

UNIVERSIDADE FEDERAL DE SANTA MARIA
CENTRO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

**CLASSIFICAÇÃO DE PADRÕES ATRAVÉS DE
WAVELETS E MÉTODOS BAYESIANOS**

DISSERTAÇÃO DE MESTRADO

Douglas Camargo Foster

Santa Maria, RS, Brasil

2011

CLASSIFICAÇÃO DE PADRÕES ATRAVÉS DE WAVELETS E MÉTODOS BAYESIANOS

por

Douglas Camargo Foster

Dissertação apresentada ao Programa de Pós-Graduação em Informática,
Área de concentração em Processamento de Sinais, da Universidade
Federal de Santa Maria (UFSM, RS), como requisito parcial para
obtenção do grau de
Mestre em Computação.

Orientadora: Prof. Dra. Alice de Jesus Kozakevicius

Co-orientadora: Prof. Dra. Patrícia Klarmann Ziegelmann

Santa Maria, RS, Brasil

2011

**Universidade Federal de Santa Maria
Centro de Tecnologia
Programa de Pós-Graduação em Informática**

A Comissão Examinadora, abaixo assinada,
aprova a Dissertação de Mestrado

**CLASSIFICAÇÃO DE PADRÕES ATRAVÉS DE WAVELETS
E MÉTODOS BAYESIANOS**

elaborada por
Douglas Camargo Foster

como requisito parcial para obtenção do grau de
Mestre em Computação

COMISSÃO EXAMINADORA:

Alice de Jesus Kozakevicius, Dra.
(Presidenta/Orientadora)

Patrícia Klarmann Ziegelmann, Dra.
(Co-orientadora)

Raul Ceretta Nunes, Dr. (UFSM)

Giovani Baratto, Dr. (UFSM)

Santa Maria, 24 de Março de 2011

AGRADECIMENTOS

Agradeço a minha família, em especial aos meus pais Paulo e Margarete, por todo apoio e confiança durante toda minha vida acadêmica, e durante curso de mestrado teve fundamental participação para que a motivação sempre me acompanhasse. Também quero agradecer a minha noiva Jacséli pela companhia, apoio, compreensão e motivação durante o desenvolvimento deste trabalho.

Agradecimento ao Programa de Pós-Graduação em Informática (PPGI) pela oportunidade de realizar este trabalho. Em especial à ex-secretária Marinelma Aimi de Carvalho e o atual secretário Josmar Nuernberg pela ajuda com a documentação e procedimentos burocráticos.

Agradecimento especial à orientadora Profa. Dra. Alice de Jesus Kozakevicius pela confiança depositada, pela orientação e, nos momento mais difíceis e críticos deste trabalho teve paciência, compreensão e ajudou a fazer as escolhas certas.

Agradecimento especial à co-orientadora Profa. Dra. Patrícia Klarmann Ziegelmann pela orientação e sugestões em todas as fases do trabalho.

Agradecimento aos colegas de GMICRO, em especial aos colegas Jaziel Souza Lobo, Renato Preigschadt e Mateus Beck Fonseca pelo companheirismo, amizade, e todas as discussões e conversas que colaboraram para a realização deste trabalho.

Agradecimento ao Prof. Dr. João Baptista Martins, que há 4 anos me convidou para trabalhar em Santa Maria. A partir deste convite, também tive a oportunidade de cursar o mestrado.

Agradecimento à todos os colegas da SMDH pela amizade e pelo apoio para minha qualificação profissional.

RESUMO

Dissertação de Mestrado
Programa de Pós-Graduação em Informática
Universidade Federal de Santa Maria, RS, Brasil

CLASSIFICAÇÃO DE PADRÕES ATRAVÉS DE WAVELETS E MÉTODOS BAYESIANOS

AUTOR: DOUGLAS CAMARGO FOSTER

ORIENTADORA: ALICE DE JESUS KOZAKEVICIUS

CO-ORIENTADORA: PATRÍCIA KLARMANN ZIEGELMANN

Local da Defesa e Data: Santa Maria, 24 de Março de 2011.

O interesse na área de classificação de padrões tem aumentado ultimamente devido a grande demanda computacional para a manipulação de grandes conjuntos de dados, e também ao aumento de aplicações desafiadoras. Métodos de classificação estatística, como os métodos baseados na teoria de decisão das regras de Bayes, aplicam a abordagem de estimação de parâmetros a partir de um conjunto de dados de treinamento que definem as diferentes classes dentro de uma base de dados. Neste trabalho é avaliada a contribuição da transformada *wavelet* discreta (TWD) na extração de variáveis para a realização de classificação. A partir dos coeficientes de escala de diferentes níveis de TWD serão compostos novos conjuntos de dados de treinamento para serem aplicados em métodos de classificação Bayesiano. Para as transformadas uni- e bidimensionais são consideradas funções da família de *wavelets* ortonormais de Daubechies (Haar, Daubechies Db2 e Db8). Também é proposta uma metodologia híbrida para o tratamento de dados bidimensionais que compreende em aplicar tanto a transformada uni- quanto a bidimensional em estágios consecutivos da análise dos dados. Para a avaliação da metodologia de classificação associada à transformada unidimensional são utilizados dois conjuntos de dados unidimensionais diferentes: o primeiro é composto por dados gerados sinteticamente, e o outro é composto por informações de tráfego de dados em rede (base DARPA 1999). Para avaliar as metodologias bidimensional e híbrida são utilizados dados bidimensionais (imagens originadas de fotografias digitais, com e sem uso de flash). Uma vantagem da aplicação da metodologia híbrida é a manutenção da regularidade de classificação e o aumento nas classificações corretas em alguns casos.

Palavras-chave: classificação de padrões; métodos bayesianos; wavelets; imagens.

ABSTRACT

Master Dissertation

Graduation Program in Computer Science
Federal University of Santa Maria, RS, Brazil

PATTERN CLASSIFICATION USING WAVELETS AND BAYESIAN METHODS

AUTHOR: DOUGLAS CAMARGO FOSTER

ADVISOR: ALICE DE JESUS KOZAKEVICIUS

CO-ADVISOR: PATRÍCIA KLARMANN ZIEGELMANN

Place and Date: Santa Maria, March 24th, 2011.

The interest in the pattern classification field has increased due to challenging applications and also due to computational demands, specially when big datasets have to be analyzed. Statistical classification methods, as the based Bayes rules decision theory, apply the parameter estimation from a training dataset for recognizing different classes inside the dataset. In this work it is investigated the contribution of using the discrete wavelet transformation (DWT) for feature extraction during the classification process. From the scale coefficients of different decomposition levels, new training datasets, which are used in Bayesian classifier, are formed. For the one and two dimensional transforms the Daubechies wavelet family is considered. Three specifically wavelet functions are analyzed (Haar, Daubechies Db2 and Db8). Also, a hybrid methodology is proposed, in which 2D and 1D wavelet transformations are applied in consecutive stages of data analysis. For the evaluation of the one dimensional transform methodology, two different unidimensional datasets are used: one is composed by synthetic data, and the other is composed by network traffic data (DARPA1999 dataset). For the evaluation of 2D and hybrid methodologies two-dimensional data are considered. The two-dimensional data are images with different digital pictures with and without using flash light. One advantage of applying the hybrid methodology is the maintenance of the classification regularity and the increase of correct classification in some cases.

Keywords: pattern classification; bayesian methods; wavelets; images.

LISTA DE FIGURAS

Figura 1	Fluxo de decomposição via transformada <i>wavelet</i>	p. 30
Figura 2	Transformada <i>wavelet</i> unidimensional: filtragem de sinal unidimensional x	p. 30
Figura 3	Transformada <i>wavelet</i> inversa: reconstrução do sinal unidimensional x	p. 31
Figura 4	Ordem de execução da transformada <i>wavelet</i> : primeiramente, linhas, e então as colunas.	p. 31
Figura 5	Resultado da transformada <i>wavelet</i> em 2 dimensões.	p. 32
Figura 6	Decomposição de imagem através de <i>wavelet Db8</i> periódica.	p. 33
Figura 7	Etapas de um sistema de classificação. Adaptado de (Pei et al. 1995), (Pittner e Kamarthi 1997)	p. 43
Figura 8	Abordagem unidimensional para dados unidimensionais.	p. 44
Figura 9	Abordagem 1D para dados bidimensionais.	p. 45
Figura 10	Abordagem 2D para dados bidimensionais.	p. 45
Figura 11	Abordagem híbrida para dados bidimensionais.	p. 46
Figura 12	Amostras de classes simuladas.	p. 51
Figura 13	TWD de amostra da classe 1.	p. 52
Figura 14	TWD de amostra da classe 2.	p. 52
Figura 15	TWD de amostra da classe 3.	p. 52
Figura 16	Amostra de cada anomalia, das diferentes classes.	p. 55
Figura 17	TWD de amostra da classe de ataques curtos, chamado <i>pod</i>	p. 55
Figura 18	TWD de amostra da classe de ataques curtos, chamado <i>satan</i>	p. 55
Figura 19	TWD de amostra da classe de ataques longos, chamado <i>mailbomb</i>	p. 56
Figura 20	TWD de amostra da classe de ataques longos, chamado <i>neptune</i>	p. 56
Figura 21	Amostras que compõem o conjunto de imagens de cartas de baralho.	p. 58
Figura 22	Obtenção de imagem do naipe a partir das imagens originais: recorte da área do naipe, gerando nova imagem com dimensão de 512 x 512 pixels.	p. 59
Figura 23	Naipe de copas.	p. 59
Figura 24	Naipe de ouros.	p. 59

Figura 25	Naipes de paus.	p. 59
Figura 26	Naipes de espadas.	p. 59
Figura 27	Representação das amostras de imagens com <i>flash</i> para as 4 classes de naipes de cartas: (a) classe Copas, (b) classe Ouros, (c) classe Espadas, (d) classes Paus.	p. 61
Figura 28	Representação das amostras de imagens sem <i>flash</i> para as 4 classes de naipes de cartas: (a) classe Copas, (b) classe Ouros, (c) classe Espadas, (d) classes Paus.	p. 61
Figura 29	Representação das 4 classes de naipes de cartas no 8 ^o nível de TWD, a partir de imagens com <i>flash</i>	p. 62
Figura 30	Representação das 4 classes de naipes de cartas no 8 ^o nível de TWD, a partir de imagens sem <i>flash</i>	p. 62
Figura 31	Representação das 4 classes de naipes de cartas no 5 ^o nível de TWD 2D, a partir de imagens com <i>flash</i> . Então, as imagens têm as linhas concatenadas em um vetor.	p. 70
Figura 32	Representação das 4 classes de naipes de cartas no 5 ^o nível de TWD 2D, a partir de imagens sem <i>flash</i> . Então, as imagens têm as linhas concatenadas em um vetor.	p. 71
Figura 33	Representação das 4 classes de naipes de cartas no 4 ^o nível de TWD 2D, a partir de imagens com <i>flash</i> . Então, as imagens têm as linhas concatenadas em um vetor.	p. 78
Figura 34	Representação das 4 classes de naipes de cartas no 4 ^o nível de TWD 2D, a partir de imagens sem <i>flash</i> . Então, as imagens têm as linhas concatenadas em um vetor.	p. 79
Figura 35	Representação das amostras para a classe Copas, a partir das imagens sem <i>flash</i> . Os 3 métodos de extração são apresentados.	p. 86
Figura 36	Representação das amostras para a classe Ouros, a partir das imagens sem <i>flash</i> . Os 3 métodos de extração são apresentados.	p. 87
Figura 37	Representação das amostras para a classe Espadas, a partir das imagens sem <i>flash</i> . Os 3 métodos de extração são apresentados.	p. 87
Figura 38	Representação das amostras para a classe Paus, a partir das imagens sem <i>flash</i> . Os 3 métodos de extração são apresentados.	p. 88
Figura 39	Erros de classificação para cenários com $n_r = 26$ amostras de treinamento por classe. Método unidimensional usado para extração de atributos.	p. 89
Figura 40	Erros de classificação para cenários com $n_r = 26$ amostras de treinamento por classe. Método bidimensional usado para extração de atributos.	p. 89
Figura 41	Erros de classificação para cenários com $n_r = 26$ amostras de treinamento por classe. Método híbrido aplicado para extração de atributos.	p. 90

Figura 42	Erros de classificação para cenários com $n_r = 130$ amostras de treinamento por classe. Método unidimensional aplicado no processo de extração de atributos.	p. 90
Figura 43	Erros de classificação para cenários com $n_r = 130$ amostras de treinamento por classe. Método bidimensional usado para extração de atributos. .	p. 90
Figura 44	Erros de classificação para cenários com $n_r = 130$ amostras de treinamento por classe. Método híbrido usado para extração de atributos.	p. 91
Figura 45	Coefficientes para filtros de decomposição para <i>Db2</i>	p. 98
Figura 46	Coefficientes para filtros de decomposição para <i>Db8</i>	p. 98
Figura 47	Funções geradoras para <i>Db2</i>	p. 99
Figura 48	Funções geradoras para <i>Db8</i>	p. 99
Figura 49	Funções geradoras para <i>Haar</i>	p. 100
Figura 50	Coefficientes para filtros de decomposição para <i>Haar</i>	p. 101
Figura 51	Aproximação para a função $f(x)$, a partir de coeficientes de escala. .	p. 102
Figura 52	Coefficientes de detalhe para a função $f(x)$	p. 102

LISTA DE TABELAS

Tabela 2	Resultados de classificação dos dados simulados, com <i>wavelet</i> de Haar.	p. 51
Tabela 3	Resultados de classificação dos dados simulados, com <i>wavelet</i> de Daubechies <i>Db2</i> .	p. 52
Tabela 4	Resultados de classificação de ataques curtos, com <i>wavelet</i> de Haar.	p. 56
Tabela 5	Resultados de classificação de ataques curtos, com <i>wavelet</i> de Daubechies <i>Db2</i> .	p. 56
Tabela 6	Resultados de classificação de ataques longos, com <i>wavelet</i> de Haar.	p. 57
Tabela 7	Resultados de classificação de ataques longos, com <i>wavelet</i> de Daubechies <i>Db2</i> .	p. 57
Tabela 8	Resultados de classificação de imagens de naipes de cartas. Imagens com rotação. Método unidimensional, <i>wavelet</i> de Haar.	p. 63
Tabela 9	Resultados de classificação de imagens de naipes de cartas. Imagens com rotação. Método unidimensional, <i>wavelet</i> de Daubecheis <i>Db2</i> .	p. 64
Tabela 10	Resultados de classificação de imagens de naipes de cartas. Imagens com rotação. Método unidimensional, <i>wavelet</i> de Daubechies <i>Db8</i> .	p. 65
Tabela 11	Resultados de classificação de imagens de naipes de cartas. Imagens com rotação. Método unidimensional, <i>wavelet</i> de Haar.	p. 67
Tabela 12	Resultados de classificação de imagens de naipes de cartas. Imagens com rotação. Método unidimensional, <i>wavelet</i> de Daubechies <i>Db2</i> .	p. 68
Tabela 13	Resultados de classificação de imagens de naipes de cartas. Imagens com rotação. Método unidimensional, <i>wavelet</i> de Daubechies <i>Db8</i> .	p. 69
Tabela 14	Resultados de classificação de imagens de naipes de cartas. Imagens com rotação. Método bidimensional, <i>wavelet</i> de Haar.	p. 71
Tabela 15	Resultados de classificação de imagens de naipes de cartas. Imagens com rotação. Método bidimensional, <i>wavelet</i> de Daubechies <i>Db2</i> .	p. 73
Tabela 16	Resultados de classificação de imagens de naipes de cartas. Imagens com rotação. Método bidimensional, <i>wavelet</i> de Daubechies <i>Db8</i> .	p. 74
Tabela 17	Resultados de classificação de imagens de naipes de cartas. Imagens com rotação. Método bidimensional, <i>wavelet</i> de Haar.	p. 75
Tabela 18	Resultados de classificação de imagens de naipes de cartas. Imagens com rotação. Método bidimensional, <i>wavelet</i> de Daubechies <i>Db2</i> .	p. 76
Tabela 19	Resultados de classificação de imagens de naipes de cartas. Imagens com rotação. Método bidimensional, <i>wavelet</i> de Daubechies <i>Db8</i> .	p. 77

Tabela 20	Resultados de classificação de imagens de naipes de cartas. Imagens com rotação. Método híbrido, <i>wavelet</i> de Haar.	p. 79
Tabela 21	Resultados de classificação de imagens de naipes de cartas. Imagens com rotação. Método híbrido, <i>wavelet</i> de Daubecheis <i>Db2</i>	p. 80
Tabela 22	Resultados de classificação de imagens de naipes de cartas. Imagens com rotação. Método híbrido, <i>wavelet</i> de Daubechies <i>Db8</i>	p. 81
Tabela 23	Resultados de classificação de imagens de naipes de cartas. Imagens com rotação. Método híbrido, <i>wavelet</i> de Haar.	p. 83
Tabela 24	Resultados de classificação de imagens de naipes de cartas. Imagens com rotação. Método híbrido, <i>wavelet</i> de Daubechies <i>Db2</i>	p. 84
Tabela 25	Resultados de classificação de imagens de naipes de cartas. Imagens com rotação. Método híbrido, <i>wavelet</i> de Daubechies <i>Db8</i>	p. 85

LISTA DE ALGORITMOS

1	Algoritmo <i>leave-one-out</i>	p. 41
2	Algoritmo de classificação.	p. 42

LISTA DE SIGLAS E ABREVIATURAS

MRA	<i>Multiresolution Analysis</i>
TWD 1D	Transformada <i>Wavelet</i> Discreta Unidimensional
TWD 2D	Transformada <i>Wavelet</i> Discreta Bidimensional

SUMÁRIO

Agradecimentos

Resumo

Abstract

1	Introdução	p. 17
1.1	Proposta	p. 18
1.2	Organização do texto	p. 20
2	<i>Wavelet</i> x Fourier	p. 22
2.1	<i>Wavelet</i> Ortonormal de Daubechies	p. 23
2.1.1	Análise multiresolução	p. 24
2.1.2	Relação de escala e equação <i>wavelet</i>	p. 25
2.1.2.1	Propriedade dos Momentos Nulos	p. 26
2.1.3	Expansão de uma função em V_j	p. 27
2.2	Transformada <i>wavelet</i> discreta unidimensional (TWD 1D)	p. 28
2.3	Transformada <i>wavelet</i> bidimensional (TWD 2D)	p. 31
2.3.1	Tratamento das fronteiras para TWD 2D	p. 32
3	Classificação de padrões	p. 34
3.1	Métodos estatísticos de classificação de padrões	p. 35
3.1.1	Classificador Bayesiano	p. 36
4	Classificação Bayesiana e Wavelets	p. 39

4.1	Algoritmo de classificação Bayesiano	p. 39
4.2	Algoritmo de classificação Bayesiano e pré-processamento <i>wavelet</i>	p. 43
5	Experimentos Numéricos	p. 48
5.1	Conjunto de dados Unidimensionais	p. 50
5.1.1	Conjunto de dados simulados	p. 50
5.1.2	Conjunto de dados de anomalias da base DARPA 1999	p. 53
5.2	Conjunto de dados bidimensionais - Conjunto de imagens	p. 58
5.2.1	Método de classificação via transformada <i>wavelet</i> 1D	p. 60
5.2.2	Método de classificação via transformada <i>wavelet</i> 2D	p. 69
5.2.3	Método de classificação via método híbrido	p. 77
5.2.4	Análise dos Resultados de Classificação de Dados Bidimensionais	p. 86
6	Conclusões	p. 92
6.1	Trabalhos Futuros	p. 93
	Referências	p. 94
	Anexo A – Funções <i>wavelets</i> da família de Daubechies	p. 97
A.1	Exemplos de funções <i>wavelet</i>	p. 97
A.2	Exemplo de transformada <i>wavelet</i> discreta	p. 101

1 INTRODUÇÃO

O interesse na área de reconhecimento e classificação de padrões tem sido renovado ultimamente devido tanto ao aumento de aplicações desafiadoras, quanto também à demanda por grande capacidade computacional. Seleção de dados, extração de características e a classificação de sinais tornaram-se problemas comuns em várias aplicações. Dentre estas aplicações, incluem-se entre outras, mineração de dados, classificação de documentos, previsão financeira, organização de base de dados multimídia e biometria (Jain, Duin e Mao 2000).

O grande desenvolvimento de recursos computacionais nos últimos anos e a disponibilidade de potência computacional habilitam o rápido processamento de grandes conjuntos de dados, assim como também facilitam o uso de diferentes métodos para análise e classificação de dados. Ao mesmo tempo, devido à disponibilidade de grandes bases de dados, aumentam-se as demandas por sistemas automáticos de reconhecimento de padrões que atendam requisitos de desempenho (velocidade, exatidão e custo).

A classificação de padrões (como sinais cardíacos, sinais meteorológicos, entre outros) pressupõe o conhecimento prévio do número de padrões existentes. A idéia básica é que partindo de um número pré-estabelecido de padrões, o interesse é classificar um objeto de padrão desconhecido em um destes padrões pré-estabelecidos. Neste trabalho estes objetos são chamados de amostras, que são representadas por um vetor p -dimensional de pontos.

Existem inúmeros métodos aplicados à classificação de padrões. Dentre estes métodos, existem os métodos de classificação estatística, sendo muito aplicado os métodos baseados na Teoria de Bayes (Duda e Hart 1973). Classificação de padrões via métodos Bayesianos são utilizados para diferentes tipos de dados e aplicações, como reconhecimento de faces em imagens (Zhu e Schwartz 2002, Liu 2003, Liping, XinYuan e Yuqiang 2009), diagnóstico em imagens biomédicas (Al-Kadi 2009) e exames clínicos (Dan, Yuan-yuan e Chen-xiang 2010), reconhecimento de assinatura

(Cavalcanti et al. 2002), identificação de padrões de imagens (Keren 2002), caracterização de padrões de sinais (Liu 2004), classificação de vinhos (Beltran et al. 2006).

Classificadores Bayesianos são classificadores estatísticos que classificam uma amostra desconhecida em uma determinada classe, baseando-se na probabilidade *a posteriori* desta amostra pertencer a esta classe. Assim como outros métodos de classificação, os métodos Bayesianos são baseados na existência de amostras de treinamento. Amostras de treinamento são necessárias para estimar o comportamento médio do padrão de cada classe. Neste sentido é fundamental a obtenção de uma relação equilibrada entre a quantidade de amostras de treinamento e o número de pontos de cada amostra. Em situações práticas é muito comum a existência de poucas amostras de treinamento tornando impraticável a aplicação destes métodos de classificação.

Uma parte importante quanto à formulação do classificador é a questão da limitação do número de amostras dentro dos conjuntos de dados, que são parte importante no procedimento de classificação. A dificuldade está em estimar os padrões que serão utilizados no método, quando as amostras são compostas por muitos pontos e existem poucas amostras de treinamento para o classificador.

É fundamental a obtenção de uma relação equilibrada entre a quantidade de amostras e o número de pontos, uma vez que para a realização desta classificação de padrões o algoritmo necessita estimar as matrizes de covariância para as classes e este precisa ser suficientemente robusto para a realização da classificação das amostras. Esse equilíbrio pode ser alcançado a partir da idéia de extração de atributos das amostras para que sejam as fontes de conhecimento para os algoritmos de classificação realizarem as decisões.

1.1 Proposta

Muitos dos trabalhos práticos com classificação de padrões encontram o problema de desequilíbrio entre a grande dimensão das amostras e o número de amostras de treinamento disponíveis. Em muitas situações o número de amostras é insuficiente para a prática do processo de classificação. Desse desequilíbrio trata a idéia chamada de "maldição da dimensionalidade" (Jain e Chandrasekaran 1982).

Afim de chegar a um equilíbrio, aplicam-se técnicas de redução de dimensionalidade. Elas podem ser agrupadas em duas categorias gerais: extração de atributos e seleção de atributos. A seleção de atributos refere-se a algoritmos que selecionam um melhor subconjunto de pontos do conjunto de entrada. A técnica de extração de atributos refere-se

a métodos que geram novos dados baseadas em transformação dos dados de entrada.

Além de aplicações em estatística, a transformada *wavelet* é uma ferramenta muito utilizada na área de processamento de sinais, pois através da decomposição de sinais unidimensionais ou bidimensionais permite a extração de características com grande representatividade dos sinais originais, possibilitando assim representações compactas destes sinais. É esta a propriedade que será explorada na área de classificação de padrões, permitindo aumentar a eficiência e desempenho de classificação. O método de extração de características via transformada *wavelet* é aplicado tanto a dados unidimensionais, através da transformada unidimensional (1D) (Yen e Lin 2000, Bruce, Koger e Li 2002, Goumas, Zervakis e Stavrakakis 2002, Subramani, Sahu e Verma 2006, Prochazka, Kukul e Vysata 2008) quanto a dados bidimensionais através da transformada *wavelet* bidimensional (2D) (Rajpoot e Rajpoot 2004, Yen e Chu 2008, Narote, Narote e Waghmare 2009, Sung, Kwak e Lyou 2010).

Entre os vários procedimentos aplicados dentro da técnica de extração de características, há os trabalhos que aplicam transformada *wavelet* para a extração de características e métodos Bayesianos para classificação de padrões (Zhu e Schwartz 2002, Beltran et al. 2006, Al-Kadi 2009, Liping, XinYuan e Yuqiang 2009, Dan, Yuan-yuan e Chen-xiang 2010). O estudo realizado aqui segue esta linha.

O trabalho que inspirou este estudo foi o artigo de Mallet et.al. (Mallet et al. 1997), no qual é aplicado classificador Bayesiano para a classificação de dados simulados e dados mineralógicos, e uma transformada *wavelet* adaptativa é construída para a redução do número de variáveis das amostras. Além disso em (Mallet et al. 1997) é aplicada uma transformada *wavelet* discreta triádica que é adaptada para otimizar critério discriminante, e a partir dos coeficientes *wavelet* é realizado o processo de classificação. Neste sentido, a aplicação de transformada *wavelet* caracteriza-se como uma ferramenta importante no processo de redução do número de variáveis, pois mantém as características da função em uma representação multiresolução, formada por componentes com um menor número de pontos a cada nível de fatoração.

No trabalho desenvolvido aqui algumas das idéias propostas em (Mallet et al. 1997) são consideradas e estudadas. A partir de uma versão simplificada e não adaptativa do algoritmo dado em (Mallet et al. 1997) e do uso da transformada *wavelet* diádica, cuja implementação é rápida e mais simples do que a em base três utilizada no artigo

de referência, é estudado o impacto desta transformada quando aplicada à extração de características das amostras, e conseqüentemente, reduzindo a dimensão das amostras que fazem parte do conjunto de dados de treinamento do classificador Bayesiano.

O objetivo deste trabalho é o estudo da contribuição da transformada *wavelet* ortonormal da família de Daubechies ao algoritmo de classificação Bayesiano quando o cenário de classificação apresenta o desequilíbrio na relação entre quantidade de amostras das matrizes de treinamento e dimensão das amostras. A aplicação de transformada *wavelet* é realizada às amostras de treinamento e objetos a classificar, e a partir dos coeficientes obtidos pela transformação são extraídas características que serão utilizadas no processo de classificação.

Para a avaliação da classificação através do uso de transformada *wavelet* e métodos Bayesianos, dados unidimensionais e bidimensionais são utilizados. Para a extração de atributos de dados unidimensionais, a primeira abordagem utiliza a transformada *wavelet* unidimensional (TWD 1D), e os coeficientes resultantes da decomposição formam o novo conjunto de atributos. Para os dados bidimensionais é aplicada uma segunda abordagem via transformada *wavelet* bidimensional (TWD 2D) para a extração de atributos dos dados. A partir de uma nova representação da imagem são extraídas as características e realiza-se a classificação.

Neste trabalho, uma terceira abordagem é proposta, denominada de método híbrido. Neste método são mescladas as aplicações das transformadas *wavelet* bi- e unidimensional aos dados provenientes de imagens. Desta forma, a transformada unidimensional consegue capturar informações existentes entre dois níveis consecutivos da transformada bidimensional, o que permite a extração de características complementares em relação aos dados bidimensionais, provenientes apenas da transformada bidimensional.

Para cada uma das abordagens mencionadas, exemplos com aplicações específicas são analisados e o processo de classificação é então avaliado.

1.2 Organização do texto

O texto desta dissertação está organizado da seguinte maneira:

Os Capítulos 2 e 3 apresentam os fundamentos teóricos deste trabalho. No Capítulo 2 encontram-se as definições sobre transformada *wavelet* discreta, considerando-se funções *wavelets* da família de funções ortonormais de Daubechies. Além disso, no Capítulo 2 são apresentadas os aspectos de multiresolução intrínsecos às decomposições uni- e

bidimensionais. No Capítulo 3 são apresentados conceitos de classificação de padrões e método de extração de características. Além disso, classificador Bayesiano é apresentado como um método estatístico de classificação de padrões.

No Capítulo 4 o algoritmo de classificação baseado na teoria de decisão pela regra de Bayes é formulado. Neste capítulo também se faz a junção das duas componentes da metodologia estudada: a transformada *wavelet* é portanto associada ao processo de extração de características para a redução de dimensão dos dados.

No Capítulo 5 são apresentados os resultados obtidos para diferentes aplicações. Para a avaliação do algoritmo de classificação, primeiramente são considerados conjuntos de dados unidimensionais. Aqui dois conjuntos de dados são analisados: um conjunto formado por dados gerados sinteticamente e outro conjunto de dados de tráfego de redes de computadores. Os dados bidimensionais de teste são formados por imagens de cartas de baralho, cujas fontes de aquisição possuem diferença de intensidade de luz.

O Capítulo 6 apresenta uma discussão sobre os resultados obtidos para os problemas unidimensionais e bidimensionais apresentados, as conclusões e algumas perspectivas para trabalhos futuros.

2 WAVELET X FOURIER

Muitas funções matemáticas podem ser representadas por somas de funções fundamentais, que são conhecidas como funções base. Estas representações são dadas como expansões em séries, e um exemplo bem conhecido é a expansão em série de Fourier (Figueiredo 1997), que é válida para qualquer função f periódica (em particular com período 1). A expansão de Fourier de uma função f neste caso é dada por:

$$f(x) = \sum_{k=-\infty}^{\infty} c_k e^{i2\pi kx} = \sum_{k=-\infty}^{\infty} a_k \cos(2\pi kx) + b_k \sin(2\pi kx), x \in \mathbb{R}, \quad (2.1)$$

que representa a decomposição de f como superposição de harmônicos com frequências k em relação à base de funções seno e cosseno. Os coeficientes c_k , e consequentemente, a_k e b_k , da expansão são dados pelas integrais

$$c_k = \int_0^1 f(x) e^{-i2\pi kx} dx, \quad a_k = \int_0^1 f(x) \cos(2\pi kx) dx, \quad b_k = \int_0^1 f(x) \sin(2\pi kx) dx. \quad (2.2)$$

A transformada rápida de Fourier (FFT) é a maneira de se calcular os coeficientes a_k e b_k para dados discretos, considerando extensões periódicas desses dados, através de algoritmos rápidos e otimizados. Cada coeficiente c_k pode ser pensado como sendo o conteúdo harmônico médio (em um período) de f na frequência k .

Analicamente, a reconstrução de f é exata caso a função seja periódica e atenda certas hipóteses (ver (Figueiredo 1997) para aspectos teóricos sobre a transformada inversa de Fourier), mas na prática isto raramente acontece, principalmente pelo fato de serem analisados sinais e funções essencialmente não periódicos.

A transformada de Fourier, apesar de ser amplamente utilizada nas mais diversas aplicações em engenharia, apresenta ainda outras restrições:

- Fenômeno de Gibbs: oscilações nas proximidades das fronteiras do sinal analisado,

devido à expansão periódica considerada pela FFT.

- Impossibilidade de representação exata de dados que apresentem descontinuidades.
- Para dados truncados e discretos, a FFT inversa pode não recuperar o sinal original, dependendo das simplificações realizadas, tipo de extensão nas fronteiras, hipótese dos dados serem provenientes de funções pares ou ímpares.
- Caso um valor da função seja alterado, toda a série de Fourier será alterada. Ou seja, perturbações locais dos dados de entrada são percebidas por todos os coeficientes da expansão.

Como uma alternativa para a expansão de uma função em série de Fourier, existem outras funções que podem ser consideradas como base para representação. Na próxima seção é apresentada a família de funções *Wavelets* Ortonormais de Daubechies e sua Transformada Discreta rápida.

2.1 *Wavelet* Ortonormal de Daubechies

Existem várias famílias de funções *wavelets*. Este trabalho considera as funções *wavelets* ortonormais de Daubechies. Para um estudo detalhado e bastante aprofundado sobre esta e as demais famílias de *wavelets* ver (Daubechies 1992).

Uma das propriedades mais exploradas dessas funções é que as *wavelets* também podem ser usadas como base para a expansão de uma função f , assim como as funções trigonométricas foram utilizadas na expansão em série de Fourier. No entanto, as *wavelets* possuem uma característica essencial que as diferem das funções trigonométricas: elas são “ondas pequenas”, ou seja, são funções oscilatórias que decaem rapidamente e que diferem da função nula apenas em um pequeno intervalo fechado da reta, ou como se diz em terminologia específica da área, *wavelets* possuem suporte compacto (Daubechies 1992).

Assim como para as funções trigonométricas, na família das *wavelets* também existem duas funções fundamentais: a função escala $\phi(x)$ e a função *wavelet* $\psi(x)$, ambas pertencendo ao espaço $\mathcal{L}^2(\mathbb{R})$ das funções de energia finita, ou seja:

$$\int_{-\infty}^{\infty} \phi(x)^2 dx = 1 \quad \int_{-\infty}^{\infty} \psi(x)^2 dx = 1.$$

A função escala $\phi(x)$ é aquela que preserva a área, como mostra a equação 2.3 e portanto será capaz de captar informações médias, quando realizada a transformação. A

função *wavelet* $\psi(x)$ é a encarregada de perceber as variações em relação à média.

$$\int_{-\infty}^{\infty} \phi(x)dx = 1 \quad \int_{-\infty}^{\infty} \psi(x)dx = 0. \quad (2.3)$$

Ambas são ortogonais entre si:

$$\int_{-\infty}^{\infty} \phi(x)\psi(x)dx = 0$$

Denotando $\phi_{j,k}(x) = 2^{j/2}\phi(2^jx - k)$ e $\psi_{j,k}(x) = 2^{j/2}\psi(2^jx - k)$, $j, k \in \mathbb{Z}$, a propriedade de ortogonalidade e norma unitária ficam preservadas. Na verdade, a propriedade de ortonormalidade pode ser expressa pelas expressões a seguir (Daubechies 1992):

$$\int_{-\infty}^{\infty} \phi_{j,k}(x)\phi_{j,l}(x)dx = \delta_{k,l}, \quad (2.4)$$

$$\int_{-\infty}^{\infty} \psi_{j,k}(x)\psi_{j,l}(x)dx = \delta_{i,j}\delta_{k,l}, \quad (2.5)$$

$$\int_{-\infty}^{\infty} \phi_{j,k}(x)\psi_{j,l}(x)dx = 0, \quad (2.6)$$

sendo $i, j, k, l \in \mathbb{Z}$ e $\delta_{k,l}$ é o delta de Kronecker.

2.1.1 Análise multiresolução

As propriedades iniciais das *wavelets* ortonormais de Daubechies podem ser expressas também em termos dos seus espaços gerados, ou seja, dos conjuntos de todas as funções obtidas através de combinações lineares envolvendo as *wavelets*, suas translações e dilatações. É conveniente portanto definir os espaços de aproximação, que serão importantes na compreensão da estrutura multiresolução (multi escala) apresentada pela transformada discreta.

Assim fixada uma escala j qualquer, V_j será o espaço gerado por todas as translações da função escala $\phi_j(x)$ e W_j será o espaço gerado por todas as translações da função *wavelet* $\psi_j(x)$. E pelas relações 2.4, 2.5 e 2.6, os espaços V_j e W_j serão também ortogonais entre si, seja qual for a escala j escolhida.

Desta forma uma análise multiresolução (*multiresolution analysis*, MRA) é caracteri-

zada pelas seguintes propriedades entre estes espaços V_j , W_j e o espaço $\mathcal{L}^2(\mathbb{R})$:

1. $\{0\} \subset \dots \subset V_{-1} \subset V_0 \subset V_1 \subset \dots \subset \mathcal{L}^2(\mathbb{R})$;
2. $\overline{\bigcup_{j=-\infty}^{\infty} V_j} = \mathcal{L}^2(\mathbb{R})$;
3. $\{\phi(x - k)\}_{k \in \mathbb{Z}}$ é base ortonormal para V_0 ;
4. $f \in V_j \iff f(2 \cdot) \in V_{j+1}$;
5. $V_{j+1} = V_j \oplus W_j$.

O item (1) significa que todos os níveis de resolução terão energia finita. Segundo o item (2) qualquer função de energia finita pode ser representada por elementos dos espaços V_j com diferentes níveis de resolução. Pelo item (3), as translações da função escala formam uma base para o espaço V_0 . O item (4) diz que dado um elemento do espaço V_j , sua dilatação estará contida no nível seguinte. Como consequência, as bases para os espaços V_j também diferirão entre si apenas por um fator de escala. E finalmente o item (5) afirma que, apesar da passagem de um nível para outro ser dado por um fator de escala, ainda assim poderá existir informação complementar entre estes níveis, esta sendo então representada por elementos de W_j (Daubechies 1992)(Nielsen 1998).

Como os espaços W_j são complementos ortogonais de V_j em V_{j+1} , essa relação pode ser estendida entre quaisquer níveis. Assim dois espaços V_{J_0} e V_J , com $J > J_0$, podem ser decompostos

$$V_J = V_{J_0} \oplus \left(\bigoplus_{j=J_0}^{J-1} W_j \right).$$

Esta é a essência da transformada discreta, na qual diferentes níveis de resolução de uma função são traduzidos como uma informação grosseira (aqui representada pelos dados em V_{J_0}) que é corrigida por vários níveis de informações complementares, aqui representados pelas informações contidas em cada um dos espaços W_j .

2.1.2 Relação de escala e equação *wavelet*

Como fixado pela estrutura de análise multiresolução, $V_0 \subset V_1$. E portanto qualquer função em V_0 pode ser expandida em termos de funções base de V_1 . Em particular, a própria função escala $\phi(x) = \phi_{0,0}(x) \in V_0$ pode ser escrita como:

$$\phi(x) = \sum_{k=-\infty}^{\infty} h_k \phi_{1,k}(x) = \sqrt{2} \sum_{k=-\infty}^{\infty} h_k \phi(2x - k), \quad (2.7)$$

sendo

$$h_k = \int_{-\infty}^{\infty} \phi(x) \phi_{1,k}(x) dx. \quad (2.8)$$

Como as funções escala possuem suporte compacto, a integral indefinida é na verdade definida, e apenas um número finito de h_k será diferente de zero. Assim

$$\phi(x) = \sqrt{2} \sum_{k=0}^{D-1} h_k \phi(2x - k). \quad (2.9)$$

A equação (2.9) é fundamental para a teoria *wavelet* e é conhecida como *relação de escala*. D é um número par inteiro positivo chamado de *ordem da wavelet*, e os números h_0, h_1, \dots, h_{D-1} são chamados *coeficientes do filtro*. A função escala é unicamente caracterizada por esses coeficientes. Para maiores detalhes sobre a obtenção destes valores, ver (Daubechies 1992).

Analogamente à equação (2.9), tem-se a relação para a *wavelet* básica ψ . Como $\psi \in W_0$, $W_0 \in V_1$ e ψ também de suporte compacto, pode-se expandir ψ como

$$\psi(x) = \sqrt{2} \sum_{k=0}^{D-1} g_k \phi(2x - k), \quad (2.10)$$

sendo os coeficientes g_k dados por:

$$g_k = \int_{-\infty}^{\infty} \psi(x) \phi_{1,k}(x) dx. \quad (2.11)$$

A equação (2.10) é chamada equação *wavelet* e os coeficientes g_k são os filtros *wavelets*. Uma propriedade importante é a possibilidade de representar g_k em termos de h_k , como segue:

$$g_k = (-1)^k h_{D-1-k}, k = 0, 1, \dots, D-1.$$

2.1.2.1 Propriedade dos Momentos Nulos

Outra importante propriedade da função escala é a capacidade de representar polinômios exatamente para um grau $N - 1$, de acordo com o número de filtros D con-

siderados na relação de escala que a define, sendo que $D = 2N$. Consequentemente esses polinômios serão ortogonais às funções wavelets. Mais precisamente,

$$\int_{-\infty}^{\infty} x^p \psi(x) dx = 0 \quad (2.12)$$

e

$$x^p = \sum_{k=-\infty}^{\infty} M_k^p \phi(x - k), \quad x \in \mathbb{R}, \quad p = 0, 1, \dots, N - 1, \quad (2.13)$$

onde

$$M_k^p = \int_{-\infty}^{\infty} x^p \phi(x - k), \quad k \in \mathbb{Z}, \quad p = 0, 1, \dots, N - 1. \quad (2.14)$$

O que em termos de espaços de multiresolução significa que os polinômios até grau $N - 1$ estão em V_j e não necessitam de informações complementares dos espaços W_j . Isto desde que a função escala e a *wavelet* tenham $D = 2N$ filtros associados.

2.1.3 Expansão de uma função em V_j

Uma função f qualquer $\in V_j$ pode ser expandida em função da base de V_j , neste caso, translações da função escala dada no nível j , como dada na expressão 2.15 a seguir:

$$f(x) = \sum_{l=-\infty}^{\infty} c_{j,l} \phi_{j,l}(x), \quad x \in \mathbb{R}, \quad (2.15)$$

sendo os coeficientes da expansão dados por:

$$c_{j,l} = \int_{-\infty}^{\infty} f(x) \phi_{j,l}(x) dx. \quad (2.16)$$

No entanto, quando pensamos em $V_j = V_{j-1} \cup W_{j-1}$, a representação de f pode ser dada agora em função das bases de V_{j-1} e W_{j-1} , cuja união também forma uma base para V_j (lembrando que isso é válido, pois V_{j-1} e W_{j-1} são complementos ortogonais em V_j). Assim a expansão para a função f é dada por:

$$f(x) = \sum_{l=-\infty}^{\infty} c_{j-1,l} \phi_{j-1,l}(x) + \sum_{l=-\infty}^{\infty} d_{j-1,l} \psi_{j-1,l}(x), \quad (2.17)$$

cujos coeficientes são dados por:

$$\begin{aligned} c_{j-1,l} &= \int_{-\infty}^{\infty} f(x)\phi_{j-1,l}(x)dx, \\ d_{j-1,l} &= \int_{-\infty}^{\infty} f(x)\psi_{j-1,l}(x)dx. \end{aligned} \quad (2.18)$$

Observa-se que V_{j-1} por sua vez pode ser decomposto em função de suas duas componentes ortogonais V_{j-2} e W_{j-2} . Assim 2.17 pode ser novamente decomposta em mais um nível de resolução, desta vez considerando-se $j-2$ o nível mais grosseiro e j o nível mais fino. Neste caso, dois níveis de detalhes são necessários para se recuperar as informações dadas no nível j , como apresentado na representação envolvendo dois níveis dada a seguir:

$$f(x) = \sum_{l=-\infty}^{\infty} c_{j-2,l}\phi_{j-1,l}(x) + \left(\sum_{l=-\infty}^{\infty} d_{j-2,l}\psi_{j-2,l}(x) + \sum_{l=-\infty}^{\infty} d_{j-1,l}\psi_{j-1,l}(x) \right), \quad (2.19)$$

cujos coeficientes são dados de forma análoga aos coeficientes em 2.18.

Sendo assim, quando um certo nível J_0 for escolhido como sendo o nível mais grosseiro para a representação da função e J for considerado o nível mais fino, a representação em multinível é dada por:

$$f(x) = \sum_{l=-\infty}^{\infty} c_{J_0,l}\phi_{J_0,l}(x) + \sum_{j=J_0}^{J-1} \sum_{l=-\infty}^{\infty} d_{j-1,l}\psi_{j-1,l}(x), \quad (2.20)$$

novamente com coeficientes da expansão dados de forma análoga aos apresentados em 2.18.

2.2 Transformada *wavelet* discreta unidimensional (TWD 1D)

Quando a função f inicialmente considerada em V_J estiver definida em um intervalo $[a, b]$ finito e for discretizada com uma quantidade inicial de pontos também finita e dada de forma diádica por $N_J = 2^{J_{max}}$ pontos dentro do intervalo $[a, b]$, então todas as expansões 2.15, 2.17, 2.19 ou 2.20 serão simplificadas, uma vez que os somatórios em relação ao parâmetro de translação serão finitos, percorrendo as N_J posições possíveis, de acordo com a escala.

Da mesma forma que as expressões para as séries em base *wavelet* serão simplificadas,

também será simplificada a maneira com que se obtém seus coeficientes. Essa simplificação no cálculo dos coeficientes da expansão para o caso de funções discretizadas e malhas diádicas é o que se denomina *transformada wavelet discreta*. Neste trabalho é utilizado o algoritmo rápido de Mallat para esta transformação inversível (*fast wavelet transform, FWT*)(Mallat 1989).

Para se obter a transformada discreta, consideram-se as relações 2.18 para os coeficientes. Escrevendo-se as relações de escala (2.9) e a equação *wavelet* (2.10) para funções nos níveis $j - 1$ e substituindo-se então estas relações nas integrais, obtém-se as seguintes expressões envolvendo coeficientes de dois níveis consecutivos:

$$\text{Para } i = 0, \dots, N_{j-1} - 1 : c_{j-1,i} = \sum_{k=0}^{D-1} h_k c_{j,2i+k}, \quad d_{j-1,i} = \sum_{k=0}^{D-1} g_k c_{j,2i+k}. \quad (2.21)$$

Lembrando que $D = 2N$ é a ordem da *wavelet* e é determinado de acordo com a escolha feita para o número de coeficientes de filtros na definição da função escala e consequentemente, da *wavelet*. Observa-se ainda que a quantidade de coeficientes escala (e *wavelet*) no nível $j - 1$ é a metade do número de coeficientes do nível j , ou seja, $N_{j-1} = N_j/2$. As relações dadas em (2.2) são então as expressões para a Transformada *wavelet* Discreta para 1° nível de decomposição.

Considerando-se o vetor inicial de coeficientes $c_J = (c_{J,0}, \dots, c_{J,N_J-1})$ com $N_J = 2^{J_{max}}$ como sendo o vetor contendo os valores discretos de uma função f na escala J , mais fina de resolução, e considerando-se ainda que o nível mais grosseiro da transformada seja J_0 , então a expansão em série neste caso discreto é dada por:

$$f = \sum_{i=0}^{N_{J_0}-1} c_{J_0,i} \phi_{J_0,i} + \sum_{j=J_0}^{J-1} \sum_{i=0}^{N_j-1} d_{j,i} \Psi_{j,i}, \quad (2.22)$$

sendo os coeficientes $c_{J_0,i}$ referentes a informação do sinal no nível mais grosseiro. A família de coeficientes *wavelet* $d_{j,i}$, $j = J_0, J_0 + 1, \dots, J - 1$ são todos detalhes em diferentes escalas necessários para a reconstrução da função no nível mais fino (função original).

Visualmente, pode-se observar o resultado de decomposição de uma função f através do fluxo de transformada *wavelet* dos coeficientes de escala resultantes a cada nível de decomposição (Figura 1).

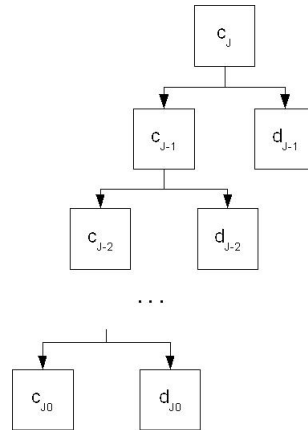


Figura 1: Fluxo de decomposição via transformada *wavelet*.

Do ponto de vista computacional, o processo de transformada *wavelet* discreta de um sinal unidimensional pode ser descrito como procedimento de filtragem deste sinal. Os coeficientes h representam coeficientes de filtro passa-baixas, os coeficientes de escala da decomposição são resultado da filtragem. Os coeficientes g representam coeficientes de filtro passa-altas, gerando os coeficientes de detalhe (coeficientes *wavelet*) do sinal, que representam o conteúdo de alta-freqüência.

O processo inverso, a transformada *wavelet* discreta inversa realiza a reconstrução do sinal a partir do mesmo conjunto de filtros. A reconstrução de c_{j-1} passa pela reconstrução do sinal a partir dos coeficientes de escala c_{j-2} e dos coeficientes de detalhe d_{j-2} . A operação de transformada *wavelet* discreta a partir de *wavelets* periodizadas é exatamente inversível (Nielsen 1998). As Figuras 2 e 3 representam a operação de filtragem de um sinal unidimensional de entrada denominado x : a Figura 2 representa a transformada *wavelet* como operação de filtragem, e a Figura 3 apresenta a reconstrução do sinal.

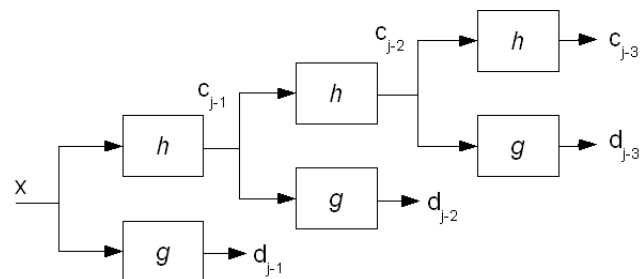


Figura 2: Transformada *wavelet* unidimensional: filtragem de sinal unidimensional x .

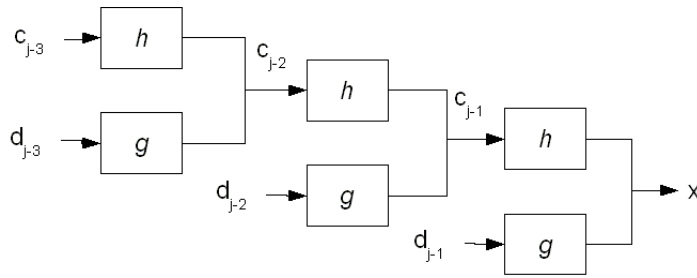


Figura 3: Transformada *wavelet* inversa: reconstrução do sinal unidimensional x .

2.3 Transformada *wavelet* bidimensional (TWD 2D)

Para a decomposição de uma função $f(x, y)$ ou uma imagem s , esta deverá ser dada de forma discreta e armazenada em uma matriz $A_{M \times N}$ cuja dimensão seja dada por valores diádicos: $M = 2^{imax}$ e $N = 2^{jmax}$. Uma possibilidade de se obter a transformada rápida bidimensional é através da execução da operação unidimensional de transformação primeiramente aplicada em todas as linha e então aplicada em todas as colunas da matriz A , como representado na Figura 4.

A transformada *wavelet* bidimensional (2D) pode ser obtida de diversas formas, basta que a ordem de execução das transformadas *wavelet* unidimensional para as linhas e colunas seja realizada em ordem diferente. Para cada maneira de se compor estas transformações unidimensionais, haverá uma transformada inversa correspondente. Neste trabalho considera-se a transformada 1D aplicada primeiro em todas as linhas e depois em todas as colunas, como apresentado em (Stollnitz, Derose e Salesin 1995, Stollnitz, Derose e Salesin 1995).

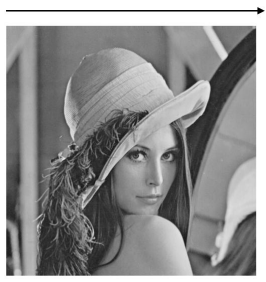


Figura 4: Ordem de execução da transformada *wavelet*: primeiramente, linhas, e então as colunas.

A partir da transformada *wavelet* discreta 2D (W) da imagem s , obtém-se a faturação da imagem em 4 sub-bandas a cada nível de decomposição. As sub-bandas CD, DC e DD, representam os detalhes resultantes da decomposição da imagem na direção horizontal, vertical e diagonal, respectivamente. A sub-banda CC representa os coeficientes de escala

do sinal de imagem (Figura 5). A imagem aproximada (sub-banda CC) é usada para obter o próximo nível de decomposição.

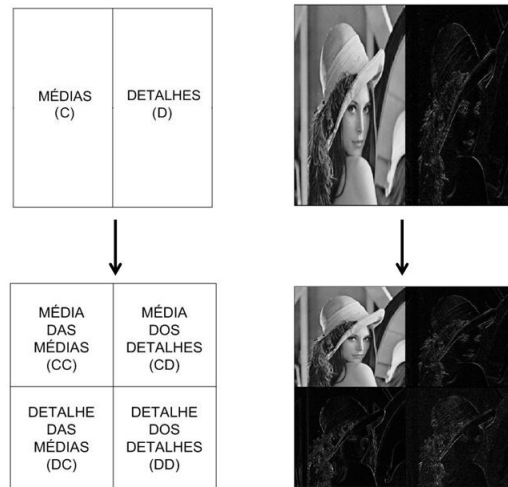


Figura 5: Resultado da transformada *wavelet* em 2 dimensões.

Para decompor a imagem em i níveis, deve-se executar a transformada *wavelet* na sub-banda CC: resulta-se em 4 novas sub-bandas, sendo $W(CC_i) = [CC_{i+1}, CD_{i+1}, DC_{i+1}, DD_{i+1}]$. A cada nível, cada sub-banda tem dimensão $N/2^i \times N/2^i$.

2.3.1 Tratamento das fronteiras para TWD 2D

Para as *wavelets* periódicas, o algoritmo de TWD periódica considera a utilização dos dados do início de cada linha (ou coluna) para calcular o valor dos coeficientes do próximo nível e no processo inverso considera estes valores novamente. Neste trabalho são consideradas as *wavelets* periódicas *Db2* e *Db8* para o processamento das imagens.

A Figura 6 apresenta a decomposição da imagem através de *wavelets Db8* periódica. Percebe-se a consequência do uso de *wavelets* periódica na observação das fronteiras das imagens resultantes em cada nível de decomposição. As imagens com menor resolução tem distorção inserida nas fronteiras, e a distorção fica maior a cada nível de decomposição principalmente para *wavelets* com maior número de coeficientes nos filtros de decomposição como é o caso da *wavelet Db8*. Observando cada nível da decomposição, nota-se que a imagem passa impressão que é deslocada para a diagonal superior esquerda, e parte da imagem reaparece na fronteira dentro das linhas brancas tracejadas. Assim como para a transformada *wavelet* unidimensional, a operação de transformada *wavelet* bidimensional a partir de *wavelets* periodizadas é exatamente inversível (Nielsen 1998).



Figura 6: Decomposição de imagem através de *wavelet Db8* periódica.

3 CLASSIFICAÇÃO DE PADRÕES

Classificação de padrões é uma área que busca a classificação de amostras em subconjuntos de elementos com atributos em comum. O estudo desta técnica busca realizar a distinção entre diferentes padrões dentro de um espaço de observação, e realizar decisões sobre as categorias de padrões observados. Exemplos deste tipo de aplicação têm-se reconhecimento de letras manuscritas, classificação de imagens nas áreas biomédicas, dados de sensoriamento remoto, bio-identificação.

O *design* de sistemas de classificação de padrões essencialmente envolve os seguintes aspectos: 1) aquisição de amostras e pré-processamento, 2) estimativa dos padrões, 3) realização de decisão para dados futuros. O domínio do problema dita a forma para aquisição de dados, pré-processamento, forma de representação. Em geral, um problema de reconhecimento bem definido, onde há pouca correlação entre classes e muita correlação dentro das classes, conduzem para uma representação compacta dos padrões e uma simples estratégia de decisão (Jain, Duin e Mao 2000).

A aprendizagem a partir de um conjunto de treinamento é uma característica importante e desejada na maioria dos sistemas de classificação de padrões. As abordagens mais conhecidas e aplicadas são: “*template matching*” (equiparação com modelos), classificação estatística e redes neurais (Jain, Duin e Mao 2000). Um fator que prejudica o processo de classificação é o desequilíbrio entre a dimensão das amostras e número de amostras de treinamento

Há dois grandes motivos para manter a dimensionalidade da representação dos padrões (número de pontos) tão pequeno quanto possível. O primeiro está relacionado ao custo de medição e exatidão na classificação. Um conjunto de características limitado, porém proeminente, simplifica tanto a representação dos padrões quanto a construção dos classificadores a partir da representação selecionada. Assim, o classificador resultante será mais rápido e usará menos recursos de memórias. Além disso, um número menor de ca-

racterísticas pode reduzir o efeito da maldição da dimensionalidade quando há limitação no número de amostras. Por outro lado, a redução do número de características pode resultar em uma perda na capacidade de discriminação e então diminuir a exatidão do resultado de classificação (Jain, Duin e Mao 2000).

As técnicas de redução de dimensionalidade podem ser agrupadas em duas categorias gerais: extração de atributos e seleção de atributos.

No primeiro caso, transformações lineares ou não-lineares são aplicadas aos dados originais, concentrando o poder de discriminação entre as classes em um número menor de pontos, permitindo desta forma reduzir a dimensionalidade dos dados originais com uma perda de informação desprezível. No processo de seleção de atributos critérios de separabilidade entre as classes são utilizados para selecionar um subconjunto ótimo dos dados originais (Moraes e Haertel 2007). É importante distinguir entre extração de atributos e seleção de atributos. O primeiro termo refere-se a métodos que criam novas características baseadas na transformação ou combinações do conjunto de dados originais. O termo seleção de atributos refere-se à algoritmos que selecionam o melhor sub-conjunto de dados do conjunto de entrada. Um conjunto de dados transformados gerados por método de extração de atributos pode prover uma melhor capacidade que o melhor subconjunto de conjunto de atributos de entrada, mas estes novos dados podem não conter um significado físico claro (Jain, Duin e Mao 2000).

3.1 Métodos estatísticos de classificação de padrões

Técnicas estatísticas têm sido utilizadas para desenvolver métodos de classificação de padrões, onde o foco está nas propriedades estatísticas dos padrões, geralmente expressas em densidades probabilísticas (Duda e Hart 1973). O padrão é representado por um conjunto de p atributos, visto como um vetor de pontos p -dimensional. Conceitos bem conhecidos da teoria de decisão são utilizados para estabelecer limites de decisão entre as classes de padrões. O sistema de classificação funciona em duas etapas: etapa de treinamento e etapa de classificação. No modo de treinamento o módulo extrai/seleciona os atributos apropriados para representar os padrões pré-estabelecidos e o classificador é treinado para representar as classes. Na etapa de classificação, o classificador treinado designa as amostras de entrada em uma das classes sob consideração baseado nos atributos medidos.

O principal objetivo é extrair/selecionar atributos que identifiquem corretamente

padrões que classifiquem de forma efetiva os dados analisados, de tal forma que classes bem definidas possam ser identificadas. Isto pode ser realizado usando teoria de classificação baseada em regras de decisão de Bayes (Mallet et al. 1997). De posse destas classes o classificador tem por objetivo estabelecer limites de decisão. No classificador Bayesiano os limites de decisão são determinados pelas probabilidades *a posteriori* de uma particular amostra pertencer à cada uma das classes.

A regra de decisão de Bayes designa uma amostra para a classe cuja probabilidade *a posteriori* seja a maior. A regra de decisão de Bayes gera um classificador ótimo, no sentido que, fixadas determinadas probabilidades *a priori*, função perda e densidades condicionais probabilísticas conhecidas, nenhuma outra regra de decisão terá um risco tão baixo. Na prática, a regra de decisão de Bayes empírica, ou regra “plug-in”, é usada: estimativas das densidades são usadas no lugar das densidades verdadeiras que são desconhecidas. Estas estimativas de densidade podem ser obtidas através de métodos paramétricos ou não paramétricos. Os métodos paramétricos são comumente aplicados sendo utilizada a distribuição normal multivariada para características contínuas e as distribuições binomial e multinomial para características discretos.

3.1.1 Classificador Bayesiano

Cada amostra representará um espectro composto por medições discretas feitas em p variáveis e representado por um vetor p -dimensional de pontos $\mathbf{x} = [x_1, \dots, x_p]^T$. Os dados de treinamento constituem uma amostra aleatória do vetor \mathbf{x} para cada classe de interesse $r = 1, \dots, R$. Estes dados são armazenados em linhas de matriz \mathbf{X} ($n \times p$), com $n = n_1 + n_2 + \dots + n_{r(r=1,2,\dots,R)}$ e n_r sendo o número de amostras de treinamento da classe r . Abaixo, é representada uma matriz de treinamento com $R = 3$ classes. Esta matriz tem o mesmo número de amostras de treinamento para cada uma das classes. Porém, o número de amostras de treinamento não precisa ser igual para todas as classes.

$$X = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} x_{1,(1)}^1 & x_{1,(1)}^2 & \cdots & x_{1,(1)}^p \\ x_{2,(1)}^1 & x_{2,(1)}^2 & \cdots & x_{2,(1)}^p \\ \dots & \dots & \dots & \dots \\ x_{n_1,(1)}^1 & x_{n_1,(1)}^2 & \cdots & x_{n_1,(1)}^p \end{bmatrix} \\ \begin{bmatrix} x_{1,(2)}^1 & x_{1,(2)}^2 & \cdots & x_{1,(2)}^p \\ x_{2,(2)}^1 & x_{2,(2)}^2 & \cdots & x_{2,(2)}^p \\ \dots & \dots & \dots & \dots \\ x_{n_2,(2)}^1 & x_{n_2,(2)}^2 & \cdots & x_{n_2,(2)}^p \end{bmatrix} \\ \begin{bmatrix} x_{1,(3)}^1 & x_{1,(3)}^2 & \cdots & x_{1,(3)}^p \\ x_{2,(3)}^1 & x_{2,(3)}^2 & \cdots & x_{2,(3)}^p \\ \dots & \dots & \dots & \dots \\ x_{n_3,(3)}^1 & x_{n_3,(3)}^2 & \cdots & x_{n_3,(3)}^p \end{bmatrix} \end{bmatrix}$$

Uma matriz de treinamento é então construída para cada classe r . Baseado nestes dados estima-se um comportamento médio com intervalos de confiança para cada classe. Para estas estimativas supõem-se distribuição normal multivariada para $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_R$, com médias distintas para cada classe e com matriz de variância/covariância igual para todas as classes.

Para se classificar uma amostra futura \mathbf{x}_0 de classe desconhecida, calcula-se as probabilidades *a posteriori* $P(r|\mathbf{x}_0)$ para $r = 1, 2, \dots, R$. A amostra \mathbf{x}_0 será classificada na classe cuja probabilidade *a posteriori* seja a maior. O cálculo de $P(r|\mathbf{x}_0)$ é realizado utilizando-se o teorema de Bayes (Duda e Hart 1973). Todos os parâmetros necessários para este cálculo são obtidos utilizando os dados da amostra \mathbf{x}_0 e as estimativas das classes já obtidas com as amostras de treinamento. O valor de $P(r|\mathbf{x}_0)$ é encontrado aplicando

$$P(r|\mathbf{x}_0) = \frac{p(\mathbf{x}_0|r)P(r)}{\sum_{r=1}^R p(\mathbf{x}_0|r)P(r)}, r = 1, \dots, R,$$

sendo $P(r)$ a probabilidade *a priori* de um objeto pertencer à classe r (em geral assume-se *priori* não informativa, ou seja, mesma probabilidade *a priori* para todas as R classes). $p(\mathbf{x}_0|r)$ é o valor da densidade estimada da classe r no ponto x_0 . Supondo matrizes de variâncias/covariâncias iguais para todas as classes (esta suposição conjuntamente com o uso da probabilidade *a posteriori* máxima para classificação caracteriza a análise discriminante linear Bayesiana) e distribuição normal multivariada, o valor $p(\mathbf{x}_0|r)$ é calculado por:

$$p(\mathbf{x}_0|r) = (2\pi)^{-\frac{p}{2}} |\mathbf{S}_{pooled}|^{-\frac{1}{2}} \exp[-0,5(\mathbf{x}_0 - \bar{\mathbf{x}}_r)\mathbf{S}_{pooled}^{-1}(\mathbf{x}_0 - \bar{\mathbf{x}}_r)^T],$$

sendo $\mathbf{S}_{pooled} = \left[\sum_{r=1}^R n_r \mathbf{S}_r \right] / n$ a matriz de covariância ponderada e \mathbf{S}_r , a matriz de variância/co-variância estimada para a classe r dada por:

$$\mathbf{S}_r = \left[\sum_{i=1}^{n_r} (\mathbf{x}_{i(r)} - \bar{\mathbf{x}}_r)^T (\mathbf{x}_{i(r)} - \bar{\mathbf{x}}_r) \right] / n_r.$$

Na equação acima, $\bar{\mathbf{x}}_r = \left[\sum_{i=1}^{n_r} \mathbf{x}_{i(r)} \right] / n_r$ é o vetor de médias estimado para a classe r e $\mathbf{x}_{i(r)}$ indica que a amostra de treinamento i pertence à classe r .

O algoritmo de classificação baseado na regra de decisão de Bayes é descrito na Seção 4.1.

4 CLASSIFICAÇÃO BAYESIANA E WAVELETS

Nesta seção é apresentada uma simplificação do algoritmo de classificação apresentado em (Mallet et al. 1997). A transformada *wavelet* é aplicada no pré-processamento dos dados de classificação, tanto para os dados de treinamento como para os objetos de classificação. O algoritmo original aplica uma família de *wavelets* triádicas e é construída de forma adaptativa, e realiza a avaliação de resultados de classificação utilizando os coeficientes das 3 bandas de decomposição para formar as matrizes de treinamento para o classificador.

No algoritmo descrito a seguir (Seção 4.1), a parte adaptativa de construção da *wavelet* foi retirada e uma *wavelet* diádica foi utilizada. Os coeficientes de escala são utilizados para formar as matrizes de treinamento e objetos de classificação.

Na Seção 4.2 são apresentadas as três abordagens para a extração de atributos a partir da transformada *wavelet*. É apresentado ainda o modo como são realizados os procedimentos para esta extração.

4.1 Algoritmo de classificação Bayesiano

O método de classificação Bayesiano considera que cada amostra a ser classificada é representada por um vetor p -dimensional, $x_F = (x_1, x_2, \dots, x_p)$, de classe desconhecida. Supondo que existam R classes, r_1, r_2, \dots, r_R classes treinadas com base em n amostras de treinamento, o objetivo é obter a qual dessas classes o vetor x_F pertence. O classificador atribuirá o vetor x_F a classe que tem a maior probabilidade *a posteriori* $P(r|x_F)$. Desse modo, x_F é associada à classe r_k , se e somente se a probabilidade $P(r_k|x_F)$ é maior que $P(r_j|x_F)$. Ou seja:

$$P(r_k|x_F) > P(r_j|x_F), \forall j, 1 \leq j \leq R, j \neq k.$$

Apesar da regra de classificação acima ser ótima, ela tem forte influência da qualidade das amostras de treinamento.

Para avaliar a qualidade das amostras de treinamento utiliza-se neste trabalho a técnica denominada *leave-one-out*. Nesta técnica uma amostra de treinamento é considerada uma amostra a ser classificada. Esta técnica está descrita no Algoritmo 1. Cada amostra de treinamento de cada classe $r = 1, 2, \dots, R$ é considerada como uma amostra de classe desconhecida e atribuída a x_0 . Esta amostra é retirada da matriz de treinamento, e são realizados os cálculos de vetor de médias para as classes, matrizes de variância/covariância para as classes, matriz de covariância ponderada, e cálculo das probabilidades a *posteriori* de classificação. Esta ação é repetida para todas as amostras de treinamento de todas as classes, para avaliar cada amostra de treinamento. A idéia básica desta técnica é construir a regra de classificação sem considerar a amostra retirada da matriz de treinamento e classificá-la em sua classe de origem. Se as classes forem bem definidas e suficientemente distintas umas das outras, este processo não produzirá falsos positivos. Caso haja um alto grau de similaridade entre as classes, nesta etapa então esta característica será percebida pelo algoritmo através de várias escolhas para a classificação e, portanto, haverá um alto grau de falsos positivos.

A classificação é executada quando tem-se uma amostra de classe desconhecida e deseja-se classificá-la em uma classe pré-estabelecida. O algoritmo de classificação é similar ao Algoritmo 1: as estimativas para as classes são realizadas com todas as amostras de treinamento de todas as classes, e x_0 recebe uma amostra de classes desconhecida x_F . Ao final são calculadas as probabilidades a *posteriori* desta amostra de classe desconhecida pertencer à cada classe. Este processo é repetido para todas as amostras de um conjunto de amostras de classe desconhecida. Este procedimento é descrito no Algoritmo 2.

Algoritmo 1: Algoritmo *leave-one-out*.

1) Entrada de matriz \mathbf{X} de amostras de treinamento, $n \times p$. $n = n_1 + n_2 + \dots + n_R$, em que n representa número de amostras, e R é o total de classes, cada amostra com p pontos. Tem-se então n_1 amostras da matriz de dados referentes à curvas da Classe 1, n_2 amostras da Classe 2, etc...

2) Retira a amostra $x_{i(r)}$ da matriz de amostras de treinamento e $x_0 = x_{i(r)}$.

Faz-se:

a) Cálculo do vetor de médias para cada classe $r = 1, \dots, R$ a partir de:

$$\bar{\mathbf{x}}_r = \left[\sum_{i=1}^{n_r} \mathbf{x}_{i(r)} \right] / n_r$$

b) Cálculo da matriz de variância/covariâncias S_r para cada classe $r = 1, \dots, R$:

$$\mathbf{S}_r = \left[\sum_{i=1}^{n_r} (\mathbf{x}_{i(r)} - \bar{\mathbf{x}}_r)^T (\mathbf{x}_{i(r)} - \bar{\mathbf{x}}_r) \right] / n_r$$

c) Cálculo da matriz de variância/covariâncias ponderada S_{pooled}

$$\mathbf{S}_{pooled} = \left[\sum_{r=1}^R n_r \mathbf{S}_r \right] / n$$

d) Para x_0 , calcula-se a verossimilhança $p(x_0|r)$, resultando em R verossimilhanças.

$$p(\mathbf{x}_0|r) = (2\pi)^{-\frac{p}{2}} |\mathbf{S}_{pooled}|^{-\frac{1}{2}} \exp[-0,5(\mathbf{x}_0 - \bar{\mathbf{x}}_r) \mathbf{S}_{pooled}^{-1} (\mathbf{x}_0 - \bar{\mathbf{x}}_r)^T].$$

e) Através dos valores determinados em 2(d), calcula-se as probabilidades a *posteriori* $P(r|x_0)$. Inicialmente, utiliza-se $P(r) = 1/R$ para todo $r = 1, \dots, R$.

Obtém-se R probabilidades a *posteriori*, e a soma destas probabilidades é igual a 1. Esta *posteriori* representa a probabilidade atualizada de uma curva pertencer a classe r .

$$P(r|\mathbf{x}_0) = \frac{p(\mathbf{x}_0|r)P(r)}{\sum_{r=1}^R p(\mathbf{x}_0|r)P(r)}, r = 1, \dots, R$$

f) Classifique x_0 na classe cuja probabilidade a *posteriori* calculada em 2(e) é a de maior valor.

3) Repita o passo 2 para todas as amostras de treinamento, ou seja, para

$$x_{i(r)} = x_{1(1)}, x_{2(1)}, \dots, x_{1(2)}, \dots, x_{1(3)}, \dots, x_{n(R)}.$$

Algoritmo 2: Algoritmo de classificação.

1) Entrada de matriz \mathbf{X} de amostras de treinamento, $n \times p$. $n = n_1 + n_2 + \dots + n_R$, em que n representa número de amostras, e R é o total de classes, cada amostra com p pontos.

2) Cálculo do vetor de médias para cada classe $r = 1, \dots, R$ a partir de:

$$\bar{\mathbf{x}}_r = \left[\sum_{i=1}^{n_r} \mathbf{x}_{i(r)} \right] / n_r$$

3) Cálculo da matriz de variância/covariâncias S_r para cada classe $r = 1, \dots, R$:

$$\mathbf{S}_r = \left[\sum_{i=1}^{n_r} (\mathbf{x}_{i(r)} - \bar{\mathbf{x}}_r)^T (\mathbf{x}_{i(r)} - \bar{\mathbf{x}}_r) \right] / n_r$$

4) Cálculo da matriz de variância/covariâncias ponderada S_{pooled}

$$\mathbf{S}_{pooled} = \left[\sum_{r=1}^R n_r \mathbf{S}_r \right] / n$$

5) x_0 recebe amostra de classe desconhecida x_F ($x_0 = x_F$):

a) Para x_0 , calcula-se a verossimilhança $p(x_0|r)$, resultando em R verossimilhanças.

$$p(\mathbf{x}_0|r) = (2\pi)^{-\frac{p}{2}} |\mathbf{S}_{pooled}|^{-\frac{1}{2}} \exp[-0,5(\mathbf{x}_0 - \bar{\mathbf{x}}_r) \mathbf{S}_{pooled}^{-1} (\mathbf{x}_0 - \bar{\mathbf{x}}_r)^T].$$

b) Calcula-se as probabilidades a *posteriori* $P(r|x_0)$. Inicialmente, utiliza-se $P(r) = 1/R$ para todo $r = 1, \dots, R$. Obtém-se R probabilidades a *posteriori*, e a soma destas probabilidades é igual a 1. Esta *posteriori* representa a probabilidade atualizada de uma curva pertencer a classe r .

$$P(r|\mathbf{x}_0) = \frac{p(\mathbf{x}_0|r)P(r)}{\sum_{r=1}^R p(\mathbf{x}_0|r)P(r)}, r = 1, \dots, R$$

f) Classifique x_0 na classe cuja probabilidade a *posteriori* calculada em 2(e) é a de maior valor.

6) Repita o passo 5 para todas as amostras de classe desconhecida.

4.2 Algoritmo de classificação Bayesiano e pré-processamento *wavelet*

Para realização da classificação é essencial que exista uma relação equilibrada entre o número de amostras que compõem as matrizes de treinamento e a dimensão das amostras. Essa relação garante o cálculo das matrizes inversas das matrizes de covariância para as classes, e viabiliza o cálculo das probabilidades de classificação. Um dos objetivos deste trabalho é o estudo da contribuição da transformada *wavelet* ao algoritmo de classificação quando o cenário de classificação apresenta desequilíbrio entre a quantidade de amostras das matrizes de treinamento e suas dimensões. A aplicação de transformada *wavelet* é realizada às amostras de treinamento e objetos à classificar.

A transformada *wavelet* é aplicada como parte da etapa de pré-processamento dos dados (Figura 7). O pré-processamento simplifica as operações subsequentes sem perda de informação relevante (Duda e Hart 1973). Este processo é executado afim de extrair os valores mais significativos, reduzir ruídos devido a aquisição dos dados, e assim processar um volume menor de dados com grande grau de representatividade dos dados originais.

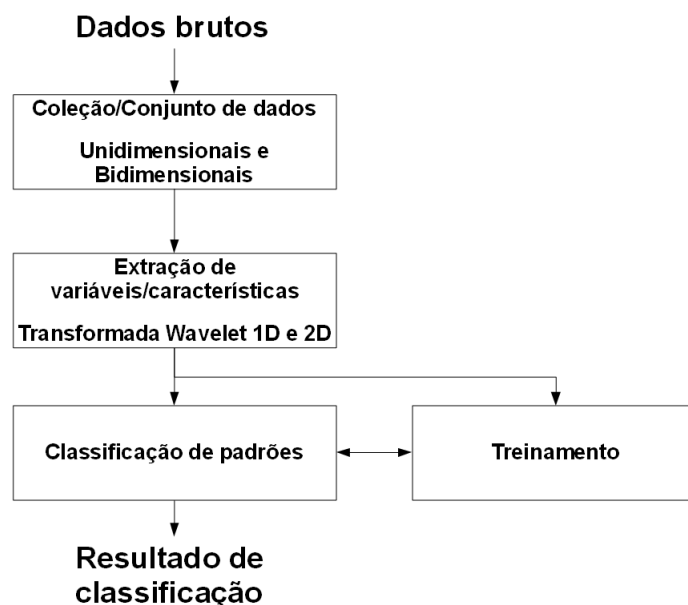


Figura 7: Etapas de um sistema de classificação. Adaptado de (Pei et al. 1995), (Pittner e Kamarthi 1997)

A aplicação da transformada *wavelet* funciona como um método de extração de atributos das amostras. A partir da transformada *wavelet* de cada amostra da matriz de treinamento, forma-se a nova matriz de treinamento com as características levantadas a partir da decomposição *wavelet*. Os resultados de classificação podem ser avaliados a cada

nível de decomposição da transformada *wavelet*, avaliando cada novo grupo de atributos extraídos. A transformada *wavelet* também é aplicada ao objeto de classificação, pois deve-se avaliar o mesmo princípio de extração de atributos que é aplicado à matriz de treinamento e amostra de classificação.

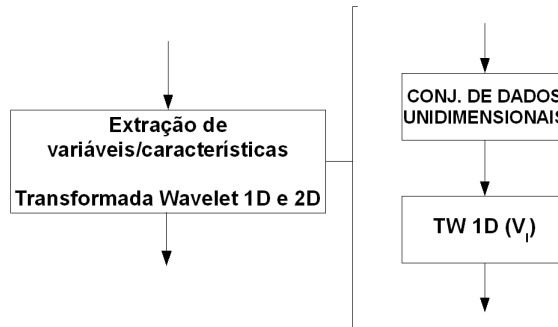


Figura 8: Abordagem unidimensional para dados unidimensionais.

O algoritmo de classificação (descrito na Seção 4.1), em modo de classificação de objetos, avalia sinais unidimensionais. Para estes sinais a transformada *wavelet* unidimensional é aplicada e os coeficientes de escala compõem as amostras de treinamento (Figura 8). O número de níveis de decomposição pela transformada *wavelet* é um parâmetro que pode ser definido no início do processo de classificação. Neste trabalho, para a transformada unidimensional praticamente todos os níveis da decomposição são avaliados. A partir dos coeficientes resultantes, obtém-se nova representação aproximada das amostras. Então, uma nova matriz de treinamento é composta. O algoritmo de classificação executa o procedimento de classificação e avalia-se a taxa de acerto de classificação. A cada nível de decomposição é feita essa análise, afim de avaliar o equilíbrio entre número de amostras e atributos extraídos das amostras.

Neste trabalho também é avaliado o comportamento do classificador para o reconhecimento de padrões de imagens, ou seja, dados bidimensionais. A partir desta avaliação são utilizadas duas abordagens para a classificação de dados bidimensionais.

A primeira abordagem aplica a transformada *wavelet* unidimensional aos dados bidimensionais, mas os dados bidimensionais devem ser passados para o formato de um vetor (devem ser unidimensionais) durante etapa de pré-processamento (bloco em destaque na Figura 9). O vetor é formado pela concatenação das linhas da matriz de dados. Tanto as amostras das matrizes de treinamento quanto os objetos de classificação devem passar por este procedimento. A partir destas novas representações unidimensionais, a transformada *wavelet* é aplicada e os coeficiente de escala compõem as novas matrizes de treinamento das classes. A cada nível de decomposição é avaliada a classificação dos objetos. O número

de níveis de decomposição que são avaliados na classificação é um parâmetro definido no começo do procedimento e neste trabalho são avaliados praticamente todos os níveis.

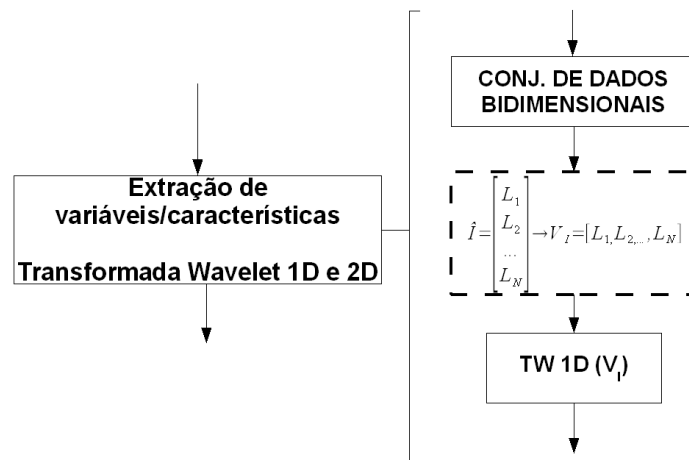


Figura 9: Abordagem 1D para dados bidimensionais.

A segunda abordagem utilizada para a análise de dados bidimensionais é através do uso da transformada *wavelet* 2D para a extração de atributos das amostras. Através da transformada *wavelet* 2D, os dados bidimensionais são reduzidos e obtém-se novas representações. Os dados da sub-banda CC, que representam os coeficientes de escala do sinal de imagem, formam as matrizes de treinamento. Assim como a abordagem que aplica a transformada unidimensional, para esta abordagem o número de níveis para a transformada bidimensional é um parâmetro definido no início do processo. Os dados bidimensionais somente são passados para formato unidimensional (concatenação das linhas da matriz, como para o método unidimensional) a partir de um determinado nível da transformada *wavelet* 2D. Este nível é definido como um parâmetro dado no início do processo de classificação para compor as matrizes de treinamento (bloco em destaque na Figura 10).

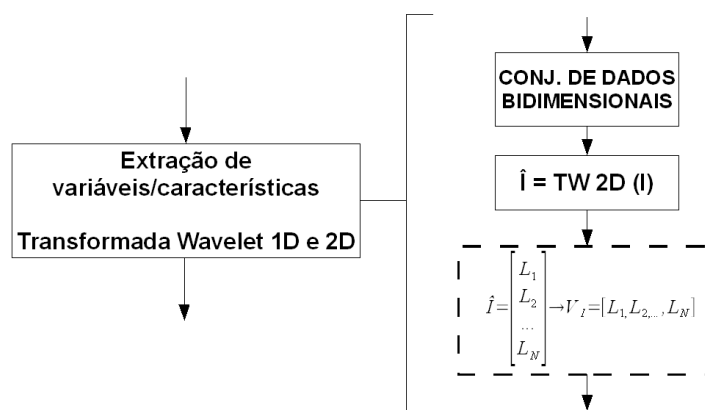


Figura 10: Abordagem 2D para dados bidimensionais.

Para a análise dos dados bidimensionais, neste trabalho é proposto um método híbrido. Nesta abordagem executam-se as transformadas *wavelet* bidimensional e unidimensional em dois processos consecutivos. Inicialmente, a transformada bidimensional é aplicada aos dados bidimensionais até um nível determinado inicialmente (parâmetro definido no início do processo de classificação). A partir da matriz de coeficientes de escala CC, que representam os coeficientes de escala da imagem, as matrizes são passadas para vetores através da concatenação das linhas da matriz. A partir desta nova representação, a transformada *wavelet* unidimensional é executada para a extração de atributos a partir dos coeficientes de escala CC da imagem original. O número de níveis de decomposição para a transformada unidimensional também é definido no início do processo. Os coeficientes de escala resultantes da transformada unidimensional formam as amostras que compõem as matrizes de treinamento. Assim como para as abordagens unidimensional e bidimensional descritas anteriormente, praticamente todos os níveis são avaliados durante a classificação dos dados. A Figura 11 ilustra a seqüência dos procedimentos para a extração de atributos pelo método híbrido. Em destaque na imagem, a etapa de concatenação das linhas da matriz dos coeficientes de escala CC da imagem.

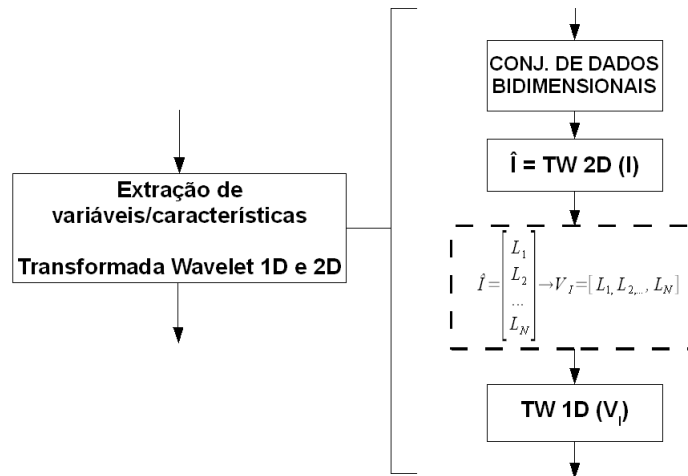


Figura 11: Abordagem híbrida para dados bidimensionais.

O algoritmo de classificação Bayesiano que utiliza-se neste trabalho foi apresentado na Seção 4.1. Na Seção 4.2 apresentou-se aplicação da transformada *wavelet* no pré-processamento dos dados unidimensionais e bidimensionais para realização de classificação de padrões.

Para a avaliação destes métodos propostos, são utilizados duas fontes de dados unidimensionais e uma fonte de dados bidimensionais (coleção de imagens). O método de pré-processamento via transformada *wavelet* 1D é aplicado para os dados unidimensionais. Os dados bidimensionais são tratados no pré-processamento de três maneiras diferentes,

tanto aplicando somente transformada *wavelet* 1D no pré-processamento, quanto também avaliando-se o método de pré-processamento via transformada *wavelet* 2D. Também propõe-se um método híbrido para o pré-processamento dos dados bidimensionais através da aplicação da transformada *wavelet* 2D em uma primeira seleção de variáveis, e partir de um determinado nível, essas novas representações a partir dos coeficientes de escala resultantes da TWD 2D são passadas para vetores (amostras unidimensionais) e aplicada a transformada *wavelet* 1D. No próximo capítulo são apresentados os resultados de experimentos numéricos realizados com estes diferentes conjuntos de dados.

5 *EXPERIMENTOS NUMÉRICOS*

Neste capítulo são apresentados os resultados obtidos através da aplicação do algoritmo de classificação de padrões na análise de dados unidimensionais e bidimensionais. A transformada *wavelet* é aplicada às amostras de treinamento. A partir dos coeficientes de escala resultantes da transformada *wavelet* são compostas as matrizes de treinamento para os algoritmos de classificação. Então é executado o algoritmo de classificação de padrões e é avaliada a relação entre os níveis de transformada e a resposta de classificação dos dados.

Neste trabalho dois pressupostos são considerados para a realização de classificação segundo análise discriminante linear: os dados representam uma distribuição normal multivariada (Mallet et al. 1997, Beltran et al. 2006), e igualdade de matrizes de covariância entre as classes (Duda e Hart 1973).

Na Seção 5.1 são apresentados os resultados para conjunto de dados unidimensionais, separados em dois conjuntos de dados. As matrizes de treinamento têm diferentes quantidades de amostras de treinamento. Os resultados são apresentados nas seguintes subseções:

- a Subseção 5.1.1 apresenta os resultados para conjunto de dados sintéticos e são classificados 3 diferentes classes de dados; e
- a Subseção 5.1.2 apresenta os resultados para conjunto de dados de tráfego de rede. Neste conjunto de dados classificam-se diferentes classes de ataques.

Os resultados para conjunto de dados bidimensionais são apresentados na Seção 5.2. As matrizes de treinamento têm diferentes quantidades para as amostras de treinamento de cada classe. Um conjunto de imagens com rotação são classificadas de acordo com três abordagens para a extração de características dos dados:

- na Subseção 5.2.1 o método unidimensional (TWD 1D) para a extração de características é aplicado aos dados;
- na Subseção 5.2.2 o método bidimensional (TWD 2D) para a extração de características é aplicado aos dados; e
- na Subseção 5.2.3 o método híbrido (TWD 2D e TWD 1D) para a extração de características é aplicado aos dados provenientes de imagens monocromáticas.

Os resultados obtidos com a aplicação do algoritmo de classificação são apresentados como porcentagens de acertos de classificação: a partir de um conjunto de amostras para classificação, cujas classes são conhecidas inicialmente, o acerto acontece quando a probabilidade $P(r_k|x_0)$ de uma amostra x_0 da classe $r = k$ é maior do que as probabilidades estimadas para as outras classes.

Para ambos os conjuntos de dados unidimensionais, foram analisadas duas funções *wavelet* da família de Daubechies: *wavelet* de Haar e *wavelet* de Daubechies *Db2*. Para o conjunto de imagens foram utilizadas 3 funções *wavelet* diferentes da família de Daubechies: *wavelet* de Haar, *wavelet* de Daubechies *Db2*, e *wavelet* de Daubechies *Db8*. A aplicação da *wavelet Db2* e *Db8* tem como finalidade avaliar o comportamento do algoritmo de classificação ao usar uma transformada periódica na decomposição das imagens.

Durante a apresentação dos resultados há a indicação *N/C* para a o resultado de classificação em determinadas condições. Esta indicação significa *Não Classifica* e faz referência quando o algoritmo não consegue convergir nas operações com as matrizes de covariância das classes e realizar os cálculos de probabilidade de objetos de classificação pertencerem a determinada classe. Esta ocorrência deve-se principalmente à quantidade de amostras de treinamento não ser suficiente para que realize-se a classificação em um cenário de amostras com grande dimensionalidade. Em um cenário em que a dimensionalidade das amostras é muito grande, deve-se ter uma quantidade suficiente de amostras em cada classe para que o processo de classificação seja realizado. Essa relação garante o cálculo das matrizes inversas das matrizes de covariância para as classes, e viabiliza o cálculo das probabilidades de classificação.

Todos os algoritmos de transformada *wavelet*, o algoritmo de classificação Bayesiano, e a manipulação, geração e tratamento dos dados unidimensionais e bidimensionais, foram implementados dentro do ambiente do *MATLAB*. É uma ferramenta de alto desempenho voltada para o cálculo numérico, que apresenta ambiente simples e linguagem de programação muito próxima da linguagem algébrica.

5.1 Conjunto de dados Unidimensionais

Os dados unidimensionais são provenientes de duas fontes distintas.

O primeiro conjunto foi extraído do artigo (Mallet et al. 1997) e é formado por amostras sintéticas. As equações geradoras para estes dados são apresentadas no Capítulo 5, Seção 5.1. Dentro deste conjunto de dados tem-se 3 classes diferentes. As amostras tem dimensão de 128 variáveis. Afim de avaliar a influência da relação entre número de amostras de treinamento para cada classe e a dimensão das amostras, o número de amostras por classe sofre variação e também são observados os resultados de classificação de acordo com a dimensão das amostras, de acordo com os níveis de decomposição das amostras via transformada *wavelet*. O algoritmo de classificação é avaliado durante o modo de treinamento, onde aplica-se o método *leave-one-out*.

O segundo conjunto de teste foi construído com informações de tráfego de dados em redes de computadores em que há ocorrências de anomalias e ataques que compõem a base DARPA, que é uma base com dados relativos a tráfego de rede de computadores. Nesta base estão identificados alguns ataques, e outros não estão identificados. Assim como para o primeiro conjunto, o algoritmo de classificação é avaliado no modo de treinamento através do método *leave-one-out*.

5.1.1 Conjunto de dados simulados

Neste conjunto de dados simulados estão definidas $R = 3$ classes, contendo curvas com $p = 128$ variáveis. Para a avaliação do algoritmo são realizados diferentes cenários para teste, e cada classe contém $n_r = 5$, $n_r = 10$, $n_r = 15$, $n_r = 25$, $n_r = 50$ e $n_r = 100$ amostras de cada classe, para cada cenário. Os conjuntos de dados simulados são gerados a partir das seguintes equações (Mallet et al. 1997):

$$\begin{aligned} x_j^{(1)} &= (6 + \eta) \cdot \chi_{[a,b]}(j) + \epsilon(j), & \text{Classe 1} \\ x_j^{(2)} &= (6 + \eta) \cdot \chi_{[a,b]}(j)(j - a)/(b - a) + \epsilon(j), & \text{Classe 2} \\ x_j^{(3)} &= (6 + \eta) \cdot \chi_{[a,b]}(j)(b - j)/(b - a) + \epsilon(j), & \text{Classe 3} \end{aligned}$$

sendo que:

- $j = 1, \dots, 128$ representa o número de pontos para o vetor

- $\eta = \epsilon(j) \sim N(0, 1)$, número real randômico entre 0 e 1, de acordo com distribuição normal
- $\chi_{[a,b]}(j) = 1$ se j está no intervalo $[a, b]$, senão $\chi(j) = 0$
- $a \sim U_Z(16, 32)$, número inteiro randômico entre 16 e 32
- $(b - a) \sim U_Z(32, 96)$, número inteiro randômico definido pela diferença $b - a$

A Figura 12 apresenta uma amostra de cada classe para este primeiro conjunto de testes.

A transformada *wavelet* foi executada até o 5^o nível, e neste nível de decomposição tem-se representação de cada amostra com 4 pontos. Desta forma, para o algoritmo de classificação foram compostas diferentes matrizes de treinamento com os coeficientes de escala de cada amostra. As Figuras 13, 14, 15 apresentam o resultado da decomposição *wavelet* de uma amostra de cada classe até 3^o nível, aplicando os filtros de Daubechies *Db2*. Observa-se que através dos níveis de decomposição, o resultado da transformada mantém grande parte de informação do sinal original em uma representação média que utiliza um número menor de pontos.

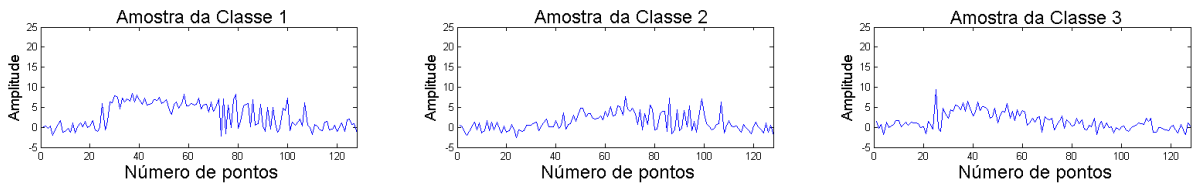


Figura 12: Amostras de classes simuladas.

As Tabelas 2 e 3 apresentam os resultados de classificação para este primeiro conjunto de dados, aplicando transformada *wavelet* de Haar e Daubechies *Db2*.

Tabela 2: Resultados de classificação dos dados simulados, com *wavelet* de Haar.

Classificação de dados simulados - Wavelet de Haar									
NÍVEL	$n_r = 5$ amostras			$n_r = 10$ amostras			$n_r = 15$ amostras		
	Cl. 1	Cl. 2	Cl. 3	Cl. 1	Cl. 2	Cl. 3	Cl. 1	Cl. 2	Cl. 3
1 ($p = 64$)	N/C	N/C	N/C	N/C	N/C	N/C	N/C	N/C	N/C
2 ($p = 32$)	N/C	N/C	N/C	N/C	N/C	N/C	100%	0%	46,67%
3 ($p = 16$)	N/C	N/C	N/C	100%	10%	50%	100%	53,33%	60%
4 ($p = 8$)	100%	20%	20%	100%	60%	70%	100%	73,33%	46,67%
5 ($p = 4$)	100%	80%	40%	100%	90%	0%	100%	93,33%	26,67%
NÍVEL	$n_r = 25$ amostras			$n_r = 50$ amostras			$n_r = 100$ amostras		
	Cl. 1	Cl. 2	Cl. 3	Cl. 1	Cl. 2	Cl. 3	Cl. 1	Cl. 2	Cl. 3
1 ($p = 64$)	80%	0%	52%	100%	8%	48%	100%	89%	49%
2 ($p = 32$)	100%	8%	60%	100%	88%	42%	100%	100%	45%
3 ($p = 16$)	100%	84%	44%	100%	98%	42%	100%	100%	42%
4 ($p = 8$)	100%	100%	48%	100%	100%	44%	100%	100%	45%
5 ($p = 4$)	100%	100%	40%	100%	100%	20%	100%	100%	40%

Para a execução da etapa de classificação com $n_r = 5$ amostras para cada classe, nos 1^o, 2^o, 3^o e 4^o níveis de decomposição, não houve convergência nos cálculos, como

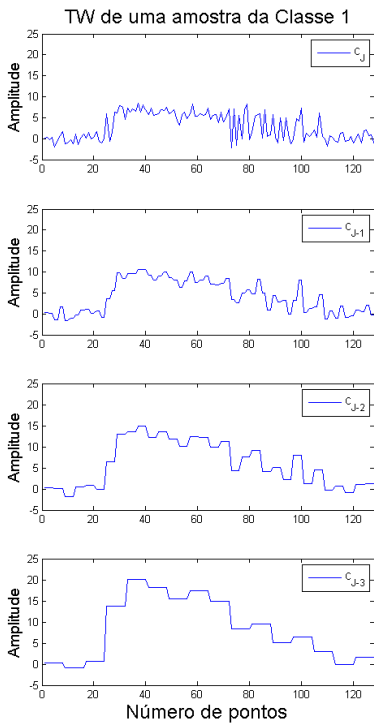


Figura 13: TWD de amostra da classe 1.

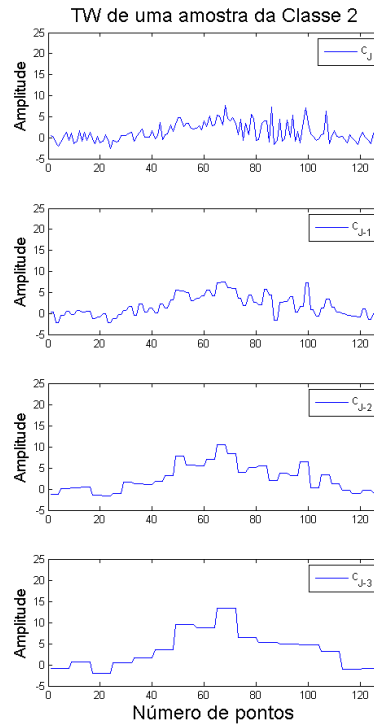


Figura 14: TWD de amostra da classe 2.

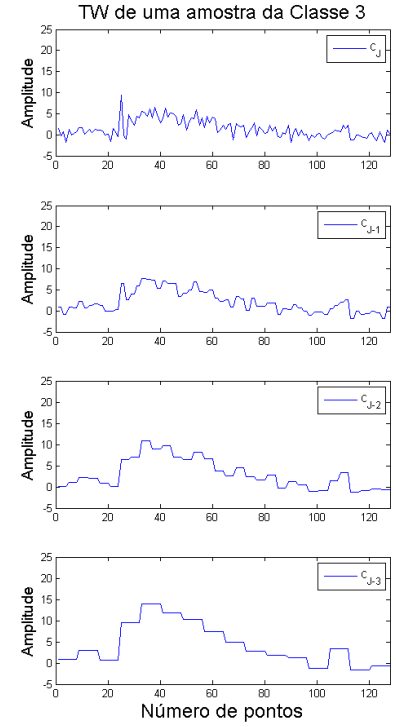


Figura 15: TWD de amostra da classe 3.

Tabela 3: Resultados de classificação dos dados simulados, com *wavelet* de Daubechies *Db2*.

Conjunto de dados simulados - <i>Wavelet</i> Daubechies <i>Db2</i>									
NÍVEL	$n_r = 5$ amostras			$n_r = 10$ amostras			$n_r = 15$ amostras		
	Cl. 1	Cl. 2	Cl. 3	Cl. 1	Cl. 2	Cl. 3	Cl. 1	Cl. 2	Cl. 3
1 ($p = 64$)	N/C	N/C	N/C	N/C	N/C	N/C	N/C	N/C	N/C
2 ($p = 32$)	N/C	N/C	N/C	N/C	N/C	N/C	100%	0%	26,67%
3 ($p = 16$)	N/C	N/C	N/C	100%	0%	50%	100%	60%	40%
4 ($p = 8$)	100%	0%	40%	100%	70%	80%	100%	80%	46,67%
5 ($p = 4$)	100%	60%	20%	100%	100%	60%	100%	86,67%	53,33%
NÍVEL	$n_r = 25$ amostras			$n_r = 50$ amostras			$n_r = 100$ amostras		
	Cl. 1	Cl. 2	Cl. 3	Cl. 1	Cl. 2	Cl. 3	Cl. 1	Cl. 2	Cl. 3
1 ($p = 64$)	88%	0%	64%	100%	16%	52%	100%	91%	46%
2 ($p = 32$)	100%	28%	64%	100%	84%	54%	100%	100%	46%
3 ($p = 16$)	100%	88%	60%	100%	100%	46%	100%	100%	50%
4 ($p = 8$)	100%	100%	36%	100%	100%	44%	100%	100%	46%
5 ($p = 4$)	100%	100%	40%	100%	100%	44%	100%	100%	44%

indicado nas Tabelas 2 e 3. A etapa de cálculo de probabilidade de classificação leva em consideração o cálculo de matrizes inversas para as matrizes de covariância das classes, e nesta etapa o algoritmo não consegue determinar as probabilidades de classificação. Apenas no 5º nível houve sucesso nos cálculos. Este fato tem relação direta ao número de amostras por classe e o número de variáveis para cada amostra, pois deve-se ter um número de amostras suficiente para compor a matriz de treinamento.

O mesmo acontece quando se executa o algoritmo com $n_r = 10$ e $n_r = 15$ amostras para cada classe, em que se obtém sucesso nos cálculos a partir dos 2º e 3º níveis, respectivamente. A partir da execução do algoritmo com mais de $n_r = 25$ amostras para cada classe obtém-se convergência nos cálculos em todos os níveis de decomposição na transformada *wavelet*.

Para a classe $r = 1$, o classificador sempre obteve alto grau de acerto na classificação, com diferentes quantidades de amostras e em diferentes níveis de decomposição. Na classificação dos elementos da classe $r = 2$, o grau de acerto cresceu de acordo com o aumento do número de amostras por classe, e também em diferentes níveis de representação. A classe $r = 3$ apresentou sempre os menores valores de acerto, devido aos dados desta classe apresentarem alto grau de correlação com as outras e podendo as amostras nesta classe não ficarem bem definidas após a transformada *wavelet*.

Em geral, nota-se o mesmo comportamento no aumento no número de acertos de classificação quando aumenta-se o número de amostras para cada classe, e através dos níveis de decomposição da transformada. Este comportamento vale para as diferentes *wavelets* aplicadas na extração de características dos dados.

Os resultados apresentados nesta subseção servem como validação para a implementação do algoritmo de classificação. Comparativamente aos resultados apresentados em (Mallet et al. 1997) para o conjunto de dados sintéticos, os resultados apresentados nesta subseção apresentam um comportamento semelhante mesmo executando um algoritmo simplificado em relação ao algoritmo de classificação apresentado em (Mallet et al. 1997). Este primeiro conjunto de dados serviu principalmente para avaliar a implementação do algoritmo de classificação e as diferentes *wavelets*.

5.1.2 Conjunto de dados de anomalias da base DARPA 1999

O segundo conjunto de dados unidimensionais é composto por dados de tráfego de redes de computadores, provenientes da base de dados da DARPA (MIT 1999). O grupo IST (*Information Systems Technology Group*) do *Lincoln Laboratory* do MIT, financiado pela DARPA e o Laboratório de Pesquisa da Força Aérea dos Estados Unidos (*Air Force Research Laboratory - ARFL*), tem coletado e distribuído documento padrão para avaliação de sistemas de detecção de intrusão IDS - (*Intrusion Detection System*) em redes de computadores. Cada conjunto de avaliações foi desenvolvido nos anos de 1998 e 1999. A partir da base DARPA 1999 é possível analisar o fluxo de dados em diferentes protocolos de comunicação (ICMP, UDP, TCP e IP) e analisar a ocorrência das anomalias em diferentes períodos. Mais detalhes sobre estes protocolos disponíveis para análise são dados em (Tanenbaum 2002).

Dentro dessas informações de tráfego de dados em rede, existem registros de eventos que são classificados como anomalias ou ataques maliciosos. Inicialmente são $R = 2$ classes de anomalias de curta duração e $R = 2$ classes de anomalias de longa duração.

São definidos dois cenários para a aplicação do algoritmo: o primeiro cenário é composto por duas classes de anomalias que apresentam característica de período de curta duração, e as curvas de observação têm $p = 128$ variáveis; o segundo cenário é composto por duas classes de anomalias que apresentam período de longa duração e as curvas de observação contém $p = 1024$ variáveis.

A primeira dificuldade ao trabalhar com esse conjunto de dados está em conseguir reunir um conjunto mínimo de amostras para compor a matriz de dados de treinamento. Através da busca dentro dos dados identificados na base DARPA, reuniram-se 3 amostras de cada anomalia e as outras amostras foram geradas/simuladas para construir as matrizes de dados de treinamento com $n_r = 5$, $n_r = 10$, $n_r = 15$, $n_r = 20$ e $n_r = 25$ amostras para cada classe. Este procedimento foi realizado a fim de viabilizar a execução do procedimento de classificação destes dados.

São dois cenários para classificação dos ataques contidos na base DARPA: o primeiro cenário é composto por ataques de curta duração (representados com até $p = 128$ variáveis) chamados de *satan* e *pod*, com $n_r = 5$, $n_r = 10$, $n_r = 15$, $n_r = 25$ amostras em cada execução do algoritmo; o segundo cenário é composto por ataques de longa duração (representados com 1024 variáveis) chamados de *neptune* e *mailbomb*, com $n_r = 5$, $n_r = 10$, $n_r = 20$, $n_r = 25$ amostras em cada execução do algoritmo. A Figura 16 apresenta amostras de cada classe de anomalias de curta e longa duração. As Figuras 17, 18, 19 e 20 apresentam os resultados de decomposição através da transformada *wavelet* de uma amostra de cada classe. Observa-se que os ataques de curta duração têm como principal característica um grande aumento na amplitude durante um curto espaço de tempo, e essa característica é preservada na decomposição *wavelet*. Já os ataques de longa duração têm como característica um aumento na amplitude, não tão grande relativamente ao tráfego anterior ou posterior ao ataque, durante um período de tempo maior em relação aos ataques curtos. Esta característica fica mais clara na decomposição *wavelet* pois o valores médios de tráfego de dados tanto durante o tráfego normal quanto do tráfego sob ataque ficam mais evidentes.

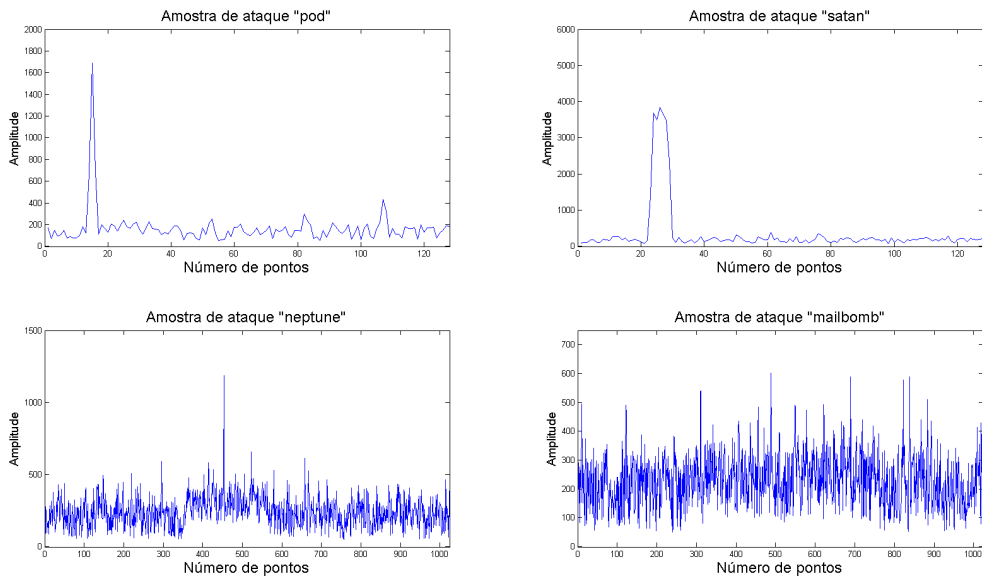


Figura 16: Amostra de cada anomalia, das diferentes classes.

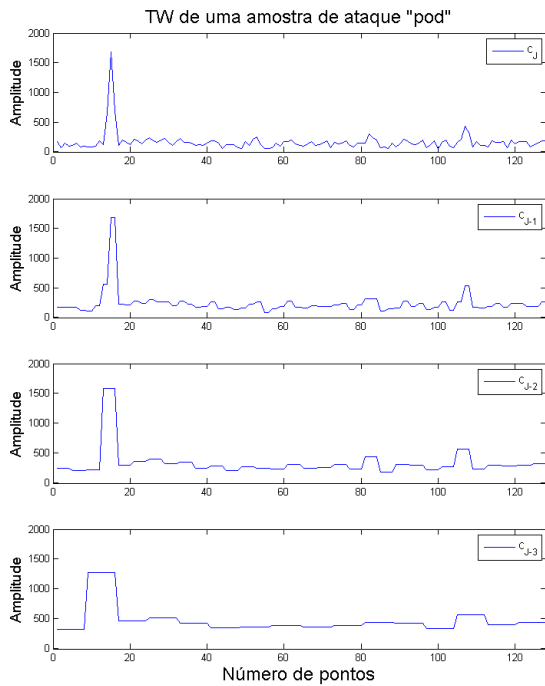


Figura 17: TWD de amostra da classe de ataques curtos, chamado *pod*.

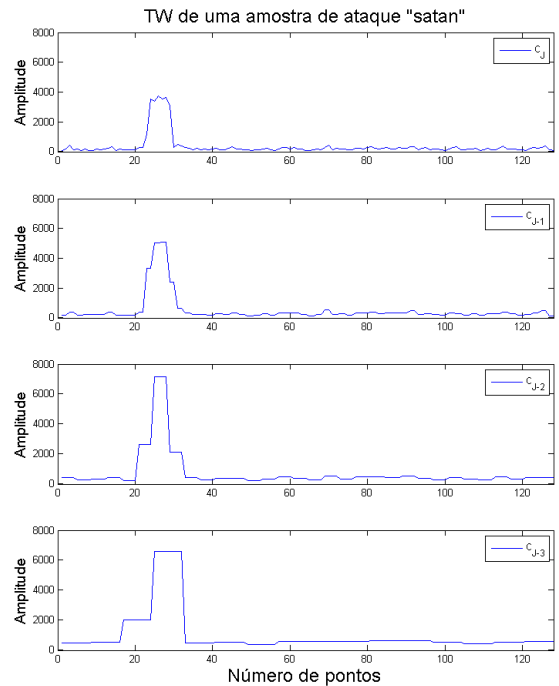


Figura 18: TWD de amostra da classe de ataques curtos, chamado *satan*.

As Tabelas 4 e 5 apresentam os resultados de classificação para o primeiro cenário de avaliação do algoritmo de classificação, avaliando diferentes famílias de *wavelets*.

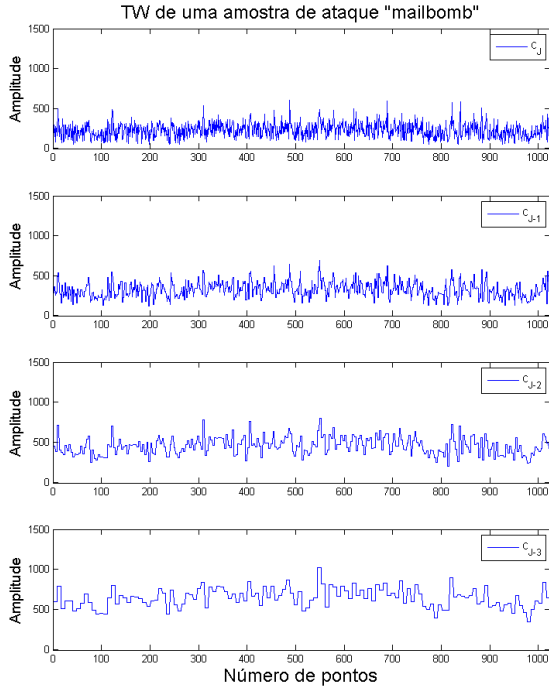


Figura 19: TWD de amostra da classe de ataques longos, chamado *mailbomb*.

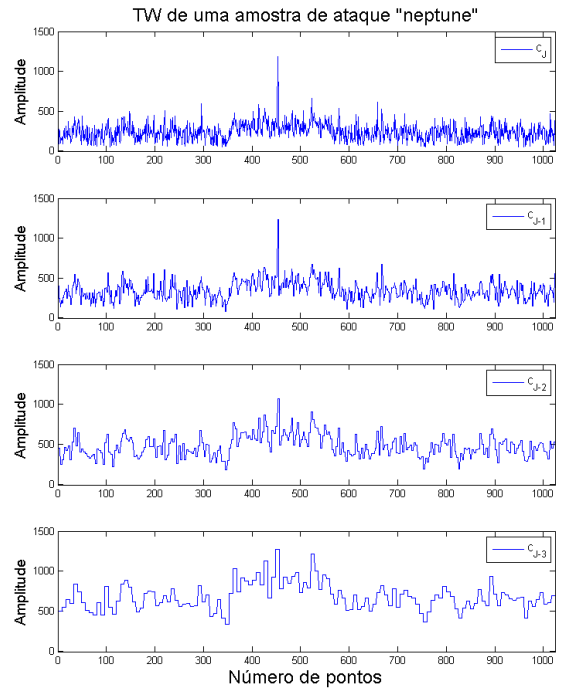


Figura 20: TWD de amostra da classe de ataques longos, chamado *neptune*.

Tabela 4: Resultados de classificação de ataques curtos, com *wavelet* de Haar.

Ataques curtos simulados - <i>Wavelet</i> de Haar								
NÍVEL	$n_r = 5$ amostras		$n_r = 10$ amostras		$n_r = 15$ amostras		$n_r = 25$ amostras	
	Cl. 1	Cl. 2	Cl. 1	Cl. 2	Cl. 1	Cl. 2	Cl. 1	Cl. 2
1 ($p = 64$)	N/C	N/C	N/C	N/C	N/C	N/C	N/C	N/C
2 ($p = 32$)	N/C	N/C	N/C	N/C	N/C	N/C	100%	100%
3 ($p = 16$)	N/C	N/C	50%	40%	66,67%	53,33%	100%	100%
4 ($p = 8$)	N/C	N/C	90%	60%	66,67%	60%	100%	100%

Tabela 5: Resultados de classificação de ataques curtos, com *wavelet* de Daubechies *Db2*.

Ataques curtos simulados - <i>Wavelet</i> Daubechies <i>Db2</i>								
NÍVEL	$n_r = 5$ amostras		$n_r = 10$ amostras		$n_r = 15$ amostras		$n_r = 25$ amostras	
	Cl. 1	Cl. 2	Cl. 1	Cl. 2	Cl. 1	Cl. 2	Cl. 1	Cl. 2
1 ($p = 64$)	N/C	N/C	N/C	N/C	N/C	N/C	N/C	N/C
2 ($p = 32$)	N/C	N/C	N/C	N/C	N/C	N/C	100%	100%
3 ($p = 16$)	N/C	N/C	50%	40%	40%	46,67%	100%	100%
4 ($p = 8$)	N/C	N/C	40%	50%	66,67%	46,67%	100%	100%

No primeiro cenário, obtiveram-se os primeiros resultados a partir de $n_r = 10$ e $n_r = 15$ amostras para cada classe e partir dos 3^o e 4^o níveis de decomposição para ambas famílias *wavelet* analisadas, conforme pode-se ver nas Tabelas 4 e 5. A partir de $n_r = 25$ amostras para cada classe obteve-se resultados a partir do 2^o nível de decomposição. O grau de acerto de classificação chegou aos melhores valores quando se tem $n_r = 25$ amostras por classe. Para a *wavelet* de Haar, obtiveram-se os melhores resultados de classificação e nota-se o aumento no grau de classificação quando se analisa os níveis mais grosseiros de decomposição.

As Tabelas 6 e 7 apresentam os resultados de classificação para o segundo cenário de avaliação do algoritmo de classificação, analisando diferentes famílias de *wavelet*.

Tabela 6: Resultados de classificação de ataques longos, com *wavelet* de Haar.

Ataques longos simulados - <i>Wavelet</i> de Haar								
NÍVEL	$n_r = 5$ amostras		$n_r = 10$ amostras		$n_r = 20$ amostras		$n_r = 25$ amostras	
	Cl. 1	Cl. 2	Cl. 1	Cl. 2	Cl. 1	Cl. 2	Cl. 1	Cl. 2
1 ($p = 512$)	N/C	N/C	N/C	N/C	N/C	N/C	N/C	N/C
2 ($p = 256$)	N/C	N/C	N/C	N/C	N/C	N/C	N/C	N/C
3 ($p = 128$)	N/C	N/C	N/C	N/C	N/C	N/C	N/C	N/C
4 ($p = 64$)	N/C	N/C	N/C	N/C	N/C	N/C	N/C	N/C
5 ($p = 32$)	N/C	N/C	N/C	N/C	100%	100%	100%	100%

Tabela 7: Resultados de classificação de ataques longos, com *wavelet* de Daubechies *Db2*.

Ataques longos simulados - <i>Wavelet</i> Daubechies <i>Db2</i>								
NÍVEL	$n_r = 5$ amostras		$n_r = 10$ amostras		$n_r = 20$ amostras		$n_r = 25$ amostras	
	Cl. 1	Cl. 2	Cl. 1	Cl. 2	Cl. 1	Cl. 2	Cl. 1	Cl. 2
1 ($p = 512$)	N/C	N/C	N/C	N/C	N/C	N/C	N/C	N/C
2 ($p = 256$)	N/C	N/C	N/C	N/C	N/C	N/C	N/C	N/C
3 ($p = 128$)	N/C	N/C	N/C	N/C	N/C	N/C	N/C	N/C
4 ($p = 64$)	N/C	N/C	N/C	N/C	N/C	N/C	N/C	N/C
5 ($p = 32$)	N/C	N/C	N/C	N/C	90%	100%	100%	100%

No segundo cenário, obteve-se resultado a partir de $n_r = 20$ e $n_r = 25$ amostras para cada classe e a partir 4º nível de decomposição, obtendo valores máximos de acerto de classificação (Tabelas 6 e 7). As classes são bem definidas, pois embora sejam representados com $p = 1024$ variáveis, os ataques tem duração diferente dentro da “janela” de observação. Com matrizes de treinamento com menos de 25 amostras para cada classe não foi possível executar o algoritmo para os primeiros níveis de decomposição, mostrando que a transformada *wavelet* contribuiu para a redução dos dados, e permitindo a execução da classificação.

Inicialmente, quando a base DARPA foi escolhida para validação do algoritmo de classificação, não se tinha a informação precisa de quantas amostras para cada tipo de ataque estavam disponíveis no banco de dados. Infelizmente, o número extremamente limitado de exemplares para cada tipo de ataque catalogado prejudicou o desempenho do algoritmo de classificação. Esta limitação de amostras para composição das matrizes de treinamento prejudicou a execução do algoritmo e a avaliação do potencial de classificação. Foi necessária a geração de dados sintéticos, com as mesmas características dos ataques identificados, para aumentar o número de amostras de treinamento. Foram geradas novas amostras para as diferentes classes de ataques para viabilizar a execução da classificação.

5.2 Conjunto de dados bidimensionais - Conjunto de imagens

Os dados bidimensionais são obtidos de conjunto de imagens de cartas de baralho, disponibilizados pelos professores André Moura e José Manuel Fonseca, do Departamento de Ciência da Computação e Eletrônica da Universidade Nova de Lisboa. São imagens coloridas capturadas por máquina digital, com dimensão de 12 megapixels. Algumas imagens tem presença de *flash*, enquanto outras não tem a presença de *flash*. A partir deste conjunto de imagens, são definidas 4 classes de imagens baseadas nos 4 diferentes naipes de cartas de baralho.

Os dados bidimensionais são formados por imagens de naipes de cartas de baralho. Os registros dos naipes são originados de imagens coloridas de cartas de baralho, em que parte destas imagens há presença de flash e outras imagens não têm presença de flash. Foram formadas $r = 4$ classes (4 naipes de cartas). Para avaliar o conjunto de dados bidimensionais, são aplicados os métodos de extração de características através de transformada *wavelet* discreta unidimensional (TWD 1D), extração de características através de *wavelet* bidimensional (TWD 2D), e extração de características através de método híbrido (TWD 2D e TWD 1D).

As imagens apresentam as cartas com recorte de apenas 25% da carta, em que aparecem a informação do naipe e número da carta e o fundo preto. A Figura 21 apresenta uma amostra destas imagens. As imagens coloridas têm dimensão de 12 megapixels.

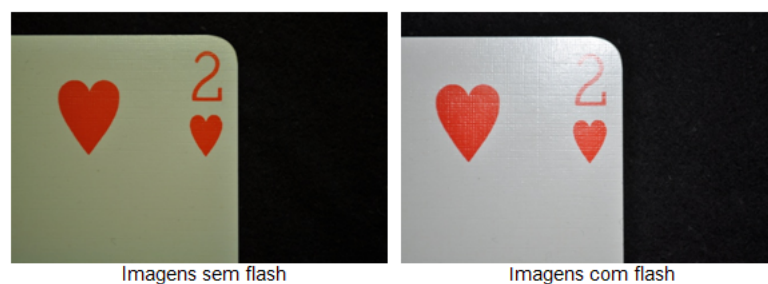


Figura 21: Amostras que compõem o conjunto de imagens de cartas de baralho.

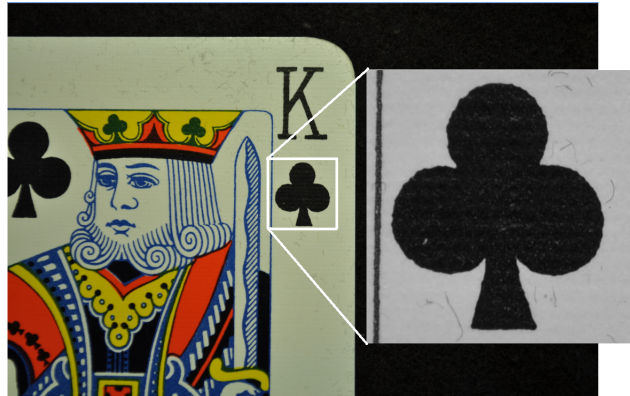


Figura 22: Obtenção de imagem do naipe a partir das imagens originais: recorte da área do naipe, gerando nova imagem com dimensão de 512 x 512 pixels.

A partir do conjunto de imagens originais, são formadas 4 classes de imagens a partir do recorte da imagem do naipe da carta em cada imagem original (como apresentado na Figura 22), obtendo novas imagens com dimensão de 512 x 512 pixels. Algumas amostras destas imagens que contém apenas os naipes são apresentadas nas Figuras 23, 24, 25 e 26.



Figura 23: Naipe de copas.



Figura 24: Naipe de ouros.



Figura 25: Naipe de paus.

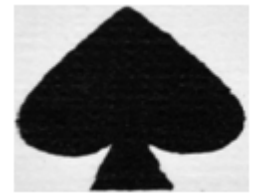


Figura 26: Naipe de espadas.

As 4 classes são: copas, ouros, paus e espadas. Obtém-se então 13 amostras para cada classe, pois são 13 cartas/números (2,3,4,5,6,7,8,9,10,J,Q,K,A) para cada naipe, de acordo com a presença ou não de *flash*. Então, somando-se todas as amostras de cada naipe de carta, cada classe tem 26 amostras para cada imagem de naipe. Todas as imagens que compõem as classes estão em escala de cinza.

Para avaliação do método de classificação aplicando *wavelets*, foram elaborados 2 cenários que diferem no número de amostras nas matrizes de treinamento. A avaliação do método nesses dois cenários tem como finalidade observar a relação entre número de amostras de treinamento para as classes e a qualidade de classificação do algoritmo, e nesse sentido observar a contribuição da transformada *wavelet*. O primeiro cenário tem $R = 4$ classes com $n_r = 26$ amostras de treinamento em cada. As 26 amostras de cada classe não tem algum tipo de variabilidade: são imagens sem rotação ou deformidades. O segundo cenário tem $n_r = 130$ amostras de treinamento para cada uma das $R = 4$ classes.

No segundo cenário, as amostras de treinamento das classes tem, além das imagens dos naipes em forma padrão, imagens dos naipes rotacionados tanto para o sentido horário quanto para o sentido anti-horário. Os ângulos de rotação aplicados foram $\pm 5^\circ$ e $\pm 10^\circ$.

5.2.1 Método de classificação via transformada *wavelet* 1D

O método de classificação Bayesiano utilizado neste trabalho tem as matrizes treinamento e amostra de entrada representada por vetores. Para aplicar esse método à classificação de imagens, deve-se, primeiramente, transformar as informações das imagens em vetores. Então, as imagens são convertidas em vetores que são formados pela concatenação das linhas das imagens, por exemplo: $imagem1_{vet} = [L_1, L_2, L_3, \dots, L_{512}]$. Realiza-se este procedimento para todas as amostras das 4 classes (4 naipes). Ao final do processo de conversão das imagens em vetores, tem-se que cada amostra é representada por 262144 pontos.

Os cenários de classificação desenvolvidos para a avaliação do algoritmo de classificação têm em comum que a relação de número de amostras de treinamento e a dimensão das amostras não permite a classificação das imagens. A aplicação da TWD 1D aos vetores das imagens tem como objetivo realizar uma extração de características das imagens afim de reduzir o número de variáveis que compõem as amostras das matrizes de treinamento. A TWD 1D é aplicada a matriz de treinamento (a cada amostra de cada classe) e também é aplicada ao objeto a ser classificado. A cada nível de decomposição da TWD é avaliada a relação entre número de variáveis extraídas e a eficiência da classificação.

As Figuras 27 e 28 apresentam as representações vetoriais das amostras das 4 classes de naipes, tanto para imagens com *flash* quanto as imagens sem *flash*. Elas apresentam as curvas para as 4 classes, com dimensão de 262144 pontos. Após a transformada *wavelet*, percebe-se que as amostras de cada classe seguem um padrão. As Figuras 29 e 30 apresentam as amostras com 1024 pontos (8^o nível de transformação). Observando as Figuras 29 e 30, as classes de imagens sem *flash* mantêm um padrão das amostras dentro de cada classe. As amostras das imagens com *flash* também seguem um padrão, mas, na média, semelhante ao padrão das imagens sem *flash*. O ruído que observa-se na representação das amostras das imagens com *flash* é decorrente da presença de *flash*. A presença de uma fonte de luz (*flash*) na imagem marca uma diferença maior entre os pixels claros (brancos) da imagem e os pixels escuros (pretos) que formam a imagem do naipe na imagem. As imagens sem presença de *flash* têm os pixels claros e escuros com valores não tão desequilibrados, e na decomposição *wavelet* têm seus valores mais equilibrados.

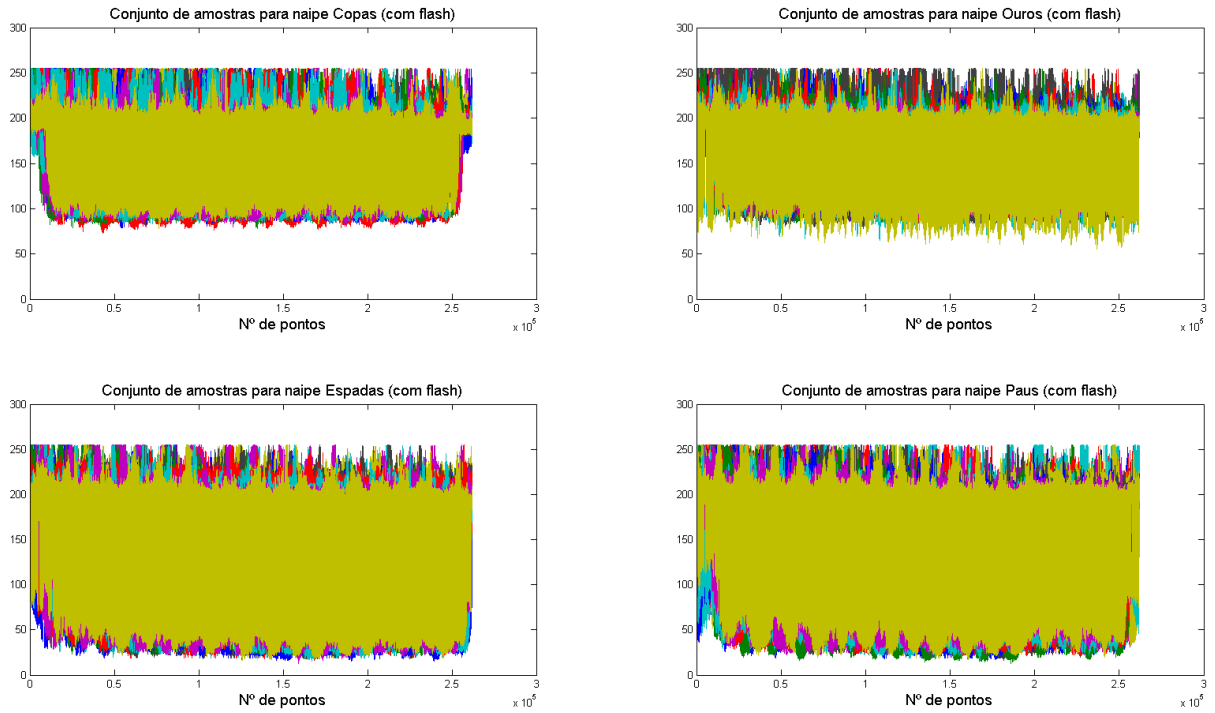


Figura 27: Representação das amostras de imagens com *flash* para as 4 classes de naipes de cartas: (a) classe Copas, (b) classe Ouros, (c) classe Espadas, (d) classes Paus.

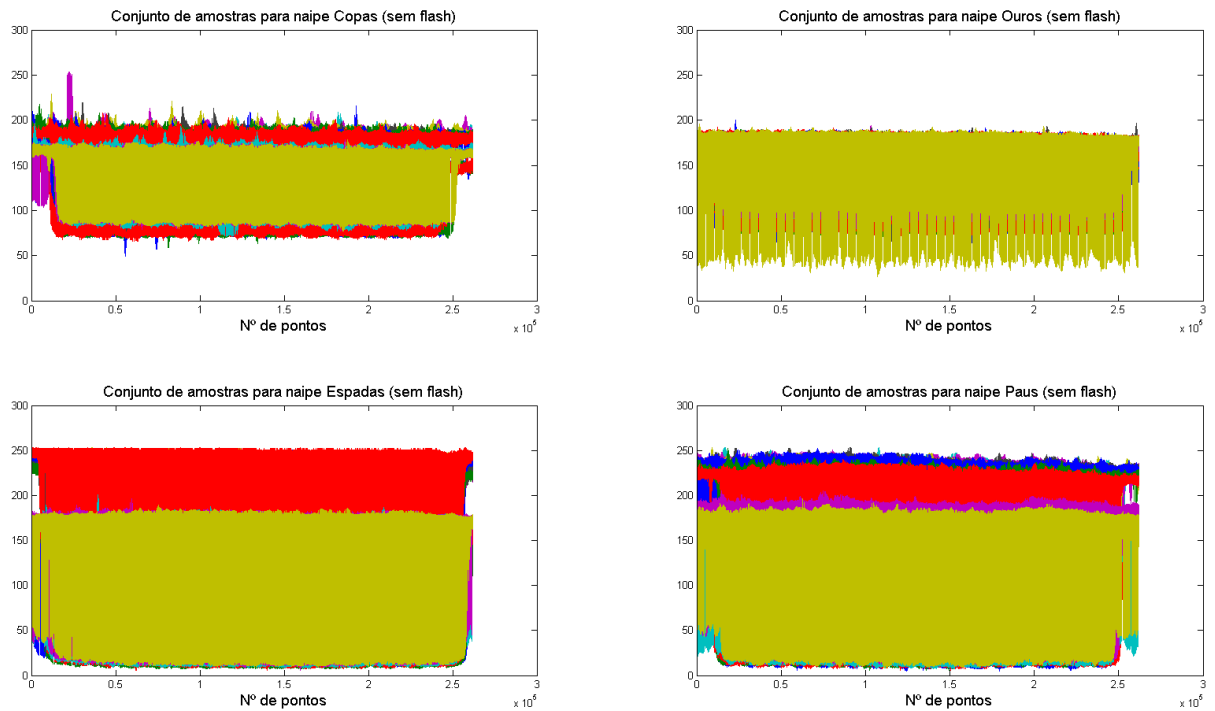


Figura 28: Representação das amostras de imagens sem *flash* para as 4 classes de naipes de cartas: (a) classe Copas, (b) classe Ouros, (c) classe Espadas, (d) classes Paus.

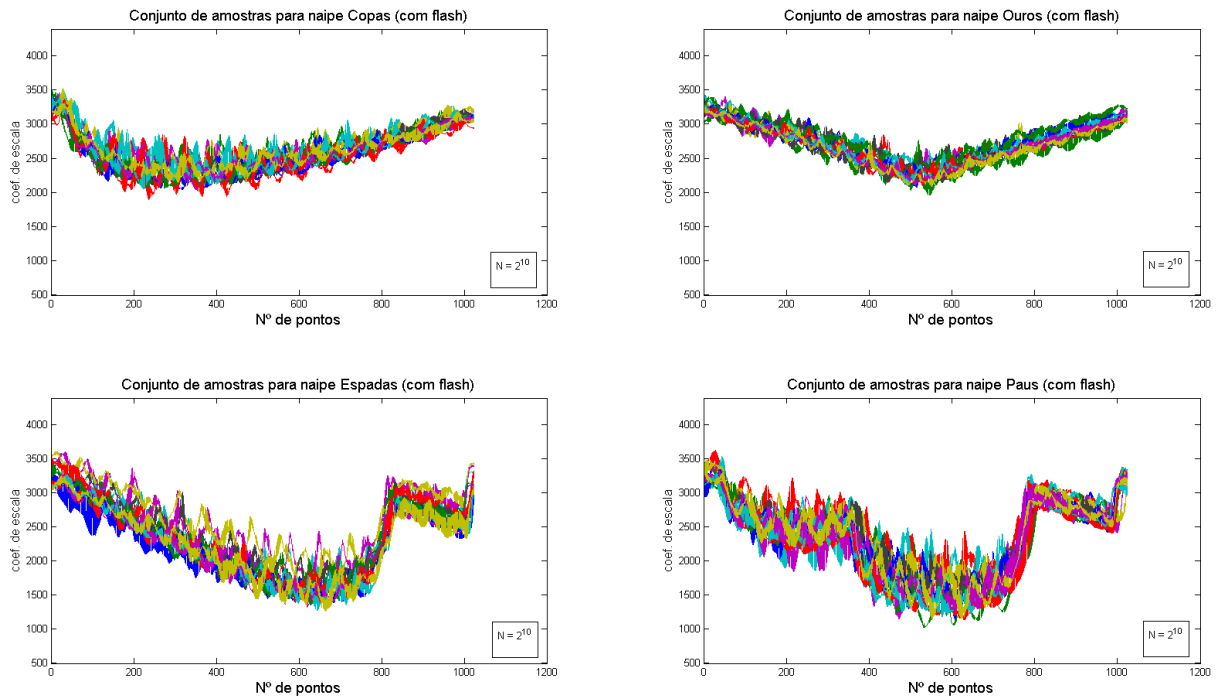


Figura 29: Representação das 4 classes de naipes de cartas no 8^o nível de TWD, a partir de imagens com *flash*.

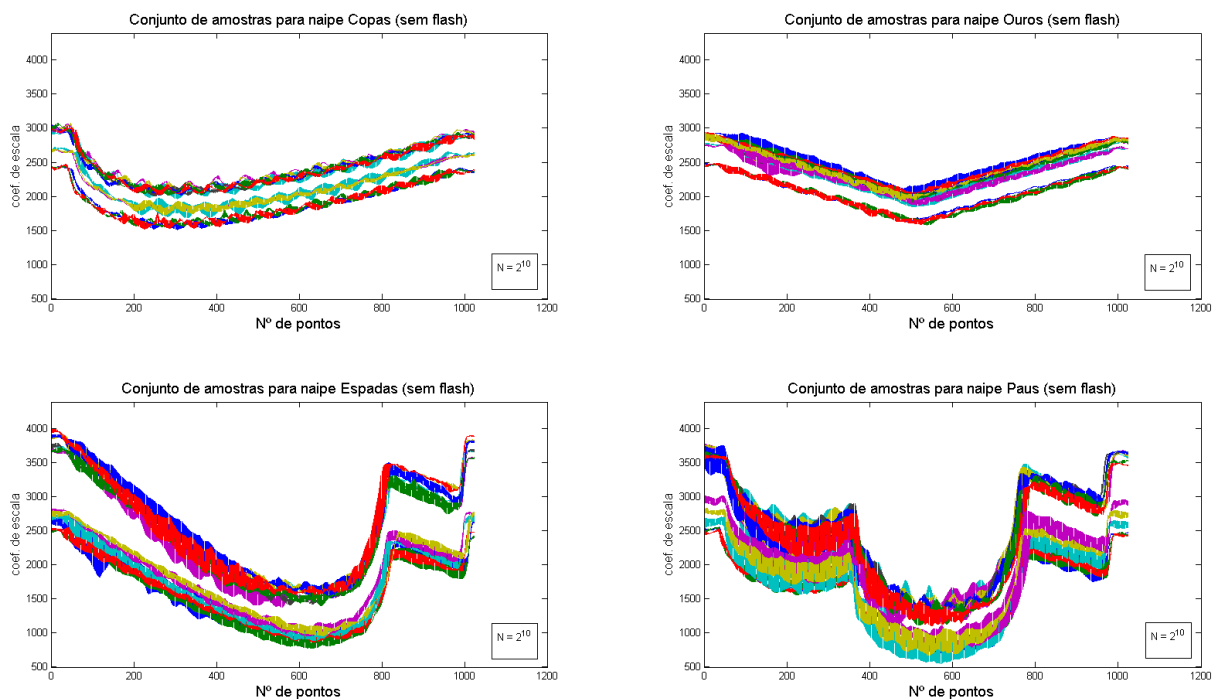


Figura 30: Representação das 4 classes de naipes de cartas no 8^o nível de TWD, a partir de imagens sem *flash*.

As amostras representadas nas Figuras 29 e 30 são referentes ao 8^o nível de decomposição e têm dimensão $p = 1024$. Os resultados apresentados nas tabelas seguintes são decorrentes da execução do algoritmo de classificação para matrizes de treinamento for-

madras com amostras inicialmente com $p = 512$ variáveis, ou seja, amostras resultantes do 9º nível de decomposição da transformada *wavelet*. Afim de avaliar a qualidade destas amostras para compor a matriz de treinamento, o algoritmo de classificação é executado primeiramente em modo *leave-one-out*, no qual cada amostra é retirada da matriz de treinamento de sua classe e deve ser classificada em sua respectiva classe. Se as amostras forem classificadas corretamente de acordo com suas classes, elas representam muito bem a classe. Os resultados obtidos durante a execução da classificação em modo *leave-one-out* foram sempre máximos. Ou seja, todas as amostras foram classificadas corretamente nas suas classes de origem.

Tabela 8: Resultados de classificação de imagens de naipes de cartas. Imagens com rotação. Método unidimensional, *wavelet* de Haar.

Classificação de Naipes - Wavelet de Haar.									
ROTAÇÃO 0°					ROTAÇÃO 10°				
NÍVEL	$n_r = 26$ amostras				NÍVEL	$n_r = 26$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
9 ($p = 512$)	N/C	N/C	N/C	N/C	9 ($p = 512$)	N/C	N/C	N/C	N/C
10 ($p = 256$)	N/C	N/C	N/C	N/C	10 ($p = 256$)	N/C	N/C	N/C	N/C
11 ($p = 128$)	100,00%	100,00%	100,00%	100,00%	11 ($p = 128$)	N/C	N/C	N/C	N/C
12 ($p = 64$)	100,00%	100,00%	100,00%	100,00%	12 ($p = 64$)	N/C	N/C	N/C	N/C
13 ($p = 32$)	100,00%	100,00%	100,00%	100,00%	13 ($p = 32$)	26,92%	96,15%	N/C	N/C
14 ($p = 16$)	100,00%	100,00%	100,00%	100,00%	14 ($p = 16$)	0,00%	42,31%	88,46%	0,00%
15 ($p = 8$)	100,00%	100,00%	100,00%	100,00%	15 ($p = 8$)	0,00%	100,00%	0,00%	0,00%
ROTAÇÃO 15°					ROTAÇÃO 20°				
NÍVEL	$n_r = 26$ amostras				NÍVEL	$n_r = 26$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
9 ($p = 512$)	N/C	N/C	N/C	N/C	9 ($p = 512$)	N/C	N/C	N/C	N/C
10 ($p = 256$)	N/C	N/C	N/C	N/C	10 ($p = 256$)	N/C	N/C	N/C	N/C
11 ($p = 128$)	N/C	N/C	N/C	N/C	11 ($p = 128$)	N/C	N/C	N/C	N/C
12 ($p = 64$)	N/C	N/C	N/C	N/C	12 ($p = 64$)	N/C	N/C	N/C	N/C
13 ($p = 32$)	3,85%	100,00%	N/C	N/C	13 ($p = 32$)	0,00%	100,00%	88,46%	0,00%
14 ($p = 16$)	0,00%	96,15%	88,46%	0,00%	14 ($p = 16$)	0,00%	80,77%	0,00%	0,00%
15 ($p = 8$)	0,00%	100,00%	0,00%	0,00%	15 ($p = 8$)	0,00%	100,00%	19,23%	0,00%
ROTAÇÃO 30°					ROTAÇÃO 45°				
NÍVEL	$n_r = 26$ amostras				NÍVEL	$n_r = 26$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
9 ($p = 512$)	N/C	N/C	N/C	N/C	9 ($p = 512$)	N/C	N/C	N/C	N/C
10 ($p = 256$)	N/C	N/C	N/C	N/C	10 ($p = 256$)	N/C	N/C	N/C	N/C
11 ($p = 128$)	N/C	N/C	N/C	N/C	11 ($p = 128$)	N/C	N/C	N/C	N/C
12 ($p = 64$)	N/C	N/C	N/C	N/C	12 ($p = 64$)	N/C	N/C	N/C	N/C
13 ($p = 32$)	0,00%	34,62%	N/C	0,00%	13 ($p = 32$)	0,00%	19,23%	N/C	N/C
14 ($p = 16$)	0,00%	7,69%	80,77%	0,00%	14 ($p = 16$)	0,00%	7,69%	73,08%	0,00%
15 ($p = 8$)	0,00%	100,00%	88,46%	0,00%	15 ($p = 8$)	0,00%	96,15%	100,00%	0,00%

A Tabela 8 apresenta os resultados de acertos de classificação para um cenário com matrizes de treinamento com $n_r = 26$ amostras para cada classe, sendo todas as amostras de treinamento sem nenhum tipo de variabilidade (rotação, corte, ou ruído). Para os resultados desta tabela, o método de transformada unidimensional utilizando a *wavelet* de Haar foi aplicado no processo de extração de características.

Os resultados são apresentados a partir do 9º nível da decomposição. Neste nível as amostras apresentam dimensão $p = 512$ variáveis, e para cada nível seguinte reduz à metade a dimensão.

Afim de avaliar a robustez do cenário de classificação, as imagens para classificação apresentam rotação em seu eixo no sentido horário, num total de 104 objetos para classificação.

Para o método unidimensional de extração de características utilizando *wavelet* de Haar, a classificação de objetos sem rotação tem acerto total a partir de amostras com 128 variáveis, mas apresenta muita irregularidade para objetos com rotação.

As Classes 1 e 4 para rotações a partir de 10° apresentam os piores resultados de classificação para qualquer nível de decomposição *wavelet* e, de acordo com o aumento no grau de rotação do objeto de classificação, os erros são máximos (nenhum acerto de classificação). Para as Classes 2 e 3, os resultados de classificação são melhores em relação às outras classes. Para estas classes, devido ao aumento da rotação do objeto, a classificação do objeto tornou-se mais eficiente quando a extração de características reduzia à um número de 16 ou 8 variáveis (níveis 14 e 15, respectivamente).

Tabela 9: Resultados de classificação de imagens de naipes de cartas. Imagens com rotação. Método unidimensional, *wavelet* de Daubechies *Db2*.

Classificação de Naipes - Wavelet de Daubechies <i>Db2</i> .									
ROTAÇÃO 0°					ROTAÇÃO 10°				
NÍVEL	$n_r = 26$ amostras				NÍVEL	$n_r = 26$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
9 ($p = 512$)	N/C	N/C	N/C	N/C	9 ($p = 512$)	N/C	N/C	N/C	N/C
10 ($p = 256$)	N/C	N/C	N/C	N/C	10 ($p = 256$)	N/C	N/C	N/C	N/C
11 ($p = 128$)	100,00%	100,00%	100,00%	100,00%	11 ($p = 128$)	N/C	N/C	N/C	N/C
12 ($p = 64$)	100,00%	100,00%	100,00%	100,00%	12 ($p = 64$)	76,92%	96,15%	15,38%	N/C
13 ($p = 32$)	100,00%	100,00%	100,00%	100,00%	13 ($p = 32$)	100,00%	96,15%	100,00%	19,23%
14 ($p = 16$)	100,00%	100,00%	100,00%	100,00%	14 ($p = 16$)	100,00%	100,00%	96,15%	30,77%
15 ($p = 8$)	100,00%	100,00%	100,00%	100,00%	15 ($p = 8$)	100,00%	100,00%	100,00%	80,77%
ROTAÇÃO 15°					ROTAÇÃO 20°				
NÍVEL	$n_r = 26$ amostras				NÍVEL	$n_r = 26$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
9 ($p = 512$)	N/C	N/C	N/C	N/C	9 ($p = 512$)	N/C	N/C	N/C	N/C
10 ($p = 256$)	N/C	N/C	N/C	N/C	10 ($p = 256$)	N/C	N/C	N/C	N/C
11 ($p = 128$)	N/C	N/C	N/C	N/C	11 ($p = 128$)	N/C	N/C	N/C	N/C
12 ($p = 64$)	57,69%	88,46%	N/C	N/C	12 ($p = 64$)	50,00%	50,00%	N/C	N/C
13 ($p = 32$)	100,00%	96,15%	65,38%	0,00%	13 ($p = 32$)	100,00%	80,77%	0,00%	0,00%
14 ($p = 16$)	100,00%	96,15%	19,23%	0,00%	14 ($p = 16$)	100,00%	84,62%	0,00%	0,00%
15 ($p = 8$)	100,00%	100,00%	100,00%	23,08%	15 ($p = 8$)	100,00%	96,15%	50,00%	0,00%
ROTAÇÃO 30°					ROTAÇÃO 45°				
NÍVEL	$n_r = 26$ amostras				NÍVEL	$n_r = 26$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
9 ($p = 512$)	N/C	N/C	N/C	N/C	9 ($p = 512$)	N/C	N/C	N/C	N/C
10 ($p = 256$)	N/C	N/C	N/C	N/C	10 ($p = 256$)	N/C	N/C	N/C	N/C
11 ($p = 128$)	N/C	N/C	N/C	N/C	11 ($p = 128$)	N/C	N/C	N/C	N/C
12 ($p = 64$)	19,23%	N/C	N/C	N/C	12 ($p = 64$)	N/C	N/C	N/C	N/C
13 ($p = 32$)	88,46%	57,69%	N/C	N/C	13 ($p = 32$)	26,92%	61,54%	N/C	N/C
14 ($p = 16$)	53,85%	69,23%	0,00%	0,00%	14 ($p = 16$)	11,54%	96,15%	N/C	N/C
15 ($p = 8$)	42,31%	92,31%	0,00%	0,00%	15 ($p = 8$)	3,85%	96,15%	0,00%	0,00%

A Tabela 9 apresenta os resultados de acertos de classificação para cenário com matrizes de treinamento com $n_r = 26$ amostras por classe. Todas as amostras de treinamento não apresentam variabilidade (rotação, corte, ruído).

Para os resultados desta tabela, o método de transformada unidimensional utilizando a *wavelet* de Daubechies *Db2* foi aplicado no processo de extração de características. Assim como na tabela anterior, são apresentados resultados de classificação a partir do 9º nível de decomposição *wavelet*. Os objetos de classificação apresentam rotação de seu eixo de até 45º no sentido horário. Os objetos de classificação totalizam 104 amostras.

Para objetos sem rotação (0º) obtém-se o número máximo de acertos para todas as classes a partir do 11º nível de decomposição.

Os resultados de classificação para as Classes 1 e 2 apresentam maior regularidade até rotação de 30º.

As Classes 3 e 4, a partir de rotações maiores que 15º apresentam muitos erros de classificação para qualquer nível de decomposição da transformada *wavelet*.

Mesmo com a diminuição dos acertos de classificação para as Classes 3 e 4 para rotações maiores que 15º, a extração de características via *wavelet* de Daubechies *Db2* apresenta melhores resultados de classificação geral quando comparado à *wavelet* de Haar.

Tabela 10: Resultados de classificação de imagens de naipes de cartas. Imagens com rotação. Método unidimensional, *wavelet* de Daubechies *Db8*.

Classificação de Naipes - Wavelet de Daubechies <i>Db8</i> .									
ROTAÇÃO 0°					ROTAÇÃO 10°				
NÍVEL	$n_r = 26$ amostras				NÍVEL	$n_r = 26$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
9 ($p = 512$)	N/C	N/C	N/C	N/C	9 ($p = 512$)	N/C	N/C	N/C	N/C
10 ($p = 256$)	N/C	N/C	N/C	N/C	10 ($p = 256$)	N/C	N/C	N/C	N/C
11 ($p = 128$)	100,00%	100,00%	100,00%	100,00%	11 ($p = 128$)	N/C	N/C	N/C	N/C
12 ($p = 64$)	100,00%	100,00%	100,00%	100,00%	12 ($p = 64$)	65,38%	92,31%	19,23%	N/C
13 ($p = 32$)	100,00%	100,00%	100,00%	100,00%	13 ($p = 32$)	100,00%	92,31%	100,00%	19,23%
14 ($p = 16$)	100,00%	100,00%	100,00%	100,00%	14 ($p = 16$)	100,00%	92,31%	100,00%	42,31%
15 ($p = 8$)	100,00%	100,00%	100,00%	100,00%	15 ($p = 8$)	100,00%	92,31%	100,00%	61,54%
ROTAÇÃO 15°					ROTAÇÃO 20°				
NÍVEL	$n_r = 26$ amostras				NÍVEL	$n_r = 26$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
9 ($p = 512$)	N/C	N/C	N/C	N/C	9 ($p = 512$)	N/C	N/C	N/C	N/C
10 ($p = 256$)	N/C	N/C	N/C	N/C	10 ($p = 256$)	N/C	N/C	N/C	N/C
11 ($p = 128$)	N/C	N/C	N/C	N/C	11 ($p = 128$)	N/C	N/C	N/C	N/C
12 ($p = 64$)	73,08%	76,92%	N/C	N/C	12 ($p = 64$)	61,54%	53,85%	0,00%	N/C
13 ($p = 32$)	100,00%	88,46%	61,54%	0,00%	13 ($p = 32$)	100,00%	61,54%	0,00%	0,00%
14 ($p = 16$)	100,00%	100,00%	76,92%	0,00%	14 ($p = 16$)	100,00%	92,31%	3,85%	0,00%
15 ($p = 8$)	100,00%	88,46%	50,00%	0,00%	15 ($p = 8$)	100,00%	50,00%	7,69%	0,00%
ROTAÇÃO 30°					ROTAÇÃO 45°				
NÍVEL	$n_r = 26$ amostras				NÍVEL	$n_r = 26$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
9 ($p = 512$)	N/C	N/C	N/C	N/C	9 ($p = 512$)	N/C	N/C	N/C	N/C
10 ($p = 256$)	N/C	N/C	N/C	N/C	10 ($p = 256$)	N/C	N/C	N/C	N/C
11 ($p = 128$)	N/C	N/C	N/C	N/C	11 ($p = 128$)	N/C	N/C	N/C	N/C
12 ($p = 64$)	3,85%	0,00%	N/C	N/C	12 ($p = 64$)	N/C	0,00%	N/C	N/C
13 ($p = 32$)	88,46%	57,69%	N/C	N/C	13 ($p = 32$)	0,00%	57,69%	N/C	N/C
14 ($p = 16$)	46,15%	80,77%	3,85%	0,00%	14 ($p = 16$)	0,00%	100,00%	N/C	N/C
15 ($p = 8$)	100,00%	11,54%	0,00%	0,00%	15 ($p = 8$)	50,00%	15,38%	0,00%	0,00%

A Tabela 10 apresenta os resultados de acertos de classificação para cenário com ma-

trizes de treinamento com $n_r = 26$ amostras por classe. Estas amostras de treinamento, assim como para os cenários anteriores, não apresentam varibilidade (corte, rotação, ruído).

Neste momento, o método de transformada unidimensional utiliza a *wavelet* de Daubechies *Db8* para processo de extração de características. Assim como na tabela anterior, são apresentados resultados de classificação para níveis de decomposição a partir do 9º nível. Os objetos de classificação apresentam rotação de seu eixo de até 45° no sentido horário, e os objetos de classificação totalizam 104 amostras.

Os resultados de classificação para as Classes 1 e 2 apresentam maior regularidade até rotação de 30°, assim como para a *wavelet* Daubechies *Db2*, principalmente para os níveis 13, 14, e 15.

As Classes 3 e 4 a partir de rotações maiores que 15° apresentam muitos erros de classificação, como visto para a *wavelet* Daubechies *Db2*.

Mesmo com diminuição dos acertos de classificação para as Classes 3 e 4 para rotações maiores que 15°, a extração de características via *wavelet* de Daubechies *Db8* apresenta os resultados semelhantes aos resultados de classificação geral quando comparado à *wavelet* *Db2*, e esses resultados também são melhores do que os resultados de classificação para *wavelet* de Haar.

A partir das próximas tabelas, os cenários de classificação apresentam aumento no número de amostras de treinamento das classes. O aumento de amostras de treinamento para cada classe tem origem na inserção de amostras com rotação em seu eixo de $\pm 5^\circ$ e $\pm 10^\circ$, assim totalizando em $n_r = 130$ amostras de treinamento por classe. Neste cenário com mais amostras por classe, observa-se os efeitos da relação de número de amostras e dimensão das amostras.

Tabela 11: Resultados de classificação de imagens de naipes de cartas. Imagens com rotação. Método unidimensional, *wavelet* de Haar.

Classificação de Naipes - <i>Wavelet</i> de Haar.									
ROTAÇÃO 15°					ROTAÇÃO 20°				
NÍVEL	$n_r = 130$ amostras				NÍVEL	$n_r = 130$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
9 ($p = 512$)	N/C	N/C	N/C	N/C	9 ($p = 512$)	N/C	N/C	N/C	N/C
10 ($p = 256$)	N/C	N/C	N/C	N/C	10 ($p = 256$)	N/C	N/C	N/C	N/C
11 ($p = 128$)	N/C	N/C	N/C	N/C	11 ($p = 128$)	N/C	N/C	N/C	N/C
12 ($p = 64$)	0,00%	0,00%	0,00%	0,00%	12 ($p = 64$)	0,00%	0,00%	0,00%	0,00%
13 ($p = 32$)	0,00%	0,00%	0,00%	0,00%	13 ($p = 32$)	0,00%	0,00%	0,00%	0,00%
14 ($p = 16$)	0,00%	0,00%	0,00%	0,00%	14 ($p = 16$)	0,00%	0,00%	0,00%	0,00%
15 ($p = 8$)	0,00%	0,00%	73,08%	0,00%	15 ($p = 8$)	0,00%	0,00%	73,08%	0,00%
ROTAÇÃO 30°					ROTAÇÃO 45°				
NÍVEL	$n_r = 130$ amostras				NÍVEL	$n_r = 130$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
9 ($p = 512$)	N/C	N/C	N/C	N/C	9 ($p = 512$)	N/C	N/C	N/C	N/C
10 ($p = 256$)	N/C	N/C	N/C	N/C	10 ($p = 256$)	N/C	N/C	N/C	N/C
11 ($p = 128$)	N/C	N/C	N/C	N/C	11 ($p = 128$)	N/C	N/C	N/C	N/C
12 ($p = 64$)	0,00%	0,00%	0,00%	0,00%	12 ($p = 64$)	0,00%	0,00%	0,00%	0,00%
13 ($p = 32$)	0,00%	0,00%	0,00%	0,00%	13 ($p = 32$)	0,00%	0,00%	0,00%	0,00%
14 ($p = 16$)	0,00%	0,00%	0,00%	0,00%	14 ($p = 16$)	0,00%	0,00%	0,00%	0,00%
15 ($p = 8$)	0,00%	0,00%	73,08%	0,00%	15 ($p = 8$)	0,00%	0,00%	69,23%	0,00%

A Tabela 11 apresenta os resultados de acertos de classificação para um cenário de classificação com matrizes de treinamento com $n_r = 130$ amostras para cada classe.

Para os resultados desta tabela, o método unidimensional para extração de características foi aplicado utilizando a *wavelet* de Haar. Os resultados são apresentados a partir do 9º nível da decomposição. Neste nível as amostras apresentam $p = 512$ variáveis. Os objetos de classificação apresentam rotação em seu eixo, no sentido horário, num total de 104 objetos para classificação.

Independente do grau de rotação do objeto de classificação e nível de decomposição, para o método unidimensional utilizando *wavelet* de Haar, a classificação apresenta muita irregularidade. Este comportamento fora observado antes para o cenário com $n_r = 26$ amostras por classe, onde os objetos com rotação a partir de 15° não apresentaram resultados de acertos de classificação.

Tabela 12: Resultados de classificação de imagens de naipes de cartas. Imagens com rotação. Método unidimensional, *wavelet* de Daubechies *Db2*.

Classificação de Naipes - Wavelet de Daubechies <i>Db2</i> .									
ROTAÇÃO 15°					ROTAÇÃO 20°				
NÍVEL	$n_r = 130$ amostras				NÍVEL	$n_r = 130$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
9 ($p = 512$)	N/C	N/C	N/C	N/C	9 ($p = 512$)	N/C	N/C	N/C	N/C
10 ($p = 256$)	N/C	N/C	N/C	N/C	10 ($p = 256$)	N/C	N/C	N/C	N/C
11 ($p = 128$)	N/C	N/C	N/C	N/C	11 ($p = 128$)	N/C	N/C	N/C	N/C
12 ($p = 64$)	100,00%	100,00%	100,00%	100,00%	12 ($p = 64$)	100,00%	100,00%	88,46%	88,46%
13 ($p = 32$)	100,00%	100,00%	100,00%	100,00%	13 ($p = 32$)	100,00%	100,00%	92,31%	92,31%
14 ($p = 16$)	100,00%	100,00%	100,00%	100,00%	14 ($p = 16$)	100,00%	100,00%	100,00%	96,15%
15 ($p = 8$)	100,00%	100,00%	100,00%	100,00%	15 ($p = 8$)	100,00%	100,00%	100,00%	76,92%
ROTAÇÃO 30°					ROTAÇÃO 45°				
NÍVEL	$n_r = 130$ amostras				NÍVEL	$n_r = 130$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
9 ($p = 512$)	N/C	N/C	N/C	N/C	9 ($p = 512$)	N/C	N/C	N/C	N/C
10 ($p = 256$)	N/C	N/C	N/C	N/C	10 ($p = 256$)	N/C	N/C	N/C	N/C
11 ($p = 128$)	N/C	N/C	N/C	N/C	11 ($p = 128$)	N/C	N/C	N/C	N/C
12 ($p = 64$)	69,23%	100,00%	N/C	N/C	12 ($p = 64$)	96,15%	73,08%	N/C	N/C
13 ($p = 32$)	96,15%	100,00%	0,00%	0,00%	13 ($p = 32$)	100,00%	80,77%	0,00%	0,00%
14 ($p = 16$)	96,15%	100,00%	96,15%	0,00%	14 ($p = 16$)	76,92%	100,00%	0,00%	0,00%
15 ($p = 8$)	100,00%	100,00%	100,00%	0,00%	15 ($p = 8$)	96,15%	50,00%	96,15%	0,00%

A Tabela 12 apresenta os resultados de acertos de classificação para um cenário com matrizes de treinamento com $n_r = 130$ amostras para cada classe.

Para os resultados desta tabela, o método unidimensional para extração de características foi aplicado utilizando a *wavelet* de Daubechies *Db2*. Os resultados são apresentados a partir do 9º nível da decomposição. São 104 objetos para classificação, com diferentes graus de rotação.

Para imagens com rotação de 15° e 20°, os resultados apresentam grande grau de acerto para todas as classes, principalmente para níveis 13 e 14.

A Classe 4 apresenta redução total nos acertos de classificação para objetos com rotação a partir de 30°, enquanto para Classe 3 o número de acertos reduz de acordo com o aumento de grau de rotação (a partir de 30°) e número maior de variáveis por amostra.

Para todos os graus de rotação as Classes 1 e 2 apresentam regularidade na classificação, com grande taxas de acerto, principalmente para os 12º, 13º e 14º níveis de decomposição *wavelet*.

Tabela 13: Resultados de classificação de imagens de naipes de cartas. Imagens com rotação. Método unidimensional, *wavelet* de Daubechies *Db8*.

Classificação de Naipes - <i>Wavelet</i> de Daubechies <i>Db8</i> .									
ROTAÇÃO 15°					ROTAÇÃO 20°				
NÍVEL	$n_r = 130$ amostras				NÍVEL	$n_r = 130$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
9 ($p = 512$)	N/C	N/C	N/C	N/C	9 ($p = 512$)	N/C	N/C	N/C	N/C
10 ($p = 256$)	N/C	N/C	N/C	N/C	10 ($p = 256$)	N/C	N/C	N/C	N/C
11 ($p = 128$)	N/C	N/C	N/C	N/C	11 ($p = 128$)	N/C	N/C	N/C	N/C
12 ($p = 64$)	100,00%	100,00%	100,00%	100,00%	12 ($p = 64$)	100,00%	100,00%	84,62%	84,62%
13 ($p = 32$)	100,00%	100,00%	100,00%	100,00%	13 ($p = 32$)	100,00%	100,00%	100,00%	92,31%
14 ($p = 16$)	100,00%	100,00%	100,00%	100,00%	14 ($p = 16$)	100,00%	100,00%	100,00%	96,15%
15 ($p = 8$)	100,00%	100,00%	100,00%	100,00%	15 ($p = 8$)	100,00%	100,00%	100,00%	65,38%
ROTAÇÃO 30°					ROTAÇÃO 45°				
NÍVEL	$n_r = 130$ amostras				NÍVEL	$n_r = 130$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
9 ($p = 512$)	N/C	N/C	N/C	N/C	9 ($p = 512$)	N/C	N/C	N/C	N/C
10 ($p = 256$)	N/C	N/C	N/C	N/C	10 ($p = 256$)	N/C	N/C	N/C	N/C
11 ($p = 128$)	N/C	N/C	N/C	N/C	11 ($p = 128$)	N/C	N/C	N/C	N/C
12 ($p = 64$)	84,62%	100,00%	N/C	N/C	12 ($p = 64$)	96,15%	69,23%	N/C	N/C
13 ($p = 32$)	92,31%	100,00%	3,85%	0,00%	13 ($p = 32$)	80,77%	80,77%	N/C	N/C
14 ($p = 16$)	96,15%	100,00%	3,85%	0,00%	14 ($p = 16$)	69,23%	100,00%	0,00%	0,00%
15 ($p = 8$)	100,00%	100,00%	80,77%	0,00%	15 ($p = 8$)	76,92%	69,23%	0,00%	0,00%

Assim como para os dois últimos cenários, a Tabela 13 apresenta os resultados de acertos de classificação para um cenário de classificação com matrizes de treinamento com $n_r = 130$ amostras por classe.

Para os resultados desta tabela, o método unidimensional para extração de características foi aplicado utilizando a *wavelet* de Daubechies *Db8*. Os resultados são apresentados a partir do 9º nível de decomposição. São 104 objetos para classificação.

Assim como observado para a aplicação da *wavelet* Daubechies *Db2*, para imagens com rotação de 15° e 20° os resultados apresentam grande número de acertos para todas as classes.

A Classe 4 também apresenta redução total nos acertos de classificação para objetos com rotação a partir de 30°, enquanto a Classe 3 reduz os acertos de classificação de acordo com o aumento de grau de rotação (a partir de 30°).

Para todos os graus de rotação as Classes 1 e 2 apresentam regularidade na classificação, com grande taxas de acerto, assim como observado para *wavelet* Daubechies *Db2*.

5.2.2 Método de classificação via transformada *wavelet* 2D

Neste método para a extração de variáveis das imagens, a transformada *wavelet* 2D é aplicada. A TWD 2D é aplicada a cada amostra de cada classe na matriz de treinamento. A classificação dos objetos é realizada para cada nível de decomposição. Cada nível de

decomposição resulta em nova representação da imagem, uma representação média da imagem original e com dimensão menor: o nível 1 tem dimensão 256x256 pixels, o nível 2 tem dimensão 128x128 pixels, nível 3 tem dimensão 64x64 pixels.

O algoritmo de classificação tem as matrizes de treinamento e objetos de classificação representados por vetores. Então, cada representação de cada nível da decomposição *wavelet* 2D deve passar para representação em um vetor, através da concatenação das linhas da imagem. O mesmo procedimento é aplicado para o objeto de classificação. As representações das amostras de treinamento após a transformada *wavelet* 2D estão na Figura 31, para imagens com *flash*, e na Figura 32, para imagens sem *flash*.

Nota-se que as curvas que representam as amostras pós-extração de características são muito diferentes quando observa-se as Figuras 29 e 30, tanto para imagens com *flash* quanto para imagens sem *flash*.

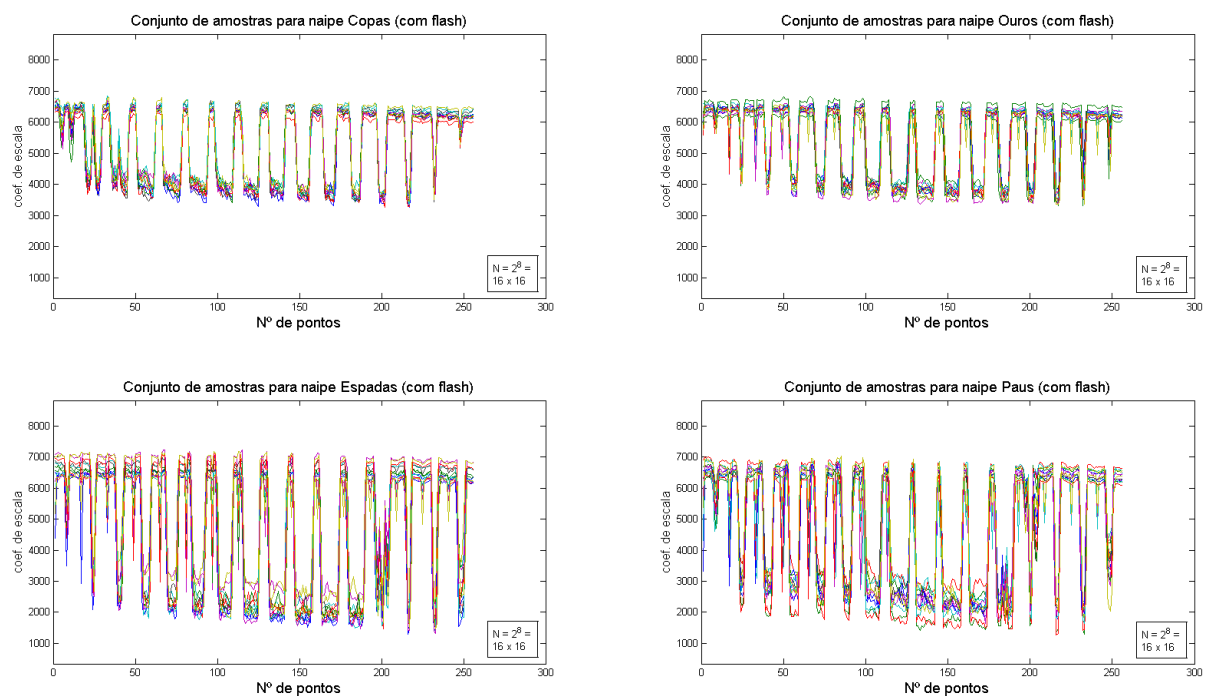


Figura 31: Representação das 4 classes de naipes de cartas no 5º nível de TWD 2D, a partir de imagens com *flash*. Então, as imagens têm as linhas concatenadas em um vetor.

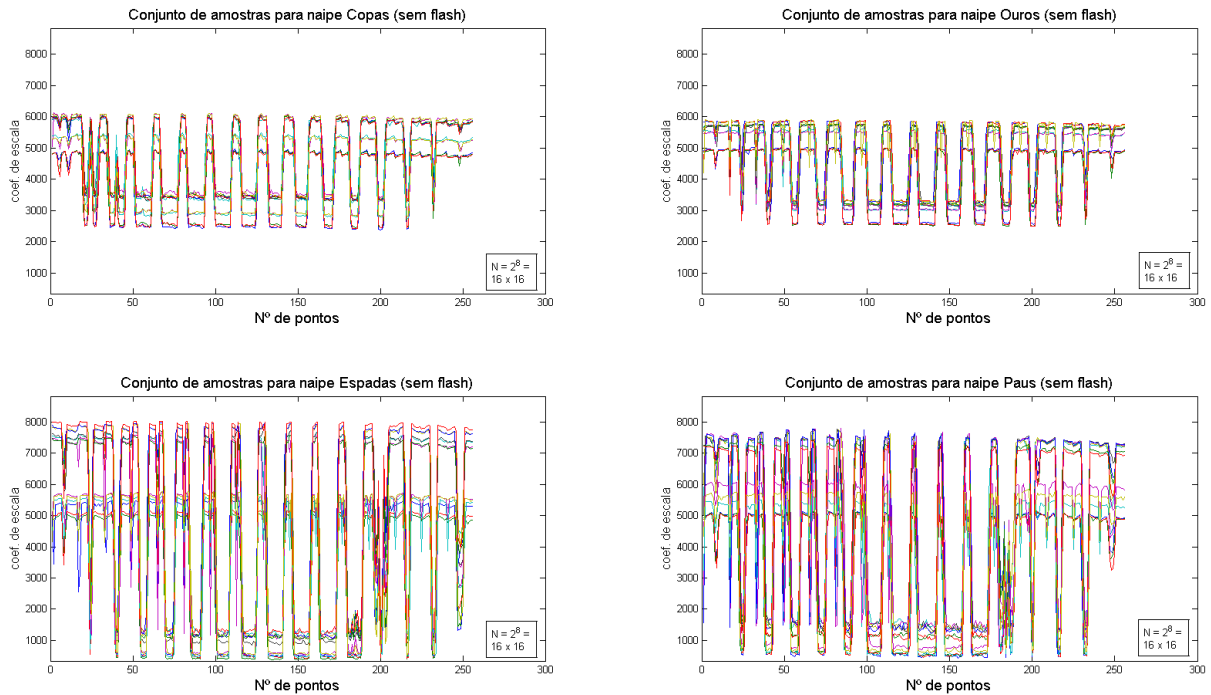


Figura 32: Representação das 4 classes de naipes de cartas no 5º nível de TWD 2D, a partir de imagens sem *flash*. Então, as imagens têm as linhas concatenadas em um vetor.

Tabela 14: Resultados de classificação de imagens de naipes de cartas. Imagens com rotação. Método bidimensional, *wavelet* de Haar.

Classificação de Naipes - Wavelet de Haar.									
ROTAÇÃO 0°					ROTAÇÃO 10°				
NÍVEL	$n_r = 26$ amostras				NÍVEL	$n_r = 26$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
1 (p=256x256)	N/C	N/C	N/C	N/C	1 (p=256x256)	N/C	N/C	N/C	N/C
2 (p=128x128)	N/C	N/C	N/C	N/C	2 (p=128x128)	N/C	N/C	N/C	N/C
3 (p=64x64)	N/C	N/C	N/C	N/C	3 (p=64x64)	N/C	N/C	N/C	N/C
4 (p=32x32)	N/C	N/C	N/C	N/C	4 (p=32x32)	N/C	N/C	N/C	N/C
5 (p=16x16)	11,54%	34,62%	N/C	N/C	5 (p=16x16)	3,85%	46,15%	100,00%	23,08%
6 (p=8x8)	100,00%	100,00%	100,00%	100,00%	6 (p=8x8)	100,00%	100,00%	100,00%	100,00%
ROTAÇÃO 15°					ROTAÇÃO 20°				
NÍVEL	$n_r = 26$ amostras				NÍVEL	$n_r = 26$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
1 (p=256x256)	N/C	N/C	N/C	N/C	1 (p=256x256)	N/C	N/C	N/C	N/C
2 (p=128x128)	N/C	N/C	N/C	N/C	2 (p=128x128)	N/C	N/C	N/C	N/C
3 (p=64x64)	N/C	N/C	N/C	N/C	3 (p=64x64)	N/C	N/C	N/C	N/C
4 (p=32x32)	N/C	N/C	N/C	N/C	4 (p=32x32)	N/C	N/C	N/C	N/C
5 (p=16x16)	N/C	N/C	N/C	N/C	5 (p=16x16)	N/C	N/C	N/C	N/C
6 (p=8x8)	100,00%	100,00%	100,00%	80,77%	6 (p=8x8)	100,00%	100,00%	100,00%	46,15%
ROTAÇÃO 30°					ROTAÇÃO 45°				
NÍVEL	$n_r = 26$ amostras				NÍVEL	$n_r = 26$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
1 (p=256x256)	N/C	N/C	N/C	N/C	1 (p=256x256)	N/C	N/C	N/C	N/C
2 (p=128x128)	N/C	N/C	N/C	N/C	2 (p=128x128)	N/C	N/C	N/C	N/C
3 (p=64x64)	N/C	N/C	N/C	N/C	3 (p=64x64)	N/C	N/C	N/C	N/C
4 (p=32x32)	N/C	N/C	N/C	N/C	4 (p=32x32)	N/C	N/C	N/C	N/C
5 (p=16x16)	N/C	N/C	N/C	N/C	5 (p=16x16)	N/C	N/C	N/C	N/C
6 (p=8x8)	80,77%	92,31%	73,08%	3,85%	6 (p=8x8)	0,00%	96,15%	0,00%	0,00%

A Tabela 14 apresenta os resultados de acertos de classificação para cenário com matrizes de treinamento com $n_r = 26$ amostras para cada classe. As amostras de treinamento não têm algum tipo de variabilidade (rotação, corte, ou ruído).

Para estes resultados, o método de transformada bidimensional utilizando a *wavelet* de Haar foi aplicado no processo de extração de características. São 104 objetos para classificação, que apresentam rotação de até 45°. Os resultados são apresentados a partir do 1º nível da decomposição, utilizando a representação da sub-banda CC da decomposição. Neste nível de decomposição, esta sub-banda tem dimensão de 256x256 variáveis. O classificador opera com amostras representadas por vetores, então as representações das imagens devem ser transformadas para vetores através da concatenação das linhas da imagem. Assim, o 1º nível tem representação em um vetor com dimensão $p = 65536$ variáveis.

A tendência de grande redução da dimensão das amostras para que seja possível a classificação, observada para o método unidimensional para a extração de características, também segue para a extração de características através do método bidimensional.

Para o método bidimensional utilizando *wavelet* de Haar, os primeiros resultados de classificação surgem quando tem dimensão $p = 16 \times 16 = 256$ variáveis para as amostras, embora não tenha resultados de classificação para todas as classes.

Para o próximo nível de decomposição, a dimensão é reduzida para $p = 8 \times 8 = 64$ e os resultados apresentam grande número de acertos de classificação para objetos com rotação até 15°.

A partir de 20° de rotação para os objetos de classificação os resultados de acertos de classificação apresentam redução pequena para as Classes 1, 2 e 3 e a Classe 4 apresenta a maior redução de acertos de classificação.

Para objetos com rotação de 45° apenas para a Classe 2 há acertos de classificação.

Tabela 15: Resultados de classificação de imagens de naipes de cartas. Imagens com rotação. Método bidimensional, *wavelet* de Daubechies *Db2*.

Classificação de Naipes - <i>Wavelet</i> de Daubechies <i>Db2</i> .									
ROTAÇÃO 0°					ROTAÇÃO 10°				
NÍVEL	$n_r = 26$ amostras				NÍVEL	$n_r = 26$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
1 (p=256x256)	N/C	N/C	N/C	N/C	1 (p=256x256)	N/C	N/C	N/C	N/C
2 (p=128x128)	N/C	N/C	N/C	N/C	2 (p=128x128)	N/C	N/C	N/C	N/C
3 (p=64x64)	N/C	N/C	N/C	N/C	3 (p=64x64)	N/C	N/C	N/C	N/C
4 (p=32x32)	N/C	N/C	N/C	N/C	4 (p=32x32)	N/C	N/C	N/C	N/C
5 (p=16x16)	26,92%	46,15%	23,08%	N/C	5 (p=16x16)	N/C	3,85%	N/C	N/C
6 (p=8x8)	100,00%	100,00%	100,00%	100,00%	6 (p=8x8)	100,00%	92,31%	100,00%	96,15%
ROTAÇÃO 15°					ROTAÇÃO 20°				
NÍVEL	$n_r = 26$ amostras				NÍVEL	$n_r = 26$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
1 (p=256x256)	N/C	N/C	N/C	N/C	1 (p=256x256)	N/C	N/C	N/C	N/C
2 (p=128x128)	N/C	N/C	N/C	N/C	2 (p=128x128)	N/C	N/C	N/C	N/C
3 (p=64x64)	N/C	N/C	N/C	N/C	3 (p=64x64)	N/C	N/C	N/C	N/C
4 (p=32x32)	N/C	N/C	N/C	N/C	4 (p=32x32)	N/C	N/C	N/C	N/C
5 (p=16x16)	N/C	N/C	N/C	N/C	5 (p=16x16)	N/C	N/C	N/C	N/C
6 (p=8x8)	100,00%	80,77%	100,00%	57,69%	6 (p=8x8)	80,77%	30,77%	100,00%	7,69%
ROTAÇÃO 30°					ROTAÇÃO 45°				
NÍVEL	$n_r = 26$ amostras				NÍVEL	$n_r = 26$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
1 (p=256x256)	N/C	N/C	N/C	N/C	1 (p=256x256)	N/C	N/C	N/C	N/C
2 (p=128x128)	N/C	N/C	N/C	N/C	2 (p=128x128)	N/C	N/C	N/C	N/C
3 (p=64x64)	N/C	N/C	N/C	N/C	3 (p=64x64)	N/C	N/C	N/C	N/C
4 (p=32x32)	N/C	N/C	N/C	N/C	4 (p=32x32)	N/C	N/C	N/C	N/C
5 (p=16x16)	N/C	N/C	N/C	N/C	5 (p=16x16)	N/C	N/C	N/C	N/C
6 (p=8x8)	50,00%	N/C	26,92%	N/C	6 (p=8x8)	0,00%	3,85%	N/C	N/C

A Tabela 15 apresenta os resultados de classificação para cenário com matrizes de treinamento com $n_r = 26$ amostras por classe. As amostras de treinamento não têm algum tipo de variabilidade (rotação, corte, ou ruído).

O método de transformada bidimensional utilizando a *wavelet* de Daubechies *Db2* foi aplicado no processo de extração de características. São 104 objetos para classificação, e apresentam rotação de até 45°.

Assim como observado na tabela anterior, a redução da dimensionalidade tornou possível a execução de classificação.

Os resultados de classificação surgem a partir de amostras com dimensão $p = 16 \times 16 = 256$ variáveis para objetos sem rotação, mas não para todas as classes. A partir do próximo nível há resultados de classificação para todas as classes.

Comparativamente, os resultados obtidos com *wavelet* Daubechies *Db2* apresentam grande número de acertos de classificação assim como a *wavelet* de Haar para objetos com até 15° de rotação.

A partir da classificação de objetos com 15° de rotação, os acertos de classificação diminuem muito em relação com rotações menores e também em relação à *wavelet* de Haar.

Para rotações a partir de 30° os acertos de classificação reduzem, e em algumas classes não há nenhum acerto de classificação.

Tabela 16: Resultados de classificação de imagens de naipes de cartas. Imagens com rotação. Método bidimensional, *wavelet* de Daubechies *Db8*.

Classificação de Naipes - <i>Wavelet</i> de Daubechies <i>Db8</i> .									
ROTAÇÃO 0°					ROTAÇÃO 10°				
NÍVEL	$n_r = 26$ amostras				NÍVEL	$n_r = 26$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
1 (p=256x256)	N/C	N/C	N/C	N/C	1 (p=256x256)	N/C	N/C	N/C	N/C
2 (p=128x128)	N/C	N/C	N/C	N/C	2 (p=128x128)	N/C	N/C	N/C	N/C
3 (p=64x64)	N/C	N/C	N/C	N/C	3 (p=64x64)	N/C	N/C	N/C	N/C
4 (p=32x32)	N/C	N/C	N/C	N/C	4 (p=32x32)	N/C	N/C	N/C	N/C
5 (p=16x16)	N/C	N/C	N/C	N/C	5 (p=16x16)	N/C	N/C	N/C	N/C
6 (p=8x8)	50,00%	46,15%	42,31%	11,54%	6 (p=8x8)	N/C	34,62%	N/C	N/C
ROTAÇÃO 15°					ROTAÇÃO 20°				
NÍVEL	$n_r = 26$ amostras				NÍVEL	$n_r = 26$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
1 (p=256x256)	N/C	N/C	N/C	N/C	1 (p=256x256)	N/C	N/C	N/C	N/C
2 (p=128x128)	N/C	N/C	N/C	N/C	2 (p=128x128)	N/C	N/C	N/C	N/C
3 (p=64x64)	N/C	N/C	N/C	N/C	3 (p=64x64)	N/C	N/C	N/C	N/C
4 (p=32x32)	N/C	N/C	N/C	N/C	4 (p=32x32)	N/C	N/C	N/C	N/C
5 (p=16x16)	N/C	N/C	N/C	N/C	5 (p=16x16)	N/C	N/C	N/C	N/C
6 (p=8x8)	N/C	N/C	N/C	N/C	6 (p=8x8)	N/C	N/C	N/C	N/C
ROTAÇÃO 30°					ROTAÇÃO 45°				
NÍVEL	$n_r = 26$ amostras				NÍVEL	$n_r = 26$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
1 (p=256x256)	N/C	N/C	N/C	N/C	1 (p=256x256)	N/C	N/C	N/C	N/C
2 (p=128x128)	N/C	N/C	N/C	N/C	2 (p=128x128)	N/C	N/C	N/C	N/C
3 (p=64x64)	N/C	N/C	N/C	N/C	3 (p=64x64)	N/C	N/C	N/C	N/C
4 (p=32x32)	N/C	N/C	N/C	N/C	4 (p=32x32)	N/C	N/C	N/C	N/C
5 (p=16x16)	N/C	N/C	N/C	N/C	5 (p=16x16)	N/C	N/C	N/C	N/C
6 (p=8x8)	N/C	N/C	N/C	N/C	6 (p=8x8)	N/C	N/C	N/C	N/C

A Tabela 16 apresenta os resultados de classificação para cenário com matrizes de treinamento com $n_r = 26$ amostras por classe. As amostras de treinamento não têm algum tipo de variabilidade.

O processo de extração de características utiliza transformada bidimensional aplicando *wavelet* de Daubechies *Db8*. São 104 objetos para classificação que apresentam rotação de até 45° em seu eixo.

Dentre as 3 funções *wavelet* avaliadas, a *wavelet* de Daubechies *Db8* apresenta os piores resultados de classificação. Apenas há resultados de classificação de objetos sem rotação e, ainda assim, têm quantidades de acertos menores quando comparadas às *wavelet* de Haar e Daubechies *Db2*.

Para os objetos com rotações a partir de 10° , não há acerto de classificação para qualquer nível de decomposição e em qualquer classe.

Para avaliação desta metodologia de classificação, que utiliza a transformada *wavelet* 2D para a extração de características, também foi avaliado o desempenho da classificação em cenários com mais amostras de treinamento. Para o aumento de amostras de treina-

mento para cada classe, são inseridas novas amostras com rotação em seu eixo de $\pm 5^\circ$ e $\pm 10^\circ$, assim totalizando em $n_r = 130$ amostras de treinamento por classe.

Tabela 17: Resultados de classificação de imagens de naipes de cartas. Imagens com rotação. Método bidimensional, *wavelet* de Haar.

Classificação de Naipes - Wavelet de Haar.									
ROTAÇÃO 15°					ROTAÇÃO 20°				
NÍVEL	$n_r = 130$ amostras				NÍVEL	$n_r = 130$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
1 (p=256x256)	N/C	N/C	N/C	N/C	1 (p=256x256)	N/C	N/C	N/C	N/C
2 (p=128x128)	N/C	N/C	N/C	N/C	2 (p=128x128)	N/C	N/C	N/C	N/C
3 (p=64x64)	N/C	N/C	N/C	N/C	3 (p=64x64)	N/C	N/C	N/C	N/C
4 (p=32x32)	N/C	N/C	N/C	N/C	4 (p=32x32)	N/C	N/C	N/C	N/C
5 (p=16x16)	100,00%	100,00%	100,00%	100,00%	5 (p=16x16)	100,00%	100,00%	92,31%	96,15%
6 (p=8x8)	100,00%	100,00%	100,00%	100,00%	6 (p=8x8)	100,00%	100,00%	100,00%	100,00%
ROTAÇÃO 30°					ROTAÇÃO 45°				
NÍVEL	$n_r = 130$ amostras				NÍVEL	$n_r = 130$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
1 (p=256x256)	N/C	N/C	N/C	N/C	1 (p=256x256)	N/C	N/C	N/C	N/C
2 (p=128x128)	N/C	N/C	N/C	N/C	2 (p=128x128)	N/C	N/C	N/C	N/C
3 (p=64x64)	N/C	N/C	N/C	N/C	3 (p=64x64)	N/C	N/C	N/C	N/C
4 (p=32x32)	N/C	N/C	N/C	N/C	4 (p=32x32)	N/C	N/C	N/C	N/C
5 (p=16x16)	N/C	N/C	61,54%	N/C	5 (p=16x16)	N/C	N/C	N/C	N/C
6 (p=8x8)	100,00%	100,00%	84,62%	53,85%	6 (p=8x8)	0,00%	100,00%	N/C	N/C

A Tabela 17 apresenta os resultados de acertos de classificação para cenário com matrizes de treinamento com $n_r = 130$ amostras para cada classe.

Para estes resultados, o método de transformada bidimensional utilizando a *wavelet* de Haar foi aplicado no processo de extração de características. São 104 objetos para classificação.

Ao comparar-se com os resultados obtidos com a *wavelet* de Haar para o cenário com $n_r = 26$ amostras por classe, observa-se que o aumento do número de amostras de treinamento por classe contribuiu para a execução de classificação quando as amostras têm dimensão $p = 16 \times 16 = 256$ variáveis. Neste nível de decomposição os resultados de classificação apresentam grande número de acertos para objetos com rotação até 15° .

Para as amostras com dimensão $p = 8 \times 8 = 64$ variáveis, os objetos com rotação até 30° apresentam grande número de acertos de classificação

Para objetos com rotação a partir de 30° , os resultados para amostras com dimensão de $p = 16 \times 16 = 256$ variáveis reduzem muito, ou não apresentam nenhum acerto de classificação.

Para os objetos com rotação de 45° os acertos de classificação diminuem muito e apenas a Classe 2 tem acertos de classificação para dimensão $p = 8 \times 8 = 64$ variáveis.

Tabela 18: Resultados de classificação de imagens de naipes de cartas. Imagens com rotação. Método bidimensional, *wavelet* de Daubechies *Db2*.

Classificação de Naipes - <i>Wavelet</i> de Daubechies <i>Db2</i> .									
ROTAÇÃO 15°					ROTAÇÃO 20°				
NÍVEL	$n_r = 130$ amostras				NÍVEL	$n_r = 130$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
1 (p=256x256)	N/C	N/C	N/C	N/C	1 (p=256x256)	N/C	N/C	N/C	N/C
2 (p=128x128)	N/C	N/C	N/C	N/C	2 (p=128x128)	N/C	N/C	N/C	N/C
3 (p=64x64)	N/C	N/C	N/C	N/C	3 (p=64x64)	N/C	N/C	N/C	N/C
4 (p=32x32)	N/C	N/C	N/C	N/C	4 (p=32x32)	N/C	N/C	N/C	N/C
5 (p=16x16)	N/C	N/C	N/C	N/C	5 (p=16x16)	N/C	N/C	N/C	N/C
6 (p=8x8)	100,00%	100,00%	100,00%	100,00%	6 (p=8x8)	100,00%	100,00%	100,00%	100,00%
ROTAÇÃO 30°					ROTAÇÃO 45°				
NÍVEL	$n_r = 130$ amostras				NÍVEL	$n_r = 130$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
1 (p=256x256)	N/C	N/C	N/C	N/C	1 (p=256x256)	N/C	N/C	N/C	N/C
2 (p=128x128)	N/C	N/C	N/C	N/C	2 (p=128x128)	N/C	N/C	N/C	N/C
3 (p=64x64)	N/C	N/C	N/C	N/C	3 (p=64x64)	N/C	N/C	N/C	N/C
4 (p=32x32)	N/C	N/C	N/C	N/C	4 (p=32x32)	N/C	N/C	N/C	N/C
5 (p=16x16)	N/C	N/C	61,54%	N/C	5 (p=16x16)	N/C	N/C	N/C	N/C
6 (p=8x8)	96,15%	100,00%	84,62%	26,92%	6 (p=8x8)	0,00%	100,00%	92,31%	0,00%

A Tabela 18 apresenta os resultados de acertos de classificação para cenário com matrizes de treinamento com $n_r = 130$ amostras para cada classe.

Para estes resultados, o método de transformada bidimensional utilizando a *wavelet* de Daubechies *Db2* foi aplicado no processo de extração de características. São 104 objetos para classificação em que apresentam rotação de até 45°.

Comparando com os resultados obtidos com a *wavelet* de Daubechies *Db2* para o cenário anterior, onde o número de amostras de treinamento por classe é $n_r = 26$, observa-se que não houve classificação para as amostras com dimensão maior que $p = 8 \times 8 = 64$ variáveis. A classificação dos objetos para dimensões maiores que $p = 8 \times 8 = 64$ foi observada para o caso que aplica a *wavelet* de Haar.

Para o 6º nível de decomposição os resultados de classificação apresentam grande número de acertos para objetos com rotação até 30°, embora a Classe 4 tenha um baixo número de acertos para a objetos com rotação de 30°.

Para objetos com rotação a partir de 30°, apenas as Classes 2 e 3 têm acertos de classificação, enquanto as outras classes não apresentam resultados de classificação.

Comparando com o cenário em que há $n_r = 26$ amostras por classe para a *wavelet* de Daubechies *Db2*, o aumento no número de amostras por classe ($n_r = 130$) contribui para o aumento no número de acertos de classificação para objetos com rotação de 20° e 30°.

Tabela 19: Resultados de classificação de imagens de naipes de cartas. Imagens com rotação. Método bidimensional, *wavelet* de Daubechies *Db8*.

Classificação de Naipes - <i>Wavelet</i> de Daubechies <i>Db8</i> .									
ROTAÇÃO 15°					ROTAÇÃO 20°				
NÍVEL	$n_r = 130$ amostras				NÍVEL	$n_r = 130$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
1 (p=256x256)	N/C	N/C	N/C	N/C	1 (p=256x256)	N/C	N/C	N/C	N/C
2 (p=128x128)	N/C	N/C	N/C	N/C	2 (p=128x128)	N/C	N/C	N/C	N/C
3 (p=64x64)	N/C	N/C	N/C	N/C	3 (p=64x64)	N/C	N/C	N/C	N/C
4 (p=32x32)	N/C	N/C	N/C	N/C	4 (p=32x32)	N/C	N/C	N/C	N/C
5 (p=16x16)	N/C	N/C	N/C	N/C	5 (p=16x16)	N/C	N/C	N/C	N/C
6 (p=8x8)	N/C	N/C	N/C	N/C	6 (p=8x8)	N/C	N/C	N/C	N/C
ROTAÇÃO 30°					ROTAÇÃO 45°				
NÍVEL	$n_r = 130$ amostras				NÍVEL	$n_r = 130$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
1 (p=256x256)	N/C	N/C	N/C	N/C	1 (p=256x256)	N/C	N/C	N/C	N/C
2 (p=128x128)	N/C	N/C	N/C	N/C	2 (p=128x128)	N/C	N/C	N/C	N/C
3 (p=64x64)	N/C	N/C	N/C	N/C	3 (p=64x64)	N/C	N/C	N/C	N/C
4 (p=32x32)	N/C	N/C	N/C	N/C	4 (p=32x32)	N/C	N/C	N/C	N/C
5 (p=16x16)	N/C	N/C	N/C	N/C	5 (p=16x16)	N/C	N/C	N/C	N/C
6 (p=8x8)	N/C	N/C	N/C	N/C	6 (p=8x8)	N/C	N/C	N/C	N/C

A Tabela 19 apresenta os resultados de acertos de classificação para cenário com matrizes de treinamento com $n_r = 130$ amostras para cada classe.

O método de transformada bidimensional utilizando a *wavelet* de Daubechies *Db8* foi aplicado no processo de extração de características. São 104 objetos para classificação em que apresentam rotação de até 45° .

O processo de classificação utilizando esta *wavelet* para a extração de características apresenta os piores resultados: nenhuma amostra foi classificada corretamente. Em nenhum dos níveis de decomposição houve classificação.

Este comportamento pode estar relacionado à decomposição *wavelet* via *Db8* descaracterizar as representações com menor resolução a cada nível. Isto acaba gerando informação que não define as características da imagem de maneira suficiente para que cada classe seja bem identificada.

5.2.3 Método de classificação via método híbrido

Uma alternativa é apresentada neste trabalho quanto à seleção das variáveis das amostras de treinamento para compor a matriz de treinamento. O método híbrido aplica TWD 2D e TWD 1D às amostras da matriz de treinamento e às amostras a serem classificadas.

O procedimento de transformada *wavelet* 2D é aplicado até o nível 4 de decomposição, quando obtém-se representação da imagem com dimensão de 32x32 pixels. Assim como no método unidimensional, deve-se, primeiramente, transformar as informações das

classes de imagens dos naipes das cartas em vetores. Então, as imagens são convertidas em vetores que são formados pela concatenação das linhas das imagens, por exemplo: $imagem1_{vet} = [L_1, L_2, L_3, \dots, L_{512}]$. Realiza-se este procedimento para todas as amostras das 4 classes (4 naipes). Através da concatenação das linhas da imagem, obtém-se os vetores de amostras que compõem as matrizes de treinamento com 1024 variáveis. A partir dessa representação, é aplicada a transformada *wavelet* unidimensional. As Figuras 33 e 34 apresentam as amostras para as imagens com *flash* e sem *flash*, respectivamente, a partir do 4º nível de TWD 2D. Os objetos de classificação também passam pelo mesmo procedimento: é realizada a transformada *wavelet* 2D do objeto de classificação e concatena-se as linhas desta representação mais grosseira da imagem. Então realiza-se a transformada *wavelet* 1D do vetor que representa o objeto de classificação.

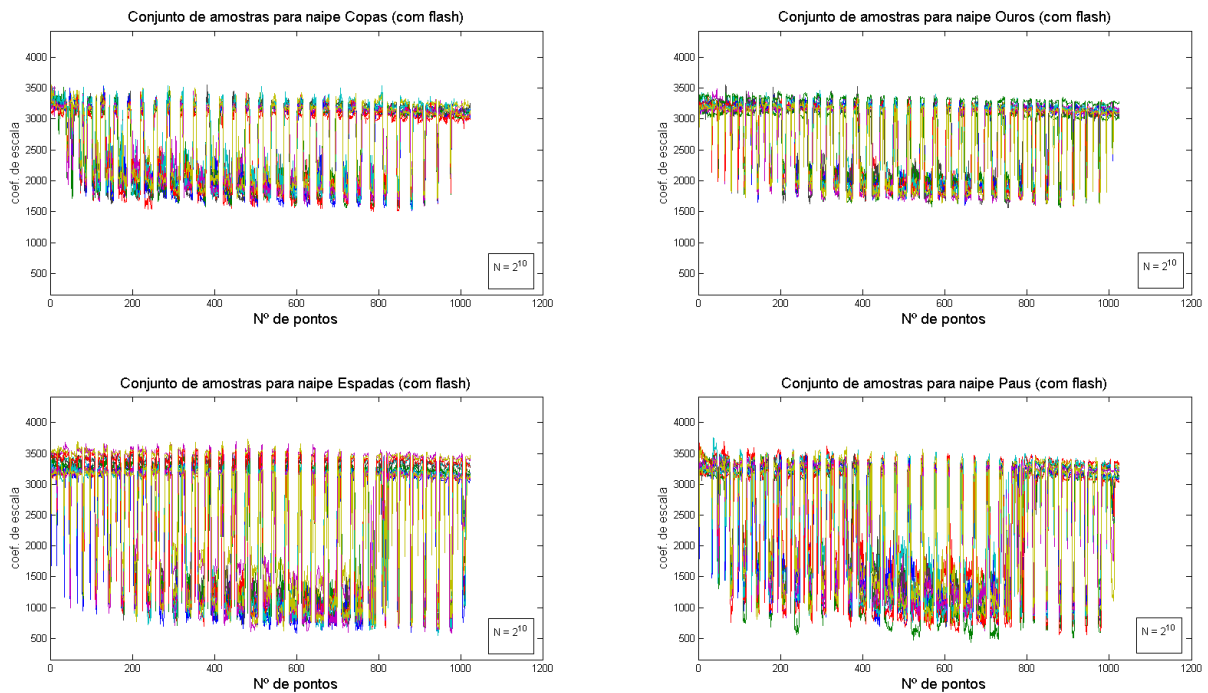


Figura 33: Representação das 4 classes de naipes de cartas no 4º nível de TWD 2D, a partir de imagens com *flash*. Então, as imagens têm as linhas concatenadas em um vetor.

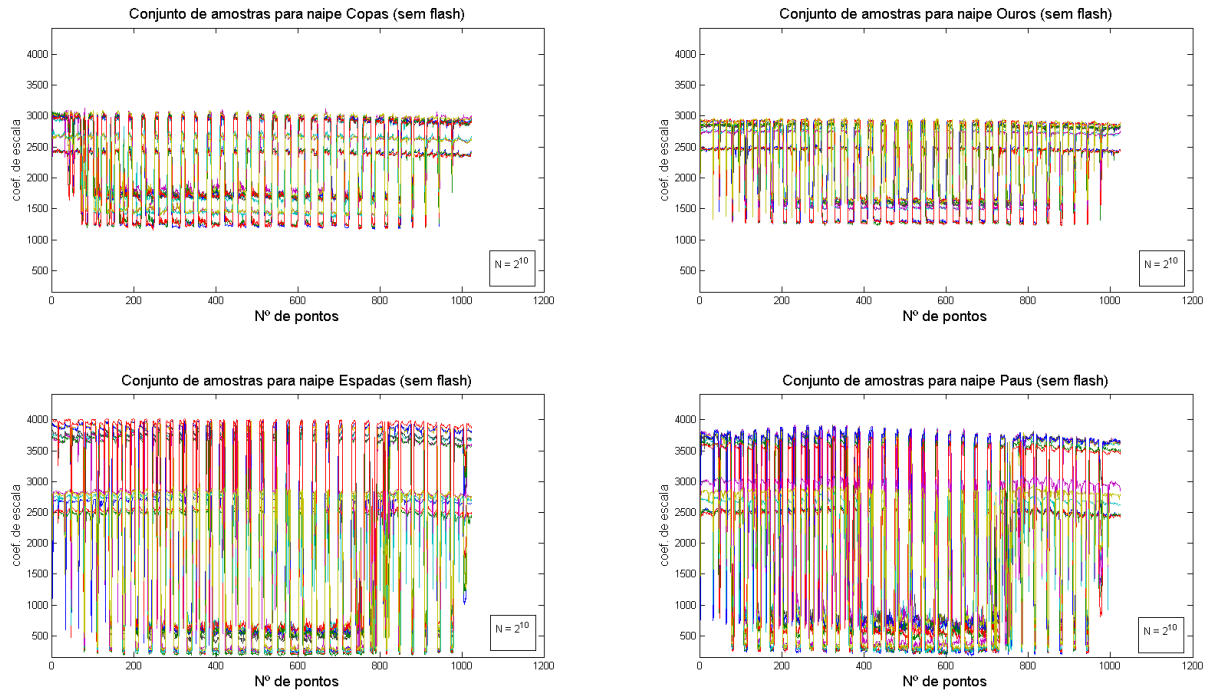


Figura 34: Representação das 4 classes de naipes de cartas no 4º nível de TWD 2D, a partir de imagens sem *flash*. Então, as imagens têm as linhas concatenadas em um vetor.

Tabela 20: Resultados de classificação de imagens de naipes de cartas. Imagens com rotação. Método híbrido, *wavelet* de Haar.

Classificação de Naipes - Wavelet de Haar.									
ROTAÇÃO 0°					ROTAÇÃO 10°				
NÍVEL	$n_r = 26$ amostras				NÍVEL	$n_r = 26$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
1 ($p = 512$)	N/C	N/C	N/C	N/C	1 ($p = 512$)	N/C	N/C	N/C	N/C
2 ($p = 256$)	N/C	N/C	N/C	N/C	2 ($p = 256$)	N/C	N/C	N/C	N/C
3 ($p = 128$)	100,00%	N/C	N/C	N/C	3 ($p = 128$)	N/C	N/C	N/C	N/C
4 ($p = 64$)	100,00%	100,00%	100,00%	100,00%	4 ($p = 64$)	26,92%	42,31%	N/C	N/C
5 ($p = 32$)	100,00%	100,00%	100,00%	100,00%	5 ($p = 32$)	100,00%	100,00%	100,00%	23,08%
6 ($p = 16$)	100,00%	100,00%	100,00%	100,00%	6 ($p = 16$)	100,00%	100,00%	100,00%	26,92%
7 ($p = 8$)	100,00%	100,00%	100,00%	100,00%	7 ($p = 8$)	100,00%	100,00%	100,00%	38,46%
ROTAÇÃO 15°					ROTAÇÃO 20°				
NÍVEL	$n_r = 26$ amostras				NÍVEL	$n_r = 26$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
1 ($p = 512$)	N/C	N/C	N/C	N/C	1 ($p = 512$)	N/C	N/C	N/C	N/C
2 ($p = 256$)	N/C	N/C	N/C	N/C	2 ($p = 256$)	N/C	N/C	N/C	N/C
3 ($p = 128$)	N/C	N/C	N/C	N/C	3 ($p = 128$)	N/C	N/C	N/C	N/C
4 ($p = 64$)	N/C	N/C	N/C	N/C	4 ($p = 64$)	N/C	N/C	N/C	N/C
5 ($p = 32$)	100,00%	84,62%	88,46%	0,00%	5 ($p = 32$)	100,00%	57,69%	0,00%	0,00%
6 ($p = 16$)	100,00%	100,00%	38,46%	0,00%	6 ($p = 16$)	100,00%	96,15%	0,00%	0,00%
7 ($p = 8$)	100,00%	100,00%	30,77%	0,00%	7 ($p = 8$)	100,00%	100,00%	0,00%	0,00%
ROTAÇÃO 30°					ROTAÇÃO 45°				
NÍVEL	$n_r = 26$ amostras				NÍVEL	$n_r = 26$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
1 ($p = 512$)	N/C	N/C	N/C	N/C	1 ($p = 512$)	N/C	N/C	N/C	N/C
2 ($p = 256$)	N/C	N/C	N/C	N/C	2 ($p = 256$)	N/C	N/C	N/C	N/C
3 ($p = 128$)	N/C	N/C	N/C	N/C	3 ($p = 128$)	N/C	N/C	N/C	N/C
4 ($p = 64$)	N/C	N/C	N/C	N/C	4 ($p = 64$)	N/C	N/C	N/C	N/C
5 ($p = 32$)	96,15%	46,15%	N/C	0,00%	5 ($p = 32$)	26,92%	50,00%	N/C	N/C
6 ($p = 16$)	42,31%	88,46%	0,00%	0,00%	6 ($p = 16$)	3,85%	100,00%	0,00%	0,00%
7 ($p = 8$)	100,00%	38,46%	0,00%	0,00%	7 ($p = 8$)	50,00%	38,46%	0,00%	0,00%

A Tabela 20 apresenta os resultados de acertos de classificação para um cenário de classificação com matrizes de treinamento com $n_r = 26$ amostras para cada classe. Um conjunto de 104 objetos são classificados e eles têm rotação de até 45°.

Para os resultados desta tabela, o método híbrido para extração de características foi aplicado utilizando a *wavelet* de Haar.

Os resultados de classificação para as Classes 1, 2 e 3 apresentam maior regularidade até rotação de 15° , principalmente para os 5° e 6° níveis de decomposição *wavelet*.

A Classe 4, a partir de rotações maiores que 15° apresenta muitos erros de classificação. O mesmo acontece para a Classe 3 para objetos com rotação maiores que 20° .

Mesmo com diminuição dos acertos de classificação para as Classes 1 e 2 para rotações maiores que 15° , o método híbrido de extração de características via *wavelet* de Haar apresenta melhores resultados de classificação em relação ao método unidimensional utilizando a mesma *wavelet*.

Observa-se esse aumento de dos acertos de classificação principalmente pela classificação de objetos com 10° de rotação, em que todas as classes apresentam resultado de classificação. No método unidimensional, a classificação das Classes 1, 2 e 4 já era ineficiente para objetos com rotação a partir de 10° . A classificação dos objetos no método bidimensional tem resultado semelhante para amostras com dimensão $p = 8 \times 8 = 64$ variáveis, ao comparar a decomposição via *wavelet* de Haar.

Tabela 21: Resultados de classificação de imagens de naipes de cartas. Imagens com rotação. Método híbrido, *wavelet* de Daubechies *Db2*.

Classificação de Naipes - Wavelet de Daubechies <i>Db2</i> .									
ROTAÇÃO 0°					ROTAÇÃO 10°				
NÍVEL	$n_r = 26$ amostras				NÍVEL	$n_r = 26$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
1 ($p = 512$)	N/C	N/C	N/C	N/C	1 ($p = 512$)	N/C	N/C	N/C	N/C
2 ($p = 256$)	N/C	N/C	N/C	N/C	2 ($p = 256$)	N/C	N/C	N/C	N/C
3 ($p = 128$)	100,00%	N/C	N/C	100,00%	3 ($p = 128$)	N/C	N/C	N/C	N/C
4 ($p = 64$)	100,00%	100,00%	100,00%	100,00%	4 ($p = 64$)	50,00%	50,00%	N/C	N/C
5 ($p = 32$)	100,00%	100,00%	100,00%	100,00%	5 ($p = 32$)	100,00%	96,15%	100,00%	42,31%
6 ($p = 16$)	100,00%	100,00%	100,00%	100,00%	6 ($p = 16$)	100,00%	96,15%	100,00%	76,92%
7 ($p = 8$)	100,00%	100,00%	100,00%	100,00%	7 ($p = 8$)	100,00%	100,00%	100,00%	84,62%
ROTAÇÃO 15°					ROTAÇÃO 20°				
NÍVEL	$n_r = 26$ amostras				NÍVEL	$n_r = 26$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
1 ($p = 512$)	N/C	N/C	N/C	N/C	1 ($p = 512$)	N/C	N/C	N/C	N/C
2 ($p = 256$)	N/C	N/C	N/C	N/C	2 ($p = 256$)	N/C	N/C	N/C	N/C
3 ($p = 128$)	N/C	N/C	N/C	N/C	3 ($p = 128$)	N/C	N/C	N/C	N/C
4 ($p = 64$)	N/C	N/C	N/C	N/C	4 ($p = 64$)	N/C	N/C	N/C	N/C
5 ($p = 32$)	100,00%	96,15%	65,38%	0,00%	5 ($p = 32$)	100,00%	96,15%	0,00%	0,00%
6 ($p = 16$)	100,00%	96,15%	88,46%	15,38%	6 ($p = 16$)	100,00%	88,46%	7,69%	0,00%
7 ($p = 8$)	100,00%	96,15%	57,69%	19,23%	7 ($p = 8$)	100,00%	65,38%	3,85%	0,00%
ROTAÇÃO 30°					ROTAÇÃO 45°				
NÍVEL	$n_r = 26$ amostras				NÍVEL	$n_r = 26$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
1 ($p = 512$)	N/C	N/C	N/C	N/C	1 ($p = 512$)	N/C	N/C	N/C	N/C
2 ($p = 256$)	N/C	N/C	N/C	N/C	2 ($p = 256$)	N/C	N/C	N/C	N/C
3 ($p = 128$)	N/C	N/C	N/C	N/C	3 ($p = 128$)	N/C	N/C	N/C	N/C
4 ($p = 64$)	N/C	N/C	N/C	N/C	4 ($p = 64$)	N/C	N/C	N/C	N/C
5 ($p = 32$)	100,00%	53,85%	N/C	N/C	5 ($p = 32$)	26,92%	38,46%	N/C	N/C
6 ($p = 16$)	88,46%	57,69%	0,00%	0,00%	6 ($p = 16$)	11,54%	65,38%	N/C	N/C
7 ($p = 8$)	100,00%	7,69%	0,00%	0,00%	7 ($p = 8$)	53,85%	0,00%	0,00%	0,00%

A Tabela 21 apresenta os resultados de acertos de classificação para um cenário de classificação com matrizes de treinamento com $n_r = 26$ amostras para cada classe.

Para obter os resultados apresentados na Tabela 21, o método híbrido para extração de características foi aplicado utilizando a *wavelet* de Daubechies *Db2*.

Os resultados de classificação para as Classes 1, 2 e 3 apresentam maior regularidade até rotação de 15°, principalmente para os níveis 7 e 8 da decomposição. A Classe 4 apresenta queda de acertos de classificação para objetos com rotação de 15° e tem poucos acertos para amostras formadas com pequeno número de variáveis (níveis 7 e 8 de decomposição *wavelet*).

A partir de rotações maiores que 15°, as Classes 3 e 4 apresentam aumento de erros de classificação, e com rotação de 30° não têm classificação correta de nenhuma amostra.

Embora há diminuição do número de acertos de classificação para as Classes 1 e 2 para rotações maiores que 30°, a extração de características via *wavelet* de Daubechies *Db2* apresenta melhores resultados de classificação em relação a *wavelet* de Haar. Ainda observa-se o aumento de acertos de classificação para objetos com rotação até 15° em relação ao método unidimensional com a *wavelet* Daubechies *Db2*. Analisando a classificação dos objetos no método bidimensional utilizando *wavelet* de Daubechies *Db2* (Tabela 15), os resultados de classificação são muito semelhantes para objetos com dimensão $p = 8 \times 8 = 64$ variáveis.

Tabela 22: Resultados de classificação de imagens de naipes de cartas. Imagens com rotação. Método híbrido, *wavelet* de Daubechies *Db8*.

Classificação de Naipes - Wavelet de Daubechies <i>Db8</i> .									
ROTAÇÃO 0°					ROTAÇÃO 10°				
NÍVEL	$n_r = 26$ amostras				NÍVEL	$n_r = 26$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
1 ($p = 512$)	N/C	N/C	N/C	N/C	1 ($p = 512$)	N/C	N/C	N/C	N/C
2 ($p = 256$)	N/C	N/C	N/C	N/C	2 ($p = 256$)	N/C	N/C	N/C	N/C
3 ($p = 128$)	N/C	N/C	N/C	N/C	3 ($p = 128$)	N/C	N/C	N/C	N/C
4 ($p = 64$)	100,00%	100,00%	100,00%	100,00%	4 ($p = 64$)	34,62%	46,15%	N/C	N/C
5 ($p = 32$)	100,00%	100,00%	100,00%	100,00%	5 ($p = 32$)	100,00%	100,00%	100,00%	15,38%
6 ($p = 16$)	100,00%	100,00%	100,00%	100,00%	6 ($p = 16$)	100,00%	96,15%	100,00%	26,92%
7 ($p = 8$)	100,00%	100,00%	100,00%	100,00%	7 ($p = 8$)	100,00%	100,00%	100,00%	38,46%
ROTAÇÃO 15°					ROTAÇÃO 20°				
NÍVEL	$n_r = 26$ amostras				NÍVEL	$n_r = 26$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
1 ($p = 512$)	N/C	N/C	N/C	N/C	1 ($p = 512$)	N/C	N/C	N/C	N/C
2 ($p = 256$)	N/C	N/C	N/C	N/C	2 ($p = 256$)	N/C	N/C	N/C	N/C
3 ($p = 128$)	N/C	N/C	N/C	N/C	3 ($p = 128$)	N/C	N/C	N/C	N/C
4 ($p = 64$)	N/C	N/C	N/C	N/C	4 ($p = 64$)	N/C	N/C	N/C	N/C
5 ($p = 32$)	100,00%	84,62%	96,15%	0,00%	5 ($p = 32$)	100,00%	61,54%	7,69%	0,00%
6 ($p = 16$)	100,00%	96,15%	76,92%	0,00%	6 ($p = 16$)	100,00%	76,92%	3,85%	0,00%
7 ($p = 8$)	100,00%	96,15%	53,85%	0,00%	7 ($p = 8$)	100,00%	57,69%	7,69%	0,00%
ROTAÇÃO 30°					ROTAÇÃO 45°				
NÍVEL	$n_r = 26$ amostras				NÍVEL	$n_r = 26$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
1 ($p = 512$)	N/C	N/C	N/C	N/C	1 ($p = 512$)	N/C	N/C	N/C	N/C
2 ($p = 256$)	N/C	N/C	N/C	N/C	2 ($p = 256$)	N/C	N/C	N/C	N/C
3 ($p = 128$)	N/C	N/C	N/C	N/C	3 ($p = 128$)	N/C	N/C	N/C	N/C
4 ($p = 64$)	N/C	N/C	N/C	N/C	4 ($p = 64$)	N/C	N/C	N/C	N/C
5 ($p = 32$)	53,85%	53,85%	N/C	N/C	5 ($p = 32$)	0,00%	7,69%	N/C	N/C
6 ($p = 16$)	42,31%	61,54%	0,00%	0,00%	6 ($p = 16$)	0,00%	76,92%	N/C	N/C
7 ($p = 8$)	100,00%	30,77%	0,00%	0,00%	7 ($p = 8$)	42,31%	7,69%	0,00%	0,00%

A Tabela 22 apresenta os resultados de acertos de classificação para um cenário com matrizes de treinamento com $n_r = 26$ amostras para cada classe.

O método híbrido para extração de características foi aplicado utilizando a *wavelet* de Daubechies *Db8*. As amostras de treinamento e objetos de classificação têm inicialmente 1024 variáveis.

Os resultados de classificação para as Classes 1, 2 e 3 apresentam regularidade dos acertos de classificação até rotação de 15° . A Classe 4 não apresenta acertos de classificação para objetos com rotação a partir de 15° .

A partir de rotação de 20° , a Classe 3 tem grande queda no número de acertos classificação. Para a classificação de objetos com rotação de 45° todas as Classes apresentam queda ou nenhum acerto de classificação.

Em comparação com a *wavelet* Daubechies *Db2*, apresenta regularidade semelhante nos acertos de classificação.

Os resultados obtidos com método híbrido para esta *wavelet* são muito semelhantes aos resultados obtidos com o método unidimensional, diferenciando-se principalmente para a classificação de objetos sem rotação que no método unidimensional apresentaram grande acerto de classificação para dimensão $p = 128$. Comparando com o método bidimensional para dimensão $p = 64$, os resultados apresentados na Tabela 22 são melhores.

Para avaliação desta metodologia de extração de características aplicando TWD 2D e TWD 1D consecutivamente, também foi avaliado o desempenho da classificação em cenários com mais amostras de treinamento. Para aumentar a quantidade de amostras de treinamento para cada classe, são inseridas novas amostras com rotação em seu eixo de $\pm 5^\circ$ e $\pm 10^\circ$, assim totalizando em $n_r = 130$ amostras de treinamento por classe.

Tabela 23: Resultados de classificação de imagens de naipes de cartas. Imagens com rotação. Método híbrido, *wavelet* de Haar.

Classificação de Naipes - <i>Wavelet</i> de Haar.									
ROTAÇÃO 15°					ROTAÇÃO 20°				
NÍVEL	$n_r = 130$ amostras				NÍVEL	$n_r = 130$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
1 ($p = 512$)	N/C	N/C	N/C	N/C	1 ($p = 512$)	N/C	N/C	N/C	N/C
2 ($p = 256$)	N/C	N/C	N/C	N/C	2 ($p = 256$)	N/C	N/C	N/C	N/C
3 ($p = 128$)	N/C	N/C	N/C	N/C	3 ($p = 128$)	N/C	N/C	N/C	N/C
4 ($p = 64$)	100,00%	100,00%	100,00%	100,00%	4 ($p = 64$)	100,00%	100,00%	100,00%	88,46%
5 ($p = 32$)	100,00%	100,00%	100,00%	100,00%	5 ($p = 32$)	100,00%	100,00%	100,00%	92,31%
6 ($p = 16$)	100,00%	100,00%	100,00%	100,00%	6 ($p = 16$)	100,00%	100,00%	100,00%	80,77%
7 ($p = 8$)	100,00%	100,00%	100,00%	100,00%	7 ($p = 8$)	100,00%	100,00%	100,00%	76,92%
ROTAÇÃO 30°					ROTAÇÃO 45°				
NÍVEL	$n_r = 130$ amostras				NÍVEL	$n_r = 130$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
1 ($p = 512$)	N/C	N/C	N/C	N/C	1 ($p = 512$)	N/C	N/C	N/C	N/C
2 ($p = 256$)	N/C	N/C	N/C	N/C	2 ($p = 256$)	N/C	N/C	N/C	N/C
3 ($p = 128$)	N/C	N/C	N/C	N/C	3 ($p = 128$)	N/C	N/C	N/C	N/C
4 ($p = 64$)	65,38%	84,62%	N/C	N/C	4 ($p = 64$)	N/C	N/C	N/C	N/C
5 ($p = 32$)	100,00%	100,00%	3,85%	0,00%	5 ($p = 32$)	100,00%	88,46%	N/C	N/C
6 ($p = 16$)	92,31%	100,00%	7,69%	0,00%	6 ($p = 16$)	38,46%	100,00%	0,00%	0,00%
7 ($p = 8$)	100,00%	100,00%	61,54%	0,00%	7 ($p = 8$)	46,15%	92,31%	11,54%	0,00%

A Tabela 23 apresenta os resultados de acertos de classificação para um cenário com matrizes de treinamento com $n_r = 130$ amostras para cada classe.

Para os resultados desta tabela, o método híbrido para extração de características foi aplicado utilizando a *wavelet* de Haar. São 104 objetos para classificação, com diferentes graus de rotação.

Para graus de rotação de 15° e 20°, os resultados apresentam grande número de acertos de classificação para todas as classes.

A partir da classificação de objetos com 30°, a Classe 4 apresenta redução total nos acertos de classificação. A Classe 3 reduz o grau de acerto de classificação a partir da rotação de 30°.

Para todos os graus de rotação as Classes 1 e 2 apresentam regularidade na classificação, com grande taxas de acerto em diferentes níveis de decomposição.

Importante observar que para o método unidimensional, a *wavelet* de Haar teve os piores resultados de classificação com muito poucos acertos para a Classe 2.

Para o método híbrido aplicando *wavelet* de Haar, os resultados são bem diferentes e atinge alto grau de acerto quando comparado com método unidimensional. Comparando com o método bidimensional para dimensão $p = 64$, os resultados são muito semelhantes, embora para a dimensão $p = 128$ variáveis o método bidimensional apresenta bons resultados para a classificação de objetos com 15° e 20°.

Tabela 24: Resultados de classificação de imagens de naipes de cartas. Imagens com rotação. Método híbrido, *wavelet* de Daubechies *Db2*.

Classificação de Naipes - Wavelet de Daubechies <i>Db2</i> .									
ROTAÇÃO 15°					ROTAÇÃO 20°				
NÍVEL	$n_r = 130$ amostras				NÍVEL	$n_r = 130$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
1 ($p = 512$)	N/C	N/C	N/C	N/C	1 ($p = 512$)	N/C	N/C	N/C	N/C
2 ($p = 256$)	N/C	N/C	N/C	N/C	2 ($p = 256$)	N/C	N/C	N/C	N/C
3 ($p = 128$)	N/C	N/C	N/C	N/C	3 ($p = 128$)	N/C	N/C	N/C	N/C
4 ($p = 64$)	100,00%	100,00%	100,00%	100,00%	4 ($p = 64$)	100,00%	100,00%	100,00%	92,31%
5 ($p = 32$)	100,00%	100,00%	100,00%	100,00%	5 ($p = 32$)	100,00%	100,00%	96,15%	92,31%
6 ($p = 16$)	100,00%	100,00%	100,00%	100,00%	6 ($p = 16$)	100,00%	100,00%	100,00%	92,31%
7 ($p = 8$)	100,00%	100,00%	100,00%	100,00%	7 ($p = 8$)	100,00%	100,00%	100,00%	57,69%
ROTAÇÃO 30°					ROTAÇÃO 45°				
NÍVEL	$n_r = 130$ amostras				NÍVEL	$n_r = 130$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
1 ($p = 512$)	N/C	N/C	N/C	N/C	1 ($p = 512$)	N/C	N/C	N/C	N/C
2 ($p = 256$)	N/C	N/C	N/C	N/C	2 ($p = 256$)	N/C	N/C	N/C	N/C
3 ($p = 128$)	N/C	N/C	N/C	N/C	3 ($p = 128$)	N/C	N/C	N/C	N/C
4 ($p = 64$)	100,00%	100,00%	N/C	N/C	4 ($p = 64$)	19,23%	0,00%	N/C	N/C
5 ($p = 32$)	96,15%	100,00%	0,00%	0,00%	5 ($p = 32$)	96,15%	80,77%	0,00%	0,00%
6 ($p = 16$)	100,00%	100,00%	65,38%	0,00%	6 ($p = 16$)	100,00%	100,00%	0,00%	0,00%
7 ($p = 8$)	100,00%	100,00%	46,15%	0,00%	7 ($p = 8$)	84,62%	61,54%	0,00%	0,00%

A Tabela 24 apresenta os resultados de acertos de classificação para um cenário de classificação com matrizes de treinamento com $n_r = 130$ amostras para cada classe.

O método híbrido para extração de características foi aplicado utilizando a *wavelet* de Daubechies *Db2*. São 104 objetos para classificação, com diferentes graus de rotação.

Para graus de rotação de 15° e 20°, os resultados apresentam grande número de acertos para classificação das amostras em todas as classes, principalmente para os 5º, 6º e 7º níveis de decomposição.

A partir de 30°, a Classe 4 não apresenta acertos de classificação para objetos com rotação.

A Classe 3 tem redução dos acertos para objetos com rotação de 45°. Para todos os graus de rotação as Classes 1 e 2 apresentam grande taxas de acerto.

Em comparação com o método unidimensional de extração de características com *wavelet Db2* tem resultados muito semelhantes para a classificação dos objetos com rotação, analisando as mesmas dimensões dos objetos. Em comparação com o método bidimensional, apenas a dimensão $p = 64$ pode ser comparada e em ambos métodos têm bons resultados.

Tabela 25: Resultados de classificação de imagens de naipes de cartas. Imagens com rotação. Método híbrido, *wavelet* de Daubechies *Db8*.

Classificação de Naipes - <i>Wavelet</i> de Daubechies <i>Db8</i> .									
ROTAÇÃO 15°					ROTAÇÃO 20°				
NÍVEL	$n_r = 130$ amostras				NÍVEL	$n_r = 130$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
1 ($p = 512$)	N/C	N/C	N/C	N/C	1 ($p = 512$)	N/C	N/C	N/C	N/C
2 ($p = 256$)	N/C	N/C	N/C	N/C	2 ($p = 256$)	N/C	N/C	N/C	N/C
3 ($p = 128$)	N/C	N/C	N/C	N/C	3 ($p = 128$)	N/C	N/C	N/C	N/C
4 ($p = 64$)	100,00%	100,00%	100,00%	100,00%	4 ($p = 64$)	100,00%	100,00%	100,00%	88,46%
5 ($p = 32$)	100,00%	100,00%	100,00%	100,00%	5 ($p = 32$)	100,00%	100,00%	100,00%	76,92%
6 ($p = 16$)	100,00%	100,00%	100,00%	100,00%	6 ($p = 16$)	100,00%	100,00%	100,00%	65,38%
7 ($p = 8$)	100,00%	100,00%	100,00%	100,00%	7 ($p = 8$)	100,00%	100,00%	100,00%	76,92%
ROTAÇÃO 30°					ROTAÇÃO 45°				
NÍVEL	$n_r = 130$ amostras				NÍVEL	$n_r = 130$ amostras			
	Cl. 1	Cl. 2	Cl. 3	Cl. 4		Cl. 1	Cl. 2	Cl. 3	Cl. 4
1 ($p = 512$)	N/C	N/C	N/C	N/C	1 ($p = 512$)	N/C	N/C	N/C	N/C
2 ($p = 256$)	N/C	N/C	N/C	N/C	2 ($p = 256$)	N/C	N/C	N/C	N/C
3 ($p = 128$)	N/C	N/C	N/C	N/C	3 ($p = 128$)	N/C	N/C	N/C	N/C
4 ($p = 64$)	57,69%	88,46%	N/C	N/C	4 ($p = 64$)	N/C	N/C	N/C	N/C
5 ($p = 32$)	84,62%	100,00%	53,85%	0,00%	5 ($p = 32$)	46,15%	76,92%	N/C	N/C
6 ($p = 16$)	84,62%	100,00%	15,38%	0,00%	6 ($p = 16$)	38,46%	88,46%	0,00%	0,00%
7 ($p = 8$)	100,00%	100,00%	84,62%	0,00%	7 ($p = 8$)	46,15%	69,23%	3,85%	0,00%

A Tabela 25 apresenta os resultados de acertos de classificação para um cenário de classificação com matrizes de treinamento com $n_r = 130$ amostras para cada classe.

O método híbrido para extração de características foi aplicado utilizando a *wavelet* de Daubechies *Db8*. São 104 objetos para classificação.

Assim como para a *wavelet* Daubechies *Db2* (resultados apresentados anteriormente), para objetos com graus de rotação de 15° e 20°, os resultados apresentam grande efetividade para classificação das amostras em todas as classes.

A Classe 4 não apresenta acertos de classificação para objetos com rotação a partir de 30°.

A Classe 3 tem redução dos acertos para objetos com rotação de 45°.

Para graus de rotação de 30° e 45°, as Classes 1 e 2 apresentam taxas de acerto menores quando comparados aos resultados com *wavelet* Daubechies *Db2*.

Comparando com o método unidimensional de extração de características com *wavelet* *Db8*, para todos os níveis em que foram possíveis realizar a classificação, os resultados são muito semelhantes para a classificação dos objetos com rotação. Comparando com o método bidimensional, analisando objetos com dimensão $p = 64$, os resultados são muito diferentes.

5.2.4 Análise dos Resultados de Classificação de Dados Bidimensionais

Foram aplicados 3 métodos de extração de características via transformada *wavelet* para a redução de dimensão: (a) aplicação de transformada unidimensional aos dados agrupados de tal modo que o vetor de entrada é composto pelas linhas da matriz da imagem, (b) aplicação de transformada bidimensional aos dados da matriz de imagem, e (c) uma abordagem híbrida, na qual os coeficientes de escala da transformada bidimensional são ainda analisados por uma transformada unidimensional.

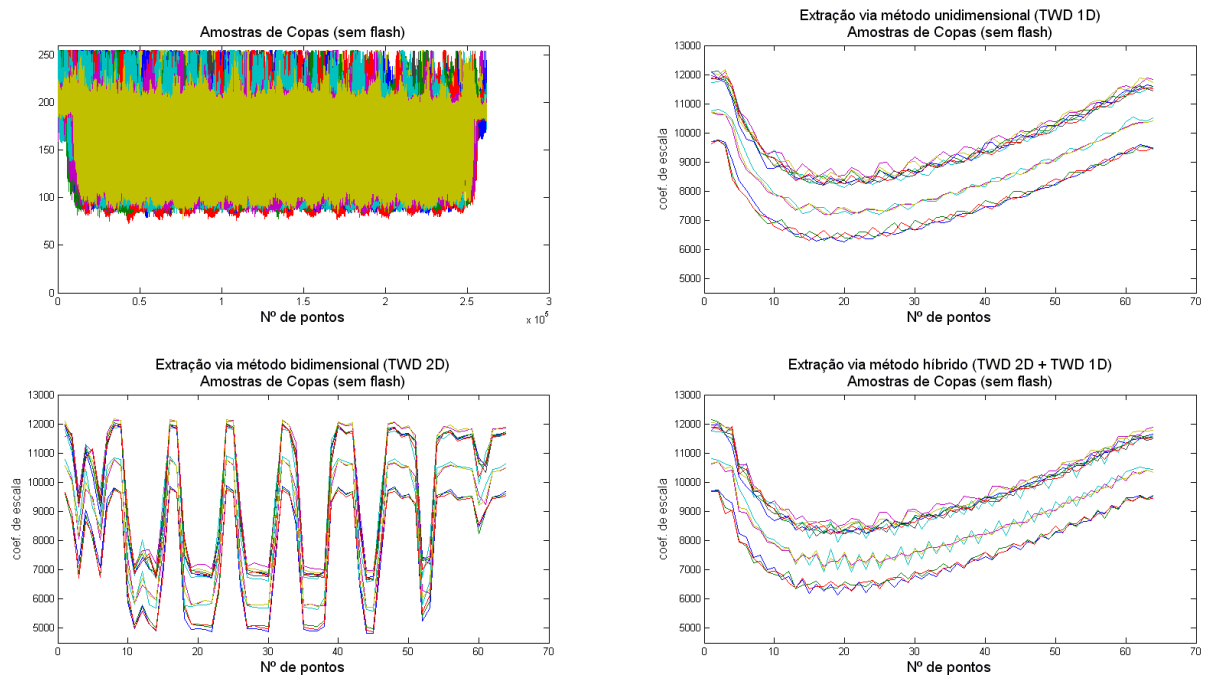


Figura 35: Representação das amostras para a classe Copas, a partir das imagens sem *flash*. Os 3 métodos de extração são apresentados.

As Figuras 35, 36, 37 e 38 mostram os resultados de extração de atributos através dos 3 métodos utilizados neste trabalho para cada conjunto de amostras das 4 diferentes classes. Nestas figuras, observa-se a representação das amostras com $p = 64$ pontos e observa-se principalmente a diferença do processo de extração de atributos via método bidimensional (TWD 2D) para os métodos unidimensional (TWD 1D) e método híbrido. Estes últimos 2 métodos, para a representação com $p = 64$ pontos, apresentam curvas de extração de atributos semelhantes, sendo que para o método unidimensional este resultado é obtido após o 12^o nível de decomposição e para o método híbrido foram necessários 4 níveis de decomposição através da TWD 2D e 4 níveis de decomposição através da TWD 1D. O método bidimensional executa somente a TWD 2D e os coeficientes resultantes de cada nível formam as matrizes de treinamento.

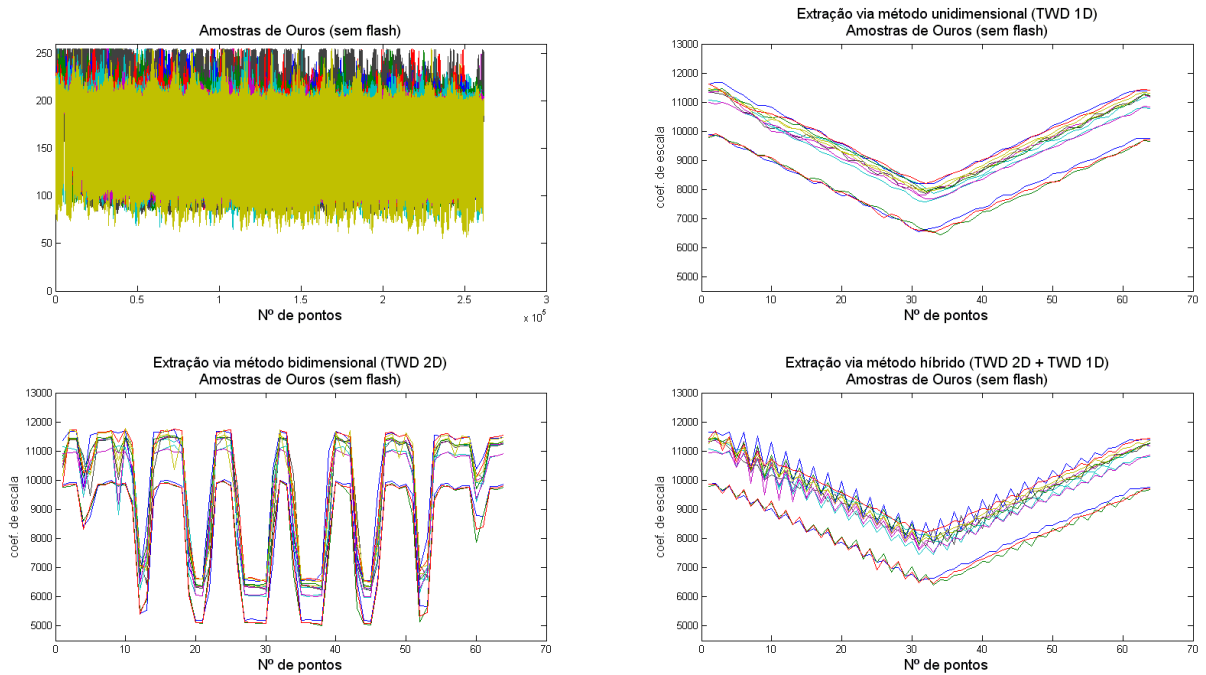


Figura 36: Representação das amostras para a classe Ouros, a partir das imagens sem *flash*. Os 3 métodos de extração são apresentados.

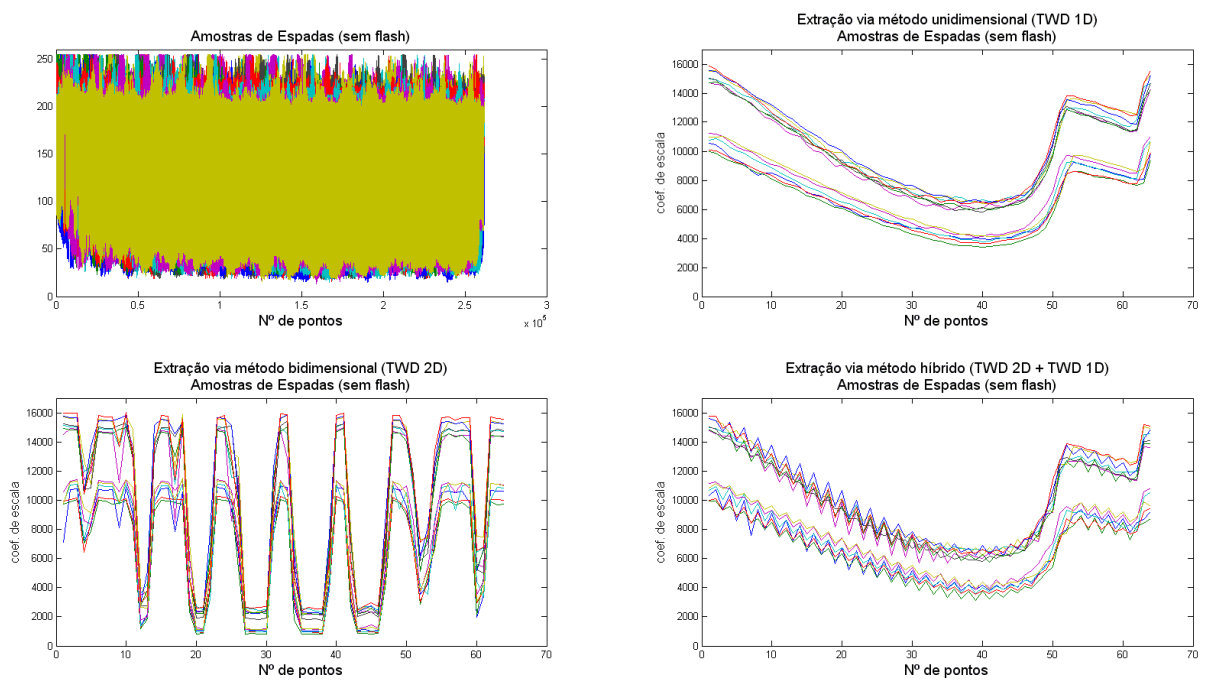


Figura 37: Representação das amostras para a classe Espadas, a partir das imagens sem *flash*. Os 3 métodos de extração são apresentados.

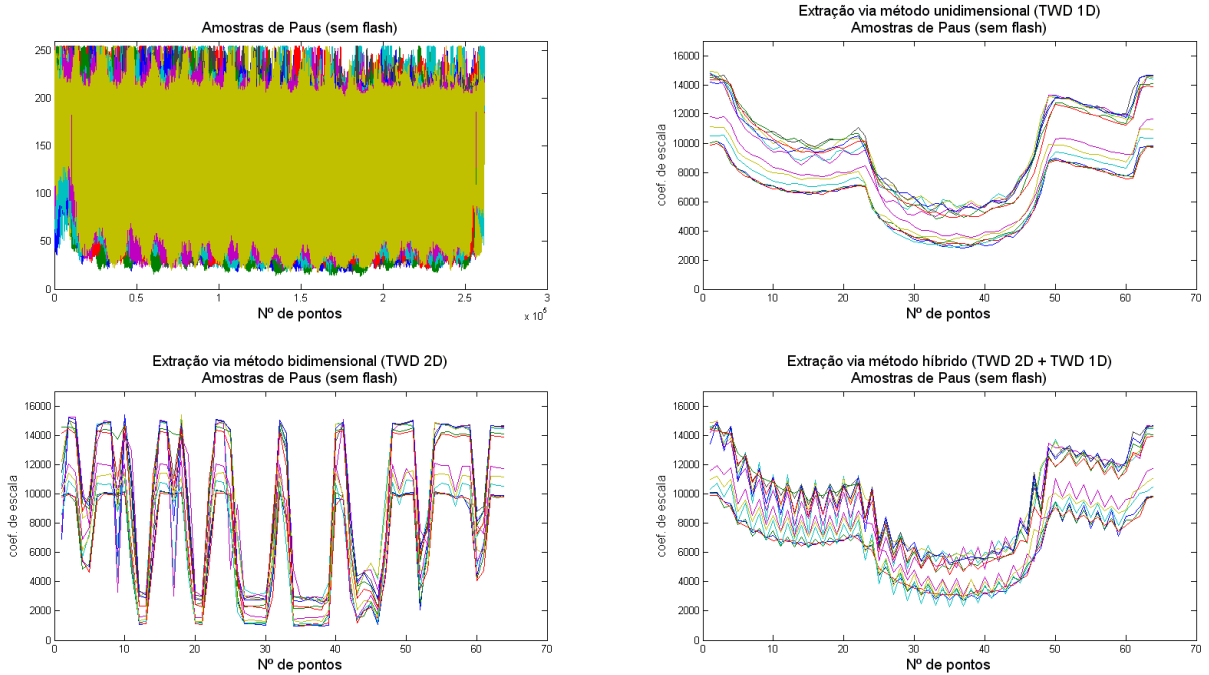


Figura 38: Representação das amostras para a classe Paus, a partir das imagens sem *flash*. Os 3 métodos de extração são apresentados.

Para o cenário com $n_r = 26$ amostras de treinamento para cada classe, os três métodos para a extração de atributos foram utilizados para a redução da dimensão das amostras. Para cada nível da transformada *wavelet* os resultados de classificação são avaliados. A Figura 39 apresenta as taxas de erros de classificação de padrões formados por atributos extraídos através do método unidimensional. Na Figura 40 são apresentados os resultados de erros de classificação de padrões resultantes da extração de atributos através do método bidimensional, e na Figura 41 são apresentadas as taxas de erros de classificação para padrões formados por características extraídas via método híbrido. Para o método unidimensional, os erros de classificação são menores para amostras com dimensões menores que 64 pontos e, em geral, a diminuição de erros de classificação está relacionada com as funções *wavelet* (*Db2* e *Db8*) que têm um número maior de coeficientes para os filtros.

Através do método bidimensional, a redução de dimensão entre os níveis de decomposição *wavelet* é maior (através de TWD 1D, a cada nível reduz-se à metade a dimensão, enquanto para TWD 2D reduz-se à razão 4) e a classificação é realizada para dimensões $p=64=8 \times 8$ e $p=4 \times 4=16$ pontos. Para este método a função Haar produz menos erros de classificação para dimensão $p = 16$ pontos para cada amostra e esta característica pode estar relacionada ao fato das transformadas com as funções *wavelet* *Db2* e *Db8* utilizarem extensões periódicas. Para estas funções, o cálculo dos coeficientes das fronteiras das ima-

gens necessita de informações fora do domínio e estes, quando aproximados por valores tomados de forma periódica, podem descaracterizar as informações originais das imagens e dificultar a classificação dos objetos.

Quanto maior o número de coeficientes para os filtros das funções *wavelet* periodizadas, maior a quantidade de informações inseridas artificialmente para o cálculo dos coeficientes das fronteiras, e conseqüentemente maior erro de classificação. E essas informações são levadas adiante em cada nível da decomposição *wavelet*. Nota-se que para a função *Db8* não é realizada a análise de resultados para a dimensão $p = 16$ pontos para as amostras, porque esta função tem 16 coeficientes nos seus filtros, e para dimensões menores do que 8 pontos não é possível realizar a operação da transformada.

Ao observar os resultados para o método híbrido, os erros de classificação entre as diferentes funções *wavelet* são semelhantes. A diminuição de erros de classificação é observada para as dimensões menores que $p = 64$ pontos. Nota-se a diminuição dos erros de classificação para a função Haar, quando comparado ao método unidimensional. No método bidimensional esta função apresenta os menores resultados de erros e, como o método híbrido aplica TWD 2D e TWD 1D (consecutivamente) às amostras do classificador, esta etapa de aplicação de TWD 2D acrescenta qualidade no processo de extração de atributos e influencia diretamente no processo de extração via TWD 1D.

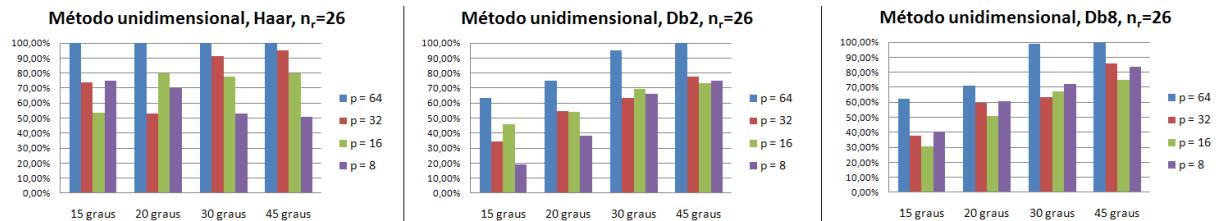


Figura 39: Erros de classificação para cenários com $n_r = 26$ amostras de treinamento por classe. Método unidimensional usado para extração de atributos.

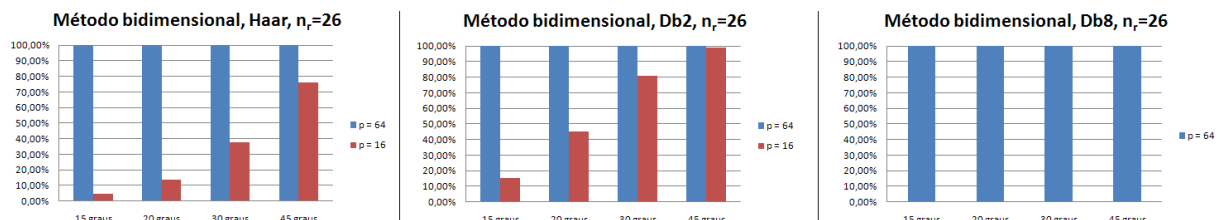


Figura 40: Erros de classificação para cenários com $n_r = 26$ amostras de treinamento por classe. Método bidimensional usado para extração de atributos.

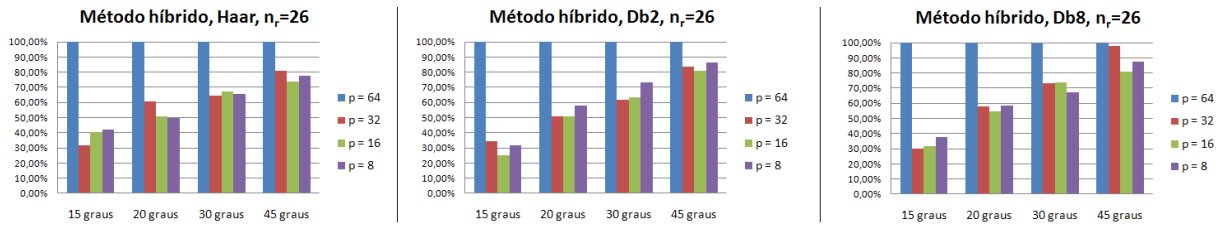


Figura 41: Erros de classificação para cenários com $n_r = 26$ amostras de treinamento por classe. Método híbrido aplicado para extração de atributos.

Os erros de classificação para cenário com $n_r = 130$ amostras para as matrizes de treinamento das classes também são avaliados. Na Figura 42 são apresentadas as taxas de erros de classificação de padrões formados por características extraídas pelo método unidimensional. Na Figura 43 são apresentados os erros de classificação de padrões representados por características extraídas via método bidimensional, e na Figura 44 são apresentadas as taxas de erros de classificação para padrões formados por atributos extraídos através do método híbrido.

O simples aumento de amostras de treinamento, com a inserção de novas amostras que apresentam variabilidade, faz com que a classificação dos dados apresente menos erros. Os conjuntos de atributos que são extraídos fazem o processo de classificação mais eficiente para as imagens com rotação. Observa-se isto na diminuição dos erros de classificação para os três métodos de extração de atributos.

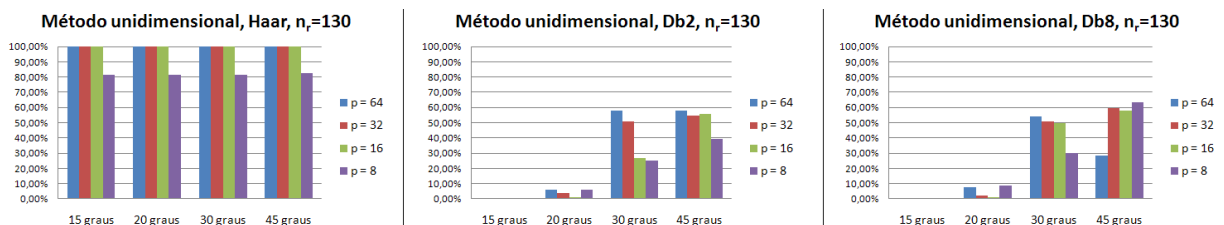


Figura 42: Erros de classificação para cenários com $n_r = 130$ amostras de treinamento por classe. Método unidimensional aplicado no processo de extração de atributos.

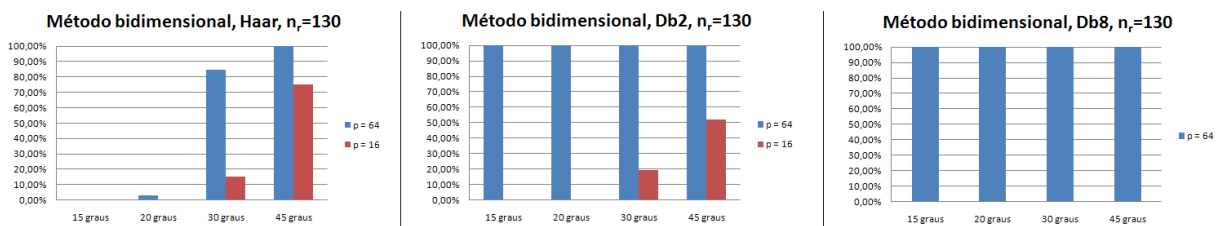


Figura 43: Erros de classificação para cenários com $n_r = 130$ amostras de treinamento por classe. Método bidimensional usado para extração de atributos.

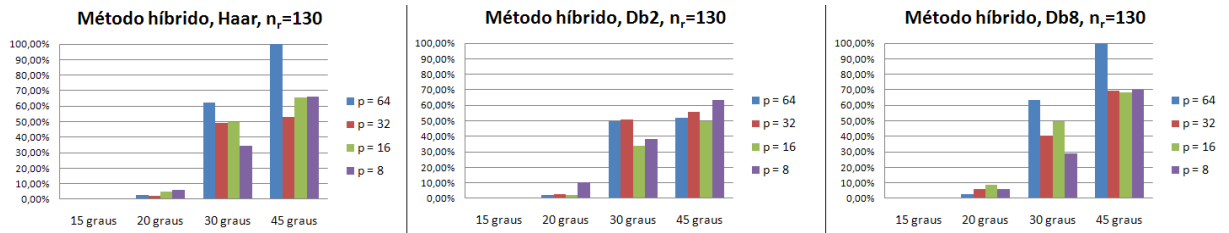


Figura 44: Erros de classificação para cenários com $n_r = 130$ amostras de treinamento por classe. Método híbrido usado para extração de atributos.

Para o método unidimensional, a função Haar produz muitos erros de classificação, enquanto que para o método bidimensional o comportamento é o contrário, inclusive apresentando muitos acertos para amostras com dimensão $p = 64$ pontos. Para os processos de classificação em que as extrações de atributos são realizadas através das funções *Db2* e *Db8* apresentam os menores erros de classificação para o método unidimensional e híbrido, embora o maior ganho de eficiência de classificação (diminuição dos erros) esteja na extração de atributos pela função Haar no método híbrido. Assim como para o cenário com $n_r = 26$ amostras de treinamento, o método híbrido de extração de atributos aplicando Haar resulta em um conjunto de características que discriminam as classes de forma mais eficiente e, conseqüentemente, contribui para que o processo de classificação tenha menos erros.

6 CONCLUSÕES

Neste trabalho foi estudado um algoritmo de classificação Bayesiano cujos dados de entrada são os coeficientes de escala da expansão em série *wavelet* do conjunto de dados originais a ser classificado. Para validação do algoritmo e verificação da contribuição da transformada *wavelet* na redução de dimensão dos dados, foram analisados problemas com dados unidimensionais e bidimensionais.

Este trabalho apresenta a transformada *wavelet* como ferramenta para o pré-processamento de dados para a classificação de padrões. Para o processo de extração de características são apresentadas 3 abordagens: o primeiro método, o método unidimensional, aplica a transformada *wavelet* 1D e é utilizado principalmente para dados unidimensionais e é adaptado para a classificação de dados bidimensionais; o segundo método, chamado método bidimensional, aplica a transformada *wavelet* 2D para a extração de características e é usado para a classificação de dados bidimensionais; o terceiro método é chamado de método híbrido e aplica a transformada *wavelet* 2D e transformada *wavelet* 1D consecutivamente para a extração das características dos dados.

Para a classificação dos dados unidimensionais são utilizadas duas diferentes fontes de dados. O primeiro conjunto de dados, formado a partir de 3 equações matemáticas apresenta grande variabilidade no conjunto de amostras, e é um conjunto totalmente gerado sinteticamente. O segundo conjunto de dados tem origem em observação de comportamento de tráfego de rede de computadores, e nesta observação buscou-se analisar a presença de anomalias no tráfego (ataques à rede). Amostras adicionais foram geradas afim de possibilitar a avaliação de classificação dos padrões. Essas novas amostras têm grande correlação com as amostras originais.

Observou-se durante a execução de classificação em ambos conjuntos de dados unidimensionais a importância em manter equilibrada a relação entre o número de amostras de treinamento e a dimensão das amostras. A transformada *wavelet* colabora para a execução do algoritmo de classificação, porque auxilia na redução da dimensão das amostras através da extração de suas características via coeficientes da transformação.

O efeito da transformada é percebido a cada nível da decomposição, quando se obtém a dimensão de amostras que permite realizar as operações com as matrizes de covariância das classes e permite executar os cálculos de probabilidade de objetos pertencentes a determinada classe. Ou seja, permitindo que a classificação seja executada com sucesso.

Na classificação dos dados bidimensionais, observou-se que os resultados de classificação apenas são obtidos a partir da grande redução na dimensão das amostras de treinamento e objetos de classificação. Inicialmente, para amostras com 262144 pontos e com poucos representantes para cada classe, a classificação não foi possível. A alternativa então analisada foi a aplicação de transformadas *wavelets* e a partir dos coeficientes produzidos pela transformada, reavaliar as etapas do classificador.

O método híbrido apresenta-se como uma alternativa promissora para a classificação de imagens, possibilitando aumento de desempenho de classificação em cenários com reduzido número de amostras de treinamento, quando comparado aos métodos unidimensional e bidimensional para a extração de características.

6.1 Trabalhos Futuros

Ao longo deste estudo apenas os coeficientes de escala foram utilizados como valores de entrada para a construção das matrizes de variância e covariância e, portanto, para o algoritmo de classificação. Uma possibilidade de investigação seria analisar o comportamento do classificador quando um certo nível de coeficientes *wavelets* (coeficientes de detalhe) fossem utilizados para gerar essas matrizes e, conseqüentemente, para alimentarem o processo de classificação.

No caso específico de serem usados coeficientes *wavelets* (coeficientes de detalhe) para serem dados de entrada do algoritmo de classificação, várias perguntas relacionadas a isso seriam relevantes, como por exemplo, se algum nível de resolução da transformada seria mais eficiente do que outro, ou se conjuntos de coeficientes truncados poderiam oferecer uma melhor maneira de cálculo para as matrizes de variância e covariância dos dados.

Neste trabalho a aplicação da família de *wavelets* ortonormais de Daubechies se mostrou bastante estável quanto a ferramenta de pré-processamento, na extração de características dos dados. Caberia investigar se outras famílias de *wavelets* aplicadas ao processo de extração de características poderiam aumentar o desempenho do classificador.

REFERÊNCIAS

- Al-Kadi 2009 AL-KADI, O. A fractal dimension based optimal wavelet packet analysis technique for classification of meningioma brain tumours. In: Image Processing (ICIP), 2009 16th IEEE International Conference on. [S.l.: s.n.], 2009. p. 4177 –4180. ISSN 1522-4880.
- Beltran et al. 2006 BELTRAN, N.; DUARTEMERMOUD, M.; BUSTOS, M.; SALAH, S.; LOYOLA, E.; PENANEIRA, A.; JALOCHA, J. Feature extraction and classification of chilean wines. Journal of Food Engineering, Vol. 75, n. No. 1, p. 1–10, July 2006.
- Bruce, Koger e Li 2002 BRUCE, L.; KOGER, C.; LI, J. Dimensionality reduction of hyperspectral data using discrete wavelet transform feature extraction. Geoscience and Remote Sensing, IEEE Transactions on, v. 40, n. 10, p. 2331 – 2338, out. 2002. ISSN 0196-2892.
- Cavalcanti et al. 2002 CAVALCANTI, G. D. C.; DÓRIA, R. C.; ADEODATO, P. J. L.; FILHO, E. C. B. C. Verificação off-line de assinaturas de tamanhos diferentes. In: Neural Networks, 2002. SBRN 2002. Proceedings. VII Brazilian Symposium on. [S.l.: s.n.], 2002. p. 30–35.
- Dan, Yuan-yuan e Chen-xiang 2010 DAN, L.; YUAN-YUAN, H.; CHEN-XIANG, M. Feature extraction and classification of proteomics data using stationary wavelet transform and naive bayes classifier. In: Bioinformatics and Biomedical Engineering (iCBBE), 2010 4th International Conference on. [S.l.: s.n.], 2010. p. 1 –4. ISSN 2151-7614.
- Daubechies 1992 DAUBECHIES, I. Ten lectures on wavelets. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 1992. ISBN 0-89871-274-2.
- Duda e Hart 1973 DUDA, R. O.; HART, P. E. Book. Pattern classification and scene analysis. [S.l.]: Wiley New York,, 1973. xvii, 482 p. p. ISBN 0471223611.
- Figueiredo 1997 FIGUEIREDO, D. G. de. Análise de Fourier e Equações Diferenciais Parciais. 4th. ed. [S.l.]: Publicação IMPA, 1997. 274 p. ISBN: 85-244-0026-9.
- Goumas, Zervakis e Stavrakakis 2002 GOUMAS, S.; ZERVAKIS, M.; STAVRAKAKIS, G. Classification of washing machines vibration signals using discrete wavelet analysis for feature extraction. Instrumentation and Measurement, IEEE Transactions on, v. 51, n. 3, p. 497 –508, jun. 2002. ISSN 0018-9456.
- Jain e Chandrasekaran 1982 JAIN, A.; CHANDRASEKARAN, B. Dimensionality and sample size considerations in pattern recognition practice. In: KRISHNAIAH, P.; KANAL, L. (Ed.). Classification Pattern Recognition and Reduction of Dimensionality. Elsevier, 1982, (Handbook of Statistics, v. 2). p. 835 – 855. Disponível em: <<http://www.sciencedirect.com/science/article/B7P6H-4FD6Y9M-1G/2/52470e655b5da3fddcdc2754d09969cf>>.

- Jain, Duin e Mao 2000 JAIN, A. K.; DUIN, R. P. W.; MAO, J. Statistical pattern recognition: A review. IEEE Transactions on Pattern Analysis and Machine Intelligence, v. 22, p. 4–37, 2000.
- Keren 2002 KEREN, D. Painter identification using local features and naive bayes. In: Pattern Recognition, 2002. Proceedings. 16th International Conference on. [S.l.: s.n.], 2002. v. 2, p. 474 – 477 vol.2. ISSN 1051-4651.
- Liping, XinYuan e Yuqiang 2009 LIPING, N.; XINYUAN, L.; YUQIANG, D. Bayesian face recognition using wavelet transform. In: Intelligent Human-Machine Systems and Cybernetics, 2009. IHMSC '09. International Conference on. [S.l.: s.n.], 2009. v. 2, p. 312–315.
- Liu 2003 LIU, C. A bayesian discriminating features method for face detection. Pattern Analysis and Machine Intelligence, IEEE Transactions on, v. 25, n. 6, p. 725 – 740, 2003. ISSN 0162-8828.
- Liu 2004 LIU, J. Process monitoring using bayesian classification on pca subspace. Industrial & Engineering Chemistry Research, v. 43, n. 24, p. 7815–7825, 2004. Disponível em: <<http://pubs.acs.org/doi/abs/10.1021/ie0498495>>.
- Mallat 1989 MALLAT, S. A theory for multiresolution signal decomposition: the wavelet representation. Pattern Analysis and Machine Intelligence, IEEE Transactions on, v. 11, n. 7, p. 674 –693, jul. 1989. ISSN 0162-8828.
- Mallet et al. 1997 MALLET, Y.; COOMANS, D.; KAUTSKY, J.; VEL, O. de. Classification using adaptive wavelets for feature extraction. IEEE Transactions on Pattern Analysis and Machine Intelligence, v. 19, p. 1058–1066, 1997.
- MIT 1999 MIT, M. I. o. T. DARPA Intrusion Detection Data Sets. 1999.
- Moraes e Haertel 2007 MORAES, D. A. de O.; HAERTEL, V. Métodos hierárquicos para redução de dimensões e classificação de imagens aviris. In: Anais XIII Simpósio Brasileiro de Sensoriamento Remoto. Florianópolis/SC, Brasil: [s.n.], 2007. p. 6481–6488.
- Narote, Narote e Waghmare 2009 NAROTE, S. P.; NAROTE, A. S.; WAGHMARE, L. M. Iris based recognition system using wavelet transform. IJCSNS International Journal of Computer Science and Network Security, Vol. 9, N.º 11, p. 101–104, 2009.
- Nielsen 1998 NIELSEN, O. M. Wavelets in Scientific Computing. Tese (Doutorado) — Informatics and Mathematical Modelling, Technical University of Denmark, DTU, 1998.
- Pei et al. 1995 PEI, M.; GOODMAN, E. D.; III, W. F. P.; DING, Y. Genetic algorithms for classification and feature extraction. In: Annual Meeting, Classification Society of North America. [S.l.: s.n.], 1995.
- Pittner e Kamarthi 1997 PITTNER, S.; KAMARTHI, S. Feature extraction from wavelet coefficients for pattern recognition tasks. In: Neural Networks,1997., International Conference on. [S.l.: s.n.], 1997. v. 3, p. 1484 –1489 vol.3.
- Prochazka, Kukal e Vysata 2008 PROCHAZKA, A.; KUKAL, J.; VYSATA, O. Wavelet transform use for feature extraction and eeg signal segments classification. In: Communications, Control and Signal Processing, 2008. ISCCSP 2008. 3rd International Symposium on. [S.l.: s.n.], 2008. p. 719 –722.

- Rajpoot e Rajpoot 2004 RAJPOOT, K.; RAJPOOT, N. Wavelets and support vector machines for texture classification. In: Multitopic Conference, 2004. Proceedings of INMIC 2004. 8th International. [S.l.: s.n.], 2004. p. 328 – 333.
- Stollnitz, Derosé e Salesin 1995 STOLLNITZ, E. J.; DEROSE, T. D.; SALESIN, D. H. Wavelets for computer graphics: A primer - part 1. IEEE Computer Graphics and Applications, v. 15, p. 76–84, 1995.
- Stollnitz, Derosé e Salesin 1995 STOLLNITZ, E. J.; DEROSE, T. D.; SALESIN, D. H. Wavelets for computer graphics: A primer - part 2. IEEE Computer Graphics and Applications, v. 15, p. 76–84, 1995.
- Subramani, Sahu e Verma 2006 SUBRAMANI, P.; SAHU, R.; VERMA, S. Feature selection using haar wavelet power spectrum. BMC Bioinformatics, v. 7, n. 1, p. 432, 2006. ISSN 1471-2105. Disponível em: <<http://www.biomedcentral.com/1471-2105/7/432>>.
- Sung, Kwak e Lyou 2010 SUNG, G.-Y.; KWAK, D.-M.; LYOU, J. Neural network based terrain classification using wavelet features. J. Intell. Robotics Syst., Kluwer Academic Publishers, Hingham, MA, USA, v. 59, p. 269–281, September 2010. ISSN 0921-0296. Disponível em: <<http://dx.doi.org/10.1007/s10846-010-9402-2>>.
- Tanenbaum 2002 TANENBAUM, A. S. Computer Networks. 4th edition. ed. [S.l.]: Prentice Hall, 2002. Paperback. ISBN 0130384887.
- Yen e Chu 2008 YEN, E.; CHU, I.-W. M. Image recognition via bayesian likelihood analysis of wavelet coefficients. Expert Syst. Appl., Pergamon Press, Inc., Tarrytown, NY, USA, v. 35, p. 556–560, July 2008. ISSN 0957-4174. Disponível em: <<http://portal.acm.org/citation.cfm?id=1379458.1379561>>.
- Yen e Lin 2000 YEN, G.; LIN, K.-C. Wavelet packet feature extraction for vibration monitoring. Industrial Electronics, IEEE Transactions on, v. 47, n. 3, p. 650 –667, jun. 2000. ISSN 0278-0046.
- Zhu e Schwartz 2002 ZHU, Y.; SCHWARTZ, S. Efficient face detection with multiscale sequential classification. In: Image Processing. 2002. Proceedings. 2002 International Conference on. [S.l.: s.n.], 2002. v. 2, p. II–121 – II–124 vol.2. ISSN 1522-4880.

ANEXO A – FUNÇÕES WAVELETS DA FAMÍLIA DE DAUBECHIES

A.1 Exemplos de funções *wavelet*

Neste anexo são apresentadas as funções escala e *wavelet* da família de Daubechies que foram utilizadas ao longo deste trabalho. A notação considerada é DbN , sendo $N = D/2$. Assim as funções $Db2$ e $Db8$ possuem filtros dados por:

- a *wavelet* Daubechies $Db2$:

$$h_0 = -0.1294095226, h_1 = 0.2241438680,$$

$$h_2 = 0.8365163037, h_3 = 0.4829629131$$

e

$$g_0 = -0.4829629131, g_1 = 0.8365163037,$$

$$g_2 = -0.2241438680, g_3 = -0.1294095226$$

- a *wavelet* Daubechies $Db8$:

$$h_0 = -0.00011747678400228192, h_1 = 0.00067544940599855677,$$

$$h_2 = -0.00039174037299597711, h_3 = -0.0048703529930106603,$$

$$h_4 = 0.0087460940470156547, h_5 = 0.013981027917015516,$$

$$h_6 = -0.044088253931064719, h_7 = -0.017369301002022108,$$

$$h_8 = 0.12874742662018601, h_9 = 0.00047248457399797254,$$

$$h_{10} = -0.28401554296242809, h_{11} = -0.015829105256023893,$$

$$h_{12} = 0.58535468365486909, h_{13} = 0.67563073629801285,$$

$$h_{14} = 0.31287159091446592, h_{15} = 0.054415842243081609,$$

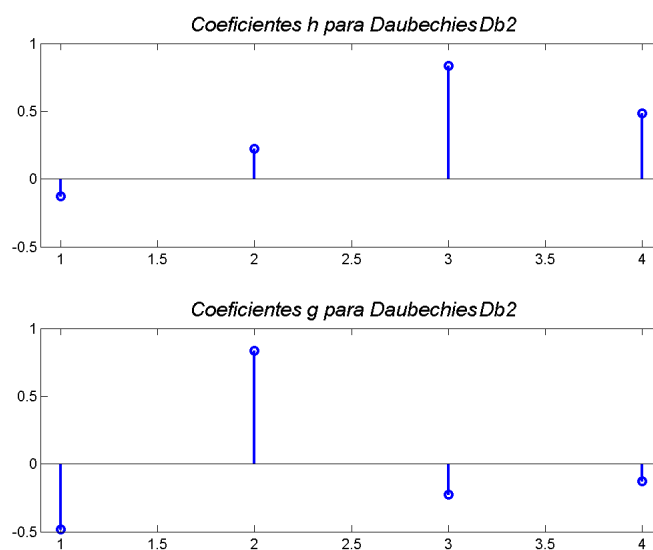


Figura 45: Coeficientes para filtros de decomposição para $Db2$.

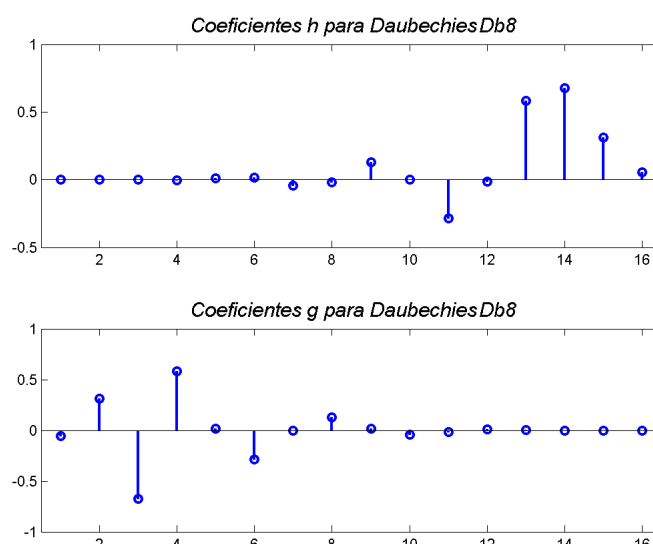


Figura 46: Coeficientes para filtros de decomposição para $Db8$.

e

$$\begin{aligned}
 g_0 &= -0,0544158422430816, & g_1 &= 0,312871590914466, \\
 g_2 &= -0,675630736298013, & g_3 &= 0,585354683654869, \\
 g_4 &= 0,0158291052560239, & g_5 &= -0,284015542962428, \\
 g_6 &= -0,000472484573997973, & g_7 &= 0,128747426620186, \\
 g_8 &= 0,0173693010020221, & g_9 &= -0,0440882539310647, \\
 g_{10} &= -0,0139810279170155, & g_{11} &= 0,00874609404701566, \\
 g_{12} &= 0,00487035299301066, & g_{13} &= -0,000391740372995977,
 \end{aligned}$$

$$g_{14} = -0,000675449405998557, g_{15} = -0,000117476784002282$$

Os gráficos para as funções escala e *wavelet Db2* são apresentados na Figura 47. Para *Db8*, na Figura 48.

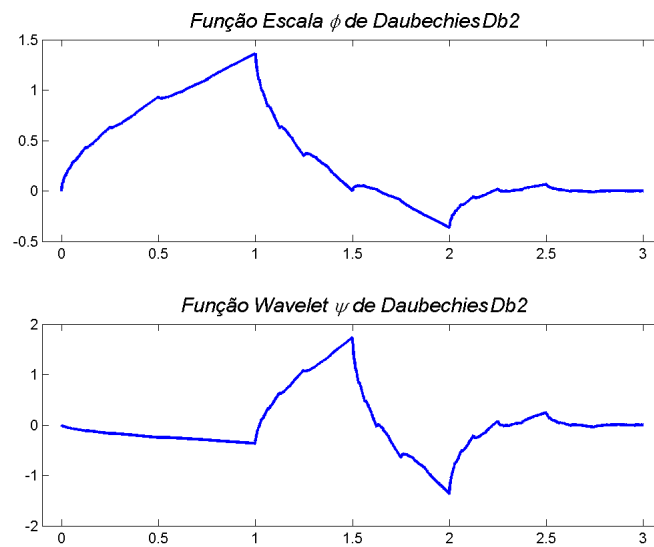


Figura 47: Funções geradoras para *Db2*.

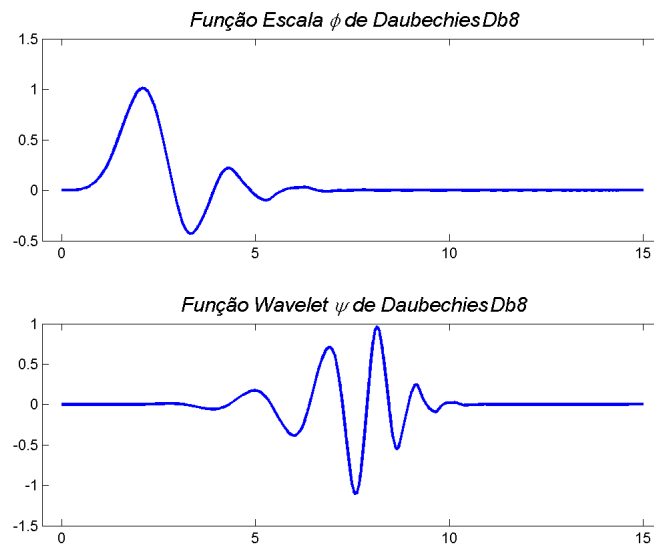


Figura 48: Funções geradoras para *Db8*.

Outro exemplo utilizado aqui são as *wavelets* de Haar, definidas em 1909 por Alfred Haar e compreendidas como *wavelet* ortonormal da família de Daubechies em 1988 (Daubechies 1992)(Nielsen 1998). É uma *wavelet* muito simples em sua implementação

e muito rápida. Ao contrário das *wavelets* de Daubechies, não necessita de nenhum tipo de tratamento especial para analisar dados finitos cujas fronteiras devam ser preservadas. Além disso, produz transformada rápida exatamente inversível, não causando efeitos indesejados nas fronteiras dos sinais. A *wavelet* de Haar opera nos dados através do cálculo de médias e diferenças de elementos adjacentes (opera a cada par de elementos). Na Figura 49 são apresentados os gráficos da função escala, que é a função constante igual à 1 no intervalo $[0,1]$ e da função *wavelet* correspondente, que é a função escada também em $[0,1]$. Observa-se que ambas funções são nulas fora do intervalo $[0,1]$.

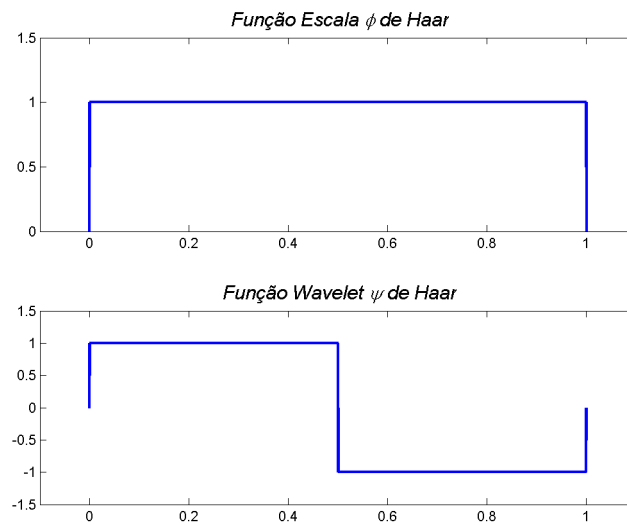


Figura 49: Funções geradoras para Haar.

Os coeficientes de Haar, considerando a correção de $\sqrt{2}$ para que em qualquer escala as funções tenham norma igual à 1, para aplicação de transformada *wavelet* discreta, são os seguintes:

• *wavelet* de Haar:

$$h_0 = 0.7071067812, \quad h_1 = 0.7071067812$$

e

$$g_0 = -0.7071067812, \quad g_1 = 0.7071067812$$

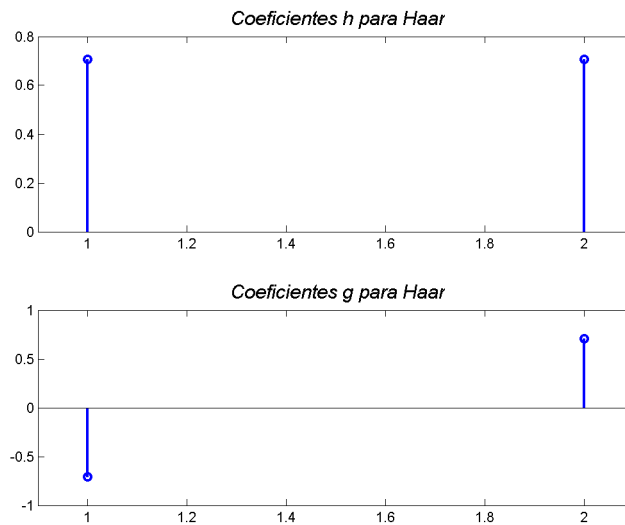


Figura 50: Coeficientes para filtros de decomposição para *Haar*.

A.2 Exemplo de transformada *wavelet* discreta

Para exemplificar o comportamento da transformada discreta *wavelet* e apresentar uma representação da funções em termos de seus coeficientes de escala e *wavelet* em diferentes níveis, consideremos a função

$$f(x) = \begin{cases} x, & 0 \leq x < 0,5 \\ x - 1, & 0,5 < x \leq 1 \end{cases}$$

que apresenta uma descontinuidade bem localizada. Abaixo, na Figura 51, está representada a aproximação por série *wavelet* desta função.

Observa-se nos detalhes (Figura 52) a posição onde a função é descontínua, o conteúdo de alta frequência no local da descontinuidade é captada pela aproximação pela expansão *wavelet*. Nas representações truncadas da função f nota-se que o erro decorrente da aproximação fica localizado na vizinhança onde a função não é contínua, sendo resultado da característica de “localidade” das *wavelets*. O comportamento não suave de f é traduzido pela grande variação dos coeficientes que estão localizados próximos a região da descontinuidade.

Também se observa que a maior parte linear da função f é representada “exatamente”, mesmo em níveis grosseiros de representação em que poucos coeficientes são utilizados para

a aproximação. Isto é manifestação de uma propriedade das *wavelets* chamada momentos nulos que significa que as funções de escala podem localmente representar polinômios de baixa ordem de forma exata.

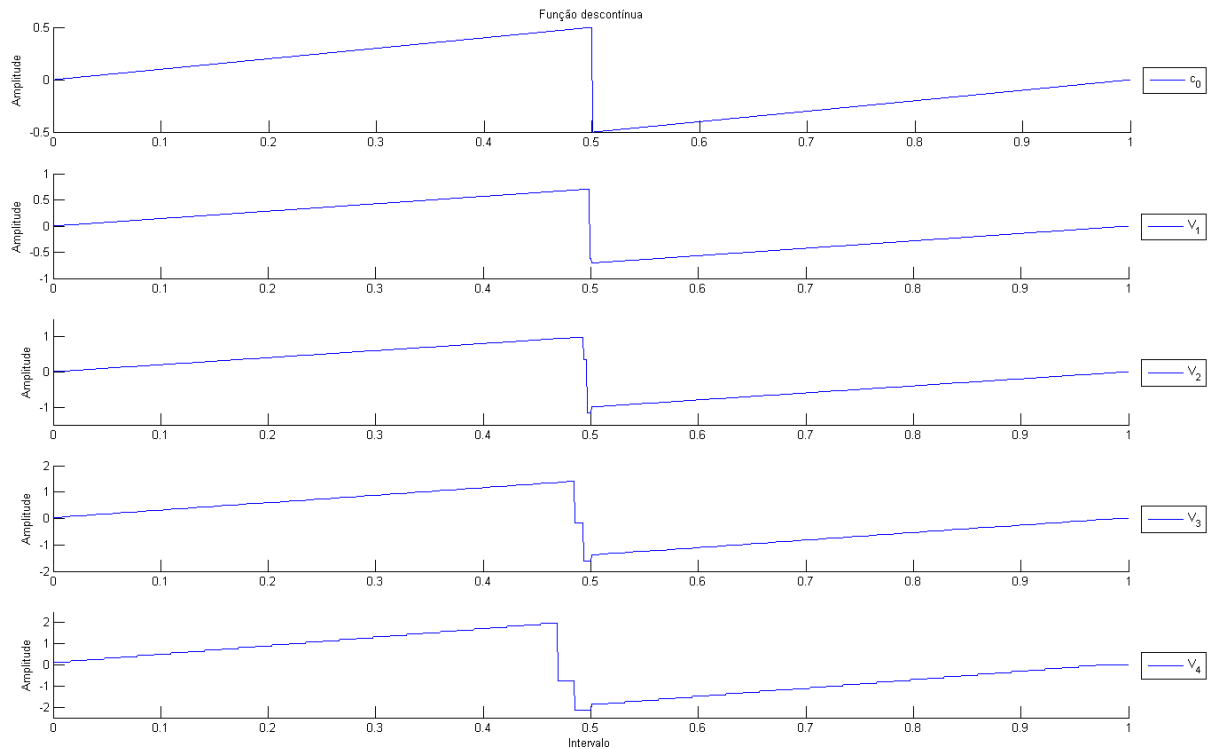


Figura 51: Aproximação para a função $f(x)$, a partir de coeficientes de escala.

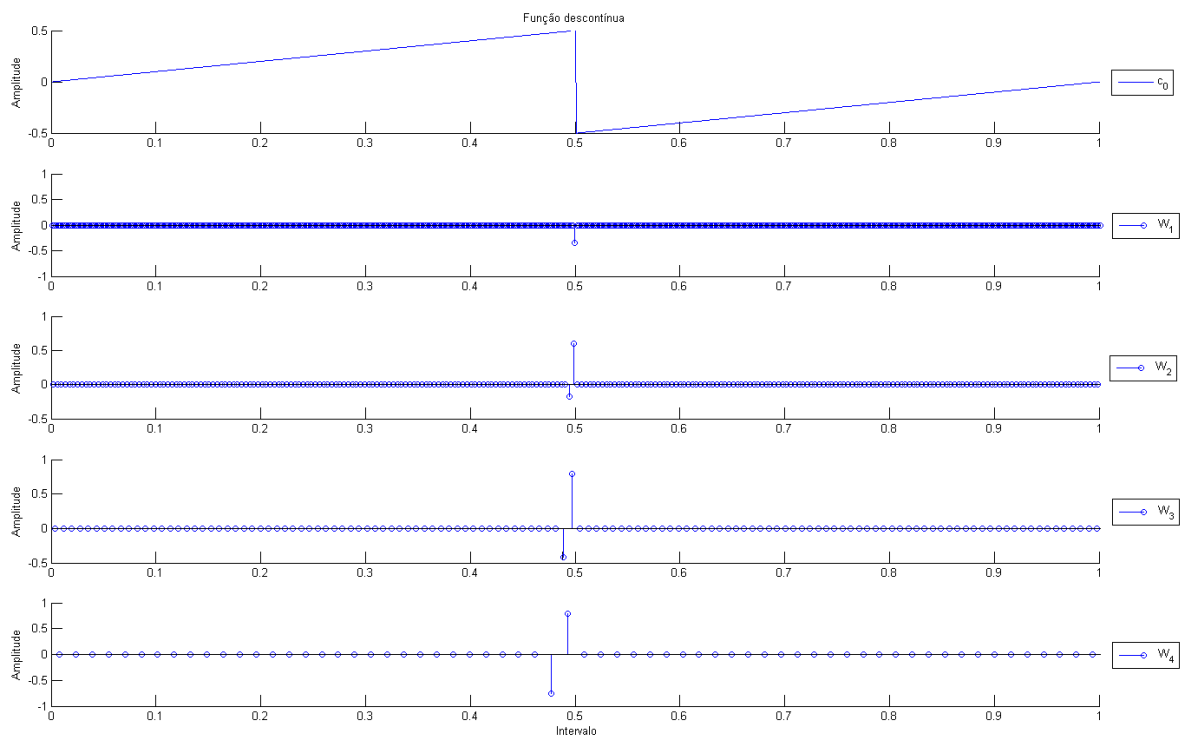


Figura 52: Coeficientes de detalhe para a função $f(x)$.