



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA DE TELEINFORMÁTICA
CURSO DE GRADUAÇÃO EM ENGENHARIA DE COMPUTAÇÃO

ARTUR RODRIGUES ROCHA NETO

EXTRAÇÃO DE ATRIBUTOS DE FACES EM NUVENS DE PONTOS USANDO
MOMENTOS INVARIANTES ADAPTADOS

FORTALEZA

2017

ARTUR RODRIGUES ROCHA NETO

EXTRAÇÃO DE ATRIBUTOS DE FACES EM NUENS DE PONTOS USANDO
MOMENTOS INVARIANTES ADAPTADOS

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Engenharia de Computação do Centro de Tecnologia da Universidade Federal do Ceará, como requisito parcial à obtenção do grau de bacharel em Engenharia de Computação.

Orientador: Prof. Dr. José Marques Soares

FORTALEZA

2017

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)

R571e Rocha Neto, Artur Rodrigues.

EXTRAÇÃO DE ATRIBUTOS DE FACES EM NUENS DE PONTOS USANDO MOMENTOS
INVARIANTES ADAPTADOS / Artur Rodrigues Rocha Neto. – 2017.
42 f. : il. color.

Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Centro de Tecnologia,
Curso de Engenharia de Computação, Fortaleza, 2017.

Orientação: Prof. Dr. José Marques Soares.

1. Nuens de pontos. 2. Momentos de Hu. 3. Momentos de Zernike. 4. Momentos de Legendre. 5.
Momentos de Tchebichef. I. Título.

CDD 621.39

ARTUR RODRIGUES ROCHA NETO

EXTRAÇÃO DE ATRIBUTOS DE FACES EM NUENS DE PONTOS USANDO
MOMENTOS INVARIANTES ADAPTADOS

Trabalho de Conclusão de Curso apresentado
ao Curso de Graduação em Engenharia de
Computação do Centro de Tecnologia da
Universidade Federal do Ceará, como requisito
parcial à obtenção do grau de bacharel em
Engenharia de Computação.

Aprovada em:

BANCA EXAMINADORA

Prof. Dr. José Marques Soares (Orientador)
Universidade Federal do Ceará (UFC)

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

Prof. Dr. Giovanni Cordeiro Barroso
Universidade Federal do Ceará (UFC)

in memoriam de Mariza Monteiro Rodrigues

AGRADECIMENTOS

Agradeço a Ageisa, Cleber, Samuel, Mariza, Artur, Armízia, Artemísia e Maria por serem a família mais carinhosa que uma pessoa pode e por me guiarem com tanto amor e respeito.

A Glícia do Nascimento Oliveira, por ser a companheira e amante mais especial que eu poderia desejar, pelo apoio e compreensão nesses e nos vários anos que virão.

Aos companheiros de pesquisa Luiz Felipe Feitosa Leite e Gil Ribeiro Alexandre por terem me ajudado a crescer como profissional no período mais rico do curso e por serem a melhor equipe com quem já trabalhei.

Ao meu orientador, professor José Marques Soares, por ter me dado uma chance há quatro anos e, desde então, ter me orientado com capacidade e amizade no meu caminho acadêmico.

Aos professores Robson da Silva Siqueira e George André Pereira Thé por compartilharem uma gama de conhecimentos valiosíssimos para a criação deste trabalho.

Aos meus amigos de longa data Hamlet Victor Oliveira, Jy Ferreira Gomes, Emanuel Carvalho e Tiago Rodrigues, por estarem ao meu lado nos momentos mais alegres e por me ajudarem nos mais difíceis.

Ao Doutorando em Engenharia Elétrica, Ednardo Moreira Rodrigues, e seu assistente, Alan Batista de Oliveira, aluno de graduação em Engenharia Elétrica, pela adequação do *template* utilizado neste trabalho para que o mesmo ficasse de acordo com as normas da biblioteca da UFC.

Por fim, gostaria de agradecer a todos aqueles com quem trabalhei e convivi nesses oito anos de Universidade Federal do Ceará (UFC) em todos os lugares por onde passei (LATIN, LESC, Seara da Ciência e RESID), por terem deixado em mim um pouco de todos vocês.

“Pelo amor de Deus, atente-se ao último passo do algoritmo: "nunca acrescente leite condensado à caipirinha".”

(Alexandre Moreira de Moraes)

RESUMO

A popularização de sensores de imagens de profundidade, hoje presentes em aparelhos celulares e dispositivos similares, tem popularizado o uso de imagens tridimensionais para identificação e reconhecimento de indivíduos e suas expressões faciais. Neste trabalho são discutidos o problema da extração de atributos de faces representadas por nuvens de pontos e as técnicas de reconhecimento de indivíduos por meio dessas estruturas de dados. São propostas quatro técnicas de extração de atributos, adaptadas a partir dos momentos invariantes de Hu, Legendre, Zernike e Tchebichef. Foi analisada a possibilidade de redução da região da face com a eliminação dos pontos mais distantes do centro geométrico da nuvem. O objetivo é avaliar se regiões menores e mais concentradas são capazes de conter suficiente descritibilidade para identificação dos indivíduos. A qualidade dos atributos extraídos foi avaliada usando-os como entrada de um sistema de classificação empregando os classificadores k-Vizinhos Mais Próximos e Máquina de Vetor de Suporte. Amostras para teste e treinamento do sistema de classificação foram extraídas da base de dados Bosphorus. Os reconhecimentos foram avaliados em termos de taxa de acerto dos classificadores e tempo de extração dos momentos. O extrator de características baseado nos momentos de Hu, aliado à extração da face completa e ao classificador k-Vizinhos mais Próximos, obteve a maior taxa de acerto, 58,53%, e tempo médio de extração/classificação de 0,1687 segundos por nuvem.

Palavras-chave: Nuvens de pontos. Momentos Invariantes. Momentos de Hu. Momentos de Zernike. Momentos de Legendre. Momentos de Tchebichef. k-Vizinhos Mais Próximos. Máquina de Vetor de Suporte.

ABSTRACT

The popularization of depth image sensors, now embedded in smartphones and similar devices, has popularized the use of tridimensional images for identification and recognition of individuals and their facial expressions. In this work are discussed the problem of featuring extraction from faces represented by point clouds and the recognition techniques by those data structures. Four featuring extraction techniques are proposed, which were adapted from the moment invariants of Hu, Legendre, Zernike and Tchebichef. A cloud sectioning approach based on the distance from the geometric center of a cloud to all of its points was studied in order to analyse the different regions of facial characterization. The quality of the extracted features was evaluated using two classification models: k-Nearest Neighbours and Support Vector Machines. The featuring extraction based on the adapted Hu Moment Invariants and the k-Nearest Neighbours classifier resulted in the best accuracy score of 58.53% with a average time of extraction/classification of 0.1687 seconds.

Keywords: Point Clouds. Moment Invariants. Hu. Zernike. Legendre. Tchebichef. k-Nearest Neighbours. Support Vector Machines.

LISTA DE ILUSTRAÇÕES

Figura 1 – Comparação entre a aquisição 3D e 2D do mesmo indivíduo	27
Figura 2 – Influência frontal entre os cortes 50, 60, 70, 80, 90 e nuvem completa	32
Figura 3 – Influência lateral entre os cortes 50, 60, 70, 80, 90 e nuvem completa	32
Figura 4 – <i>Pipeline</i> de extração-classificação	32

LISTA DE TABELAS

Tabela 1 – Subconjuntos de extração	28
Tabela 2 – Número de atributos extraídos por cada classificador	33
Tabela 3 – Tempo médio de extração e classificação (em segundos)	33
Tabela 4 – Classificação: Neutro <i>vs</i> Neutro	35
Tabela 5 – Classificação: Neutro <i>vs</i> Não-Neutro	35
Tabela 6 – Classificação: Neutro <i>vs</i> Completa	36
Tabela 7 – Média das taxas de acerto para cada corte no Teste Neutro	36

LISTA DE ABREVIATURAS E SIGLAS

k-NN	k-Vizinhos mais Próximos
k-NN-E	k-Vizinhos mais Próximos com distância euclidiana
k-NN-M	k-Vizinhos mais Próximos com distância <i>Manhattam</i>
SVM	Máquina de Vetor de Suporte
SVM-P	Máquina de Vetor de Suporte com <i>kernel</i> polinomial
SVM-R	Máquina de Vetor de Suporte com <i>kernel</i> radial

LISTA DE SÍMBOLOS

H_i	Momento de Hu
I_{xy}	Função-imagem
L_{pq}	Momentos de Legendre 2D
L_{pqr}	Momentos de Legendre 3D
m_{pq}	Momentos Geométricos/Regulares de uma Imagem
m_{pqr}	Momentos Geométricos/Regulares de uma Nuvem de Pontos
M, N	Dimensões de uma Imagem em <i>pixels</i>
m	Repetição dos Momentos de Zernike
n	Ordem dos Momentos de Zernike
p	Ordem da Dimensão x
P_p	Polinômio de Legendre
q	Ordem da Dimensão y
r	Ordem da Dimensão z
R_{nm}	Polinômios Radiais de Zernike
t_p	Polinômios de Tchebichef
T_{pq}	Momentos de Tchebichef 2D
T_{pqr}	Momentos de Tchebichef 3D
V_{nm}	Polinômios Complexos de Zernike
Z_{nm}	Momentos de Zernike 2D/3D
λ_{pq}	Constante de Normalização dos Momentos de Legendre 2D
λ_{pqr}	Constante de Normalização dos Momentos de Legendre 3D
Λ	Número de Pontos de uma Nuvem de Pontos
μ_{pq}	Momentos Centrais de uma Imagem
μ_{pqr}	Momentos Centrais de uma Nuvem de Pontos
η_{pq}	Momentos Normalizados de uma Imagem
η_{pqr}	Momentos Normalizados de uma Nuvem de Pontos

θ	Ângulo azimutal da Distância Radial de Zernike 2D/3D
ρ_{xy}	Distância Radial de Zernike 2D
ρ_{xyz}	Distância Radial de Zernike 3D
ϕ	Número de <i>pixels</i> /pontos no interior do círculo/esfera unitário(a)
$\tau(p, \Lambda)$	Constante de Normalização dos Momentos de Tchebichef 3D
$\tau(p, N)$	Constante de Normalização dos Momentos de Tchebichef 2D

SUMÁRIO

1	INTRODUÇÃO	16
2	EXTRAÇÃO DE ATRIBUTOS	17
2.1	Imagens Tridimensionais	17
2.2	Momentos Invariantes	18
2.2.1	<i>Momentos de Hu</i>	18
2.2.2	<i>Momentos de Legendre</i>	20
2.2.3	<i>Momentos de Zernike</i>	21
2.2.4	<i>Momentos de Tchebichef</i>	22
2.3	Considerações Finais deste Capítulo	23
3	CLASSIFICADORES	24
3.1	Classificador k-Vizinhos mais Próximos (k-NN)	24
3.1.1	<i>Distância Euclidiana</i>	25
3.1.2	<i>Distância Manhattan</i>	25
3.2	Classificador Máquina de Vetor de Suporte	25
3.3	Considerações Finais deste Capítulo	26
4	ADAPTAÇÃO DOS MOMENTOS INVARIANTES E METODOLOGIA	27
4.1	Base de dados	27
4.2	<i>Hardware e Software</i> utilizados	28
4.3	Adaptação dos Momentos Invariantes	28
4.3.1	<i>Hu 3D</i>	29
4.3.2	<i>Legendre 3D</i>	29
4.3.3	<i>Zernike 3D</i>	30
4.3.4	<i>Tchebichef 3D</i>	30
4.4	Seccionamento das Nuvens de Pontos	31
4.5	<i>Pipeline</i> de extração-classificação	31
4.6	Considerações Finais deste Capítulo	32
5	RESULTADOS E DISCUSSÃO	33
5.1	Resultados Experimentais I - Extração de Características	33
5.2	Resultados Experimentais II - Classificação	34
5.3	Discussão dos Resultados	36

6	CONCLUSÃO	37
	REFERÊNCIAS	39
	APÊNDICES	41
	APÊNDICE A – Momentos de Hu 3D	41

1 INTRODUÇÃO

A capacidade do ser humano de identificar seus semelhantes é de grande importância na interação dele com o ambiente. O avanço da Inteligência Artificial e dos métodos de reconhecimento de padrões já permite modelarmos técnicas capazes de identificar indivíduos através de imagens 2D com alta precisão. É natural que, com o advento de novas tecnologias para aquisição de imagens em novos formatos, novas técnicas de reconhecimento sejam necessárias.

Este trabalho propõe a adaptação de técnicas de extração de atributos usadas em imagens 2D para o uso em nuvens de pontos. O objeto de estudo para a análise dessas técnicas de extração foram faces de indivíduos humanos armazenadas em uma base de dados de ampla utilização na literatura. Além disso, o trabalho apresenta a avaliação de técnicas de classificação em busca da melhor combinação para a identificação de indivíduos.

Algumas regiões da face humana guardam uma carga maior de descritibilidade do que outras (EKMAN; FRIESEN, 1976). Para avaliar o efeito dessas regiões no problema de classificação, uma técnica simples de seccionamento baseado no centro geométrico da nuvem de pontos foi usada e estudada.

Os conceitos que serviram de base teórica e que respaldam a abordagem adotada neste trabalho são: a noção de imagens tridimensionais, o estudo dos momentos invariantes e os modelos de classificação adotados em reconhecimento de padrões. O objetivo deste trabalho é avaliar a qualidade das modificações realizadas em técnicas de extração de atributos baseadas em momentos invariantes e a capacidade de reconhecimento de indivíduos usando dois métodos de classificação.

O Capítulo 2 apresenta a teoria básica sobre Momentos Invariantes e descrições mais detalhadas dos quatro momentos escolhidos para este trabalho: Hu, Legendre, Zernike e Tchebichef.

O Capítulo 3 faz uma breve introdução sobre os métodos de classificação escolhidos: k-Vizinhos mais Próximos e Máquina de Vetor de Suporte.

O Capítulo 4 traz as adaptações propostas para os quatro momentos invariantes e a metodologia de extração de atributos e classificação de indivíduos. Um *pipeline* é descrito de forma a facilitar a reprodução e análise dos resultados.

O Capítulo 5 revela os resultados gerados neste trabalho com vistas ao tempo de extração e a taxa de acerto de classificação.

O Capítulo 6 apresenta as conclusões do trabalho e sugestões de trabalhos futuros.

2 EXTRAÇÃO DE ATRIBUTOS

O uso de estruturas de dados 3D para o armazenamento de faces vem ganhando popularidade nos últimos anos. As tecnologias para esta finalidade vem se tornando mais acessíveis e já estão presentes em vários dispositivos, inclusive com sensores embarcados em aparelhos celulares e similares. Abordagens baseadas em imagens 3D possuem vantagens em relação ao uso de imagens 2D convencionais, como independência de iluminação e de *viewpoint*, e vêm sendo muito usada em processos de reconhecimento e autenticação de indivíduos (SAVRAN *et al.*, 2008). Os dados 3D podem ser armazenados em diversos formatos, mais comumente como *Point Clouds*, ou nuvens de pontos. O objeto de estudo deste trabalho são faces representadas por nuvens de pontos.

Assim como as imagens 2D, as nuvens de pontos também guardam uma quantidade consideravelmente grande de informação. No contexto de visão computacional, é altamente custoso executar operações sobre um conjunto de dados dessa natureza. Isso requer o uso de técnicas que reduzam o conjunto de informações, mas sem perda de descritibilidade do mesmo. A extração de atributos pode ser vista como uma forma de reduzir a dimensionalidade de um conjunto complexo, preservando suas características principais em um único vetor que chamamos de vetor de atributos.

O objetivo deste capítulo é apresentar as técnicas de extração de atributos usadas neste trabalho para a classificação de faces armazenadas como nuvens de pontos. Os momentos invariantes mostrados estão em suas formas originais definidas para o espaço 2D.

2.1 Imagens Tridimensionais

Uma imagem bidimensional, ou imagem 2D, é definida por Gonzalez e Woods (2000) como uma função $f(x, y)$, em que x e y são coordenadas espaciais e f é a amplitude ou intensidade da imagem naquela coordenada. Quando x , y e $f(x, y)$ são valores finitos e discretos, nos referimos à imagem como digital. Imagens digitais 2D são, portanto, projeções no espaço bidimensional de objetos tridimensionais. A notação usada para imagens digitais 2D é I_{xy} .

Uma outra forma de descrever objetos é por meio de nuvens de pontos. Nuvens de pontos são uma estrutura de dados utilizada para representar pontos multidimensionais - mais comumente tridimensionais. Cada ponto da nuvem de pontos 3D corresponde ao conjunto de coordenadas geométricas X , Y e Z de uma certa amostra de superfície (RUSU; COUSINS, 2011).

Uma nuvem de pontos representa, portanto, o que chamamos de imagem tridimensional ou imagem 3D e pode ser capturada por meio de dispositivos como câmeras estéreo, sensores 3D e câmeras *time-of-flight* (ToF).

A crescente disponibilidade de dispositivos de aquisição 3D e as vantagens das imagens tridimensionais em relação às imagens 2D (independência das condições de iluminação e de *viewpoint*) justificam a popularização de técnicas de reconhecimento de indivíduos e de expressões faciais (SAVRAN *et al.*, 2008).

2.2 Momentos Invariantes

Momentos invariantes estão entre as técnicas de reconhecimento de padrões mais utilizadas para classificação de entidades. Chamam-se invariantes porque uma de suas propriedades principais é a invariância às transformações básicas de rotação, translação e escala. Momentos são valores escalares usados para caracterizar uma função e capturar suas características principais. Eles podem ser vistos como projeções de uma função em uma base polinomial ou como o acúmulo de energia em uma região do espaço (FLUSSER *et al.*, 2009).

Quatro momentos invariantes foram escolhidos para estudo neste trabalho: Momentos de Hu (HU, 1962), Momentos de Legendre (TEAGUE, 1980), Momentos de Zernike (TEAGUE, 1980) e Momentos de Tchebichef (MUKUNDAN *et al.*, 2001). Tais técnicas foram originalmente concebidos como ferramentas de extração de atributos para o domínio de duas dimensões. Neste trabalho são propostas adaptações dessas técnicas para o espaço tridimensional.

2.2.1 Momentos de Hu

O conceito de momento de imagem foi concebido por Hu (1962) como uma forma de reconhecer padrões visuais independente de tamanho, orientação e posição e tem sido amplamente utilizado em visão computacional (KEYES; WINSTANLEY, 2001; ZHANG *et al.*, 2015).

Hu (1962) partiu do conceito de momento regular (ou momento geométrico), sendo m_{pq} o momento regular de ordem $(p + q)$ de uma imagem $M \times N$ caracterizada por I_{xy} :

$$m_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} x^p y^q I_{xy} \quad (2.1)$$

Os momentos regulares nos permitem encontrar informações importantes sobre uma

imagem. O momento de ordem 0 (m_{00}) calcula o número de pixels da imagem, ou área da imagem. Os momentos de ordem 1 (m_{10} e m_{01}) retornam, respectivamente, a densidade de pixels nas direções x e y (no caso de uma imagem binária).

A partir das informações de número de pixels e densidade relativa, podemos encontrar o centro de massa (\bar{x}, \bar{y}) de uma imagem usando:

$$(\bar{x}, \bar{y}) = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right) \quad (2.2)$$

De posse do centro de massa da imagem (\bar{x}, \bar{y}) , podemos calcular os chamados momentos centrais μ_{pq} de ordem $(p + q)$:

$$\mu_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (x - \bar{x})^p (y - \bar{y})^q I_{xy} \quad (2.3)$$

Os momentos centrais nada mais são que a aplicação dos momentos geométricos em relação ao centro de massa, ou seja, eles são o somatório dos pesos dos pixels (x, y) em função das distâncias desses pixels ao centro de massa (\bar{x}, \bar{y}) de uma imagem I_{xy} . O uso dessa distância relativa como métrica torna os momentos centrais invariantes à translação.

Para alcançarmos a invariância à escala, precisamos adicionar uma medida de normalização aos momentos centrais, assim chegando aos momentos normalizados η_{pq} de ordem $(p + q)$ definidos por:

$$\eta_{pq} = \frac{\mu_{pq}}{(\mu_{00})^\gamma} \quad \text{na qual:} \quad \gamma = \frac{p+q}{2} + 1, \quad p+q \geq 2 \quad (2.4)$$

Por fim, os sete momentos de Hu são calculados a partir de combinações dos momentos normalizados η_{pq} até ordem 3:

$$H_1 = \eta_{20} + \eta_{02} \quad (2.5)$$

$$H_2 = (\eta_{20} - \eta_{02})^2 + 4(\eta_{11})^2 \quad (2.6)$$

$$H_3 = (\eta_{30} - 3\eta_{12})^2 + 3(\eta_{03} - 3\eta_{21})^2 \quad (2.7)$$

$$H_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{03} + \eta_{21})^2 \quad (2.8)$$

$$H_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{03} + \eta_{21})^2] \\ + (3\eta_{21} - \eta_{03})(\eta_{03} + \eta_{21})[3(\eta_{30} + \eta_{12})^2 - (\eta_{03} + \eta_{21})^2] \quad (2.9)$$

$$H_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - 7(\eta_{03} + \eta_{21})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{03} + \eta_{21}) \quad (2.10)$$

$$H_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{03} + \eta_{21})^2] \\ + (\eta_{30} - 3\eta_{12})(\eta_{03} + \eta_{21})[3(\eta_{30} + \eta_{12})^2 - (\eta_{03} + \eta_{21})^2] \quad (2.11)$$

2.2.2 Momentos de Legendre

Os momentos de Legendre foram introduzidos por Teague (1980). Sua base são os polinômios de Legendre $P_p(x)$, uma relação recursiva da forma (CHONG *et al.*, 2004):

$$P_p(x) = \frac{(2p-1)xP_{p-1}(x) - (p-1)P_{p-2}(x)}{p}, \quad \text{na qual: } P_0(x) = 1, \quad P_1(x) = x, \quad p > 1 \quad (2.12)$$

Os polinômios de Legendre são definidos no intervalo $[-1, 1]$. Dessa forma, para uma imagem de $M \times N$ *pixels*, precisamos normalizar o espaço da imagem para esse mesmo intervalo $[-1, 1]$. Fazemos isso usando uma constante de normalização λ_{pq} , calculada em função da área da imagem e da ordem do momento:

$$\lambda_{pq} = \frac{(2p+1)(2q+1)}{MN} \quad (2.13)$$

e, depois, normalizando cada coordenada de *pixels* (x, y) da imagem em função da ordem do momento:

$$x_i = \frac{2i}{M-1} - 1 \quad \text{e} \quad y_j = \frac{2j}{N-1} - 1 \quad (2.14)$$

Usando o momento regular 2.1, a expressão dos polinômios de Legendre 2.12 e as métricas de normalização 2.13 e 2.14, obtemos a forma dos momentos discretos de Legendre L_{pq} de ordem $(p + q)$:

$$L_{pq} = \lambda_{pq} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} P_p(x_i) P_q(y_j) I_{ij} \quad (2.15)$$

2.2.3 Momentos de Zernike

Os momentos de Zernike, assim como os de Legendre, também foram propostos por Teague (1980). Eles também são calculados a partir de uma relação polinomial: os polinômios complexos de Zernike.

Os polinômios complexos de Zernike, como o nome sugere, são definidos no espaço do círculo unitário complexo. É necessário, então, remapear as coordenadas da imagem para dentro do círculo unitário.

Seja uma imagem de $M \times N$ *pixels*. A origem do círculo será o centro da imagem (\bar{x}, \bar{y}) . Seja o raio do círculo r . Os *pixels* fora do círculo são desprezados e os *pixels* internos terão suas coordenadas normalizadas em função das distâncias de cada pixel à origem do círculo. O primeiro passo para a obtenção dos momentos de Zernike é o cálculo dos polinômios radiais de Zernike $R_{nm}(\rho)$:

$$R_{nm}(\rho_{xy}) = \sum_{s=0}^{\frac{n-|m|}{2}} c(n, m, s) \rho^{n-2s}, \quad \text{na qual: } c(n, m, s) = (-1)^s \frac{(n-s)!}{s! \left(\frac{n+|m|}{2} - s\right)! \left(\frac{n-|m|}{2} - s\right)!} \quad (2.16)$$

n e m são inteiros chamados de ordem e repetição, com $n > 0$ e $0 \leq |m| \leq n$. ρ é o tamanho do vetor que vai do centro do círculo unitário (\bar{x}, \bar{y}) até o pixel (x, y) , normalizado pelo tamanho r do raio do círculo:

$$\rho_{xy} = \frac{\sqrt{(x - \bar{x})^2 + (y - \bar{y})^2}}{r}, \quad \text{na qual: } 0 \leq \rho_{xy} \leq 1 \quad (2.17)$$

A partir dos polinômios radiais de Zernike, Teague (1980) mostrou uma nova formulação polar que representa um conjunto ortogonal definido no interior do círculo unitário

$x^2 + y^2 = 1$, e a batizou de polinômios complexos de Zernike:

$$V_{nm}(x, y) = V_{nm}(\rho_{xy}, \theta) = R_{nm}(\rho_{xy}) \exp(jm\theta), \quad \text{na qual: } j = \sqrt{-1}, \quad \theta = \arctan\left(\frac{y}{x}\right) \quad (2.18)$$

Assim, partindo da fórmula do momento regular 2.1 e da expressão dos polinômios complexos de Zernike 2.18, encontramos os momentos de Zernike Z_{nm} de ordem $(n + m)$:

$$Z_{nm} = \frac{n+1}{\phi} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} V_{nm}(\rho_{xy}, \theta) I_{xy} \quad (2.19)$$

na qual $\frac{n+1}{\phi}$ é o fator de normalização de Zernike com ϕ igual ao número de *pixels* localizados no interior do círculo unitário.

2.2.4 Momentos de Tchebichef

Os momentos de Zernike e Legendre mostrados nas seções anteriores são robustos e amplamente usados em diversas aplicações de processamento digital de imagens e reconhecimento de padrões (HADDADNIA *et al.*, 2002; RAO *et al.*, 2010; ZHANG *et al.*, 2011). Entretanto, ambos apresentam limitações quanto ao espaço onde são definidos.

Os momentos de Legendre são válidos apenas no intervalo $[-1, 1]$ e os momentos de Zernike apenas no interior do círculo unitário. O cálculo desses momentos requer transformações de coordenadas e aproximação de integrais contínuas, o que podem levar a erros numéricos no processo de discretização. O núcleo dos momentos de Legendre e Zernike apresentam potências de alto grau $(p + q)$ ou $(m + n)$, causando grandes variações para diferentes ordens.

Para superar essas limitações, Mukundan *et al.* (2001) introduziu um conjunto de funções de base ortogonais discretos com base nos polinômios de Tchebichef, cuja implementação não envolve nenhuma aproximação numérica, uma vez que o seu conjunto de base é ortogonal no domínio discreto das coordenadas da imagem. Os polinômios de Tchebichef $t_p(x)$ de uma imagem I_{xy} com $N \times N$ *pixels* são definidos como:

$$t_p(x) = \frac{(2p-1)t_1(x)t_{p-1}(x) - (p-1)\left(1 - \frac{(p-1)^2}{N^2}\right)t_{p-2}(x)}{p} \quad (2.20)$$

na qual $p = 2, 3, 4, \dots, N-1$, $t_0(x) = 1$ e $t_1(x) = \frac{(2p+1-N)}{N}$.

O fator de normalização de Tchebichef é escrito como:

$$\tau(p, N) = N \left[\frac{\prod_{k=1}^p (1 - \frac{k^2}{N^2})}{2p+1} \right] \quad (2.21)$$

Então, usando da fórmula do momentos regular 2.1, os polinômios de Tchebichef 2.20 e o fator de normalização 2.21, definimos os momentos de Tchebichef T_{pq} de ordem $(p+q)$ como:

$$T_{pq} = \frac{1}{\tau(p, N) \tau(q, N)} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} t_p(x) t_q(y) I_{xy} \quad (2.22)$$

2.3 Considerações Finais deste Capítulo

Este capítulo apresentou a teoria dos momentos invariantes e os extratores de atributos escolhidos para este trabalho: momentos de Hu, Legendre, Zernike e Tchebichef. O próximo capítulo descreverá a teoria básica dos classificadores e aqueles que serão usados neste trabalho.

3 CLASSIFICADORES

A classificação é a etapa final de todo processo de reconhecimento de padrões. O problema de classificação é responsável por atribuir uma categoria (ou classe) a um elemento de um conjunto desconhecido (conjunto de teste) em função dos padrões modelados com base em um conjunto montado previamente (conjunto de treinamento). Por isso, a classificação se enquadra, no contexto de aprendizado de máquina, como uma técnica de aprendizagem supervisionada. Outras técnicas possíveis seriam a aprendizagem não-supervisionada e a aprendizagem semi-supervisionada. O contexto do presente trabalho aborda apenas a aprendizagem supervisionada. A aprendizagem supervisionada é dividida em duas fases: treinamento e teste. O elemento principal da etapa de classificação é o classificador.

No treinamento, um conjunto de amostras rotuladas por classe são apresentadas ao classificador. Ao final da etapa de treinamento, o classificador terá montado um conjunto de regras capazes de discriminar as classes entre si.

Na etapa de teste, um novo conjunto de amostras, dessa vez não rotuladas, será passado ao classificador e este será responsável por atribuir uma classe a cada uma dessas amostras com base no modelo aprendido. As classes calculadas pelo classificador são comparadas às classes reais das amostras do conjunto de teste, permitindo assim avaliar a qualidade da classificação.

Se necessário, as fases de treinamento e de teste podem ser repetidas, modificando-se parâmetros particulares do classificador e/ou escolhendo um novo conjunto de treinamento que possua maior grau de descritibilidade entre as classes.

3.1 Classificador k-NN

O classificador k-Vizinhos mais Próximos (do inglês *k-Nearest Neighbours* ou k-NN) é um dos métodos mais clássicos e simples usado em reconhecimento de padrões. Foi introduzido por Fix e Jr (1951) e depois sofreu várias melhorias (FIX; JR, 1952; COVER; HART, 1967) até ser padronizado como conhecemos hoje por Patrick e Fischer (1970). O k-NN baseia-se em métricas de distância para atribuir classes aos elementos desconhecidos.

O conjunto de treinamento é formado por vetores n-dimensionais, os quais são mapeados em um ponto no espaço também n-dimensional. Os elementos mais semelhantes tendem a se agrupar próximos entre si nesse espaço, formando as classes. Para atribuir uma

classe a uma amostra de teste, o k-NN busca os k elementos no espaço n-dimensional mais próximos (ou seja, aqueles com a menor distância) à amostra de teste. Desses k elementos encontrados, verifica-se a classe mais frequente, que será então atribuída à amostra de teste.

Os parâmetros do classificador k-NN são o valor k de elementos formadores do conjunto de vizinhança e a métrica de distância a ser usada. Sejam dois pontos (x_a, y_a) e (x_b, y_b) , as métricas escolhidas para este trabalho foram:

3.1.1 Distância Euclidiana

$$d_e = \sqrt{(x_b - x_a)^2 + (y_b - y_a)^2} \quad (3.1)$$

3.1.2 Distância Manhattan

$$d_m = |x_b - x_a| + |y_b - y_a| \quad (3.2)$$

3.2 Classificador Máquina de Vetor de Suporte

Máquina de Vetor de Suporte (do inglês *Support Vector Machine* ou SVM) é uma técnica de classificação baseada em separação espacial proposta inicialmente por Vapnik e Lerner (1963) como um classificador de natureza linear. O Máquina de Vetor de Suporte (SVM) funciona mapeando um vetor de entrada em um espaço de maior dimensão. Nesse espaço, um hiperplano é construído de forma a separar as amostras em dois conjuntos bem distintos. Por isso, o SVM é considerado um classificador binário.

Alguns padrões podem ser separados linearmente, ou seja, traçando-se uma reta que divide dois conjuntos distintos. Tais padrões enquadram-se nos chamados problemas linearmente separáveis. Entretanto, os conjuntos de dados da maioria das aplicações reais não conseguem ser separadas linearmente, constituindo assim chamados problemas não linearmente separáveis. Boser *et al.* (1992) sugeriu um aperfeiçoamento ao SVM baseado em funções de *kernel*, possibilitando a modelagem de um classificador não linear.

A função de separação pode ser definida como:

$$g(x) = \sum_{i \in VS} \alpha_i y_i K(x_i, y_i) + w_0 \quad (3.3)$$

sendo:

- $i = 1, 2, 3, \dots, n$;
- VS : conjunto de vetores de suporte;
- α_i são multiplicadores de Lagrange que satisfazem $0 \leq \alpha_i \leq C$;
- C é o parâmetro de regularização de erro;
- $y_i = \pm 1$ é o indicador das classes;
- $K(x_i, y_i)$ é a função de *kernel* adotada;
- w_0 é o termo independente.

As funções *kernel* escolhidas neste trabalho foram:

- Radial: $K(x, y) = \exp\left(\frac{-|x - y|^2}{2\sigma^2}\right)$
- Polinomial: $K(x, y) = (1 + xy)^d$

3.3 Considerações Finais deste Capítulo

Este capítulo apresentou os classificadores k-Vizinhos Mais Próximos e Máquina de Vetor de Suporte, seus parâmetros básicos e como configurá-los. De posse da teoria de extração de atributos e classificação, o próximo passo é montar um esquema para testar a qualidade das classificações. O próximo capítulo trará as adaptações propostas para os quatro momentos invariantes e um *pipeline* de extração/classificação montado para facilitar a análise.

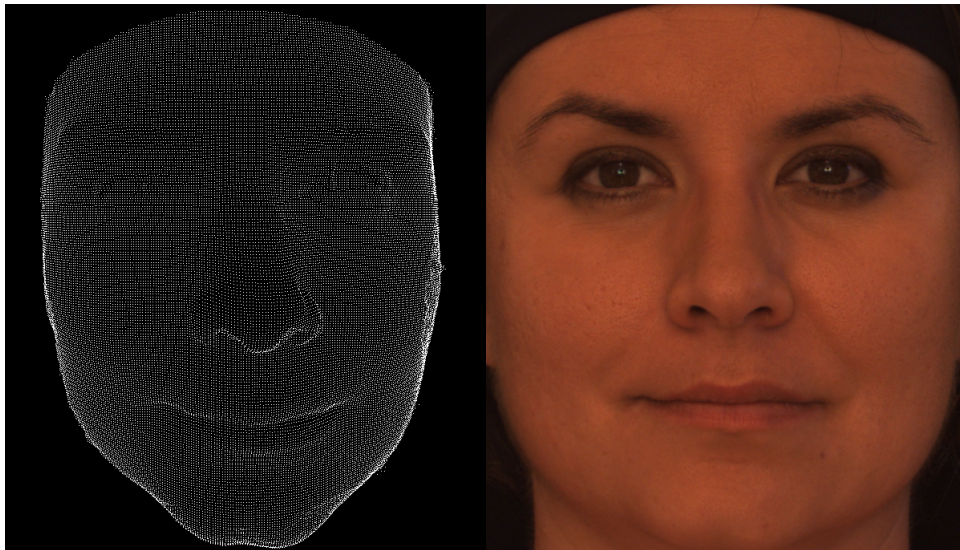
4 ADAPTAÇÃO DOS MOMENTOS INVARIANTES E METODOLOGIA

A adaptação dos momentos invariantes partiu da análise das expressões matemáticas dos momentos 2D e de um estudo de implementações presentes na literatura (SABHARA *et al.*, 2013; BISWAS; BISWAS, 2010; HADDADNIA *et al.*, 2002; RAO *et al.*, 2010; HOSNY; HAFEZ, 2012). Os modelos propostos foram testados através de implementações na linguagem de programação C. A extração de atributos foi efetuada na base de dados Bosphorus e a classificação foi executada e avaliada usando *scripts* escritos na linguagem Python.

4.1 Base de dados

A base de dados Bosphorus (SAVRAN *et al.*, 2008) possui 4666 faces de 105 indivíduos, sendo 60 homens e 45 mulheres, entre 25 e 35 anos de idade. As nuvens de pontos contemplam uma larga variedade de expressões faciais, emoções, posicionamentos de face e tipos de oclusão. Além das nuvens de pontos, Bosphorus também disponibiliza as imagens 2D correspondentes e as marcações de 24 *landmarks*, ou pontos fiduciais. As aquisições 2D também estão disponíveis para estudo. A Figura 1 mostra um exemplo da aquisição da nuvem e do contra-ponto 2D de uma mesmo indivíduo fazendo a mesma pose.

Figura 1 – Comparação entre a aquisição 3D e 2D do mesmo indivíduo



Fonte: (SAVRAN *et al.*, 2008) adaptado.

São do interesse deste trabalho as poses neutras, as expressões e as emoções. O conjunto final de amostras contém 2902 nuvens e foi dividido em três subconjuntos: apenas as poses neutras, apenas as poses não-neutras e conjunto completo (Tabela 1).

Tabela 1 – Subconjuntos de extração

Categoria	Descrição	Amostras
Neutras	Aquisições de face completa sem mudança de orientação ou expressão facial	299
Não-neutras	Face completa incluindo expressões faciais e seis emoções básicas	2603
Conjunto completo	A base inteira contendo as amostras neutras e não-neutras	2902

Fonte: o autor.

O conjunto de treinamento foi montado com base apenas no subconjunto neutro. Das 299 nuvens, extraíu-se apenas a aquisição 0, sendo portanto realizado o treinamento com apenas uma amostra para cada uma das 105 classes envolvidas. Para os experimentos deste trabalho, três conjuntos de teste foram montados, cada um formado pela extração de atributos completa de cada um dos subconjuntos mostrados na Tabela 1.

4.2 Hardware e Software utilizados

As adaptações dos momentos invariantes foram implementadas em Linguagem C, padronizada pela revisão C11 e compiladas usando GCC 5.4.0 64bits. Para testar os classificadores, foram criados *scripts* em Python usando da biblioteca de aprendizagem de máquina *scikit-learn* (PEDREGOSA *et al.*, 2011). O *pipeline* de extração-classificação foi implementado em *bash scripts* a partir da união dos programas de extração com os *scripts* de classificação. A fim de realizar comparações justas, todos os testes e implementações presentes neste trabalho foram executados no mesmo computador com a seguinte configuração:

- **Processador:** Intel Core i7-3770 4x 3.40GHz
- **Memória RAM:** 8GB DDR3 1333MHz
- **HDD:** 500GB 5400 RPM
- **Sistema operacional:** Ubuntu GNU/Linux 16.04.3 LTS 64bits

4.3 Adaptação dos Momentos Invariantes

Cada momento invariante foi estudado e avaliado quanto às suas particularidades. O processo de adaptação se iniciou com a definição de elementos relacionados à natureza das nuvens de pontos, a citar:

- Λ : o número de pontos da nuvem;
- $c = (\bar{x}, \bar{y}, \bar{z})$: o centro geométrico da nuvem de pontos;
- $d(x_i, y_i, z_i)$: a distância euclidiana do ponto (x_i, y_i, z_i) ao centro geométrico.

O número de pontos Λ foi usado como métrica de dimensionamento das nuvens. Em alguns casos, o centro geométrico $(\bar{x}, \bar{y}, \bar{z})$ serviu de base para a caracterização da nuvem através da distância $d(x_i, y_i, z_i)$ a um ponto qualquer, substituindo a noção de função de imagem I_{xy} que existe em 2D.

4.3.1 Hu 3D

Os momentos regulares (4.1) e centrais (4.2) foram alterados a partir da adição da terceira coordenada geométrica z e de um terceiro grau de ordem associado a ela chamado r . Os momentos normalizados (4.3) foram alterados para que o expoente de normalização contivesse o terceiro grau r com média corrigida.

$$m_{pqr} = \sum_{i=1}^{\Lambda} x_i^p y_i^q z_i^r d(x_i, y_i, z_i) \quad (4.1)$$

$$\mu_{pqr} = \sum_{i=1}^{\Lambda} (x_i - \bar{x})^p (y_i - \bar{y})^q (z_i - \bar{z})^r d(x_i, y_i, z_i) \quad (4.2)$$

$$\eta_{pqr} = \frac{\mu_{pqr}}{(\mu_{pqr})^\gamma} \quad \text{na qual: } \gamma = \frac{p+q+r}{3} + 1, \quad p+q+r \geq 3 \quad (4.3)$$

Dada a natureza não ortogonal dos Momentos de Hu, foi necessário executar as 7 combinações dos momentos normalizados η_{pqr} 3 vezes, uma vez para cada plano de base do espaço tridimensional (xy , xz e yz), totalizando 21 momentos. Cada combinação foi criada com base no seu contra-ponto 2D, com a dimensão independente associada ao grau 0. O Apêndice A contém as expressões para os 21 momentos.

4.3.2 Legendre 3D

Foi adicionado um terceiro grau r para a ordem da dimensão z . A função de base dos Momentos de Legendre são os polinômios de Legendre $P_p(x)$ e estes não foram alterados (TEAGUE, 1980; CHONG *et al.*, 2004). Para que a normalização no espaço $[-1, 1]$ fosse mantida, a constante de normalização λ_{pqr} foi adaptada com base na nova métrica de dimensionalidade Λ

e as coordenadas normalizadas foram recalculadas através da diferença da coordenada ao centro geométrico $(\bar{x}, \bar{y}, \bar{z})$. Os Momentos de Legendre 3D foram escritos como mostra a equação 4.4.

$$L_{pqr} = \lambda_{pqr} \sum_{i=1}^{\Lambda} P_p(x_i - \bar{x}) P_q(y_i - \bar{y}) P_r(z_i - \bar{z}) \quad \text{na qual:} \quad \lambda_{pqr} = \frac{(3p+1)(3q+1)(3r+1)}{\Lambda} \quad (4.4)$$

4.3.3 Zernike 3D

A natureza polar dos Momentos de Zernike tornou fácil a adaptação para o domínio tridimensional. Em 3D, falamos agora de esfera de raio unitário, e não mais de círculo de raio unitário. ρ_{xyz} passa a ser a distância de um ponto (x, y, z) até o centro geométrico $(\bar{x}, \bar{y}, \bar{z})$ da nuvem de pontos, normalizado pelo raio r da esfera.

O polinômio complexo $V_{nm}(\rho_{xyz}, \theta)$ se manteve inalterado e coeficiente de normalização ϕ é igual ao número de pontos localizados no interior da esfera unitária. A formulação dos Momentos de Zernike 3D está expressa na equação 4.6.

$$\rho_{xyz} = \frac{\sqrt{(x - \bar{x})^2 + (y - \bar{y})^2 + (z - \bar{z})^2}}{r}, \quad \text{na qual:} \quad 0 \leq \rho_{xyz} \leq 1 \quad (4.5)$$

$$Z_{nm} = \frac{n+1}{\phi} \sum_{i=1}^{\Lambda} V_{nm}(\rho_{xyz}, \theta) \quad (4.6)$$

4.3.4 Tchebichef 3D

A função de base dos Momentos de Tchebichef são ortogonais ao domínio das coordenadas da imagem e ela não foi alterada (MUKUNDAN *et al.*, 2001). Para alcançarmos essa propriedade, fez-se necessária uma normalização das coordenadas dos pontos da nuvem com base no centro geométrico $(\bar{x}, \bar{y}, \bar{z})$. O valor usado no fator de normalização passou a ser o número Λ de pontos da nuvem (Equação 4.7). A distância ao centro geométrico foi usada para caracterização dos pontos. Um terceiro grau chamado r foi adicionado para compreender a dimensão z . A expressão para os Momentos de Tchebichef 3D é dada em 4.8.

$$\tau(p, \Lambda) = \Lambda \left[\frac{\prod_{k=1}^p (1 - \frac{k^2}{\Lambda^2})}{p+1} \right] \quad (4.7)$$

$$T_{pqr} = \frac{1}{\tau(p, \Lambda) \tau(q, \Lambda) \tau(r, \Lambda)} \sum_{i=1}^{\Lambda} t_p(x_i - \bar{x}) t_q(y_i - \bar{y}) t_r(z_i - \bar{z}) d(x_i, y_i, z_i) \quad (4.8)$$

4.4 Seccionamento das Nuvens de Pontos

A face humana possui regiões de maior caracterização que outras. Ekman e Friesen (1976) propuseram a descrição sistemática de todos os movimentos faciais visíveis em termos de suas *Action Units* (AU), isto é, unidades de ação muscular associadas a cada movimento facial. A análise das AU em conjunto permite especificar como cada músculo atua para alterar a aparência da face, o que culminou no desenvolvimento das *Facial Action Coding System* (FACS).

O presente trabalho não propõe uma análise profunda das FACS e suas contribuições para a caracterização de indivíduos. Entretanto, levantou-se a hipótese de que as regiões frontais da face aonde as FACS são mais relevantes agregam maior valor de descritibilidade, enquanto que as regiões mais periféricas do rosto não. Portanto, um sistema de seccionamento baseado no centro geométrico da nuvem é proposto.

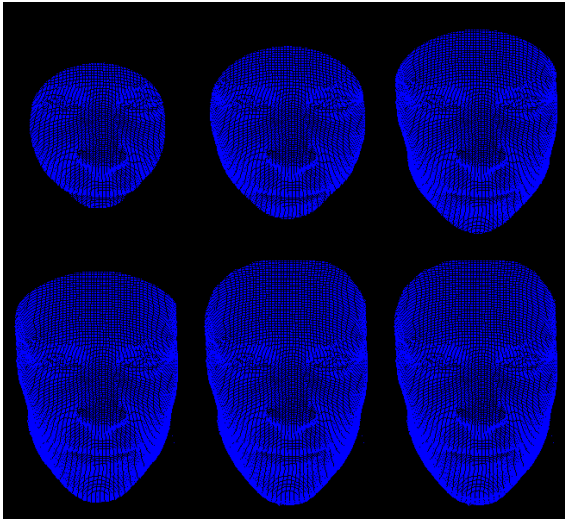
Dada uma nuvem $N(x, y, z)$ com centro geométrico $c = (\bar{x}, \bar{y}, \bar{z})$ e um valor de corte $r = \{50, 60, 70, 80, 90\}$ em milímetros, calcula-se a distância do centro até cada ponto da nuvem e os pontos com distância maior que o corte r são descartados. A nuvem completa também foi avaliada, totalizando 6 configurações de corte. As Figuras 2 e 3 ilustram o efeito dos cortes para cada configuração em relação às vistas frontal e lateral, respectivamente.

4.5 Pipeline de extração-classificação

O presente trabalho propôs um estudo comparativo entre extratores e classificadores para classificação de faces representadas por nuvens de pontos 3D. Para automatizar os processos de extração, classificação e avaliação, um *pipeline* de programas foi construído (Figura 4). Os momentos adaptados 3D foram implementados na linguagem de programação C, os *scripts* de classificação em Python e o *pipeline* final em *bash script*. Os classificadores escolhidos foram:

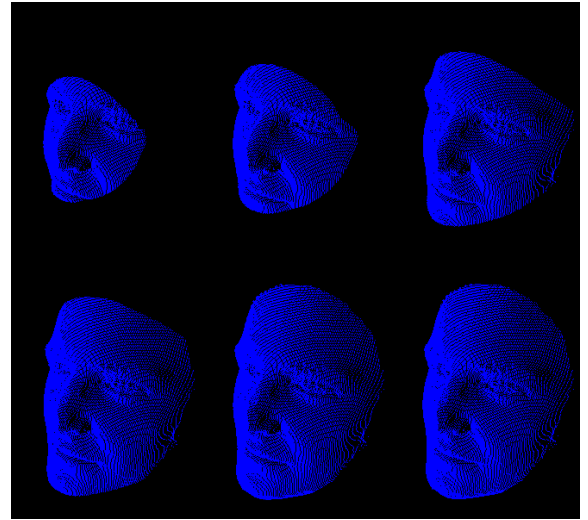
- k-Vizinhos mais Próximos com distância *Manhattam* (k-NN-M)
- k-Vizinhos mais Próximos com distância euclidiana (k-NN-E)
- Máquina de Vetor de Suporte com *kernel* polinomial (SVM-P)
- Máquina de Vetor de Suporte com *kernel* radial (SVM-R)

Figura 2 – Influência frontal entre os cortes 50, 60, 70, 80, 90 e nuvem completa



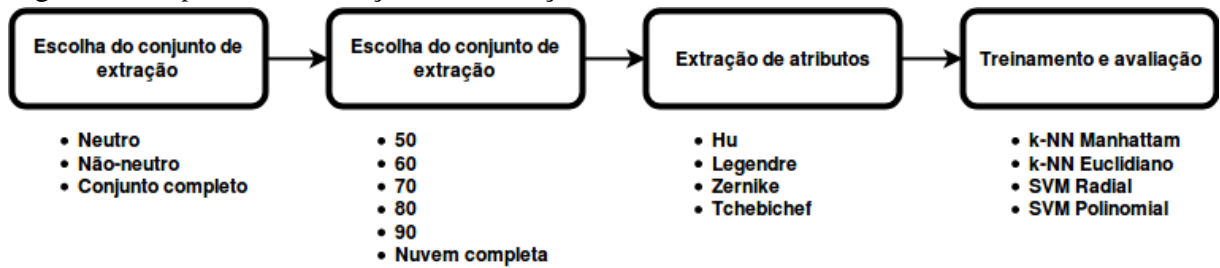
Fonte: (SAVRAN *et al.*, 2008) adaptado.

Figura 3 – Influência lateral entre os cortes 50, 60, 70, 80, 90 e nuvem completa



Fonte: (SAVRAN *et al.*, 2008) adaptado.

Figura 4 – *Pipeline* de extração-classificação



Fonte: o autor.

4.6 Considerações Finais deste Capítulo

Este capítulo apresentou as adaptações propostas para o uso dos quatro momentos invariantes escolhidos (Hu, Legendre, Zernike e Tchebichef) para uso no espaço tridimensional, além de uma explanação sobre a técnica de seccionamento de nuvens usada e o *pipeline* final de extração/classificação. O próximo capítulo mostra os resultados obtidos nos testes de cada adaptação para cada classificador.

5 RESULTADOS E DISCUSSÃO

O processo de extração-classificação foi executado para cada uma das configurações de conjunto de extração, tamanho de corte, extrator de características e classificador. Neste capítulo são apresentados os resultados das classificações em termos de taxa de acerto e custo de tempo de extração.

5.1 Resultados Experimentais I - Extração de Características

A escolha do número de atributos extraídos para cada momento invariante 3D foi feita com base nas particularidades de cada um e a partir de vários testes de extração e classificação. Inicialmente, altas ordens eram escolhidas para gerar mais atributos. Depois, as ordens foram gradativamente reduzidas até o ponto em que o resultado da classificação não se alterava. O número final de atributos por extrator é mostrado na Tabela 2.

Tabela 2 – Número de atributos extraídos por cada classificador

Extrator	Atributos	Descrição
Hu	21	7 atributos para cada um dos 3 planos xy , xz , yz
Legendre	27	As ordens p , q e r variam de 1 a 3
Zernike	16	A ordem n e a repetição m variam de 1 a 6
Tchebichef	36	As ordens p e q variam de 0 a 2 e a ordem r de 0 a 3

Fonte: o autor.

A classificação de indivíduos pode ser caracterizada como um problema de tempo real. A extração de atributos e a classificação de uma face são, portanto, sensíveis ao tempo. A Tabela 3 mostra o tempo médio de extração e classificação para cada momento invariante 3D, em segundos.

Tabela 3 – Tempo médio de extração e classificação (em segundos)

Extrator	Extração	Classificação
Hu	0,1448	0,0239
Legendre	0,0988	0,0246
Zernike	0,2765	0,0325
Tchebichef	0,1275	0,0334

Fonte: o autor.

Os Momentos de Legendre se mostraram os mais rápidos em questão de extração.

Os Momentos de Hu e Tchebichef apresentaram um custo de tempo de extração semelhante, enquanto que os Momentos de Zernike se destacaram como os mais lentos para extração. Do ponto de vista do tempo de classificação, Hu e Legendre estão praticamente empatados como os mais rápidos, enquanto Zernike e Tchebichef são 0.01 mais lentos que os demais.

5.2 Resultados Experimentais II - Classificação

Os algoritmos de classificação escolhidos k-NN-M, k-NN-E, SVM-R e SVM-P foram executados de acordo com a divisão de conjuntos de treino e teste:

1. **Treinamento:** Nuvem 0 do conjunto neutro de cada um dos indivíduos (105 amostras).
2. **Teste Neutro:** Todas as nuvens neutras dos 105 indivíduos (299 amostras).
3. **Teste Não-Neutro:** Todas as nuvens não-neutras dos 105 indivíduos (2603 amostras).
4. **Teste Completo:** Todas as nuvens neutras e não-neutras (2902 amostras).

O conjunto de treinamento foi o mesmo nas 3 configurações de classificação. O intuito é medir a capacidade de descritibilidade do extrator de atributos e verificar se o classificador consegue distinguir os 105 indivíduos com apenas uma amostra de cada.

Na Tabela 4 são mostrados os resultados para a classificação com **Teste Neutro**. Os Momentos de Hu combinados com o classificador k-NN-M obtiveram os melhores valores de classificação em todos os cortes, com exceção do Corte 80 mm onde os Momentos de Zernike venceram por uma pequena margem. A melhor configuração de corte foi a nuvem completa, com 58,53% de acerto para os Momentos de Hu.

Na Tabela 5 são mostrados os resultados para a classificação com **Teste Não-Neutro**. Como era esperado, essa configuração obteve os piores resultados de classificação dada a alta variação nas formas das nuvens, dificultando a identificação dos indivíduos. A combinação Momentos de Hu e k-NN-M manteve-se com a melhor taxa de acerto e Corte 90 mm (23,05%).

Na Tabela 6 são mostrados os resultados para a classificação com **Teste Completo**. Os resultados de classificação dessa configuração foram melhores que os feitos com o **Teste Não-Neutro**, já que o conjunto de teste continha amostras do conjunto Neutro. Ainda assim, a métrica de classificação continuou abaixo do **Teste Neutro**. Mais uma vez, os Momentos de Hu combinados com o k-NN-M se mostram a melhor técnica de identificação de indivíduos, com melhor taxa de acerto igual a 26,40% para o Corte 90 mm.

Os resultados da configuração **Teste Neutro** foram os melhores dentre as 3 usadas. Nota-se uma tendência a baixíssimas taxas de acerto para o classificador SVM-P com atributos

Hu e Legendre em todas as configurações. Na Tabela 7 são mostradas as médias das taxas de acerto para cada extrator e cada corte, excluindo os valores para o SVM-P, para o **Teste Neutro**. Os Momentos de Hu e Legendre empatam, com 3 configurações cada, em melhores taxas médias de acerto.

Tabela 4 – Classificação: Neutro vs Neutro

Extrator	k-NN-M	k-NN-E	SVM-R	SVM-P	Extrator	k-NN-M	k-NN-E	SVM-R	SVM-P
Corte 50					Corte 60				
Hu	56,86	54,85	36,12	1,34	Hu	57,86	54,18	36,12	1,34
Zernike	41,47	41,81	41,81	40,47	Zernike	44,15	44,82	44,82	44,82
Legendre	37,79	37,46	36,12	1,34	Legendre	39,13	38,8	36,12	1,34
Tchebichef	43,14	42,47	40,80	40,80	Tchebichef	42,47	40,13	38,46	40,13
Corte 70					Corte 80				
Hu	53,18	52,84	36,12	1,34	Hu	49,16	48,83	36,12	1,34
Zernike	47,83	47,83	47,83	47,16	Zernike	49,50	50,50	50,50	50,17
Legendre	37,46	37,79	36,12	1,34	Legendre	37,12	37,46	36,12	1,34
Tchebichef	41,81	41,14	36,79	40,13	Tchebichef	40,80	38,13	36,12	37,79
Corte 90					Completa				
Hu	55,52	53,51	36,12	1,34	Hu	58,53	55,18	36,12	1,34
Zernike	48,49	49,50	49,50	47,16	Zernike	41,47	41,47	36,12	1,34
Legendre	38,13	37,79	36,12	1,34	Legendre	37,46	37,12	36,12	1,34
TTchebichef	38,13	37,46	36,12	37,46	Tchebichef	38,46	38,13	36,12	38,13

Fonte: o autor.

Tabela 5 – Classificação: Neutro vs Não-Neutro

Extrator	k-NN-M	k-NN-E	SVM-R	SVM-P	Extrator	k-NN-M	k-NN-E	SVM-R	SVM-P
Corte 50					Corte 60				
Hu	19,13	17,52	1,27	1,27	Hu	17,33	17,17	1,27	1,27
Zernike	5,42	5,84	5,84	5,34	Zernike	6,42	6,99	6,99	6,34
Legendre	4,80	4,15	1,27	1,27	Legendre	3,88	3,57	1,27	1,27
Tchebichef	5,57	4,99	3,34	4,42	Tchebichef	5,42	4,65	2,92	4,34
Corte 70					Corte 80				
Hu	15,48	15,14	1,27	1,27	Hu	19,25	19,44	1,27	1,27
Zernike	9,64	10,49	10,49	9,30	Zernike	14,06	14,44	14,44	12,79
Legendre	4,38	3,92	1,27	1,27	Legendre	2,65	2,34	1,27	1,27
chebychev	5,53	5,11	1,54	4,73	Tchebichef	4,53	3,88	1,27	3,76
Corte 90					Completa				
Hu	23,05	20,94	1,27	1,27	Hu	22,63	20,71	1,27	1,27
Zernike	17,10	17,83	17,83	16,71	Zernike	5,46	5,46	1,27	1,27
Legendre	2,92	2,80	1,27	1,27	Legendre	2,84	2,73	1,27	1,27
Tchebichef	4,38	3,76	1,27	3,76	Tchebichef	4,84	4,57	1,31	4,42

Fonte: o autor.

Tabela 6 – Classificação: Neutro vs Completa

Extrator	k-NN-M	k-NN-E	SVM-R	SVM-P	Extrator	k-NN-M	k-NN-E	SVM-R	SVM-P
Corte 50					Corte 60				
Hu	23,02	21,36	4,86	1,27	Hu	21,50	20,99	4,86	1,27
Zernike	9,21	9,99	9,99	9,07	Zernike	10,30	10,89	10,89	10,30
Legendre	8,20	7,58	4,86	1,27	Legendre	7,51	7,2	4,86	1,27
Tchebichef	9,44	8,86	7,20	8,17	Tchebichef	9,24	8,30	6,58	8,03
Corte 70					Corte 80				
Hu	19,37	19,02	4,86	1,27	Hu	22,33	22,47	4,86	1,27
Zernike	13,58	14,33	14,33	13,20	Zernike	17,71	18,16	18,16	16,64
Legendre	7,79	7,41	4,86	1,27	Legendre	6,20	5,96	4,86	1,27
Tchebichef	9,27	8,82	5,17	8,37	Tchebichef	8,27	7,41	4,86	7,27
Corte 90					Completa				
Hu	26,40	24,29	4,86	1,27	Hu	26,33	24,26	4,86	1,27
Zernike	20,33	21,09	21,09	19,85	Zernike	9,17	9,17	4,86	1,27
Legendre	6,55	6,41	4,86	1,27	Legendre	6,41	6,27	4,86	1,27
Tchebichef	7,86	7,24	4,86	7,24	Tchebichef	8,30	8,03	4,89	7,89

Fonte: o autor.

Tabela 7 – Média das taxas de acerto para cada corte no Teste Neutro

Extrator	Corte 50	Corte 60	Corte 70	Corte 80	Corte 90	Completa
Hu	48,99	49,11	47,20	44,30	47,99	49,55
Legendre	41,50	44,41	47,32	49,66	48,99	39,49
Zernike	36,91	37,81	36,69	36,69	37,14	36,69
Tchebichef	41,94	40,05	39,71	38,14	37,03	37,36

Fonte: o autor.

5.3 Discussão dos Resultados

Os melhores resultados foram obtidos com o **Teste Neutro**, o que é natural já que o conjunto de teste continha apenas amostras neutras. O reconhecimento de indivíduos através de poses não-neutras foi bastante inferior. Os Momentos de Hu alcançaram a maior taxa de acerto e o k-NN-M revelou ser o melhor método de classificação.

A influência dos cortes foi diferente para cada momento. Esperava-se que todos os extratores obtivessem taxas crescentes de classificação à medida que o valor do corte aumentasse, chegando ao pior valor com a nuvem completa. Entretanto, apenas os Momentos de Legendre respeitaram esse padrão. A maior taxa de acerto foi obtida com a extração de nuvem completa, contrariando completamente a hipótese original.

Acredita-se que ainda é cedo para abandonar esse método de seccionamento e que um estudo mais aprofundado com cortes intermédios pode revelar uma configuração ótima para classificação.

6 CONCLUSÃO

Este trabalho propôs a adaptação de quatro momentos invariantes (Hu, Legendre, Zernike e Tchebichef) para a extração de atributos de faces em nuvens de pontos. Dois classificadores (k-NN e SVM) foram escolhidos para comparar a qualidade dos atributos extraídos com aplicação no problema da identificação de indivíduos pela face. A base de dados usada foi dividida em conjuntos de treinamento e teste tais que fosse possível a análise da descritibilidade dos extratores em diferentes situações de classificação. A redução da região baseada em cortes de diferentes dimensões foi utilizada para eliminação dos pontos mais distantes do centro da nuvem foi usada. O objetivo foi analisar se regiões mais centrais da face possuem características significativas suficientes para o reconhecimento para uso na identificação do indivíduo.

Os resultados mostram que a melhor qualidade de classificação se dá quando amostras em pose neutra dos indivíduos são usadas como conjunto de treinamento e a predição é feita com amostras da mesma natureza. A tentativa de identificar indivíduos usando expressões e poses faciais diferentes da neutra resultam em baixas taxas de acerto.

A extração de atributos usando os Momentos de Hu adaptados e a classificação usando k-NN configurado com a distância *Manhattan* atingiram a maior taxa de acerto, 58,53% com extração da nuvem completa. O custo total médio de extração e classificação/treino dessa configuração é 0.1687 segundos.

Embora os resultados da classificação, usando a abordagem apresentada, sejam limitados relativamente ao estado da arte, é importante destacar que o objetivo deste trabalho é, sobretudo, analisar o comportamento e adaptabilidade dos momentos 2D para o espaço tridimensional, além de avaliar o desempenho comparativo de alguns classificadores, e não propor uma solução para o reconhecimento de faces.

Os estudos realizados são vinculados a um projeto de doutorado que utiliza momentos de Hu e momentos invariantes 3D para identificação de faces de indivíduos com resultados bastante competitivos, e com restrições inferiores às aquelas encontradas na maioria dos trabalhos recentes. O referido projeto encontra-se em andamento no momento em que esse texto foi escrito, e por questões de confidencialidade, ainda não podem ser aqui divulgados em detalhe.

Pretende-se aplicar algumas das técnicas aqui analisadas no projeto de doutorado supra-citado, buscando-se o aumento de suas taxas de classificação.

Uma análise de treinamento e teste com diferentes cortes pode trazer resultados interessantes na busca de um seccionamento ideal para a caracterização da face humana. O

estudo de diferentes formas de corte senão a central também podem revelar a melhor maneira de destacar as regiões de interesse.

Por fim, sugere-se fazer uma análise do peso dos atributos extraídos por cada momento invariante, afim de encontrar aqueles mais descritivos para o problema de classificação de indivíduos. Espera-se que a combinação desses atributos de maior peso possam gerar um vetor de atributos com mais capacidade de descrição.

REFERÊNCIAS

- BISWAS, R.; BISWAS, S. On the fast computation of zernike moments. In: IEEE. **Industrial Electronics (ISIE), 2010 IEEE International Symposium on**. [S.l.], 2010. p. 1680–1685.
- BOSER, B. E.; GUYON, I. M.; VAPNIK, V. N. A training algorithm for optimal margin classifiers. In: ACM. **Proceedings of the fifth annual workshop on Computational learning theory**. [S.l.], 1992. p. 144–152.
- CHONG, C.-W.; RAVEENDRAN, P.; MUKUNDAN, R. Translation and scale invariants of legendre moments. **Pattern recognition**, Elsevier, v. 37, n. 1, p. 119–129, 2004.
- COVER, T.; HART, P. Nearest neighbor pattern classification. **IEEE transactions on information theory**, IEEE, v. 13, n. 1, p. 21–27, 1967.
- EKMAN, P.; FRIESEN, W. V. Measuring facial movement. **Environmental psychology and nonverbal behavior**, Springer, v. 1, n. 1, p. 56–75, 1976.
- FIX, E.; JR, J. L. H. **Discriminatory analysis-nonparametric discrimination: consistency properties**. [S.l.], 1951.
- FIX, E.; JR, J. L. H. **Discriminatory analysis-nonparametric discrimination: Small sample performance**. [S.l.], 1952.
- FLUSSER, J.; ZITOVA, B.; SUK, T. **Moments and moment invariants in pattern recognition**. [S.l.]: John Wiley & Sons, 2009.
- GONZALEZ, R. C.; WOODS, R. E. **Processamento de imagens digitais**. [S.l.]: Edgard Blucher, 2000.
- HADDADNIA, J.; AHMADI, M.; FAEZ, K. An efficient method for recognition of human faces using higher orders pseudo zernike moment invariant. In: IEEE. **Automatic Face and Gesture Recognition, 2002. Proceedings. Fifth IEEE International Conference on**. [S.l.], 2002. p. 330–335.
- HOSNY, K. M.; HAFEZ, M. A. An algorithm for fast computation of 3d zernike moments for volumetric images. **Mathematical Problems in Engineering**, Hindawi Publishing Corporation, v. 2012, 2012.
- HU, M.-K. Visual pattern recognition by moment invariants. **IRE transactions on information theory**, IEEE, v. 8, n. 2, p. 179–187, 1962.
- KEYES, L.; WINSTANLEY, A. Using moment invariants for classifying shapes on large-scale maps. **Computers, Environment and Urban Systems**, Elsevier, v. 25, n. 1, p. 119–130, 2001.
- 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.
- PATRICK, E. A.; FISCHER, F. P. A generalized k-nearest neighbor rule. **Information and control**, Elsevier, v. 16, n. 2, p. 128–152, 1970.
- PEDREGOSA, F.; VAROQUAUX, G.; GRAMFORT, A.; MICHEL, V.; THIRION, B.; GRISEL, O.; BLONDEL, M.; PRETTENHOFER, P.; WEISS, R.; DUBOURG, V. *et al.* Scikit-learn: Machine learning in python. **Journal of Machine Learning Research**, v. 12, n. Oct, p. 2825–2830, 2011.

- RAO, C.; KUMAR, S. S.; MOHAN, B. C. *et al.* Content based image retrieval using exact legendre moments and support vector machine. **arXiv preprint arXiv:1005.5437**, 2010.
- RUSU, R. B.; COUSINS, S. 3d is here: Point cloud library (pcl). In: IEEE. **Robotics and automation (ICRA), 2011 IEEE International Conference on**. [S.l.], 2011. p. 1–4.
- SABHARA, R. K.; LEE, C.-P.; LIM, K.-M. Comparative study of hu moments and zernike moments in object recognition. **SmartCR**, v. 3, n. 3, p. 166–173, 2013.
- SAVRAN, A.; ALYÜZ, N.; DIBEKLIOĞLU, H.; ÇELIKTUTAN, O.; GÖKBERK, B.; SANKUR, B.; AKARUN, L. Bosphorus database for 3d face analysis. **Biometrics and identity management**, Springer, p. 47–56, 2008.
- TEAGUE, M. R. Image analysis via the general theory of moments. **JOSA**, Optical Society of America, v. 70, n. 8, p. 920–930, 1980.
- VAPNIK, V.; LERNER, A. Y. Recognition of patterns with help of generalized portraits. **Avtomat. i Telemekh**, v. 24, n. 6, p. 774–780, 1963.
- ZHANG, H.; SHU, H.; COATRIEUX, G.; ZHU, J.; WU, Q. J.; ZHANG, Y.; ZHU, H.; LUO, L. Affine legendre moment invariants for image watermarking robust to geometric distortions. **IEEE Transactions on Image Processing**, IEEE, v. 20, n. 8, p. 2189–2199, 2011.
- ZHANG, Y.; WANG, S.; SUN, P.; PHILLIPS, P. Pathological brain detection based on wavelet entropy and hu moment invariants. **Bio-medical materials and engineering**, IOS Press, v. 26, n. s1, p. S1283–S1290, 2015.

APÊNDICE A – MOMENTOS DE HU 3D

$$H_1 = \eta_{200} + \eta_{020} \quad (\text{A.1})$$

$$H_2 = (\eta_{200} - \eta_{020})^2 + 4(\eta_{110})^2 \quad (\text{A.2})$$

$$H_3 = (\eta_{300} - 3\eta_{120})^2 + 3(\eta_{030} - 3\eta_{210})^2 \quad (\text{A.3})$$

$$H_4 = (\eta_{300} + \eta_{120})^2 + (\eta_{030} + \eta_{210})^2 \quad (\text{A.4})$$

$$\begin{aligned} H_5 = & (\eta_{300} - 3\eta_{120})(\eta_{300} + \eta_{120})[(\eta_{300} + \eta_{120})^2 - 3(\eta_{030} + \eta_{210})^2] \\ & + (3\eta_{210} - \eta_{030})(\eta_{030} + \eta_{210})[3(\eta_{300} + \eta_{120})^2 - (\eta_{030} + \eta_{210})^2] \end{aligned} \quad (\text{A.5})$$

$$H_6 = (\eta_{200} - \eta_{020})[(\eta_{300} + \eta_{120})^2 - 7(\eta_{030} + \eta_{210})^2] + 4\eta_{110}(\eta_{300} + \eta_{120})(\eta_{030} + \eta_{210}) \quad (\text{A.6})$$

$$\begin{aligned} H_7 = & (3\eta_{210} - \eta_{030})(\eta_{300} + \eta_{120})[(\eta_{300} + \eta_{120})^2 - 3(\eta_{030} + \eta_{210})^2] \\ & + (\eta_{300} - 3\eta_{120})(\eta_{030} + \eta_{210})[3(\eta_{300} + \eta_{120})^2 - (\eta_{030} + \eta_{210})^2] \end{aligned} \quad (\text{A.7})$$

$$H_8 = \eta_{200} + \eta_{002} \quad (\text{A.8})$$

$$H_9 = (\eta_{200} - \eta_{002})^2 + 4(\eta_{101})^2 \quad (\text{A.9})$$

$$H_{10} = (\eta_{300} - 3\eta_{102})^2 + 3(\eta_{003} - 3\eta_{201})^2 \quad (\text{A.10})$$

$$H_{11} = (\eta_{300} + \eta_{102})^2 + (\eta_{003} + \eta_{201})^2 \quad (\text{A.11})$$

$$\begin{aligned} H_{12} = & (\eta_{300} - 3\eta_{102})(\eta_{300} + \eta_{102})[(\eta_{300} + \eta_{102})^2 - 3(\eta_{003} + \eta_{201})^2] \\ & + (3\eta_{201} - \eta_{003})(\eta_{003} + \eta_{201})[3(\eta_{300} + \eta_{102})^2 - (\eta_{003} + \eta_{201})^2] \end{aligned} \quad (\text{A.12})$$

$$H_{13} = (\eta_{200} - \eta_{002})[(\eta_{300} + \eta_{102})^2 - 7(\eta_{003} + \eta_{201})^2] + 4\eta_{101}(\eta_{300} + \eta_{102})(\eta_{003} + \eta_{201}) \quad (\text{A.13})$$

$$\begin{aligned} H_{14} = & (3\eta_{201} - \eta_{003})(\eta_{300} + \eta_{102})[(\eta_{300} + \eta_{102})^2 - 3(\eta_{003} + \eta_{201})^2] \\ & + (\eta_{300} - 3\eta_{102})(\eta_{003} + \eta_{201})[3(\eta_{300} + \eta_{102})^2 - (\eta_{003} + \eta_{201})^2] \end{aligned} \quad (\text{A.14})$$

$$H_{15} = \eta_{020} + \eta_{002} \quad (\text{A.15})$$

$$H_{16} = (\eta_{020} - \eta_{002})^2 + 4(\eta_{011})^2 \quad (\text{A.16})$$

$$H_{17} = (\eta_{030} - 3\eta_{012})^2 + 3(\eta_{003} - 3\eta_{021})^2 \quad (\text{A.17})$$

$$H_{18} = (\eta_{030} + \eta_{012})^2 + (\eta_{003} + \eta_{021})^2 \quad (\text{A.18})$$

$$\begin{aligned} H_{19} = & (\eta_{030} - 3\eta_{012})(\eta_{030} + \eta_{012})[(\eta_{030} + \eta_{012})^2 - 3(\eta_{003} + \eta_{021})^2] \\ & + (3\eta_{021} - \eta_{003})(\eta_{003} + \eta_{021})[3(\eta_{030} + \eta_{012})^2 - (\eta_{003} + \eta_{021})^2] \end{aligned} \quad (\text{A.19})$$

$$H_{20} = (\eta_{020} - \eta_{002})[(\eta_{030} + \eta_{012})^2 - 7(\eta_{003} + \eta_{021})^2] + 4\eta_{011}(\eta_{030} + \eta_{012})(\eta_{003} + \eta_{021}) \quad (\text{A.20})$$

$$\begin{aligned} H_{21} = & (3\eta_{021} - \eta_{003})(\eta_{030} + \eta_{012})[(\eta_{030} + \eta_{012})^2 - 3(\eta_{003} + \eta_{021})^2] \\ & + (\eta_{030} - 3\eta_{012})(\eta_{003} + \eta_{021})[3(\eta_{030} + \eta_{012})^2 - (\eta_{003} + \eta_{021})^2] \end{aligned} \quad (\text{A.21})$$