**, Elsevier, v. 91, p. 100–110, 2019.
- BESL, P. J.; MCKAY, N. D. Method for registration of 3-d shapes. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. **Sensor fusion IV: control paradigms and data structures**. [S. l.], 1992. v. 1611, p. 586–606.
- BUSSAB, W. d. O.; MORETTIN, P. A. Estatística básica. In: **Estatística básica**. [S. l.: s. n.], 2010. p. xvi–540.
- CANTERAKIS, N. 3d zernike moments and zernike affine invariants for 3d image analysis and recognition. In: **In 11th Scandinavian Conf. on Image Analysis**. [S. l.: s. n.], 1999. p. 85–93.
- DAVE, J.; ARMSTRONG, B. Computations of high-order associated legendre polynomials. **Journal of Quantitative Spectroscopy and Radiative Transfer**, Elsevier, v. 10, n. 6, p. 557–562, 1970.
- EMAMBAKSHH, M.; EVANS, A. Nasal Patches and Curves for Expression-Robust 3D Face Recognition. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 39, n. 5, p. 995–1007, may 2017. ISSN 01628828. Disponível em: <http://ieeexplore.ieee.org/document/7467565/>.
- FLUSSER, J.; ZITOVA, B.; SUK, T. **Moments and moment invariants in pattern recognition**. [S. l.]: John Wiley & Sons, 2009.
- GONZALEZ, R. C.; WOODS, R. C. **Processamento digital de imagens**. [S. l.]: Pearson Educación, 2009.
- GUO, X. Three dimensional moment invariants under rigid transformation. In: SPRINGER. **International Conference on Computer Analysis of Images and Patterns**. [S. l.], 1993. p. 518–522.
- HARRIS, C. R.; MILLMAN, K. J.; WALT, S. J. van der; GOMMERS, R.; VIRTANEN, P.; COURNAPEAU, D.; WIESER, E.; TAYLOR, J.; BERG, S.; SMITH, N. J.; KERN, R.; PICUS, M.; HOYER, S.; KERKWIJK, M. H. van; BRETT, M.; HALDANE, A.; R'IO, J. F. del; WIEBE,

M.; PETERSON, P.; G'ERARD-MARCHANT, P.; SHEPPARD, K.; REDDY, T.; WECKESSER, W.; ABBASI, H.; GOHLKE, C.; OLIPHANT, T. E. Array programming with NumPy. **Nature**, Springer Science and Business Media LLC, v. 585, n. 7825, p. 357–362, set. 2020. Disponível em: <https://doi.org/10.1038/s41586-020-2649-2>.

HAUSDORFF, F. Summationsmethoden und momentfolgen. i. **Mathematische Zeitschrift**, Springer, v. 9, n. 1-2, p. 74–109, 1921.

HJOUJI, A.; JAOUA, E.-M.; JOURHMANE, M. Object recognition and classification by mixed finite element method and invariants of orthogonal adapted-legendre moments. **International Journal on Interactive Design and Manufacturing (IJIDeM)**, Springer, v. 14, n. 3, p. 923–938, 2020.

HOSNY, K. M. Fast and low-complexity method for exact computation of 3d legendre moments. **Pattern Recognition Letters**, Elsevier, v. 32, n. 9, p. 1305–1314, 2011.

HOSNY, K. M.; SALAH, A.; SALEH, H. I.; SAYED, M. Fast computation of 2d and 3d legendre moments using multi-core cpus and gpu parallel architectures. **Journal of Real-Time Image Processing**, Springer, p. 1–15, 2019.

HUNTER, J. D. Matplotlib: A 2d graphics environment. **Computing in Science & Engineering**, IEEE COMPUTER SOC, v. 9, n. 3, p. 90–95, 2007.

KAUR, H.; PANNU, H. S. Zernike moments-based fingerprint recognition using weighted-support vector machine. **Modern Physics Letters B**, World Scientific, v. 33, n. 21, p. 1950245, 2019.

KOBBELT, L.; BOTSCH, M. A survey of point-based techniques in computer graphics. **Computers & Graphics**, Elsevier, v. 28, n. 6, p. 801–814, 2004.

LEVOY, M.; GERTH, J.; CURLESS, B.; PULL, K. The stanford 3d scanning repository. v. 5, 2005. Disponível em: <http://www-graphics.stanford.edu/data/3dscanrep>.

LI, H.; HUANG, D.; MORVAN, J.-M.; WANG, Y.; CHEN, L. Towards 3d face recognition in the real: a registration-free approach using fine-grained matching of 3d keypoint descriptors. **International Journal of Computer Vision**, Springer, v. 113, n. 2, p. 128–142, 2015.

LIU, J.; ZHANG, Q.; TANG, C. Nose tip detection and face localization from face range image based on multi-angle energy. In: SPRINGER. **International Conference on Technologies for E-Learning and Digital Entertainment**. [S. l.], 2016. p. 136–147.

MALLAHI, M. E.; MESBAH, A.; QJIDAA, H.; ZENKOUAR, K.; FADILI, H. E. Translation and scale invariants of three-dimensional tchebichef moments. In: IEEE. **2015 Intelligent Systems and Computer Vision (ISCV)**. [S. l.], 2015. p. 1–5.

MALLAHI, M. E.; ZOUHRI, A.; AFFAR, A. E.; TAHIRI, A.; QJIDAA, H. Radial hahn moment invariants for 2d and 3d image recognition. **International Journal of Automation and Computing**, Springer, v. 15, n. 3, p. 277–289, 2018.

MCKINNEY Wes. Data Structures for Statistical Computing in Python. In: WALT Stéfan van der; MILLMAN Jarrod (Ed.). **Proceedings of the 9th Python in Science Conference**. [S. l.: s. n.], 2010. p. 56 – 61.

- MESBAH, A.; MALLAHI, M. E.; LAKHILI, Z.; QJIDAA, H.; BERRAHOU, A. Fast and accurate algorithm for 3d local object reconstruction using krawtchouk moments. In: **IEEE. 2016 5th International Conference on Multimedia Computing and Systems (ICMCS)**. [S. l.], 2016. p. 1–6.
- Ming-Kuei Hu. Visual pattern recognition by moment invariants. **IEEE Transactions on Information Theory**, v. 8, n. 2, p. 179–187, feb 1962. ISSN 0018-9448. Disponível em: <http://ieeexplore.ieee.org/document/1057692/>.
- MUKUNDAN, R.; ONG, S.; LEE, P. A. Image analysis by tchebichef moments. **IEEE Transactions on image Processing**, IEEE, v. 10, n. 9, p. 1357–1364, 2001.
- NOVOTNI, M.; KLEIN, R. Shape retrieval using 3d zernike descriptors. **Computer-Aided Design**, Elsevier, v. 36, n. 11, p. 1047–1062, 2004.
- PANDEY, V. K.; SINGH, J.; PARTHASARATHY, H. Translation and scale invariance of 2d and 3d hahn moments. In: **IEEE. 2016 3rd International Conference on Signal Processing and Integrated Networks (SPIN)**. [S. l.], 2016. p. 255–259.
- PEDREGOSA, F.; VAROQUAUX, G.; GRAMFORT, A.; MICHEL, V.; THIRION, B.; GRISEL, O.; BLONDEL, M.; PRETTENHOFER, P.; WEISS, R.; DUBOURG, V.; OTHERS. Scikit-learn: Machine learning in Python. **Journal of machine learning research**, v. 12, n. Oct, p. 2825–2830, 2011.
- SAVRAN, A.; ALYÜZ, N.; DIBEKLIOĞLU, H.; ÇELIKTUTAN, O.; GÖKBERK, B.; SANKUR, B.; AKARUN, L. Bosphorus database for 3D face analysis. In: **Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)**. [S. l.: s. n.], 2008. v. 5372 LNCS, p. 47–56. ISBN 3540899901.
- SIQUEIRA, R. S.; ALEXANDRE, G. R.; SOARES, J. M.; THE, G. A. P. Triaxial Slicing for 3-D Face Recognition From Adapted Rotational Invariants Spatial Moments and Minimal Keypoints Dependence. **IEEE Robotics and Automation Letters**, v. 3, n. 4, p. 3513–3520, oct 2018. ISSN 2377-3766. Disponível em: <https://ieeexplore.ieee.org/document/8408720/>.
- SONG, Y.; WANG, J.; SHANG, J. Estimating effective leaf area index of winter wheat using simulated observation on unmanned aerial vehicle-based point cloud data. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, IEEE, v. 13, p. 2874–2887, 2020.
- STRANG, G. **Introduction to Linear Algebra**. Wellesley-Cambridge Press, 2016. ISBN 9780980232776. Disponível em: <https://books.google.com.br/books?id=efbxjwEACAAJ>.
- SZELISKI, R. **Computer vision: algorithms and applications**. [S. l.]: Springer Science & Business Media, 2010.
- TAHRI, O. A framework for moment invariants. **arXiv preprint arXiv:1807.06644**, 2018.
- TEAGUE, M. R. Image analysis via the general theory of moments. **J. Optical Soc. Am**, v. 70, n. 8, p. 920–930, 1980.

VIRTANEN, P.; GOMMERS, R.; OLIPHANT, T. E.; HABERLAND, M.; REDDY, T.; COURNAPEAU, D.; BUROVSKI, E.; PETERSON, P.; WECKESSER, W.; BRIGHT, J.; van der Walt, S. J.; BRETT, M.; WILSON, J.; MILLMAN, K. J.; MAYOROV, N.; NELSON, A. R. J.; JONES, E.; KERN, R.; LARSON, E.; CAREY, C. J.; POLAT, İ.; FENG, Y.; MOORE, E. W.; VanderPlas, J.; LAXALDE, D.; PERKTOLD, J.; CIMRMAN, R.; HENRIKSEN, I.; QUINTERO, E. A.; HARRIS, C. R.; ARCHIBALD, A. M.; RIBEIRO, A. H.; PEDREGOSA, F.; van Mulbregt, P.; SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. **Nature Methods**, v. 17, p. 261–272, 2020.

WANG, K.; ZHU, T.; GAO, Y.; WANG, J. Efficient Terrain Matching With 3-D Zernike Moments. **IEEE Transactions on Aerospace and Electronic Systems**, v. 55, n. 1, p. 226–235, feb 2019. ISSN 0018-9251. Disponível em: <https://ieeexplore.ieee.org/document/8395033/>.

WU, G.; XU, L. Shape description and recognition by implicit chebyshev moments. **Pattern Recognition Letters**, Elsevier, v. 128, p. 137–145, 2019.

XU, D.; LI, H. Geometric moment invariants. **Pattern Recognition**, v. 41, n. 1, p. 240–249, jan 2008. ISSN 00313203.

XU, H.; LIAO, S. Image reconstruction by chebyshev-fourier moments. In: IEEE. **2019 IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)**. [S. l.], 2019. p. 0314–0319.

YAMNI, M.; DAOUI, A.; KARMOUNI, H.; SAYYOURI, M.; QJIDAA, H. *et al.* Influence of krawtchouk and charlier moment's parameters on image reconstruction and classification. **Procedia computer science**, Elsevier, v. 148, p. 418–427, 2019.

ZERNIKE, v. F. Beugungstheorie des schneidenver-fahrens und seiner verbesserten form, der phasenkontrastmethode. **physica**, v. 1, p. 689–704, 1934.

ZHOU, S.; XIAO, S. 3d face recognition: a survey. **Human-centric Computing and Information Sciences**, Springer, v. 8, n. 1, p. 35, 2018.