

Kaiser Magalde Costa Magalhães

**Identificação de golfinhos através de Técnicas de
Processamento Inteligente de Imagens**

**Natal/RN – Brasil
Março de 2003**

Universidade Federal do Rio Grande do Norte
Centro de Tecnologia
Programa de Pós-Graduação em Engenharia Elétrica

Identificação de golfinhos através de Técnicas de Processamento Inteligente de Imagens

Kaiser Magalde Costa Magalhães

Dissertação submetida ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Rio Grande do Norte como parte dos requisitos para a obtenção do grau de Mestre em Ciências.

Orientador: Prof. Dr. Adelardo Adelino Dantas de Medeiros

Co-orientador: Prof. Dr. Pablo Javier Alsina

Natal/RN - Brasil

Março de 2003

MEC UFRN Biblioteca Central "Zila Mamede"
004.932 187.038104
Nº de Chamada Registro
M1981 12102104
Fornecedor
DISSERT
Forma de Aquisição
Dinheiro Par.
Empenho Preço
R\$ 40,00



L187038

Divisão de Serviços Técnicos

Catálogo da Publicação na Fonte. UFRN / Biblioteca Central Zila Mamede

Magalhães, Kaiser Magalde Costa.

Identificação de golfinhos através de técnicas de processamento inteligente de imagens / Kaiser Magalde Costa Magalhães. – Natal, RN, 2003.

99 p.

Orientador : Adelardo Adelino Dantas de Medeiros.

Co-orientador : Pablo Javier Alsina.

Dissertação (Mestrado) – Universidade Federal do Rio Grande do Norte. Centro de Tecnologia. Programa de Pós-Graduação em Engenharia Elétrica.

1. Processamento de imagens - Tese. 2. Golfinhos - Tese. 3. Redes neurais artificiais - Tese. 4. Boto-cinza - Tese. I. Medeiros, Adelardo Adelino Dantas de. II. Alsina, Pablo Javier. III. Título.

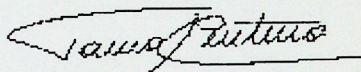
RN/UF/BCZM

CDU 004.932(043.2)

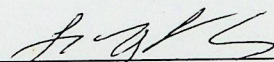
Identificação de Golfinhos através de Técnicas de Processamento Inteligente de Imagens

Kaiser Magalde Costa Magalhães

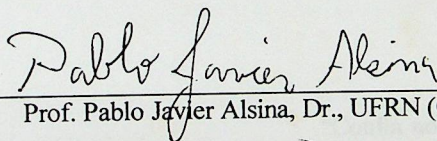
Aprovada em 05 de março de 2003 pela Comissão Examinadora formada pelos seguintes membros:



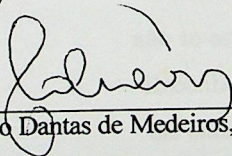
Profª Tania Mezzadri Centeno, Drª, CEFET/PR



Prof. Luiz Marcos Garcia Gonçalves, Dr., UFRN



Prof. Pablo Javier Alsina, Dr., UFRN (Co-orientador)



Prof. Adelardo Adelino Dantas de Medeiros, Dr., UFRN (Orientador)

NATAL, RN, BRASIL
Março de 2003

Agradecimentos

O livro deste trabalho deve agradecer primordialmente ao meu Senhor Jesus por toda a vida que me deu.

A minha esposa que esteve comigo durante todo o período de dedicação a este trabalho.

Às mães, Maria Filipe e Maria Magalhães, que nasceram durante a mesma época de dedicação a este trabalho.

Às mães que me ajudaram durante todas as etapas da minha vida.

Às mães que me ajudaram durante a dedicação a este trabalho: Adelaide e Fátima pela dedicação e apoio durante este trabalho.

Às mães que me ajudaram durante a dedicação a este trabalho: Adelaide e Fátima pela dedicação e apoio durante este trabalho.

Às mães que me ajudaram durante a dedicação a este trabalho: Adelaide e Fátima pela dedicação e apoio durante este trabalho.

Às mães que me ajudaram durante a dedicação a este trabalho: Adelaide e Fátima pela dedicação e apoio durante este trabalho.

Às mães que me ajudaram durante a dedicação a este trabalho: Adelaide e Fátima pela dedicação e apoio durante este trabalho.

Às mães que me ajudaram durante a dedicação a este trabalho: Adelaide e Fátima pela dedicação e apoio durante este trabalho.

“Confia no Senhor de todo o teu coração, e não te estribes no teu próprio entendimento. Reconhece-o em todos os teus caminhos, e ele endireitará as tuas veredas. Não sejas sábio a teus próprios olhos; teme ao Senhor e aparta-te do mal.” **Provérbios 3:5-7**

Agradecimentos

O fruto deste trabalho devo agradecer primordialmente ao meu Senhor Jesus por mais essa vitória que me concedeu.

À minha esposa que soube compreender-me durante todo o período de dedicação a esse trabalho.

Ao meu filho, Mateus Filipe Dantas Magalhães, que nasceu durante a escrita dessa dissertação e trouxe ainda mais regozijo ao nosso lar.

Aos meus pais que sempre me apoiaram durante todas as etapas de minha vida.

Aos meus amigos orientador e co-orientador Adelardo e Pablo pela paciência em ajudar-me durante todo esse trabalho.

Aos professores Adrião e Luiz Marcos pelo apoio em diversos momentos sugerindo técnicas e incentivando meu trabalho.

Ao meu amigo Carlos que juntos pudemos estudar e avançar nos conhecimentos passados por nossos mestres.

Aos meus amigos Diogo e Clauber pela prestividade quando solicitados por mim.

Ao Adolfo pelo grande apoio prestado em conseguir as imagens de golfinhos e dar-me algumas orientações também.

À CAPES, pelo suporte financeiro.

Finalmente a todos que, de alguma forma, contribuíram com este trabalho, mas que aqui não foram citados.

Conteúdo

Lista de Figuras	v
Resumo	xi
Abstract	xiii
1 Introdução	1
1.1 Proposta e Objetivos	2
1.2 Estado da Arte	3
1.3 Motivação	4
1.4 Estrutura do sistema proposto	5
1.5 Organização do trabalho	7
2 Golfinhos	9
2.1 Cetáceos	9
2.2 Boto-cinza	10
2.3 Foto-identificação	12
2.4 Identificação dos golfinhos	12
3 Processamento de Imagens	15
3.1 Representação de imagens	15
3.1.1 Digitalização	16
3.2 Realce de Imagens	17
3.2.1 Histograma	18
3.2.2 Transformada Karhunen-Loeve (Hotelling)	19

3.2.3	Auto-Contraste	21
3.2.4	Limiarização ou Binarização	22
3.3	Segmentação	24
3.3.1	Segmentação por textura	25
3.3.2	Segmentação por quantização vetorial	26
3.4	Morfologia Matemática	27
3.4.1	Teoria de Conjuntos	28
3.4.2	Elemento Estruturante	29
3.4.3	Dilatação	30
3.4.4	Erosão	30
3.4.5	Abertura	30
3.4.6	Extração de Fronteiras	31
3.4.7	Esqueleto	31
3.5	Extração de características	32
3.5.1	Descritores de Textura	32
3.5.2	Descritores de forma	33
4	Redes Neurais Artificiais	51
4.1	Introdução	51
4.2	Redes Neurais Artificiais	52
4.2.1	Arquitetura das Redes Neurais	54
4.2.2	Técnicas de aprendizagem	55
4.3	Modelos de redes neurais	58
4.3.1	Algoritmo de aprendizagem competitiva	58
4.3.2	Quantização Vetorial	61
4.4	Perceptrons de múltiplas camadas	61
4.4.1	Algoritmo de retropropagação	62
4.5	Identificação e reconhecimento	64
5	Sistema para Identificação de Golfinhos	67
5.1	Introdução	67
5.2	Aquisição de Imagens	68

5.3	Pré-processamento	68
5.3.1	Visualização da imagem	68
6	Experimentos e Resultados	75
6.1	Exemplo de processamento	75
6.1.1	Transformada KLT	76
6.1.2	Comparação com o processo do gradiente	77
6.1.3	Auto-contraste	78
6.2	Segmentação por textura	79
6.3	Limiarização ou Binarização	80
6.4	Cortar a imagem	81
6.5	Reflexão da imagem	82
6.6	Rotulação	82
6.7	Operadores Morfológicos	82
6.7.1	Esqueletização	83
6.7.2	Extração de bordas	84
6.8	Extração do Pico da Nadadeira	84
6.9	Sequenciação de Pontos	85
6.10	Obtenção das curvas paramétricas	85
6.11	Detecção de fendas	86
6.12	Normalização	87
6.13	Descritores de Curvatura	87
6.14	Classificação	88
7	Discussões, conclusões e perspectivas	93
	Referências Bibliográficas	94

Lista de Figuras

1.1 Diagrama de bloco do sistema proposto	5
2.1 Golfinho 3D-2D	11
3.1 Imagem de um golfinho	17
3.2 Exemplo de uma transformação KIT	20
3.3 Processo de Auto-contorno aplicado à imagem original	23
3.4 Imagem binária	24
3.5 Sequenciamento do processamento realizado	35
3.6 Golfinho com antena elevada	35
3.7 Extração do pico da nadadeira na imagem	37
3.8 Extraindo a sequência de pontos a partir da fronteira da imagem	38
3.9 Modelo gráfico das curvas paramétricas usadas	39
3.10 Detecção de pontos de inflexão	41
3.11 Duas nadadeiras com raios de curvatura diferentes	43
3.12 Duas nadadeiras com índices de dispersão distintos	44
3.13 Eixos coordenados escalados	44
3.14 Mapeamento entre planos 3D e 2D	45
3.16 Sugestão para normalização dos polígonos	49
4.1 Neurônio biológico	53
4.2 Modelo não-linear de um neurônio	53
4.3 Rede neural tipo feedforward com uma camada	56
4.4 Processo de aprendizagem supervisionada	57
4.5 Processo de aprendizagem não-supervisionada	57

Lista de Figuras

1.1	Diagrama de blocos do sistema proposto	5
2.1	Golfinho boto-cinza	11
3.1	Imagem de um golfinho	17
3.2	Exemplo de uma transformação KLT	21
3.3	Processo de Auto-contraste aplicado à imagem original	23
3.4	Imagem binária	24
3.5	Sequenciamento do processamento realizado	35
3.6	Golfinho com rotação elevada	35
3.7	Extração do pico da nadadeira na imagem	37
3.8	Extraindo a sequência de pontos a partir da fronteira da imagem	38
3.9	Modelo gráfico das curvas paramétricas usadas	39
3.10	Detecção de fendas da nadadeira	41
3.11	Duas nadadeiras com raios descritores distintos	43
3.12	Duas nadadeiras com índices de discrepância distintos	44
3.13	Eixos coordenados escolhidos	44
3.14	Mapeamento entre planos 3D e 2D	45
3.15	Seqüência para normalização dos polinômios	49
4.1	Neurônio biológico	53
4.2	Modelo não-linear de um neurônio	53
4.3	Rede neural tipo feedforward com uma camada	56
4.4	Processo de aprendizagem supervisionada	57
4.5	Processo de aprendizagem não-supervisionado	57

4.6	Modelo de uma rede competitiva com conexões diretas e laterais	59
4.7	Rede perceptron de múltiplas camadas	62
4.8	Topologia da BPNN usada.	65
5.1	Janela principal do sistema	69
5.2	Descrição da janela principal do sistema	69
5.3	Fluxograma para processamento de uma imagem	71
6.1	Imagem original já delimitada	76
6.2	Resultado da aplicação da KLT na imagem	77
6.3	Comparação: pré-processamento com KLT e com gradiente	77
6.4	Aplicação do auto-contraste, após a KLT	78
6.5	Imagem limiarizada	81
6.6	Imagens limiarizada (a), cortada (b) e refletida (c)	82
6.7	Resultados da esqueletização (a) e extração de borada (b)	84
6.8	Seqüenciação de pontos da fronteira da imagem	85
6.9	Representação das curvas polinomiais encontradas	86
6.10	Detecção de fendas	86
6.11	Resultado do processo de normalização	88
6.12	Apresentando o raio de curvatura	88
6.13	Imagens distintas de um mesmo golfinho.	89
6.14	Tempo versus número de golfinhos.	91
6.15	Épocas versus número de golfinhos.	92

Lista de Tabelas

6.1	Ativação dada pela última camada da rede e ativação ideal.	90
6.2	Identificação errada devido ao ruído.	90

Resumo

O presente trabalho investiga o uso da imagem digital aplicado à identificação de golfinhos boto, e as técnicas utilizadas como tratamento e segmentação de imagens, redes neurais, métodos de estatística e transformações geométricas, que são combinadas de forma a possibilitar que imagens de golfinhos sejam tratadas de tal forma a identificar diferentes indivíduos.

Os golfinhos estudados neste trabalho são da espécie *S. fluviatilis*, mais conhecidos como botos-cinza, e habitam uma área de mar compreendida entre as praias de Pipa (RN) e Tabatinga (RN). Eles se caracterizam principalmente por possuírem uma grande nadadeira dorsal. Esta nadadeira funciona como uma espécie de identidade do golfinho, pois são elas responsáveis pelas marcas que estes animais sofrem, formando marcas facilmente distintas. Outra característica importante é a própria forma da curvatura da nadadeira dorsal curvada.

Neste trabalho, foram implementados algoritmos capazes de tratar inicialmente a imagem, através da teoria da KLT e do auto-correlação. Usamos técnicas de morfologia matemática para auxiliar no processo de extração de características e algoritmos de redes neurais para a classificação dos animais cujas imagens foram submetidas ao sistema desenvolvido. Para a identificação da curvatura da nadadeira foram utilizados dois polinômios cúbicos paramétricos associados a cada lado da nadadeira, partindo do ponto de pico. Para isso determinamos em cada curva o comprimento relativamente a essas curvas paramétricas. Alguns pontos são informados ao sistema desenvolvido para que os polinômios, ao final, sejam normalizados. Esses dados são armazenados em um banco de dados para posterior identificação dos diferentes animais.

Os resultados experimentais obtidos mostram que as características aqui selecionadas

Resumo

O processamento inteligente de imagens digitais aplicado à identificação de golfinhos baseia-se em diversas técnicas como tratamento e segmentação de imagens, redes neurais, morfologia matemática e transformações geométricas, que são combinadas de forma a possibilitar que imagens de golfinhos sejam tratadas de tal forma a identificar diferentes indivíduos.

Os golfinhos analisados neste trabalho são da espécie *S. fluviatilis*, mais conhecidos como boto-cinza, e habitam numa área de mar compreendida entre as praias de Pipa(RN) e Tabatinga (RN). Eles se caracterizam principalmente por possuírem uma grande nadadeira dorsal. Esta nadadeira funciona como uma espécie de identidade do golfinho, pois ali são registradas agressões que estes animais sofrem, formando fendas facilmente distintas. Outra característica importante é a própria forma da curvatura da nadadeira desses animais.

Neste trabalho, foram implementados algoritmos capazes de tratar inicialmente a imagem, através da teoria da KLT e do auto-contraste. Usamos técnicas de morfologia matemática para auxiliar no processo de extração de características e algoritmos de redes neurais para a classificação dos animais cujas imagens foram submetidas ao sistema desenvolvido. Para a identificação da curvatura da nadadeira foram utilizados dois polinômios cúbicos paramétricos associados a cada lado da nadadeira, partindo do ponto de pico. Fendas são detectadas em cada curva e mapeadas relativamente a essas curvas paramétricas. Alguns pontos são informados ao sistema desenvolvido para que os polinômios, ao final, sejam normalizados. Esses dados são mantidos em um banco de dados para posterior identificação dos diferentes animais.

Os resultados experimentais obtidos mostraram que as características aqui selecio-

nadas funcionaram como bons identificadores, sendo capazes de classificar corretamente entre animais distintos e apresentar poucas falhas quando submetido a animais idênticos, mas em condições diferentes.

Abstract

The digital processing of digital images applied to the identification of dolphins is based on several techniques as preprocessing and segmentation of images, neural networks, mathematical morphology and geometrical transformations; they are combined to allow the digital images to be treated to identify different specimens.

The subject dolphins in this work are from the *S. frontalis* species usually known as bottlenose or Portuguese, and live in a sea area between Ilha and Taboão, our beaches in Rio Grande do Norte - Brazil. One of their main characteristics is a large dorsal fin. This fin acts as a dolphin identity card, since the expressions suffered by the animal may represent their. Another important characteristic is the own shape of the fin.

In this work, we have implemented several algorithms to preprocess the image using the Hough theory and edge-detection. We have used mathematical morphology techniques to help extracting the characteristics and neural networks to classify the specimen. To identify the fin curvature we used two parametric cubic polynomials associated to the two sides of the fin and both originating at the fin peak. Arches are detected at each curve and mapped to the parametric space of the polynomials. These remarkable points are informed by the user to normalize the polynomials. The final data are maintained in a database and used to identify new images.

Experimental results show that the selected characteristics are good identifiers, correctly classifying different specimens and recognizing with low error the same specimen in different images.

Abstract

The inteligente processing of digital images applied to the identification of dolphins is based on several tecniques as preprocessing and segmentation of images, neural networks, mathematical morphology and geometrical transformation; they are combined to allow the dolphin images to be treated to identify different specimen.

The analyzed dolphins in this work are from the *S. fluviatilis* species usually known as *boto-cinza* in Portuguese, and live in a sea area between Pipa and Tabatinga, two beaches in Rio Grande do Norte - Brazil. One of their main characteristics is a large dorsal fin. This fin acts as a dolphin identity card, since the aggressions suffered by the animal stay registered there. Another important characteristic is the own shape of the fin.

In this work, we have implemented several algorithms to preprocess the image using the KLT theory and auto-contrast. We have used mathematical morphology techniques to help extracting the characteristics and neural networks to classify the specimen. To identify the fin curvature we used two parametric cubic polynomials, associated to the two sides of the fin and both originating at the fin peak. Notches are detected at each curve and mapped to the parametric space of the polynomials. Some remarkable points are informed by the user to normalize the polynomials. The final data are maintained in a database and used to identify new images.

Experimental results show that the selected characteristics are good identifiers, correctly classifying different specimens and recognizing with few errors the same specimen in different images.

Capítulo 1

Introdução

Os golfinhos são mamíferos que integram a família dos cetáceos e podem ser encontrados em águas salgadas ou águas doce. Atualmente, existem cerca de 37 espécies de golfinhos catalogados no Brasil, entre os de água salgada e água doce. O tempo de vida médio destes animais é de 35 anos. Seu período de gestação é de 10 meses.

Os golfinhos são mundialmente conhecidos principalmente por suas habilidades de mergulho e saltos ao nível da água que encantam as pessoas. Suas habilidade e velocidade de natação, que pode chegar a 40Km/h, são principalmente devidas a um efeito aerodinâmico que eles alcançam contraindo a pele e formando dobras que diminuem as turbulências. Esses animais podem atingir, na fase adulta, desde um metro e meio até três metros de comprimento. Os mais conhecidos, os de focinho longo, tem cerca de dois metros de comprimento. Uma importante característica na distinção dos golfinhos é a sua nadadeira dorsal. O tamanho desta nadadeira varia de espécie para espécie e ela não possui ossos. Nela são registradas marcas de agressores que os distinguem entre os diferentes animais. Isso funciona exatamente como se fosse uma impressão digital.

Muitas pesquisas são realizadas através da contagem desses animais em ambientes parcialmente “demarcados” em praias ou rios com diversos objetivos, dentre eles o estudo comportamental entre grupos de animais [Lin00]. A maior dificuldade encontrada reside exatamente no fato de que essas áreas não possuem um limite exato; além disso, ainda existe a dificuldade observacional humana. Aquela dificuldade é parcialmente eliminada devido à permanência quase constante de uma mesma espécie de animais em uma mesma

região. Os golfinhos analisados neste trabalho residem numa área costeira do leste do estado do Rio Grande do Norte.

A outra dificuldade comentada reside exatamente na aquisição de imagens dos golfinhos. A maior parte das espécies de golfinhos, ao notarem a presença de embarcações, procuram manter sempre uma distância considerável do barco ou simplesmente “desaparecem”. Isso dificulta as observações por pesquisadores que necessitam colher dados através da identificação a olho nu ou através de fotos para posterior identificação. Esse último processo é conhecido como foto-identificação e é praticado por diversos biólogos. Normalmente, tem-se dezenas ou centenas de fotos e deseja-se usá-las de forma comparativa, uma a uma, para identificação entre os diversos indivíduos. Como estas fotos são tiradas em dias, horários e condições naturais as mais diversas, a foto-identificação, através da observação puramente humana, fica sujeita a muitas falhas e à subjetividade entre os diversos analisadores.

Os golfinhos são primariamente reconhecidos pelos padrões de fendas formadas na nadadeira dorsal por mordidas ou outros traumas a que o animal é submetido; essas marcas são permanentes no tempo [HKA⁺98]. Especialistas humanos podem sistematicamente comparar padrões de fendas em diversas fotografias daquelas bases de dados e fazer identificações confiáveis [DSW90], mas o processo é muito lento e exaustivo, quando deseja-se analisar centenas de fotos.

1.1 Proposta e Objetivos

A proposta deste trabalho é construir um sistema que seja capaz de extrair descritores eficientes para posterior identificação de diferentes golfinhos em fotos obtidas por equipes de biólogos e estudantes em praias próximas à costa marítima. Essas fotos são caracterizadas por uma grande área de mar e a presença do golfinho, normalmente, de forma centralizada no frame da foto. A eficiência aqui comentada refere-se à reunião de vários descritores capazes de identificar os diversos animais em um único sistema integrado.

Nota-se aqui que deve haver uma certa padronização do momento fotográfico, com o objetivo de se ressaltar a nadadeira dorsal do animal observado. Fotos de animais mesclados com outros animais ou barcos ao fundo prejudicarão fortemente a identificação,

assim como foto de animais cuja nadadeira está ainda parcialmente submersa.

1.2 Estado da Arte

Poucos trabalhos têm sido desenvolvidos sobre análise da barbatana de cetáceos. A baixa produção de artigos a respeito desse assunto talvez possa ser atribuída à complexidade envolvida na identificação de indivíduos em diversas fotos, devido a sua grande variedade (pose, iluminação e tamanho). Ou simplesmente, podemos atribuir essa escassez de artigos apenas ao desconhecimento dos pesquisadores da área de computação dessa aplicação de processamento de imagens.

Araabi e colaboradores [AKHW01] realizaram uma avaliação de modelos invariantes na foto-identificação de golfinhos. Três famílias de modelos invariantes diferencial, algébrico e semi-diferencial são estudados e avaliados para sua aplicabilidade e performance na foto-identificação de golfinhos. Neste artigo, observa-se que, sobre condições práticas, para curvas quase rígidas encontradas em nadadeiras dorsais, métodos baseados no casamento de curvatura mostram-se superiores a quase todas as outras técnicas, por oferecer um alto nível de invariância. Notou-se também que erros de digitalização comprometem sensivelmente os resultados, caso um modelo de invariância algébrica seja usado. O uso de modelos de invariância semi-diferencial é praticável, desde que o número de derivadas requeridas, assim como o número de pontos na correspondência, seja limitado a dois: isso se deve principalmente ao ruído sempre presente na imagem.

Hillman e colaboradores [HKA⁺98] analisam essa questão com a finalidade de descobrir descritores que sejam capazes de funcionar como índices em bancos de dados de fotos de nadadeiras de golfinhos. Neste trabalho, observou-se o uso de duas técnicas: transformadas invariantes de Fourier e *templates* deformáveis seguido de adaptação a uma curva Gaussiana. No caso do método usando as transformadas de Fourier, a curva da nadadeira é reduzida a um conjunto de coeficientes de Fourier e são expressos alguns parâmetros de similaridades invariantes, resultando em invariância em relação a posição da câmera, ângulo e distância do golfinho. Mas essa técnica não mostrou uma diferenciação significativa entre as diversas fotos de tal forma a identificar corretamente os diversos indivíduos. O segundo método baseia-se na localização de fendas, usando-se de uma curva rígida (*stiff*

snake), através de um polinômio de quarto grau que é uma aproximação da curvatura da nadadeira. Em seguida, é computada a diferença desta curva para os pontos reais da nadadeira dorsal. Essa técnica produziu um conjunto de características mais promissoras.

Cesar e Costa [CJC97] explicam como diagramas de energia de curvatura em multi-escala podem ser facilmente obtidos a partir de *curvogramas* [CJC96] e usadas como uma característica robusta geral para a caracterização morfométrica de células neurais. A energia de curvatura é uma interessante característica global para a caracterização de formas, expressando a quantidade de energia necessária para transformar uma forma específica dentro do seu estado de energia mais baixo (isto é, um círculo). O *curvograma*, que pode ser obtido precisamente através de técnicas de processamento de imagens, mais especificamente, através das transformadas de Fourier e sua inversa, provê uma representação em multi-escala da curvatura de contornos digitais. Eles discutem também que, pela normalização da energia de curvatura com respeito ao círculo padrão de perímetro unitário, essa característica torna-se eficiente para expressar formas complexas de modo a ser invariantes a rotação, translação e escala, além de ser robusta a ruídos e outros artefatos que aparecem na aquisição da imagem.

1.3 Motivação

A maior parte da dificuldade enfrentada pelos pesquisadores é a comparação entre os diversos indivíduos fotografados em diferentes situações: horários diversos, climas diversos, diferentes fotógrafos. Esses e outros fatores inserem uma certa subjetividade quando da análise de foto a foto para uma correta classificação de indivíduos.

A análise das fotos baseia-se principalmente nos padrões apresentados pela nadadeira dorsal do animal. Assim, identificadas características marcantes na nadadeira de um indivíduo, é feita a comparação com outras fotos, seguindo-se uma reunião de características afins. O esforço desse processo é diretamente proporcional ao número de fotos colecionadas pelos pesquisadores para esse fim, dissipando assim bastante tempo.

Note-se ainda que a análise das fotos realizada pelos pesquisadores, através de régua e outras formas de identificação e medida, pode classificar indivíduos ligeiramente distintos como iguais, ou ainda pior, gerar diversos indivíduos quando se tem apenas um. Esse

tipo de problema afetará diretamente o estudo populacional dos animais de uma região, levando a dados imprecisos.

1.4 Estrutura do sistema proposto

O diagrama seguinte mostra o fluxograma do sistema proposto. O sistema é dividido em 6 módulos, descritos a seguir:

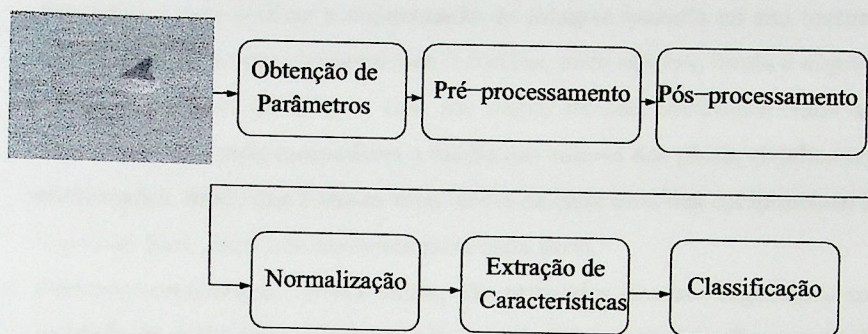


Figura 1.1: Diagrama de blocos do sistema proposto

- 1. Obtenção de Parâmetros.** Neste módulo, captura-se e digitaliza-se a imagem. Em seguida, essa imagem é apresentada ao usuário para que sejam informados alguns parâmetros de entrada necessários ao processamento. Esta etapa é responsável pela apresentação adequada das imagens original ou produzidas pela ferramenta. Neste sistema, é exigido que a imagem original seja de alta resolução. Sendo assim, as imagens apresentam-se com dimensões médias de 1800 por 1800 píxeis. É aplicado então um algoritmo de redução da imagem original apenas para apresentação ao usuário da ferramenta, mas o processamento é realizado na imagem original, evitando assim perdas de precisão de dados. Nesta etapa, também é empregado um algoritmo para delimitação manual da região de interesse. Em seguida, a região delimitada é mapeada para a imagem original.
- 2. Pré-processamento.** Neste módulo são aplicadas algumas técnicas visando melhorar a qualidade da imagem digitalizada, entre elas a transformada Karhunen-Loève

(KLT) e o auto-contraste. A transformada KLT é aplicada com o objetivo de obter uma transformação que descorrelaciona as variáveis do processo, levando a imagem transformada para um novo espaço de cores. A aplicação do auto-contraste na imagem consiste em elevar o contraste da imagem até obter uma certa distribuição de níveis de cinza ou das componentes RGB, no caso de imagens coloridas.

Neste módulo também é feita a auto-segmentação da imagem. Foi utilizada uma metodologia de segmentação não supervisionada através de uma rede neural artificial competitiva para realizar a segmentação da imagem baseada na sua textura. Esta técnica gera uma outra imagem com 2 rótulos, onde teremos fundo e objetos. Para rotular cada píxel da imagem com um destes rótulos, utilizamos como atributos para entrada na rede competitiva a média dos valores dos píxeis vizinhos aos píxeis selecionados, sendo que é tirada uma média de cada uma das componentes R, G e B o que, ao final, dará três atributos para cada píxel.

3. **Pós-processamento.** Nesta etapa, são utilizados diversos algoritmos com a finalidade de padronizar a imagem advinda da etapa anterior. É importante que a imagem conduzida para a etapa seguinte seja padronizada, afim de que sejam obtidos descritores corretos. Um caso claro da real necessidade desta etapa é que as imagens podem vir com golfinhos em diferentes rotações, por exemplo, imagens com golfinhos voltados para a esquerda do vídeo e outras com golfinhos voltados à direita do vídeo. Na etapa inicial é empregada uma ferramenta onde o usuário alinha um golfinho virtual ao golfinho da imagem atual. Esse alinhamento é feito no plano tridimensional. Após obtido o ângulo de rotação é feito o mapeamento correto para a apresentação 2D no plano do vídeo. As técnicas de morfologia matemática são utilizadas para extração de borda, esqueleto e auxílio em outros algoritmos. Esses procedimentos serão úteis para a extração de descritores de forma da barbatana do golfinho para a próxima etapa, além de servirem de base também para auxiliar em algoritmos de análise de curvatura, detecção de pico, entre outros. A forma da nadadeira é encaixada por duas curvas cúbicas paramétricas e, em seguida, são aplicadas as normalizações adequadas nesses polinômios.
4. **Extração de características.** Neste módulo são extraídas algumas características da forma da barbatana dorsal do golfinho, tais como: largura e altura da nadadeira,

localização do pico e número posição de fendas presentes na nadadeira.

5. **Classificação.** Após todo o processamento realizado são gerados descritores que representam a nadadeira do golfinho. O sistema é habilitado para armazenar diversos animais para serem posteriormente classificados usando uma rede neural de retropropagação.

1.5 Organização do trabalho

Este documento está organizado em 7 capítulos. Nesses capítulos, apresentamos uma introdução ao assunto aqui desenvolvido, técnicas utilizadas durante as etapas de processamento, alguns experimentos e o sistema desenvolvido.

No capítulo 1 é apresentada uma introdução sobre o trabalho, motivação, proposta e como ele está organizado. No capítulo 2, apresentam-se os conceitos correlatos aplicados ao objeto de estudo, o golfinho. Analisa-se a forma, tamanho, e classificação da espécie de golfinhos tratados neste trabalho. No capítulo 3, apresentam-se os aspectos gerais relativos às técnicas de processamento digital de imagens utilizadas nesse trabalho. Nele são descritas as técnicas e considerações teóricas tomadas para implementação do sistema. No capítulo 4, apresentam-se os conceitos gerais das redes neurais artificiais e, mais especificamente, são analisadas as redes competitivas e as redes de retropropagação. No capítulo 5, apresentam-se os detalhes da ferramenta desenvolvida. No capítulo 6 são apresentados alguns experimentos realizados e como é feita a identificação do animal submetido ao sistema. No capítulo 7 são feitas as conclusões, discussões gerais e perspectivas futuras deste trabalho.

Capítulo 2

Golfinhos

2.1 Cetáceos

Cetáceos são mamíferos aquáticos que apresentam a forma do corpo hidrodinâmica (para facilitar a natação) e vivem exclusivamente na água. Toda a sua morfologia e fisiologia estão adaptadas à sobrevivência no ambiente aquático [Mar93]. A origem desses mamíferos ainda não é totalmente conhecida, acredita-se que as espécies atuais evoluíram a partir de um grupo terrestre, há cerca de 5 a 25 milhões de anos atrás, mas os registros fósseis de seus antepassados datam de 40 a 60 milhões de anos [BDR03]. Os cetáceos representam os mais modificados mamíferos que existem [HBS9].

No Brasil, foram registradas em total de 27 espécies de cetáceos que habitam a costa, sendo 30 espécies de odontocetos e 7 espécies de Mysticetus [GTE97, HWS9, Sal99]. Os golfinhos, que fazem parte da sub-ordem Odontoceti, estão incluídos na família Delphinidae [Wil93]. No Rio Grande do Norte existem evidências de quatro espécies de golfinhos que habitam a região: golfinho-cara-de-guaxima (*Twinsia macrotis*), golfinho-pintado-do-Atlântico (*Stenella frontalis*), golfinho-Clowne (*Stenella clymene*) e, em maior número, o bato-cara (*Sotalia fluviatilis*), que habita as águas costeiras de ambas as margens do litoral do Estado [SQMB09].

Os cetáceos apresentam um papel importante no equilíbrio ecológico dos ecossistemas aquáticos, podendo ser indicadores da qualidade do ambiente onde são encontrados [Sch99]. Entretanto, até cerca dos anos 70, os cetáceos eram considerados um dos grupos

Capítulo 2

Golfinhos

2.1 Cetáceos

Cetáceos são mamíferos aquáticos que apresentam a forma do corpo hidrodinâmica (para facilitar a natação) e vivem exclusivamente na água. Toda a sua morfologia e fisiologia estão adaptadas à sobrevivência no ambiente aquático [Mar90]. A origem desses mamíferos ainda não é totalmente conhecida, acredita-se que as espécies atuais evoluíram a partir de um grupo terrestre, há cerca de 5 a 25 milhões de anos atrás, mas os registros fósseis de seus antepassados datam 40 a 50 milhões de anos [BDR85]. Os cetáceos representam os mais modificados mamíferos que existem [Hil95].

No Brasil, foram registradas em torno de 37 espécies de cetáceos que habitam a costa, sendo 30 espécies de odontocetos e 7 espécies de Mysticetos [GTE97], [PM92], [San99]. Os golfinhos, que fazem parte da sub-ordem Odontoceti, estão incluídos na família Delphinidae [Wil98]. No Rio Grande do Norte existem evidências de quatro espécies de golfinhos que habitam a região: golfinho-nariz-de-garrafa (*Tursiops truncatus*), golfinho-pintado-do-Atlântico (*Stenella frontalis*), golfinho Clímene (*Stenella clymene*) e, em maior número, o boto-cinza (*Sotalia fluviatilis*), que habita em águas costeiras de baías do litoral do Estado [SQMB98].

Os cetáceos apresentam um papel importante no equilíbrio ecológico dos ecossistemas aquáticos, podendo ser indicadores da qualidade do ambiente onde são encontrados [Sch90]. Entretanto, até cerca dos anos 70, os cetáceos eram considerados um dos grupos

de animais menos conhecidos pelo homem e até mesmo nos dias atuais. Carências de estudos desses animais ocorre principalmente devido ao seu *habitat*, sua rápida movimentação e também pela dificuldade de avistagem e reavistagem e do acompanhamento dos grupos e indivíduos [RJ96] [W79].

O sucesso do gerenciamento de áreas ambientais saudáveis depende da qualidade das informações disponíveis. Em locais habitados pelos cetáceos, é necessária não só a definição de limites de área para preservação, mas também o conhecimento de como esses locais são usados pelos animais e que fatores afetam a sua distribuição e abundância. Através desses dados pode-se entender outros aspectos da ecologia populacional como a estrutura da população, estratégias de forrageio e até mesmo processos de doenças [WTH97]. Um conhecimento correto da variação espacial e subdivisões de populações de espécies de cetáceos é freqüentemente necessária para desenvolver um programa prático de gerenciamento [DSH84].

Neste trabalho, vamos investigar o golfinho *Sotalia fluviatilis*, mais conhecido como boto-cinza. Eles podem ser vistos principalmente nas prais de Tabatinga e Pipa - RN. Aproveitaremos uma característica importante desse animal que é sua grande nadadeira dorsal. Nela são registradas marcas que distinguem entre os diversos animais.

2.2 Boto-cinza

O boto-cinza é parecido com o golfinho-flíper, exceto pela nadadeira dorsal triangular e pelo tamanho pequeno. A imagem 2.1 mostra a foto de um boto-cinza tirada no litoral do Rio Grande do Norte.

Existem duas formas de espécies: uma fluvial (comum na Amazônia, onde é conhecido como boto-tucuxi) e outra marinha.

Descrição

O boto-cinza é um dos menores cetáceos, e não apresenta dimorfismo sexual significativo. Seu comprimento máximo é de 2.2m, e o comprimento médio é de 1.7m. A forma fluvial é menor, com comprimento máximo de cerca de 1.6m. O comprimento dos filhotes ao nascer é de aproximadamente 1m para a forma marinha, e de 75cm a 85cm para a



Figura 2.1: Golfinho boto-cinza

forma fluvial.

O peso máximo já registrado para a forma fluvial é de 53Kg, e o peso da forma marinha ainda é desconhecido [Het93].

Apesar de relativamente pequeno, o boto-cinza é um animal robusto e de nadadeiras grandes. A nadadeira dorsal tem forma triangular, e localiza-se próximo ao centro do dorso.

O bico tem tamanho médio, e a separação entre este e a cabeça não é tão visível quanto no golfinho flíper.

A coloração da região dorsal e das nadadeiras é geralmente cinza (como indica o nome), podendo também ser amarronzada. Na região da barriga, a coloração é mais clara, chegando a ser branca ou cor-de-rosa. A forma fluvial tem coloração geralmente mais clara do que a forma marinha.

Distribuição

O boto-cinza é encontrado em águas superficiais e protegidas de estuários, baías e bocas de grandes rios. Aparentemente, as populações de *S. fluviatilis* não realizam grandes deslocamentos, permanecendo em uma mesma área durante todo o ano [Het93].

Neste trabalho, estaremos analisando imagens capturadas em uma área de mar compreendida entre as praias de Pipa(RN) e Tabatinga (RN). Nesta região, é mais comum o aparecimento destas espécies de golfinhos. Portanto, a maior parte das fotos analisadas neste trabalho e tiradas por pesquisadores, conterão basicamente os animais com as características destas espécies encontradas nesta região.

2.3 Foto-identificação

A nadadeira dorsal de um golfinho é altamente sujeita à cicatrização e o tecido não se regenera, ficando assim marcas que não mudam [Orr86]. A partir desta constatação, as espécies de pequenos cetáceos tem sido estudadas através do método de identificação individual por fotos, a foto-identificação que utiliza-se principalmente de fotos de marcas de cicatrizes da nadadeira dorsal e também da pigmentação dos animais [WJ90]. A identificação individual de cetáceos possibilita o encontro repetido de um indivíduo reconhecido por suas marcas naturais, potencializando assim o desenvolvimento de estudos ecológicos e comportamentais [MCTW00].

A foto-identificação pode fornecer informações sobre os parâmetros do movimento dos animais, e também tamanho e dinâmica das populações [MCTW00]. Este método vem sendo aplicado há quase três décadas, principalmente em golfinhos-nariz-de-garrafa (*Tursiops truncatus*) e em orcas (*Orcinus orca*), além de muitas outras espécies de cetáceos [DSW90], [Flo99], [MCTW00].

Após a identificação, busca-se a reavistagem desses animais em épocas e locais diferentes. A utilização eficiente deste método requer o estudo de animais que apresentem marcas ou cicatrizes fáceis de serem observadas. Além disso, as marcas devem ser estáveis o suficiente para permitir a identificação individual durante o seu ciclo de vida [San99]. A identificação fotografada é uma técnica bastante usada para constatar a estrutura de grupo, tamanho de populações, padrões de residência e de deslocamento em determinadas áreas [DW99] [RH99] [WW77].

2.4 Identificação dos golfinhos

A identificação das diferentes espécies de golfinhos pode ser feita visualmente através da observação de seu tamanho, forma, textura da pele e principalmente da nadadeira dorsal. Como neste trabalho estamos interessados apenas na análise de uma espécie para posterior identificação entre indivíduos, resumimos as características distintivas em duas: a textura da pele do animal e a forma da nadadeira dorsal. A análise da textura da pele do animal é particulamente útil apenas para a distinção não entre animais, mas para separá-

los de outros objetos presentes na imagem, como ondas do mar, barcos, entre outros. Além do que, como as fotos são normalmente tiradas de cima de uma embarcação, a parte realmente visível do animal constitui-se essencialmente de sua nadadeira dorsal. Como a nadadeira dorsal apresenta-se como um fator decisivo na distinção entre indivíduos de uma mesma espécie, então essa é a característica mais importante a ser verificada pelas técnicas de processamento de imagens apresentadas neste trabalho.

Capítulo 3

Técnicas de Processamento Digital de Imagens Aplicadas

Durante este capítulo, os fundamentos de processamento digital de imagens aplicadas. Os métodos de processamento digital de imagens têm duas principais áreas de aplicação: melhoramento de imagens, objetivando aumentar o nível de informação presente para posterior interpretação humana, e o processamento de dados de uma cena para interpretação automática em tempo. A primeira é aplicada para a resolução de uma variedade de problemas em diversas áreas, tais como: medicina, biologia, agricultura, física, astronomia, entre outras. A segunda área tem como principal foco de interesse procedimentos para extração de informações de imagens em uma forma adequada ao procedimento computacional, como reconhecimento e classificação de padrões automaticamente.

3.1 Representação de imagens

O processamento digital de imagens envolve alguns tópicos como: a digitalização, codificação, filtragem, segmentação e descrição. A digitalização envolve aspectos relacionados à conversão de imagens contínuas para uma forma discreta, ou seja, a amostragem das coordenadas espaciais e a quantização das níveis de cores. Codificação compreende as técnicas utilizadas para representar os dados através de conjuntos finitos de símbolos. O processo de filtragem procura extrair certas características de imagens para posterior análise ou visualização. A segmentação de uma imagem consiste em separar regiões pro-

Capítulo 3

Técnicas de Processamento Digital de Imagens Aplicadas

Discutiremos, neste capítulo, os fundamentos de processamento digital de imagens aplicados. Os métodos de processamento digital de imagens têm duas principais áreas de aplicação: melhoramento de imagens, objetivando aumentar o nível de informação presente para posterior interpretação humana, e o processamento de dados de uma cena para percepção autônoma em robôs. A primeira é aplicada para a resolução de uma variedade de problemas em diversos campos, tais como: medicina, biologia, agricultura, física, astronomia, entre outros. A segunda área tem como principal foco de interesse procedimentos para extração de informações de imagens em uma forma adequada ao processamento computacional, como: reconhecimento e classificação de padrões automaticamente.

3.1 Representação de imagens

O processamento digital de imagens envolve alguns tópicos como: a digitalização, codificação, filtragem, segmentação e descrição. A digitalização cobre os aspectos relacionados à conversão de imagens contínuas para uma forma discreta, ou seja, a amostragem das coordenadas espaciais e a quantificação dos níveis de cinza. Codificação compreende as técnicas utilizadas para representar os níveis através de conjuntos finitos de símbolos. O processo de filtragem procura acentuar certas características da imagem para posterior análise ou visualização. A segmentação de uma imagem consiste em separar regiões pre-

sententes na mesma com base em alguma característica comum aos seus píxeis, subdividindo-a em suas partes ou objetos constituintes [GWG92]. As técnicas de restauração ou filtragem procuram reverter o processo de degradação sofrido por uma imagem com ruído, distorção ou outros fenômenos que causam degradação.

3.1.1 Digitalização

A digitalização corresponde à conversão de imagens contínuas para uma forma discreta. Na forma contínua ela é representada por uma função bidimensional $I(u, v)$, sendo que o $I \in \mathbb{R}$ representa a intensidade ou brilho e $u, v \in \mathbb{R}$ as coordenadas espaciais. Uma imagem monocromática é representada desta forma, onde cada coordenada espacial representa um ponto e este ponto é identificado pela intensidade da luz ou valor do brilho na sua localização.

Imagens coloridas podem ser caracterizadas por 3 planos de imagens monocromáticas que representam, em cada ponto, as intensidades das componentes R, G e B, ou seja, vermelho, verde e azul. Cada uma destas componentes RGB são chamadas de planos.

Um computador digital processa dados discretos representados em forma binária. Então é necessário que as imagens sejam digitalizadas e armazenadas para posteriormente serem trabalhadas através das técnicas de processamento digital de imagens. Uma imagem, quando digitalizada, passa a ser representada por uma função discreta $I(x, y)$ em que x e y representam variáveis discretas e o I assumirá valores discretos em um intervalo finito. Por exemplo, podemos ter uma imagem com 400 pontos de largura e 300 pontos de altura (x variando de 0 a 399 e o y variando de 0 a 299), e os níveis de intensidade de brilho ou cor para cada ponto nas coordenadas x, y , variando de 0 à 255 (sendo que 0 representa o preto e o 255 representa o branco).

Desta forma, podemos pensar numa imagem digitalizada como uma matriz onde cada posição i, j tem um elemento que caracteriza a intensidade do brilho dos píxeis nessa posição.

O píxel é identificado como um ponto discreto em uma imagem, e o número de píxeis em uma imagem está associado a sua resolução. No geral, podemos ter imagens com resoluções de 300 píxeis por polegada (ppp) para largura e altura. Outro conceito impor-

tante é o de bits por píxel, que informa o número de bits que serão usados para codificar a cor ou intensidade de brilho de cada píxel na imagem. Em uma imagem de 8 bits por píxel cada ponto, em cada plano, possui 256 valores possíveis (8 bits podem representar valores de 0 a 255) [Mar01].

Atualmente existem diversos dispositivos para capturar e digitalizar uma imagem: câmera digital, scanner, filmadoras digitais, etc.

A figura 3.1 mostra uma imagem digitalizada. A imagem tem dimensões de 377 de largura por 133 de altura.

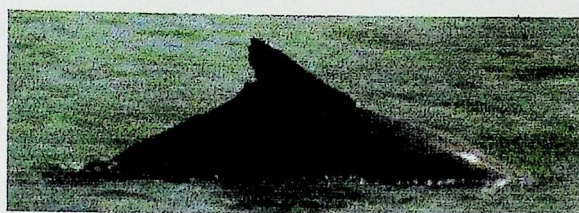


Figura 3.1: Imagem de um golfinho

3.2 Realce de Imagens

O objetivo das técnicas de realce de imagens é processar uma imagem tal que o resultado seja uma outra imagem mais apropriada para uma determinada aplicação, seja ela visualizada por observadores humanos ou para servir a um outro sistema de processamento de imagens. Os métodos de realce de imagens são bem específicos e, um método que é bastante apropriado ao realce de determinadas imagens não é, necessariamente, o melhor a ser utilizado no realce de outros determinados tipos de imagens. Uma dificuldade em relação ao realce é a questão de quantificar o critério a ser utilizado para a avaliação da qualidade da imagem. Dessa forma, a maioria das técnicas de realce são empíricas, simples e requerem procedimentos iterativos para a obtenção de melhores resultados. Pode-se aplicar também operações pontuais em imagens que são técnicas de realce baseadas nas transformações diretas da intensidade dos píxeis de uma imagem, de acordo com

a equação:

$$S = T(r) \quad (3.1)$$

onde r representa o valor de um píxel na imagem original, S representa o valor dos píxeis na imagem de saída e $T(\cdot)$, a transformação sobre os píxeis na imagem original. As transformações pontuais também são chamadas de mapeamento. Nas próximas seções serão descritas algumas técnicas pontuais e espaciais, que englobam uma região ao redor de um píxel. Neste trabalho, estamos utilizando as transformações do histograma para realce da imagem. Essa transformação é aplicada nos métodos de auto-contraste e transformada KLT, explicados mais adiante.

3.2.1 Histograma

O histograma de uma imagem digital com níveis de cinza no intervalo $[0, L - 1]$ é uma função discreta:

$$p(r_k) = \frac{n_k}{n} \quad (3.2)$$

em que r_k é o k -ésimo nível de cinza, n_k é o número de píxeis na imagem com esse nível de cinza, n é o número total de píxeis na imagem e $k = 0, 1, 2, \dots, L - 1$ [GWG92]; isto é, o histograma de uma imagem representa a frequência de ocorrência dos níveis de cinza. O histograma nos dá informações úteis de uma determinada imagem, que podem ser aplicadas no realce da imagem, na segmentação e etc. Uma operação muito comum no processamento de imagens é a transformação do histograma com o objetivo de realizar alguma modificação na mesma. Na seção seguinte, será apresentado o método para transformação usado neste trabalho. Lembramos que para imagens coloridas, imagens com 03 planos por exemplo, tem-se um histograma para cada plano (histograma do vermelho, verde e azul) cada um trabalhando o respectivo histograma como se fossem imagens em tons de cinza.

Através do histograma é possível extrair informações que auxiliam a segmentação, realce de imagens e o estudo de textura. Informações como: média, variância e desvio padrão são extraídas do histograma. Na parte inferior da figura 3.3 são mostrados dois histogramas referentes as imagens apresentadas.

3.2.2 Transformada Karhunen-Loeve (Hotelling)

A transformada de Karhunen-Loeve (KLT) ou transformação por Componentes Principais, também conhecida como transformada de Hotelling, é um método oriundo da área de processos estocásticos, utilizada em diversas aplicações na área de processamento de imagens [JAI89]. Neste trabalho, ela é aplicada como ferramenta auxiliar na extração de características lineares da imagem.

A análise das componentes principais (PCA, em inglês) consiste no cálculo dos autovalores e seus respectivos autovetores de uma matriz de variâncias-covariâncias. Onde, a diagonal principal tem as variâncias das variáveis e os demais elementos da matriz são valores de covariâncias entre pares de variáveis.

Podemos dizer que esta transformada é uma técnica que transforma um grupo de variáveis correlacionadas em um grupo de variáveis não correlacionadas em que cada variável medida pode ser considerada como um eixo de variabilidade, estando correlacionado com outras variáveis. A transformada de Hotelling transforma os dados de modo a demonstrar a mesma variabilidade total existente, com o mesmo número de eixos originais, no entanto, não mais correlacionados entre si. Quando não é possível descorrelacionar as informações, a PCA pelo menos minimiza esta correlação.

A teoria da KLT garante um mínimo (zero no caso ideal) de correlação se a seguinte transformação é usada:

$$y_i = A \times (x_i - m_x) \quad (3.3)$$

onde:

- x_i são os valores dos píxeis originais na base RGB;
- m_x é o valor médio dos píxeis x_i dado por:

$$m_x = E(x) = \frac{1}{N} \sum_{i=1}^N x_i \quad (3.4)$$

- as colunas da matriz A são os autovetores da matriz de covariâncias C_x ;
- C_x é a matriz de covariâncias dos píxeis x_i , dada por:

$$C_x = E[(x - m_x) \times (x - m_x)^T]$$

- y_i são os píxeis transformados (descorrelacionados).

A transformação é realizada de forma que o conjunto de dados possa ser representado por um número de elementos “efetivos” e que ainda retenha a maioria intrínseca do conteúdo da informação.

Quando se está trabalhando com imagens, é desejável que a representação seja bastante enxuta (compacta), com sua informação bem discriminada, isto é, que não exista redundância entre as diferentes características dos padrões. É fácil notar que o espaço de imagens é altamente redundante, pois cada píxel é muito correlacionado com outros píxeis.

Neste trabalho, a idéia é utilizar o princípio básico da transformada KLT e mudar de um espaço de cores para outro. Isto é, obter uma transformação, uma base nova, que descorrelacione as variáveis do processo levando cada píxel, ou melhor vetor, para esta nova base. Cada píxel é representado por várias componentes como um vetor. Por exemplo, uma imagem colorida com os planos R, G e B [GWG92]. Daí concluímos que cada píxel de uma imagem RGB pode ser visto como um vetor ou um processo aleatório e que cada componente R, G e B é vista como uma variável aleatória dentro do processo aleatório, ou um elemento deste vetor.

Os resultados práticos da transformada KLT são tais que a descorrelação das componentes da imagem levam a uma transformação que rotaciona os pontos (píxeis como vetores) da imagem para a direção dos eixos principais (eixos que deixaram ortogonais os vetores resultantes) [Mar01].

A figura 3.2 mostra um exemplo da aplicação da teoria da KLT em uma imagem da nadadeira do golfinho, com a imagem original do lado esquerdo e a imagem transformada no lado direito. O diagrama abaixo de cada imagem representa os pontos de cada plano de imagem R, G e B plotados em um único gráfico de três dimensões, as coordenadas x , y e z estão representando as coordenadas R, G e B. Desta forma, é possível representar cada píxel de uma imagem colocando um ponto nas coordenadas que representam a cor deste píxel. É observável que os píxeis esboçados no diagrama da imagem original formaram um eixo inclinado em relação aos eixos do diagrama, isto demonstra um alto grau de correlação existente entre as cores dos píxeis da imagem. No diagrama da imagem após a Klt, observa-se uma maior espalhamento dos pontos, ou seja, os píxeis ficaram sem uma relação específica.

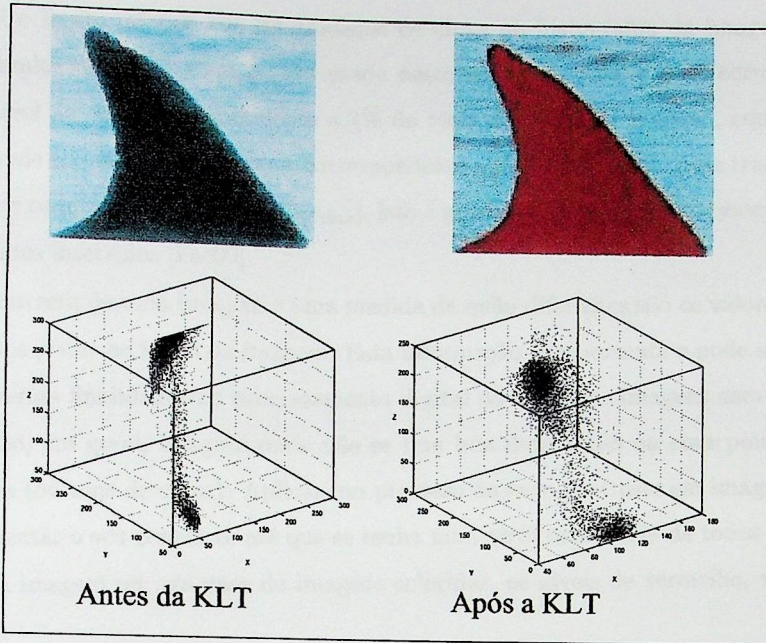


Figura 3.2: Exemplo de uma transformação KLT

3.2.3 Auto-Contraste

Para atenuar os efeitos de pouca iluminação é necessário aumentar a discriminação visual entre os objetos presentes na imagem. O contraste pode ser modelado como uma transformação $t(\cdot)$ tal como:

$$z' = t(z) \tag{3.5}$$

onde z é a intensidade luminosa antes da transformação e z' é a intensidade luminosa após a transformação para um mesmo ponto na imagem. O valor de z' fica no intervalo

$$z_{min} < z' < z_{max} \tag{3.6}$$

onde z_{min} e z_{max} são o limite inferior e superior do nível de cinza da imagem transformada, respectivamente. Neste trabalho usamos o valor “0” para z_{min} e “255” para z_{max} .

Para aumentar o contraste, com $z \in [a, b] \subset [z_{min}, z_{max}]$ faz-se

$$z' = \left[\frac{(z_{max} - z_{min})}{(b - a)} \right] (z - a) + z_{min} \tag{3.7}$$

onde a é o limite inferior das intensidades de cinza no histograma da imagem original e b é o limite superior. Para o caso usado neste trabalho, o valor de a corresponde ao menor nível de cinza correspondente a 1% do total de píxeis da imagem, enquanto que b corresponde ao maior nível de cinza correspondente a 99% desse total. Essa transformação comprime os intervalos $[z_{min}, a]$ e $[b, z_{max}]$, isto é aceitável quando se tem poucos níveis de cinza nestes intervalos [Fac00].

O contraste de uma imagem é uma medida de quão diferentes são os valores de intensidade dos píxeis ao longo da imagem. Esta informação é importante e pode ser utilizada para diversas finalidades no processamento digital de imagens. Imagens com pouco contraste são, em geral, imagens onde não se tem boa iluminação ou com pouca variação nas cores (ou tons de cinza). Aplicar um processo de auto-contraste em imagens consiste em aumentar o seu contraste até que se tenha uma certa ocorrência de todos os níveis de cinza na imagem ou, em caso de imagens coloridas, os níveis de vermelho, verde e azul [Mar01].

No caso do histograma, uma imagem com pouco contraste é aquela que possui uma grande porcentagem dos píxeis concentrados em uma determinada região do seu histograma.

O auto-contraste objetiva uniformizar a distribuição do histograma da imagem sem que com isso afete a sua forma.

A figura 3.3 mostra o resultado do processo de auto-contraste. Note-se que no histograma da imagem original há uma concentração de píxeis em torno do valor médio de intensidade. Já no histograma da imagem após o auto-contraste, nota-se um espalhamento das intensidades presentes na imagem.

A principal característica observada quando se faz um contraste linear é que as barras verticais que formam o histograma ficam espaçadas igualmente, dando a idéia de que o histograma teve um simples esticamento.

3.2.4 Limiarização ou Binarização

Limiarização é um caso especial de manipulação de contraste. É um processo extremamente veloz e, em geral, em um único passo. O objetivo desta transformação é alterar

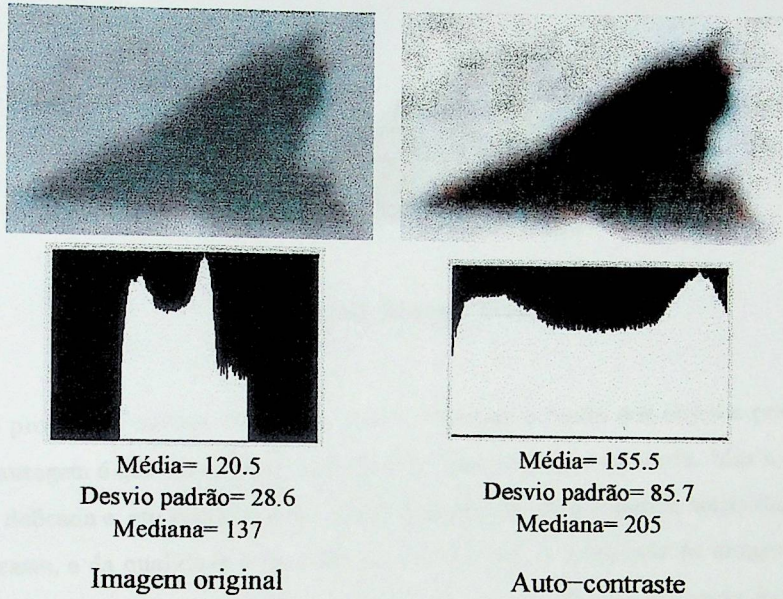


Figura 3.3: Processo de Auto-contraste aplicado à imagem original

uma imagem com vários níveis de cinza para uma imagem binária, ou seja, uma imagem que possui somente o preto e o branco. A intenção desta transformação é eliminar da imagem tudo aquilo que não for relevante à análise [GWG92]. Para imagens sem ruído, este limiar pode ser escolhido de modo que todas as discontinuidades de amplitude a partir de um nível de contraste mínimo sejam reconhecidas como bordas. A imagem com ruído torna esta seleção difícil, pois o valor do limiar tanto pode deixar despercebidas bordas existentes, quanto pode designar falsas bordas induzidas pelo ruído [Pra89]. Vários métodos estão disponíveis para a escolha do limiar. Um deles, bastante usado, consiste em suavizar o histograma e tomar o mínimo entre os dois picos encontrados.

A figura 3.4 mostra uma imagem limiarizada.

Limiarizar consiste, no caso de um objeto mais claro que o fundo, em escolher no histograma um valor T , tal que para um píxel $P(x,y)$:

$$\begin{array}{ll} \text{se } I [P(x,y)] > T & \Rightarrow P(x,y) \in \text{objeto} \\ \text{senão} & \Rightarrow P(x,y) \in \text{fundo} \end{array} \tag{3.8}$$



Figura 3.4: Imagem binária

Este processo é muitas vezes usado para dissociar o fundo dos objetos presentes. A maior vantagem é que o princípio é simples e precisa de pouca memória. Mas a escolha do limiar é delicada e, em geral, se usam técnicas empíricas. Ela depende muito da qualidade do contraste, e da qualidade e da natureza da imagem. A qualidade da imagem é função do fator de refletância e do fator de iluminação. Esse último influencia fortemente a natureza dos picos do histograma. No caso de uma iluminação não uniforme sobre toda a imagem, a separação dos picos e o estabelecimento do limiar são mais difíceis [Fac00].

Quando os valores dos níveis de cinza dos píxeis dos objetos e do fundo são razoavelmente consistentes em seus valores na imagem como um todo, somente um único valor de limiar é necessário para efetuar a separação dos objetos da imagem do fundo. Esse método é conhecido como limiarização global.

No entanto, para muitas aplicações, por causa da falta de iluminação e, até mesmo, a falta de qualidade da imagem, um simples limiar não é o suficiente para se obter uma limiarização de boa qualidade. Existem casos em que limiares diferentes são necessários para posições diferentes da imagem. A esta limiarização dá-se o nome de limiarização adaptativa ou local.

3.3 Segmentação

Segmentação de uma imagem é vista como a extração ou identificação de objetos contidos na imagem, onde objeto é toda característica com conteúdo semântico relevante para a aplicação desejada. As descrições são constituídas basicamente de uma lista de

objetos, seus rótulos e relações entre si. A segmentação é uma tarefa complexa porque tenta traduzir para o computador um processo cognitivo extremamente sofisticado realizado através da visão humana. A segmentação procura distinguir as partículas uma das outras e do fundo da imagem. Assim, permitirá interpretar píxeis contíguos e agrupá-los em regiões. Normalmente o processo de segmentação não obedece a um modelo formal, mas ele é dependente da aplicação e deverá se ajustar a diferentes tipos de imagens. O processo de segmentação consiste, portanto, em uma divisão ou separação de uma imagem em regiões de atributos similares.

Os atributos básicos utilizados são a luminância, isto quando se trata de imagem monocromática, e componentes de cores (brilho, contraste, etc.) para imagens coloridas. Contornos (bordas) e texturas também são atributos úteis à segmentação. Além disso, este processo não envolve classificar cada segmento, apenas subdivide a imagem, não se dedicando a reconhecer os segmentos individuais ou relações (similaridades) de um píxel com uma certa região ou a outros píxeis. Portanto, a segmentação faz parte, em geral, de um processo maior que é o de obter uma descrição da imagem [VM84, Pra89]. Um dos passos na análise da imagem é a identificação dos objetos simples que correspondem, em geral, a linhas ou regiões (grupos de pontos conectados). A detecção das regiões pode ser feita de dois modos: através do agrupamento de pontos vizinhos com características semelhantes, ou através da determinação da fronteira (borda) da região. Uma abordagem mais completa sobre segmentação pode ser encontrados em livros e artigos específicos, como por exemplo, [GWG92, Par97]. Neste trabalho foi utilizado um método não supervisionado de segmentação que se utiliza de redes neurais competitivas. Vale aqui aprofundarmos um pouco em 2 (dois) métodos de segmentação bastante utilizados, os quais são: segmentação por textura e segmentação por quantização vetorial. Detalharemos a seguir cada um desses métodos.

3.3.1 Segmentação por textura

Um dos fatores que motivam o uso de textura como elemento base na descrição de uma imagem é o fato de que grande parte das imagens naturais possuem um certo perfil textural. [RCA+00, GWG92, Par97, Fac00, JAI89]. Segundo Facon [Fac00], a segmentação por

textura deve permitir a colocação em evidência dos motivos da imagem considerando-se as propriedades de regularidade e repetição.

As texturas são quantificadas avaliando-se aspectos relativos à variação de tons de cinza em seções contidas na imagem, e tornam-se assim características de interesse,[Cas96].

Os três principais métodos utilizados para segmentação de imagens por textura são: estatísticos, estruturais e espectrais [GWG92, Par97, Fac00]

1. Estatístico: É considerado um bom método e trabalha com a distribuição da intensidade numa imagem e as posições dos píxeis com níveis de cinza iguais ou semelhantes.
2. Estruturais: O objetivo é, a partir de uma primitiva simples, construir texturas mais complexas, utilizando-se de regras que permitam limitar o número de arranjos possíveis da primitiva.
3. Espectrais: O espectro de fourier é muito utilizado, pois permite detectar a direcionalidade de motivos periódicos em uma imagem. As principais características do espectro de fourier são:
 - Picos proeminentes dão a direção principal dos motivos de textura;
 - A posição desses picos no plano de frequências indica o período fundamental espacial dos motivos;
 - Descartando os componentes periódicos pela filtragem, ficam os componentes não periódicos que podem ser analisados pelas técnicas estatísticas

3.3.2 Segmentação por quantização vetorial

A quantização vetorial é um método muito utilizado na auto-segmentação de imagens [Hay99]. Os algoritmos de redes neurais utilizados na segmentação diferem dos métodos de quantização utilizados na obtenção de conjuntos de dados que representam as características de uma imagem.

Os métodos de segmentação baseados em redes neurais auto-organizáveis tem como princípio a geração de agrupamentos, através da rede neural, a partir dos dados de entrada que são atributos da imagem. Estes atributos da imagem podem ser:

- Informações de textura da imagem, tais como: média, variância, contraste, entropia;

- Amostras da média da intensidade dos níveis de cinza.

No artigo de Rabelo [RCA⁺00] apresenta-se uma segmentação cujos atributos de textura são extraídos através da transformada wavelet e o processo de segmentação é baseado no algoritmo de agrupamento (clustering) K-médias. Este algoritmo é muito utilizado na segmentação de imagens por quantização vetorial.

Diversos métodos supervisionados e não-supervisionados tem sido utilizados na segmentação de imagens. A segmentação supervisionada necessita que classes significantes da imagem sejam conhecidas previamente. Já na abordagem não supervisionada, não é necessário um conhecimento prévio das classes que serão apresentadas à rede, pois em alguns tipos de imagens é muito complexo especificar as classes, e importantes informações estruturais talvez sejam perdidas, caso isso seja feito.

Como há uma grande variedade de algoritmos para segmentação de imagens, a utilização deles depende basicamente do tipo de imagem, ou que tipo de região queremos segmentar.

Neste trabalho é utilizado um algoritmo não supervisionado, denominado redes competitivas, que precede à segmentação baseada na textura da imagem. A rede neural utilizada será detalhada no capítulo 4. São selecionados atributos de cada píxel da imagem e classificados de acordo com uma determinada textura a qual ele pertence, esta classificação é realizada através de técnicas de quantização vetorial. As características utilizadas são a média dos valores R, G e B dos píxeis vizinhos. Sendo uma média para cada uma das componentes R, G e B, formando-se assim três médias para cada píxel.

3.4 Morfologia Matemática

O princípio básico da morfologia matemática consiste em extrair uma informação relativa à geometria e à topologia de um conjunto desconhecido de uma imagem pela transformação a partir de um outro conjunto bem definido. Em morfologia matemática ele é conhecido como elemento estruturante.[Fac00]

Existem dois tipos de morfologia matemática: a morfologia binária que trabalha com imagens binárias e a morfologia em níveis de cinza que trabalha com imagens em níveis

de cinza. Neste trabalho, foram utilizados os conceitos de morfologia binária. A seguir, veremos apenas as operações de morfologia matemática utilizadas no trabalho.

3.4.1 Teoria de Conjuntos

Começaremos nossas explicações pelas definições básicas da teoria de conjuntos aplicada a imagens digitais, após estas explanações passaremos a definir dilatação, erosão, extração de fronteiras (bordas), esqueletização e o processo de abertura.

De forma geral, um conjunto é definido como um grupo de objetos. Quando um conjunto não é muito grande, podemos demonstrá-lo através de uma listagem de seus elementos. No caso de serem conjuntos muito grandes utilizamos propriedades matemáticas para demonstrá-los.

Por exemplo,

$$Z = \{x | x \text{ é um número inteiro, negativo e ímpar}\}$$

$$x \in Z, \text{ se } x \text{ está em } Z, \text{ caso contrário } x \notin Z$$

Como neste trabalho ... De acordo com Parker [Par97] a *união* de dois conjuntos A e B é o conjunto de píxeis que pertencem a A e B ou aos dois conjuntos. Denotamos este conceito simbolicamente da seguinte forma:

$$A \cup B = \{c | (c \in A) \vee (c \in B)\} \quad (3.9)$$

A *reflexão* de um conjunto A é definida como:

$$\hat{A} = \{c | c = -a, \forall a \in A\} \quad (3.10)$$

Isso é o mesmo que um giro de 180° em torno de um eixo y.

A *translação* do conjunto A pelo ponto x é definida como:

$$(A)_x = \{c | c = a + x, a \in A\} \quad (3.11)$$

Podemos explicar a translação da seguinte forma, se x é o ponto (1,3) então todos os píxeis de A, serão deslocados de uma linha e três colunas.

O *complemento* do conjunto A é o conjunto de píxeis que não pertencem a A . Na teoria de conjuntos corresponderia a:

$$A^c = \{c | c \notin A\} \quad (3.12)$$

A *intersecção* de dois conjuntos A e B é o conjunto de píxeis que pertencem a A e B ao mesmo tempo:

$$A \cap B = \{c | (c \in A) \wedge (c \in B)\} \quad (3.13)$$

Finalizando estas definições básicas, a *diferença* entre os conjuntos A e B , é:

$$A - B = \{c | ((c \in A) \wedge (c \notin B))\} \quad (3.14)$$

o conjunto de píxeis que pertencem somente a A , é o mesmo que a intersecção de A com o complemento de B , ou seja, $A \cap B^c$.

3.4.2 Elemento Estruturante

O elemento estruturante é um conjunto conhecido e a partir dele busca-se extrair informações sobre a topologia e a geometria de um conjunto desconhecido de uma imagem. A seguir, utilizaremos A para representar uma imagem binária e B para o elemento estruturante aplicado à imagem.

Neste trabalho, foi bastante utilizado o elemento estruturante de tamanho 3×3 com valores de "1" formando uma cruz, o restante é preenchido com o elemento "0". Confira no gráfico mostrado:

0	1	0
1	1	1
0	1	0

Houve a necessidade também, neste trabalho, de usarmos um elemento estruturante variável com o tamanho da imagem. Usamos esse processo na detecção das fendas na nadadeira do golfinho. Mais a frente detalharemos esse algoritmo.

3.4.3 Dilatação

Conforme Parker, a *dilatação* de um conjunto A por um conjunto B é:

$$A \oplus B = \{c | c = a + b, a \in A, b \in B\} \quad (3.15)$$

Onde A representa a imagem e B é o segundo conjunto de píxeis, conhecido como elemento estruturante [Par97].

Para esclarecermos esta idéia, seja A um conjunto (o conjunto da imagem) e B o conjunto do elemento estruturante $\{(0,0)(0,1)\}$. Os píxeis no conjunto $C = A + B$ são computados usando a equação acima, esta equação pode ser escrita, neste caso, da seguinte forma:

$$A \oplus B = (A + \{(0,0)\}) \cup (A + \{(0,1)\}) \quad (3.16)$$

Podemos dizer que o cálculo da dilatação é a união de todas as translações especificadas pelo elemento estruturante. E vista pela comutatividade, é considerada como a união de todas as translações do elemento estruturante por todos os píxeis na imagem.

3.4.4 Erosão

A *erosão* tem um efeito inverso da dilatação, a erosão retira píxeis da camada externa de um objeto. A erosão da imagem A pelo elemento estruturante B , é equacionada como:

$$A \ominus B = \{c | (B)_c \subseteq A\} \quad (3.17)$$

A erosão é o conjunto de todos os píxeis " c ", tais que o elemento estruturante B transladoado por " c " corresponda a um determinado conjunto de píxeis em A . Seja uma imagem $A = \{(3,3), (3,4), (4,3)(4,4)\}$ e o elemento estruturante $B = \{(0,0), (0,1)\}$. O conjunto $A \ominus B$ é o conjunto de translações de B que o alinham sobre um determinado conjunto de píxeis em A . Isto significa que nem todas as translações precisam ser consideradas, mas só aquelas que inicialmente coloquem a origem de B sobre algum píxel de A [Par97].

3.4.5 Abertura

A operação de um processo de *Abertura* em uma imagem suaviza os contornos, suprime pequenos pontos isolados e aguça as suas extremidas. Esse processo é ideal para um estudo

da distribuição dos tamanhos dos objetos.

O processo de abertura é uma erosão seguida de dilatação e é equacionado da seguinte forma:

$$X_B = (X \ominus B) \oplus B \quad (3.18)$$

Após o processo de erosão em uma imagem, obtém-se uma nova imagem X_B onde, dependendo do tamanho e forma do elemento estruturante usado, pequenos aglomerados de pontos são suprimidos. Esse processo será particularmente útil, neste trabalho, na detecção de fendas presentes na nadadeira do golfinho.

3.4.6 Extração de Fronteiras

A extração da fronteira de um conjunto A , denotado por $\beta(A)$, pode ser obtida através da erosão de A por B , seguida da diferença de conjuntos entre A e sua erosão [GWG92].

Ou seja:

$$\beta(A) = A - (A \ominus B) \quad (3.19)$$

Com essa operação obtemos a delimitação ou borda de um objeto na imagem.

3.4.7 Esqueleto

O esqueleto de uma imagem é a sua redução até a forma de um grafo. O que pode ser obtido através de operadores morfológicos, ou seja, utilizando-se a abertura e a erosão. Imagine um objeto A de uma imagem, o seu esqueleto $S(A)$ é obtido utilizando estas técnicas explicadas a seguir. Matematicamente define-se como [GWG92]:

$$S(A) = \bigcup_{k=0}^K S_k(A) \quad (3.20)$$

com

$$S_k(A) = \bigcup_{k=0}^K \{(A \ominus kB) - [(A \ominus kB) \circ B]\} \quad (3.21)$$

Onde B é o elemento estruturante, enquanto $(A \ominus kB)$ indica k sucessivas erosões de A . K é o ultimo passo antes que A seja erodido a um conjunto vazio. Isto quer dizer que:

$$K = \max \{k | (A \ominus kB) \neq \emptyset\} \quad (3.22)$$

Pelo união dos subconjuntos de esqueletos $S_k(A)$ podemos obter o esqueleto $S(A)$ do conjunto A . Podemos ainda reconstruir o conjunto A a partir desses subconjuntos utilizando a equação:

$$A = \bigcup_{k=0}^K (S_k(A) \oplus kB) \quad (3.23)$$

onde $(S_k(A) \oplus kB)$ consiste em k dilatações sucessivas de $S_k(A)$.

As operações de extração de fronteiras (bordas) e esqueletização são processos fundamentais para o funcionamento de outros algoritmos demonstrados aqui.

3.5 Extração de características

Reconhecimento é a parte do processamento que vai classificar os objetos a partir de informações encontradas na imagem, geralmente tendo como apoio uma base de conhecimento previamente estabelecida. Esta fase é normalmente aplicada após uma fase de segmentação e extração de atributos (parametrização) da imagem. A fase de extração de atributos identifica e calcula parâmetros nos objetos segmentados. Exemplo de atributos pertinentes a um objeto são o seu perímetro, área ou informação estatística de textura.

3.5.1 Descritores de Textura

A textura é um descritor que fornece medidas de propriedades como suavidade, rugosidade e regularidade [JAI89]. As principais abordagens em processamento de imagens para a descrição de textura são as estatísticas, que caracterizam textura suave, áspera, granular, sendo estas técnicas utilizadas neste trabalho; as técnicas estruturais tratam de arranjos de primitivas de imagens, como a descrição da textura baseada em linhas paralelas regularmente espaçadas; e as técnicas espectrais que são baseadas nas propriedades do espectro de Fourier, por exemplo, na detecção de periodicidade global em uma

imagem através da identificação de picos de alta-energia no espectro. A seguir, fazemos um breve comentário sobre as abordagens estatísticas; pois as abordagens estruturais e espectrais são consideradas importantes, mas não adequadas para os objetivos deste trabalho. Para maiores detalhes em relação a essas abordagens, elas podem ser encontradas em [GWG92, JAI89, Fac00]. Neste trabalho, utilizamos os momentos estatísticos para descrever a textura.

Momentos Estatísticos

Para descrição de textura, esta abordagem é uma das mais simples, pois é através dos momentos do histograma de níveis de cinza de uma imagem ou de uma região que se consegue determinar a distribuição de luminosidade de uma imagem. O momento de ordem n é dado por:

$$\mu_n(z) = \sum_{i=1}^L (z_i - m)^n p(z_i) \quad (3.24)$$

onde m é a intensidade média (primeiro momento) que é dada pela equação:

$$m = \sum_{i=1}^L z_i p(z_i) \quad (3.25)$$

onde z é uma variável aleatória denotando a intensidade discreta de uma imagem; $p(z_i)$, $i=1, 2, \dots, L$ corresponde às intensidades de brilho presentes no histograma, sendo que L é o número de níveis distintos de intensidade.

O segundo momento é conhecido como variância (denotado por $\sigma^2(z)$), pois é uma medida de contraste de nível de cinza, podendo ser usado para descritores de suavidade relativa.

O terceiro momento é uma medida de anti-simetria do histograma, enquanto o quarto momento é uma medida de seu achatamento ou planaridade. Os momentos de ordem superiores fornecem informações quantitativas adicionais para o conteúdo de textura [GWG92].

3.5.2 Descritores de forma

Uma vez que a imagem encontra-se segmentada, os grupos de píxeis resultantes são descritos e representados.

A representação de uma região envolve duas escolhas:

- Uma região pode ser representada em termos de suas fronteiras (características externas);
- Ou ser representada em termos dos píxeis que compõem a região (características internas).

Uma representação externa é escolhida quando a atenção primária está voltada para características de forma, por outro lado, uma representação interna é selecionada quando a atenção estiver voltada para propriedades de cor e textura [GWG92].

Conforme mostrado na introdução, Araabi e colaboradores [AKHW01], em seu artigo sobre identificação de golfinhos, chegam a conclusão que métodos baseados na curvatura da nadadeira dorsal são bem mais eficientes dentre os pesquisados por eles (modelos invariantes diferencial, algébrico e semi-diferencial). Hillman e colaboradores reforçam essa afirmativa ao testar as transformadas invariantes de Fourier e *templates* deformáveis [HKA⁺98], chegando a conclusão que esse último método, baseado na curvatura da nadadeira, mostrou-se superior ao primeiro. Esse método, utilizado por Hillman e sua equipe, usa um polinômio de quarto grau para encaixar a curvatura da nadadeira do animal. É proposto aqui um método diferenciado, baseado em dois polinômios de terceiro grau, cada um representando um lado da nadadeira. A junção desses dois polinômios formam a nadadeira. Esse método apresentou-se robusto no ajuste de curvatura da nadadeira. Além disso, podemos usar esses polinômios para detecção das possíveis fendas presentes na nadadeira do animal com mais precisão. A detecção de fendas é um descritor importantíssimo para distinção entre animais de uma mesma espécie, conforme descrito no capítulo 2. Nas seções seguintes mostraremos como foram detectadas as fendas e o pico da nadadeira de um golfinho, além de outros descritores.

Para extrair as características procuradas nessas imagens de golfinhos foi necessário aplicar uma sequência de processamento que preserva ao máximo essas características. O diagrama 3.5 demonstra a sequência de processamento utilizada.

Decidiu-se direcionar esta pesquisa na extração dessas características, pois são características mais marcantes e distintivas entre os diversos animais. Uma característica importante, mas que não foi levada em consideração nesse trabalho, é a textura. Acon-

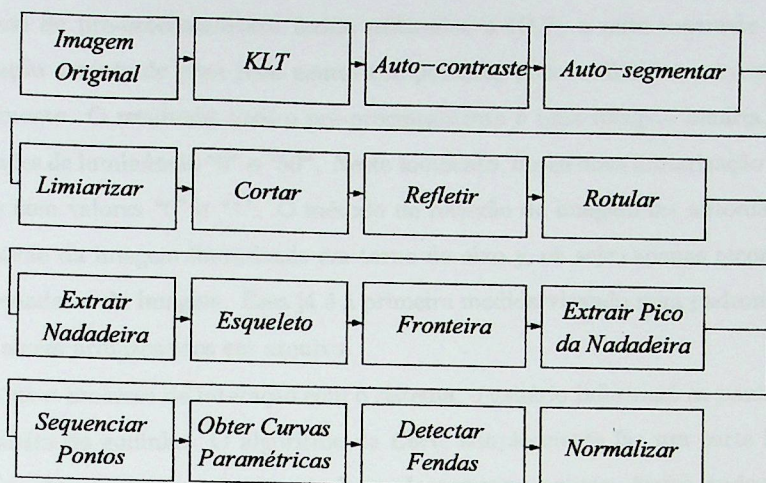


Figura 3.5: Sequenciamento do processamento realizado

tece que, devido à variabilidade luminosa presente nas fotos, acarretada por diferentes condições climáticas, essa característica é difícil de ser analisada para todos os casos. Tanto a forma da nadadeira como seu número de fendas são características menos sensíveis a variações, mas há restrições. A figura 3.6 mostra um animal que apresenta uma rotação relativamente grande em relação ao eixo y (a seção 3.5.2 explica os eixos adotados neste trabalho), neste caso, mesmo que o animal possua várias fendas e essas de tamanho considerável, a maior parte delas simplesmente não são detectadas. Além disso, a extração de curvas também será fortemente afetada.

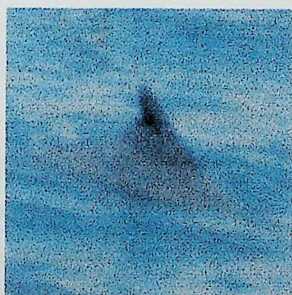


Figura 3.6: Golfinho com rotação elevada

Na fase de pré-processamento foram utilizados a KLT, o auto-contraste e a auto-segmentação através de uma rede neural competitiva, pois facilitam o processo de pós-processamento. O resultado após o pré-processamento é uma imagem binária, mas com intensidades de luminância "0" e "50". Neste momento, faz-se nova limiarização tornando-a apenas com valores "0" e "1". O método de reflexão de imagem faz automaticamente uma reflexão da imagem limiarizada em torno do eixo y, ou seja, apenas reconfigurando as coordenadas x da imagem. Essa já é a primeira medida visando uma padronização dos dados a serem armazenados em arquivo.

Durante o processo de interação com o sistema, o usuário informará os pontos da base da nadadeira do golfinho. O algoritmo de *Corte* simplesmente faz um corte inferior da nadadeira, através da equação da reta formada por esses pontos. Assim, todos os pontos inferiores a essa reta são colocados para "0", ou seja, farão parte do fundo da imagem.

O processo de rotulação da imagem é fundamental para a identificação de aglomerados de pontos. Devido a ruídos presentes na imagem após o pré-processamento, é comum existirem vários objetos presentes, mas acontece que só nos interessa a nadadeira, por isso, é feita uma filtragem de forma a eliminar ruídos remanescentes, ficando apenas a nadadeira. Esse processo de filtragem elimina objetos com área de píxeis menores que um certo limiar. Neste ponto, temos uma imagem binária apenas com a nadadeira cortada e ainda não rotacionada.

Para os processos de extração de fronteira e esqueletização usam-se as técnicas anteriormente descritas nas seções 3.4.6 e 3.4.7, respectivamente. É de fundamental importância essas novas imagens no pós-processamento, pois outros algoritmos dependerão delas.

Para extração do pico da nadadeira na imagem são utilizadas as imagens do esqueleto e da fronteira. Inicialmente, são tomados dois pontos superiores do esqueleto da nadadeira espaçados de alguns píxeis. É feita então uma reta a partir desses pontos e prolongada até o encontro do ponto mais extremo da fronteira da nadadeira que obedeça a angulação da reta obtida. Esse processo foi feito através dessas imagens, pois a forma superior do esqueleto aponta para o pico da nadadeira. A imagem 3.7 demonstra graficamente como esse algoritmo é determinado.

Para efeitos de validação do algoritmo aqui descrito é importante ressaltar que o usuário deve informar um ponto próximo ao pico da nadadeira. Caso o ponto encontrado

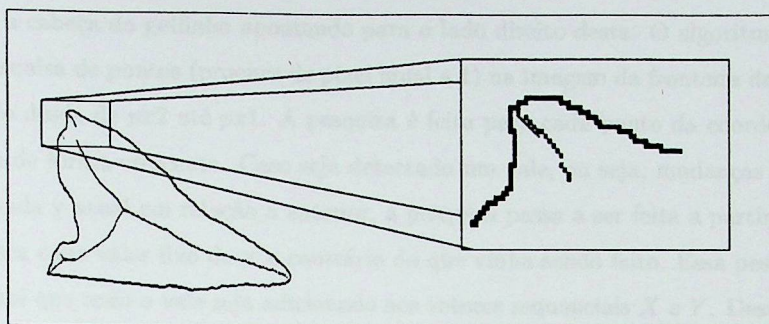


Figura 3.7: Extração do pico da nadadeira na imagem

esteja muito aquém do esperado, é feita uma pesquisa na vizinhança desse ponto e o ponto de pico passa agora a ser este novo ponto.

Decidiu-se usar esse algoritmo baseado em um prolongamento de reta, pois, em algumas imagens, não é possível usar alguns pontos do esqueleto para aproximá-los a uma curva e, em seguida, prolongá-la e achar o pico. Nesses esqueletos aparecem pontos bem distantes da curva procurada, inviabilizando assim uma curva paramétrica adequada.

Sequenciamento dos pontos da nadadeira

Os descritores usados nesse trabalho para a identificação do animal baseiam-se basicamente em características de forma. Para obtermos esses descritores é necessário primeiramente ordenar os píxeis da imagem de fronteira do objeto. A seção 3.4.6 demonstra como é obtida a fronteira de uma imagem binária.

O objetivo desse algoritmo é armazenar em dois vetores X e Y a seqüência adequada de pontos da nadadeira. Adota-se aqui pontos iniciando do lado esquerdo da imagem e indo até o lado direito.

Durante a fase de interação com usuário do sistema são fornecidos três pontos distintos:

1. Posição inicial da nadadeira (mais próximo da cabeça): $px1, py1$
2. Pico da nadadeira: $picoX, picoY$
3. Posição final da nadadeira (mais próximo da cauda): $px2, py2$

A partir desses pontos, é determinado se deve haver uma reflexão da imagem deixando

sempre a cabeça do golfinho apontando para o lado direito desta. O algoritmo faz então uma pesquisa de pontos (procura de píxel igual a 1) na imagem da fronteira da nadadeira varrendo desde de $px2$ até $px1$. A pesquisa é feita para cada ponto da coordenada x da imagem de forma crescente. Caso seja detectado um vale, ou seja, mudanças bruscas da coordenada y atual em relação a anterior, a pesquisa passa a ser feita a partir de valores de x , para cada valor fixo de y , o contrário do que vinha sendo feito. Essa pesquisa em y é feita até que todo o vale seja adicionado aos vetores sequenciais X e Y . Desta forma, a maior parte dos pontos de uma fronteira são sequenciados corretamente nos vetores X e Y . A figura 3.8 demonstra a fronteira de uma nadadeira e a sequência de pontos obtidos e plotados novamente.

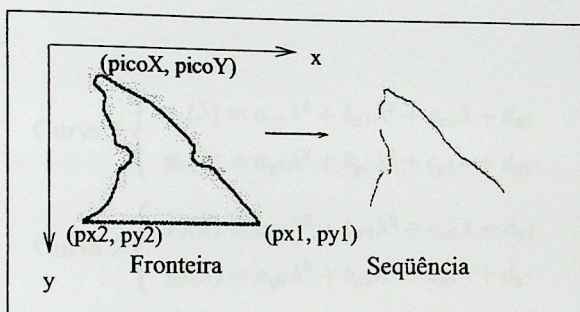


Figura 3.8: Extraíndo a sequência de pontos a partir da fronteira da imagem

Nota-se que apenas os pontos “externos” da fronteira da nadadeira são adicionados aos vetores de sequência. Percebe-se então que a imagem da fronteira parece ter linhas mais sólidas, enquanto que a imagem sequenciada possui linhas mais estreitas.

O algoritmo de sequenciamento de pontos da fronteira da nadadeira atua como se fosse um observador que varre todo o lado externo da imagem da fronteira passada, enquanto adiciona cada ponto aos dois vetores X e Y .

Representação polinomial da nadadeira do golfinho

Neste trabalho, representamos a nadadeira do golfinho através de dois polinômios paramétricos de terceiro grau. A figura 3.9 demonstra como é feita a parametrização das curvas. Note que as duas curvas são coincidentes em $\lambda = 0$. Esse artifício foi feito para

que tivéssemos curvas que se tocassem em um ponto conhecido, neste caso é exatamente o pico da dadadeira. As equações paramétricas para as duas curvas são expressas da seguinte forma:

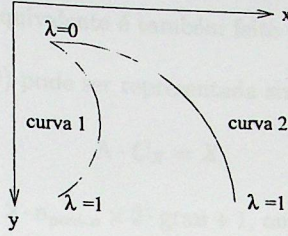


Figura 3.9: Modelo gráfico das curvas paramétricas usadas

$$\text{Curva 1} \begin{cases} x_1(\lambda) = a_{x1}\lambda^3 + b_{x1}\lambda^2 + c_{x1}\lambda + d_{x1} \\ y_1(\lambda) = a_{y1}\lambda^3 + b_{y1}\lambda^2 + c_{y1}\lambda + d_{y1} \end{cases} \quad (3.26)$$

$$\text{Curva 2} \begin{cases} x_2(\lambda) = a_{x2}\lambda^3 + b_{x2}\lambda^2 + c_{x2}\lambda + d_{x2} \\ y_2(\lambda) = a_{y2}\lambda^3 + b_{y2}\lambda^2 + c_{y2}\lambda + d_{y2} \end{cases} \quad (3.27)$$

$$\text{Restrições:} \begin{cases} x_1(0) = x_2(0) \Rightarrow x_1(0) = d_{x1} \quad e \\ x_2(0) = d_{x2} \Rightarrow d_{x1} = d_{x2} = ? \end{cases} \quad (3.28)$$

Podemos resolver essas equações de forma matricial, onde montamos matrizes para resolução dos coeficientes x para as duas curvas, e depois matrizes para resolução dos coeficientes y também das duas curvas. Matricialmente, a solução dos coeficientes x podem ser organizados da seguinte forma:

$$\begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ \lambda_1^3 & \lambda_1^2 & \lambda_1 & 1 & 0 & 0 & 0 \\ & & \dots & & & & \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \lambda_1^3 & \lambda_1^2 & \lambda_1 \\ & & \dots & & & & \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} a_{x1} \\ b_{x1} \\ c_{x1} \\ d_{x1} \\ a_{x2} \\ b_{x2} \\ c_{x2} \end{bmatrix} = \begin{bmatrix} x_1(0) \\ x_1(\lambda_1) \\ \dots \\ x_1(1) \\ x_2(0) \\ x_2(\lambda_1) \\ \dots \\ x_2(1) \end{bmatrix} \quad (3.29)$$

onde

- $0, \lambda_1, \lambda_2, \dots, 1$ são os pontos de λ normalizados de 0 a 1.
- $x_1(0), x_1(\lambda_1), \dots, x_1(1)$ são os pontos correspondentes da coordenada x na curva 1 em $\lambda = 0, \lambda_1, \dots, 1$. O equivalente é também feito para a curva 2.

A equação matricial (3.29) pode ser representada simplesmente por:

$$\Lambda \cdot C_X = X \quad (3.30)$$

A dimensão da matriz Λ é $2 \cdot n_{\text{pontos}} \times 2 \cdot \text{grau} + 1$, onde $\text{grau} = 3$ e n_{pontos} é o número de pontos x mapeados em $\lambda = 0, \dots, 1$. A dimensão de C_X (coeficiente x das duas curvas) é 1×7 (já que o grau do polinômio é 3, o que nos dá 8 coeficientes, mas apenas 7 para calcular). A dimensão de X é $2 \cdot n_{\text{pontos}} \times 1$.

Resolvendo esta equação para C_X , obtemos:

$$C_X = \text{pseudoInv}(\Lambda) \cdot X \quad (3.31)$$

onde $\text{pseudoInv}(\Lambda)$ é a pseudoinversa de Λ , dada por:

$$\text{pseudoInv}(\Lambda) = (\Lambda^t \Lambda)^{-1} \Lambda^t$$

Após estas operações, obteremos os 4 coeficientes x da curva 1 e 3 coeficientes da curva 2, sabendo que o quarto coeficiente é igual ao quarto coeficiente da curva 1. Do mesmo modo procede-se para os coeficientes paramétricos das duas curvas em y .

Deteção de fendas

Outro descritor de forma que foi definido neste trabalho foi a quantidade e posição das fendas presentes na nadadeira do animal. É importante considerar não apenas se uma determinada curva da nadadeira possui fendas e quantas são, mas também onde elas estão. Essas fendas são produzidas na nadadeira do animal principalmente por agressores marinhos. O tipo de cartilagem resistente que forma a nadadeira permite que essa fenda se torne uma marca para o golfinho durante muitos anos. É claro que pode ocorrer deste animal sofrer novas agressões descaracterizando-o da identificação anterior, mas isso não é um problema das técnicas de detecção, mas natural do ambiente. Neste caso, os biólogos devem reconsiderar este animal.

É fácil prever casos em que dois animais diferentes possuem a mesma quantidade de fendas em um mesmo lado da nadadeira. Mas se for considerada não apenas a quantidade, mas também a posição relativa destas fendas na nadadeira, diminuiremos bastante os casos de semelhança.

A figura 3.10 demonstra o processo realizado com as imagens. O algoritmo aqui implementado funciona da seguinte forma:

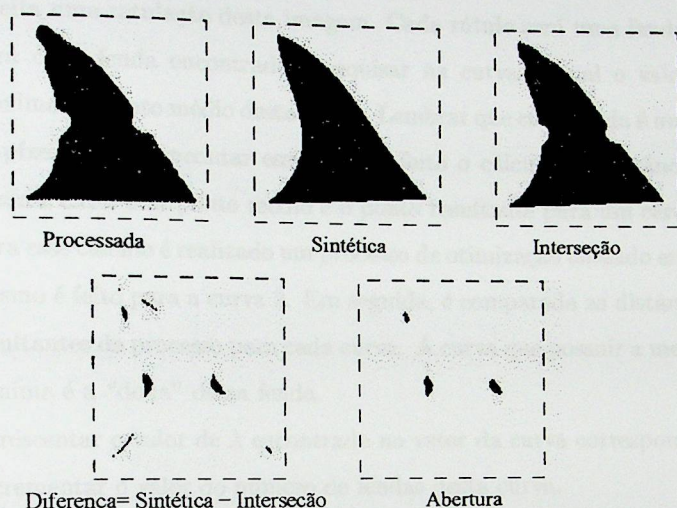


Figura 3.10: Detecção de fendas da nadadeira

1. Considerar como entrada a imagem limiarizada resultante dos processos anteriores e os coeficientes das duas curvas cúbicas encontradas.
2. A saída é a quantidade de fendas em cada curva e os vetores contendo a localização correspondente de cada fenda encontrada em cada curva. É retornada também uma imagem binária mostrando as fendas detectadas.
3. O processamento é realizado da seguinte forma:
 - Gerar a imagem de uma nadadeira a partir dos polinômios passados - a imagem sintética;
 - Fazer a interseção desta imagem resultante com a imagem processada;
 - Fazer a diferença entre a imagem sintética e o resultado da interseção.

- Criar um elemento estruturante proporcional ao tamanho da imagem para ser usado no processo seguinte;
- Proceder o processo morfológico de *abertura*, explicado anteriormente. Neste instante, os objetos de pequeno tamanho resultantes do passo anterior são eliminados, apenas os maiores permanecerão na imagem. É exatamente essa a função do processo de *abertura*.
- É feita uma rotulação desta imagem. Cada rótulo será uma fenda detectada.
- Para cada fenda encontrada, pesquisar na curva 1 qual o valor de λ mais próximo do ponto médio desta fenda. Lembrar que cada fenda é um aglomerado de píxeis. Para encontrar esse valor é feito o cálculo da distância euclidiana mínima entre esse ponto médio e o ponto resultante para um certo valor de λ . Para esse cálculo é realizado um processo de otimização baseado em bisseção. O mesmo é feito para a curva 2. Em seguida, é comparada as distâncias mínimas resultantes do processo para cada curva. A curva que possuir a menor distância mínima é a “dona” desta fenda.
- Acrescentar o valor de λ encontrado no vetor da curva correspondente.
- Incrementar o valor do número de fendas desta curva.

Descritores de Curvatura

Uma importante característica que distingue os animais é o formato da sua nadadeira. Dentre as imagens de animais analisadas notou-se diferentes curvaturas, facilmente distinguíveis ao olho humano. Para tentar capturar esta noção de formato da nadadeira, foram definidos dois descritores: os raios de curvatura dos dois lados da barbatana e os índices de discrepância ao círculo que medem a semelhança entre um lado da barbatana e um arco de círculo.

Estes descritores de curvatura baseiam-se na obtenção do centro do arco de círculo que melhor se ajusta aos polinômios descritores de um lado da barbatana do animal. Supondo que o centro do círculo está no ponto (x_c, y_c) e que são escolhidos n pontos ao longo da curva paramétrica que descreve um lado da barbatana, este ponto central é calculado de

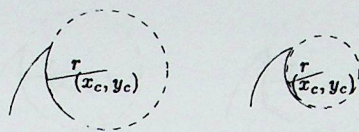


Figura 3.11: Duas nadadeiras com raios descritores distintos

tal forma a minimizar a função de custo $J(x_c, y_c)$ definida por

$$J(x_c, y_c) = \sum_{i=1}^n (l_i^2 - l_{i-1}^2) \quad (3.32)$$

onde $l_i^2 = (x_i - x_c)^2 + (y_i - y_c)^2$. Resolvendo-se esta equação pode-se chegar a seguinte forma matricial:

$$Ax = B \quad (3.33)$$

que é expandida da forma:

$$\begin{bmatrix} 2(x_0 - x_1) & 2(y_0 - y_1) \\ 2(x_1 - x_2) & 2(y_1 - y_2) \\ \vdots & \vdots \\ 2(x_{n-1} - x_n) & 2(y_{n-1} - y_n) \end{bmatrix} \begin{bmatrix} x_c \\ y_c \end{bmatrix} = \begin{bmatrix} (x_0^2 - x_1^2) + (y_0^2 - y_1^2) \\ (x_1^2 - x_2^2) + (y_1^2 - y_2^2) \\ \vdots \\ (x_{n-1}^2 - x_n^2) + (y_{n-1}^2 - y_n^2) \end{bmatrix} \quad (3.34)$$

Após a obtenção do ponto (x_c, y_c) é efetuado o cálculo das distâncias euclidianas de cada um dos n pontos de um lado da barbatana ao ponto (x_c, y_c) . O raio é dado pelo valor médio destas distâncias. A figura 3.11 mostra duas nadadeiras que diferem no tamanho do raio para aquele lado.

Para o cálculo do índice de discrepância, efetua-se o somatório do quadrado das diferenças entre o valor do raio médio e a distância a cada um dos n pontos considerados. Note-se que, se um lado da nadadeira fosse exatamente um arco de um círculo, o índice de discrepância seria zero. As duas nadadeiras da figura 3.11, por exemplo, teriam índice de discrepância bastante próximos de zero para aquele lado. Já a figura 3.12 mostra duas nadadeiras com o mesmo raio, mas que diferem no índice de discrepância para aquele lado.

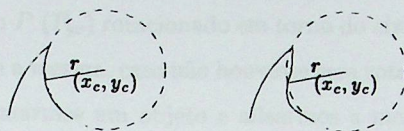


Figura 3.12: Duas nadadeiras com índices de discrepância distintos

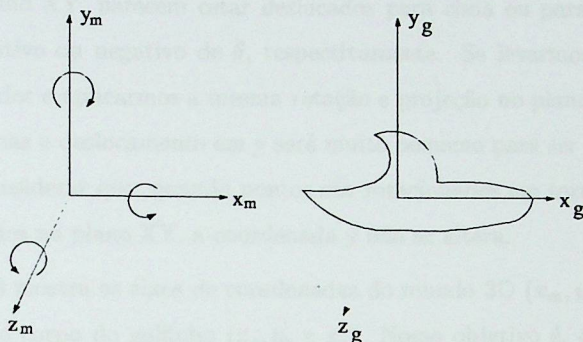


Figura 3.13: Eixos coordenados escolhidos

Obtenção das dimensões reais do golfinho mapeadas no plano de imagem

Sabe-se que a imagem de uma câmera é um mapeamento da imagem real do mundo 3D em um plano, ou seja, representação 2D. Quando obtém-se uma dessas imagens para usar no processamento, tem-se que considerar que no mundo real o golfinho pode possuir rotações nos três eixos: x , y e z . Acontece que essas imagens são obtidas normalmente de embarcações a uma certa distância dos golfinhos, exatamente porque esses animais evitam estarem muito próximos dessas embarcações. Sendo assim, podemos fazer algumas simplificações importantes. A primeira é que o ângulo de rotação sobre o eixo x é desprezível. Variações maiores seriam consideráveis se a foto fosse tirada muito próxima do animal. Para melhor entendimento sobre os ângulos de rotação nos três eixos, observe a figura 3.13. Consideramos que o corpo do golfinho está distribuído ao longo do eixo x_g . Sua cabeça aponta na direção positiva desse eixo. A largura do animal está distribuída sobre o eixo z_g , com seu centro em $z_g = 0$. A altura do golfinho está ao longo do eixo y_g , com o pico da nadadeira apontando para valores positivos desse eixo.

A segunda simplificação é que, como o animal está a uma distância considerável, a

cordenada y e de um ponto P (P_{yr}) rotacionado em torno do eixo y e projetada no plano de imagem, será exatamente a mesma, caso não houvesse essa rotação (P_y). É fácil identificar essa simplificação se tormarmos um objeto e alisarmos a mesma rotação em dois casos distintos: próximo ao observador e longe dele. Quando fazemos uma rotação θ em torno do eixo y e em um objeto próximo ao observador, notamos que alguns pontos, quando mapeados no plano XY , parecem estar deslocados para cima ou para baixo, de acordo com o valor positivo ou negativo de θ , respectivamente. Se levarmos esse objeto para longe do observador e aplicarmos a mesma rotação e projeção no plano XY irá acontecer a mesma coisa, mas o deslocamento em y será muito pequeno para ser considerado. Com isso, podemos considerar que, quando pontos são rotacionados em torno do eixo y e, em seguida, projetados no plano XY , a coordenada y não se altera.

A figura 3.13 mostra os eixos de coordenadas do mundo 3D (x_m, y_m e z_m) e os eixos colocados sobre o corpo do golfinho (x_g, y_g e z_g). Nosso objetivo é, dada uma imagem de uma foto expressa no plano XY , com $z = 0$, obter uma aproximação dos pontos reais da nadadeira no mundo mapeados de volta no plano 2D, considerando agora que a nadadeira esteja paralela ao observador sem qualquer rotação, conforme mostrado na figura 3.13. Vamos usar a figura 3.14 para melhor compreender as equações seguintes. Nesta figura mostrada, consideramos o observador colocado sobre a origem dos eixos do mundo. Note que temos uma vista superior da nadadeira do golfinho. Tomemos um ponto P da nadadeira para análise. Dado um ângulo de rotação θ em torno do eixo y , podemos encontrar as seguintes equações:

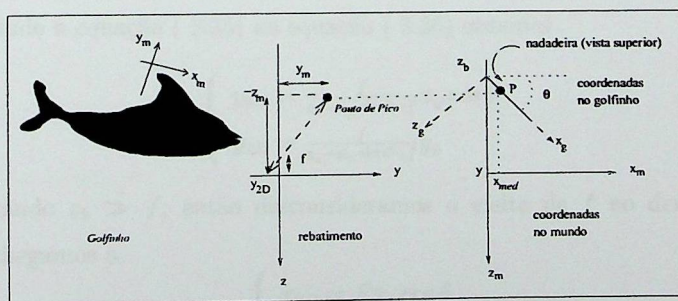


Figura 3.14: Mapeamento entre planos 3D e 2D

$$\begin{cases} x_m = x_g \cos \theta \\ y_m = y_g \\ z_m = -z_b + x_g \sin \theta \end{cases} \quad (3.35)$$

onde

- x_m, y_m, z_m são as coordenadas do ponto P no mundo.
- z_b é a coordenada da base da nadadeira no mundo.
- $x_b = 0$.

Podemos tirar outras relações a partir do rebatimento desse ponto no plano YZ . A partir da descrição desse gráfico e fazendo o mesmo para o plano XZ podemos extrair as seguintes relações:

$$\begin{cases} \frac{y_m}{-z_m - f} = \frac{y_{2D}}{f} \\ \frac{x_m}{-x_m - f} = \frac{y_{2D}}{f} \end{cases} \quad (3.36)$$

A partir da equação (3.36) podemos chegar a

$$\begin{cases} y_{2D} = \frac{f}{-z_m - f} y_m \\ x_{2D} = \frac{f}{-z_m - f} x_m \end{cases} \quad (3.37)$$

onde

- x_{2D}, y_{2D}, z_{2D} são as coordenadas do ponto P na imagem.
- f é a distância focal.

Substituindo a equação (3.35) na equação (3.36) obtemos

$$\begin{cases} y_{2D} = \frac{f}{z_b - x_g \sin \theta - f} x_g \cos \theta \\ y_{2D} = \frac{f}{z_b - x_g \sin \theta - f} y_g \end{cases} \quad (3.38)$$

Considerando $z_b \gg f$, então desconsideramos o efeito de f no denominador das equações e chegamos a

$$\begin{cases} x_{2D} = F x_g \cos \theta \\ y_{2D} = F y_g \end{cases} \quad (3.39)$$

onde $F = \frac{f}{z_b - x_g \sin \theta}$.

Isolando x_g e y_g na equação anterior, ficamos com

$$\begin{cases} x_g = \frac{1}{F \cos \theta} x_{2D} \\ y_g = \frac{1}{F} y_{2D} \end{cases} \quad (3.40)$$

Sabendo que F é apenas um fator de escala entre os planos da imagem e o plano do mundo, então chegamos a

$$\begin{cases} x_g \cong \frac{x_{2D}}{\cos \theta} \\ y_g \cong y_{2D} \end{cases} \quad (3.41)$$

Isso faz sentido, pois podemos notar que para $\theta = 0$ temos que $x_g \cong x_{2D}$, ou seja, se não houver rotação em torno do eixo y , x_g mapeia diretamente em x_{2D} , a diferença entre eles é apenas um fator de escala F . Para valores de $0 < \theta < 90$ o valor de x_g visto no plano de imagem sofre um esticamento.

Note que consideramos apenas uma rotação em relação ao eixo y . Caso os pontos da base da nadadeira ainda estejam desalinhados é porque falta uma rotação no eixo z . Sendo assim, após a rotação em y considerada até agora, calculamos o ângulo α de inclinação da base da nadadeira e aplicamos a seguinte transformação:

$$\begin{pmatrix} x_P \\ y_P \\ z_P \end{pmatrix} = \begin{pmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_g \\ y_g \\ 0 \end{pmatrix} \quad (3.42)$$

onde

- x_P, y_P, z_P são os pontos da nadadeira do golfinho mapeados em um plano XY com o observador paralelo a ele e olhando-o.
- α é o ângulo de rotação em torno do eixo z colocado no golfinho que deve ser feito para que a nadadeira fique com a base alinhada com o eixo x do golfinho.

Resolvendo a equação matricial (3.42) chegamos a

$$\begin{cases} x_P = x_g \cos \alpha - y_g \sin \alpha \\ y_P = x_g \sin \alpha + y_g \cos \alpha \end{cases} \quad (3.43)$$

Deve-se lembrar que essas transformações são aplicadas às equações paramétricas dos dois polinômios que representam a nadadeira do golfinho. Assim, chegamos às equações

que representam a nadadeira sem qualquer rotação, ou seja, alinhada ao eixo x . Após essas transformações é necessária a normalização das equações, pois poderemos receber imagens de tamanhos variados. A seção seguinte demonstra os passos necessários para normalização dos polinômios representativos da nadadeira do golfinho.

Normalização dos polinômios

As imagens advindas dos experimentos realizados por biólogos e estudantes apresentam diversas variações como:

- Iluminação: as imagens são registradas em condições climáticas as mais diversas. A presença de reflexos do sol na nadadeira do golfinho prejudicam fortemente o processamento correto da imagem;
- Escala: as imagens são registradas a distâncias variáveis. Consideramos aqui que essa distância está em torno de 10 metros do animal;
- Rotação: o efeito de rotação pode esconder uma característica muito importante na distinção de animais - o número de fendas. Imagens com rotações em torno do eixo y próximas de 90 graus prejudicarão fortemente o processamento adequado.

Dentre estes itens citados, usa-se os dois últimos para o processo de normalização dos polinômios.

Após todo o pré-processamento e o pós-processamento deve ser realizada uma normalização nos dados obtidos. Vale lembrar aqui que essa normalização não é aplicada a uma imagem, mas aos coeficientes que descrevem as duas curvas cúbicas paramétricas. Para efeitos de extração de características e posterior comparação com outros animais essa etapa é de fundamental importância.

O diagrama 3.15 mostra a sequência de etapas aplicadas aos polinômios originais até que se chegue a dados normalizados.

Inicialmente é aplicada uma rotação em torno do eixo z (veja a figura 3.15(b)). Essa rotação é feita através dos pontos $px1$ e $px2$ fornecidos. Esses pontos delimitam o corte inferior da nadadeira, separando-a do corpo do animal. A partir desses pontos é determinado o ângulo de rotação necessário para que a nadadeira fique com a base paralela ao eixo x . Em seguida é feita a primeira transladação dos pontos de forma que

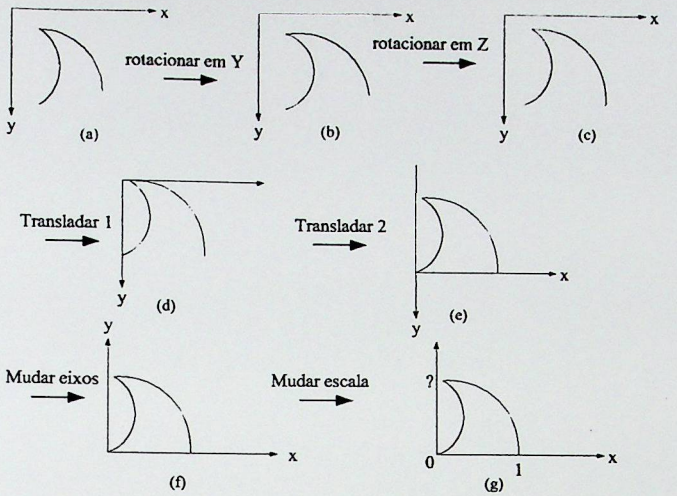


Figura 3.15: Sequência para normalização dos polinômios

$px2$ (ponto x mais próximo da cauda) fique em $x = 0$ e o ponto superior de y fique em $y = 0$ (figura 3.15(c)).

Note que até agora estávamos trabalhando com o eixo y com valores positivos para “baixo”. Neste ponto, fazemos uma nova transladação dos pontos de forma que passemos todos esses pontos para “cima”, neste caso os valores ficarão negativos. Isso é feito através da subtração de todos os valores das coordenadas y pelo maior valor de y , deixando esse último em $y = 0$ (figura 3.15(d)). Em seguida, fazemos a mudança de eixos para que o eixo y agora “aponte para cima”.

O último passo é proceder uma transformação de escala nos pontos. Assumimos que o maior valor na coordenada x será “1”, desta forma o ponto $px1$ estará em $x = 1$ e o ponto $px2$ estará em $x = 0$. O valor máximo de y fica proporcional a esse fator de escala aplicado em x (3.15(f)).

No próximo capítulo serão vistas as técnicas de redes neurais aplicadas no trabalho e como é feita a classificação dos indivíduos.

Capítulo 4

Redes Neurais Artificiais

4.1 Introdução

Neste capítulo, apresentaremos uma descrição de um modelo neural artificial utilizado neste trabalho. Inicialmente, discutiremos alguns aspectos de desenvolvimento de redes neurais artificiais e, logo após, apresentaremos os métodos de aprendizado, e em seguida, a Rede Competitiva utilizada neste trabalho, que é utilizada na segmentação de imagens. Explicaremos também o processo de classificação neural. Além da rede competitiva explicaremos também o funcionamento de uma rede de retropropagação que é usada para o processo de classificação das partes de cada imagem submetida ao sistema.

Uma rede neural artificial é um processador computacional paralelo distribuído, constituição de unidades de processamento simples, que imita a função natural de armazenar conhecimento e tomar as decisões para o uso (Fig. 34). Ela se assemelha ao cérebro humano em duas partes:

1. Através de um processo de aprendizagem, o conhecimento é adquirido pela rede.
2. Os pesos das conexões entre as neurônios, conhecidos como pesos sinápticos, são utilizados para armazenar o conhecimento.

As redes neurais artificiais são sistemas computacionais que operam em um modelo matemático da estrutura neural dos organismos. Os pesos e conexões que uma rede neural artificial adquirem o conhecimento de experiências passadas e presentes. Elas possuem uma característica de serem estruturas capazes de aprender comportamentos baseados em um

Reg. 187.038104

Capítulo 4

Redes Neurais Artificiais

4.1 Introdução

Neste capítulo, apresentaremos uma descrição sobre as redes neurais artificiais utilizadas neste trabalho. Inicialmente, faremos um breve histórico do desenvolvimento de redes neurais artificiais e, logo após, serão descritos os métodos de aprendizado, e em seguida, a Rede Competitiva utilizada neste trabalho, que é utilizada na segmentação da imagem. Explicaremos também o processo de quantização vetorial. Além da rede competitiva explicaremos também o funcionamento de uma rede de retropropagação que é usada para o processo de classificação dos padrões de nadadeiras submetidos ao sistema.

Uma rede neural artificial é um processador maciçamente paralelo distribuído constituído de unidades de processamento simples, que têm a função natural de armazenar conhecimento e torna-lo acessível para o uso [Hay99]. Ela se assemelha ao cérebro humano em dois pontos:

1. Através de um processo de aprendizagem, o conhecimento é adquirido pela rede.
2. Os pesos das conexões entre os neurônios, conhecidos como pesos sinápticos, são utilizados para armazenar o conhecimento.

As redes neurais artificiais são técnicas computacionais que apresentam um modelo matemático da estrutura neural dos organismos inteligentes e considera-se que redes neurais artificiais adquirem conhecimento de experiências passadas e presentes. Elas possuem uma característica de serem estruturas capazes de aprender comportamentos baseados em um

treinamento com ou sem professor. A estrutura de uma RNA é composta por várias unidades de processamento, os neurônios artificiais, que trabalham interconectados de forma paralela e distribuída [Hay99].

Neste trabalho está sendo usada uma rede neural artificial para a segmentação da imagem.

4.2 Redes Neurais Artificiais

As redes neurais surgiram por volta dos anos 50, a idéia era simular a atividade cognitiva de um ser humano, definindo componentes lógicos semelhantes aos neurônios humanos. O cérebro é o local no corpo humano onde se encontra a inteligência e os neurônios ali instalados são considerados entidades básicas, os quais são interligados a outros neurônios, permitindo desta forma a troca de informação entre eles, criando a inteligência humana ou biológica. Os cérebros são formados por um conjunto extremamente complexo de células, os conhecidos neurônios biológicos.

O neurônio biológico é uma célula com núcleo e corpo aonde acontecem as reações químicas e elétricas que representam o processamento da informação. Na figura 4.1 apresentamos um modelo do neurônio biológico. A informação que sai do corpo (soma) é realizada por impulsos elétricos que se propagam através do axônio. No final do axônio existem diversas ramificações que distribuem a informação para outros neurônios. A comunicação entre os neurônios biológicos se dá através das sinapses. Estas sinapses estão conectadas a um dendrito do neurônio receptor, que é a região onde dois neurônios entram em contato e, através da qual, os impulsos nervosos são transmitidos entre eles. É através da sinapse que é disparada uma substância neurotransmissora que flui do corpo celular, isto acontece quando a sinapse é excitada pelo impulso do axônio. A substância se propaga entre a sinapse e o dendrite realizando desta forma a comunicação entre os dois neurônios.

A partir dessa idéia do neurônio biológico o homem sempre teve o anseio de copiar ou imitar o funcionamento do cérebro com essa troca de informações que existe entre os neurônios. McCulloch e Pitts apresentaram um modelo de neurônio artificial como sendo uma unidade básica de processamento da informação de uma rede neural, onde

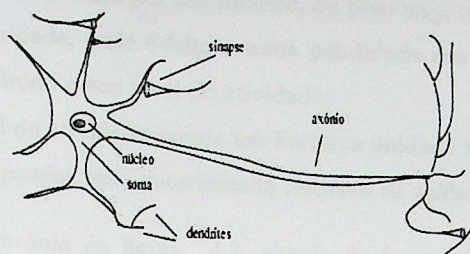


Figura 4.1: Neurônio biológico

uma rede neural pode ser formada por várias unidades de processamento. Essas unidades estão ligadas por canais de comunicação através de seus respectivos pesos. As unidades realizam operações apenas em seus dados, que são entradas recebidas pelas suas conexões. Dá interação entre as unidades de processamento provém o comportamento inteligente de uma rede neural.

Basicamente um neurônio artificial recebe estímulos como entrada e gera um sinal de saída como resultado somatório destes estímulos.

Na figura 4.2, apresenta-se o modelo de um neurônio artificial. Este modelo imita a capacidade que as células do sistema nervoso tem de realizar procedimentos biológicos.

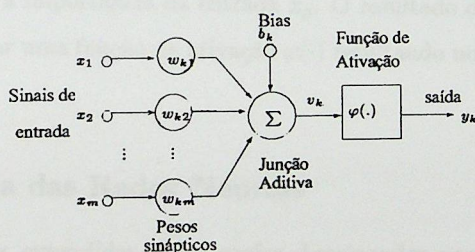


Figura 4.2: Modelo não-linear de um neurônio

O funcionamento do neurônio proposto por McCulloch e Pitts, pode ser resumido da seguinte forma:

- Os sinais são apresentados à entrada, representados por x_p na figura;

- Cada sinal é multiplicado por um número, ou peso $w_{k,j}$, que indica a sua influência na saída da unidade, onde é feita a soma ponderada dos sinais que produzem um valor que é definido como nível de atividade;
- Quando o nível de atividade excede um limite, a unidade representada como função de ativação, produz uma determinada resposta de saída.

O modelo de neurônio da figura 4.2, simula de forma simplificada as realidades biológicas que ocorrem dentro de uma célula do sistema nervoso. Matematicamente o modelo pode ser descrito da seguinte forma:

$$u_k = \sum_{j=0}^p w_{k,j} x_j \quad (4.1)$$

$$Y_k = \varphi(V_k) \quad (4.2)$$

Em que $x_1, x_2, \dots, x_p$ são os sinais de entrada; $w_{k1}, w_{k2}, \dots, w_{kp}$ são os pesos sinápticos do neurônio k ; V_k é a saída da combinação linear; b_k é o bias (limiar de ativação); $\varphi(\cdot)$ é a função de ativação e Y_k é o sinal de saída. Assim, as informações fornecidas por outros neurônios entram em p entradas do neurônio processador e o processamento consiste de uma combinação linear das entradas x_j (sinapses) onde a cada entrada está associado um peso w_{kj} que reflete a importância da entrada x_p . O resultado dessa combinação é o valor V_k , que passará por uma função de ativação $\varphi(\cdot)$ resultando no sinal de saída Y_k do neurônio.

4.2.1 Arquitetura das Redes Neurais

As RNA's podem ser entendidas como grafos direcionados nos quais os neurônios artificiais são os nós e os arcos direcionados (com pesos) são as conexões entre as saídas dos neurônios e as entradas dos neurônios. A arquitetura geralmente é organizada em forma de camadas, com unidades que podem estar conectadas às unidades da camadas posterior. As RNA's são compostas de diversos neurônios e podem possuir uma, duas ou diversas camadas. Os neurônios podem ser arranjados em várias formas de conexão. Algumas redes permitem apenas o fluxo de sinais para frente "feedforward", por exemplo perceptron,

enquanto outras podem possuir conexões retroativas “feedback”, perceptron de múltiplas camadas. Há também as redes recorrentes as quais diferenciam-se das redes feedforward pela existência de pelo menos um laço (loop) de recorrência (feedback), exemplo redes de Hopfield. A arquitetura de uma RNA esta intimamente relacionada com a sua função. A saída da rede requer a ativação de um ou mais neurônios de saída. Estes resultados podem ser identificados como uma classificação de padrões (heteroassociação) ou como uma versão do padrão de entrada (autoassociação), livre de distorções e ruídos. Em geral diferencia-se as arquiteturas das redes neurais em três classes:

- redes feedforward de uma única camada, tem-se uma camada de entrada com neurônios cujas saídas alimentam a última camada da rede;
- redes feedforward de múltiplas camadas, diferencia-se da primeira pela existência de uma ou mais camadas entre a camada de entrada e a de saída;
- redes recorrentes, identifica-se pela existência de pelo menos um laço de recorrência (realimentação).

A adição de uma ou mais camadas ocultas habilita a rede a extrair estatísticas de mais alta ordem [Hay99]. Entretanto o aumento do número de camadas tem como consequência o aumento do tempo de aprendizado, desta forma dificultando a aplicação de redes com várias camadas em problemas onde a quantidade de dados seja elevada, como é o caso de imagens.

4.2.2 Técnicas de aprendizagem

A característica mais importante das redes neurais é a capacidade de aprender, através de um processo iterativo de atualização de pesos e limiares, e dessa forma melhorar o seu desempenho. A aprendizagem é o processo pelo qual os parâmetros livres de uma rede neural são adaptados através de um mecanismo de apresentação de estímulos fornecidos pelo ambiente no qual a rede esta inserida [Hay99]. Este conceito nos leva aos seguintes aspectos:

- O ambiente estimula a rede;
- Ocorre alteração nos parâmetros livres como resultado deste estímulo;

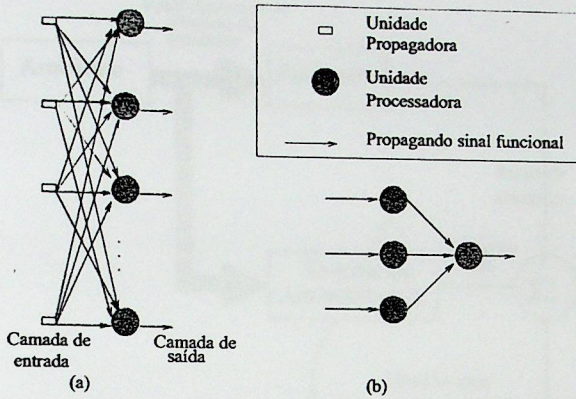


Figura 4.3: Rede neural tipo feedforward com uma camada

- Em consequência das modificações a rede responde ao ambiente com um novo padrão de resposta.

O tipo de treinamento é definido pela forma na qual os parâmetros são modificados. É necessário distinguir o processo de aprendizado utilizados em sistemas com capacidade de adaptação. Basicamente existem dois tipos de paradigmas de aprendizagem : a aprendizagem supervisionada e a aprendizagem auto-supervisionada.

Método de aprendizagem supervisionado

O método supervisionado tem como principal característica um “professor” que supervisiona externamente o aprendizado. Isto é, cada exemplo de treinamento é formado por um vetor de entrada e um vetor desejado, que é fornecido pelo supervisor. Os parâmetros da rede são ajustados pela combinação do sinal de entrada com um sinal de erro, que é a diferença entre a saída desejada e a fornecida pela rede [Hay99]. A figura 4.4 demonstra esse processo.

Método aprendizagem não-supervisionado

No aprendizado não-supervisionado ou auto-organizado não existe um “professor” para controlar o processo de treinamento. Isto é, não existem exemplos específicos a serem

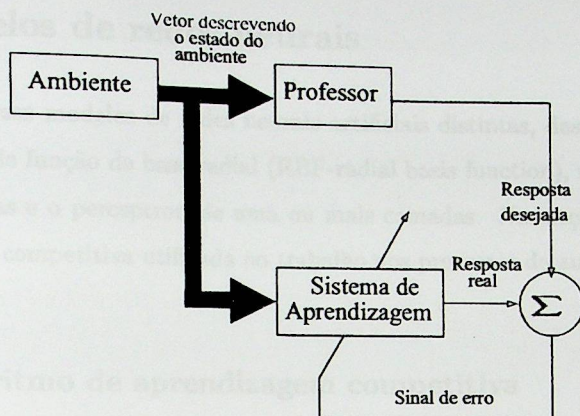


Figura 4.4: Processo de aprendizagem supervisionada

apresentados à rede. O treinamento é realizado através de uma medida da tarefa independentemente da qualidade da representação que a rede é requerida a aprender, isto é, a rede se adapta a regularidades estatísticas dos dados de entrada. Uma vez ajustada estará apta a formar representações internas para codificar características dos padrões de entrada, e desta forma criar novas classes automaticamente. A rede tem que achar atributos estatísticos relevantes. E ela mesma terá que desenvolver uma representação própria dos estímulos que entram na rede [Hay99]. Rotineiramente os métodos de aprendizagem não-supervisionada utilizam-se das regras da aprendizagem competitiva. A figura 4.5 mostra uma rede não supervisionada “vencedor leva tudo”.

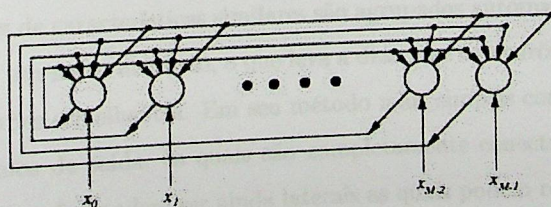


Figura 4.5: Processo de aprendizagem não-supervisionado

4.3 Modelos de redes neurais

Existem diversos modelos de redes neurais artificiais distintas, das quais podem ser citadas as redes de função de base radial (RBF-radial basis function), redes de Hopfield, redes competitivas e o perceptron de uma ou mais camadas. Nas seções seguintes será detalhada a rede competitiva utilizada no trabalho nos processos de auto-segmentação.

4.3.1 Algoritmo de aprendizagem competitiva

A característica marcante do aprendizado competitivo é que as unidades de saída da rede competem entre si para se tornarem ativas. Como resultado, somente um único neurônio de saída é ativado. Este método é apropriado para descobrir características estatísticas que podem ser utilizadas para classificar um conjunto de padrões de entrada. Este método é conhecido como o *vencedor-leva-tudo* ou "*winner take all*". A aprendizagem competitiva tem três elementos básicos:

- Um conjunto de neurônios iguais entre si, exceto pelos pesos das conexões;
- Um limite sobre o peso de cada neurônio;
- Um método que permite a competição entre os neurônios. O neurônio vencedor leva tudo.

A aprendizagem competitiva agrupa os dados de entrada, com base nas correlações dos dados. Padrões de características similares são agrupados automaticamente pela rede e identificados por um único neurônio, o que leva a dizer que os neurônios se especializam em grupos de padrões semelhantes. Em seu método mais simples consiste de uma única camada de neurônios de saída, os quais são completamente conectados com os outros neurônios. Estas conexões podem ser ainda laterais as quais podem realizar uma inibição entre os neurônios vizinhos. Os neurônios de saída i da rede conectam-se a todos os neurônios de entrada (x_j) através de pesos $w_{i,j}$, sendo $j = 1, 2, \dots, n$. Ao final da competição somente um neurônio i^* com a maior entrada total é o vencedor, isto é, para um determinado neurônio k ser o vencedor, seu campo local induzido v_k deve ser o maior

entre todos os outros neurônios, em relação a um determinado padrão de entrada x .

$$y_k = \begin{cases} 1 & \text{se } v_k > v_j \quad \forall j, j \neq k \\ 0 & \text{em todos os outros casos} \end{cases} \quad (4.3)$$

O sinal de saída y_k do neurônio vencedor é atualizado e setado com o valor 1. O sinal de saída de todos os outros neurônios são setados com 0. Uma regra simples de aprendizagem pode ser descrita como:

$$w_{k,j}(n+1) = w_{k,j}(n) + \Delta w_{k,j}(n) \quad (4.4)$$

$$\Delta w_{k,j} = \begin{cases} \eta(x_j - w_{k,j}) & \text{se } k \text{ vencer a competição,} \\ 0 & \text{se } k \text{ perder a competição} \end{cases} \quad (4.5)$$

O neurônio considerado vencedor sofre uma atualização $\Delta w_{k,j}$ no vetor de pesos $w_{k,j}$ e no caso de um neurônio não responder a um determinado vetor de entrada nenhuma aprendizagem é efetuada. Fica claro que somente os pesos do neurônio vencedor são atualizados. η é a taxa de aprendizagem. Este método tem a capacidade de mover o vetor de pesos do neurônio vencedor na direção do vetor padrão de entrada. Observa-se que, ao utilizarmos uma regra de aprendizagem competitiva, a rede não irá parar de aprender a menos que a taxa de aprendizagem η seja zero.

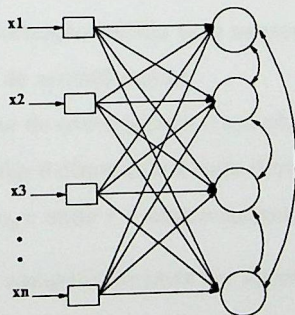


Figura 4.6: Modelo de uma rede competitiva com conexões diretas e laterais

Neste trabalho, a rede neural competitiva foi utilizada na quantização vetorial. Inicialmente os centros $w_{k,j}$ são encontrados de forma adaptativa. Estes N centros são os pesos da única camada da rede. A rede é formada de uma única camada com a dimensão

dos dados D e o número de centros N . São usadas 2 classes ou texturas fazendo-se que a rede tenha 2 saídas, cada uma representando uma textura nos dados da imagem a ser segmentada pela rede neural. A rede tem 3 entradas tendo em vista a escolha de 3 atributos para representar uma textura. No nosso sistema de detecção de golfinhos, esses atributos são a média dos pixels vizinhos para cada uma das componentes R G B. Os seguintes passos são utilizados para treinamento e calculo dos centros:

1. Inicializam-se os pesos da rede aleatoriamente;
2. São apresentados os padrões de entrada à rede;
3. Para cada padrão verifica-se qual neurônio possui a menor distância euclidiana entre seus pesos e o padrão apresentado. O que tiver a menor distância será considerado o vencedor;
4. Atualiza-se o neurônio vencedor através da seguinte equação:

$$w_{kj}(n+1) = w_{kj}(n) + \Delta w_{kj}(n) \quad (4.6)$$

$$\Delta w_{kj} = \alpha \eta (x_j - w_{kj}) \quad (4.7)$$

Em que :

- O $w_{kj}(n)$ é o neurônio vencedor na iteração n ;
- O $\Delta w_{kj}(n)$ é a atualização sofrida pelo neurônio vencedor;
- O η é o coeficiente de aprendizagem;
- O α é um coeficiente de convergência o seu objetivo e evitar oscilações na convergência. O seu valor é diminuído a cada iteração, matematicamente acontece o seguinte $\alpha_{n+1} = \alpha_n \tau$ onde τ é uma constante menor que 1.

Finalizando um algoritmo competitivo pode ser descrito como, sendo x_j o padrão de entrada:

1. Escolhem-se os N centros de maneira aleatória;
2. Escolhe-se um elemento x do conjunto de dados;
3. Verifica-se qual o neurônio vencedor w_{kj} para o elemento x ;
4. atualiza-se o neurônio vencedor;
5. Volta-se para o passo 2 até que um critério de parada seja atingido.

4.3.2 Quantização Vetorial

Um importante uso do algoritmo de treinamento competitivo é na quantização vetorial. Quantização vetorial é uma técnica em que um conjunto de dados formado por M vetores de dimensão D é separado em regiões, onde cada região é representada por um único vetor. Na quantização vetorial, escolhem-se N_p representantes para quantizar todo o conjunto, o resultado será outro conjunto C com N_p elementos (vetores) da mesma dimensão que os vetores do conjunto de dados inicial. Os vetores do conjunto X de dados é organizado em forma de tabela onde cada coluna é uma dimensão dos vetores do conjunto e cada linha representa um vetor completo. Os centros das regiões em que o conjunto foi dividido serão os N_p vetores. Após a obtenção dos centros é efetuada a classificação de cada vetor do conjunto de dados. Classificar um vetor de um conjunto de dados significa atribuir a este vetor uma única classe dentre as N escolhidas. Neste trabalho foi utilizado o classificador de distância mínima (distância euclidiana) para fazer esta atribuição. Este classificador atribui uma classe p dentre todas as N_p classes ao vetor i dentre todos os x_i do conjunto de dados X onde tenha-se uma classe $C_p \in C$ de modo que a distância $D(c_p, x_j)$ seja a menor dentre todos os elementos em C . Matematicamente isto é feito da seguinte forma:

$$p_{x_j} = \min_i \{D(c_i, x_j)\} \quad (4.8)$$

Desta forma podemos dizer que p_{x_j} é a classe atribuída ao vetor X_i e $D(x, y)$ é a distância euclidiana entre os vetores x e y dada por:

$$D(c_i, x_j) = \sqrt{(c_i - x_j) \times (c_i - x_j)} \quad (4.9)$$

4.4 Perceptrons de múltiplas camadas

A figura 4.7 mostra o grafo arquitetural de um perceptron de múltiplas camadas com uma camada oculta e uma camada de saída. A rede aqui mostrada é totalmente conectada. Isto significa que um neurônio em qualquer camada da rede está conectado a todos os nós/neurônios da camada anterior. O fluxo de sinal através da rede progride para frente, da esquerda para a direita e de camada em camada. As redes de perceptrons não necessariamente só possuem essa configuração. Podem existir redes com várias ca-

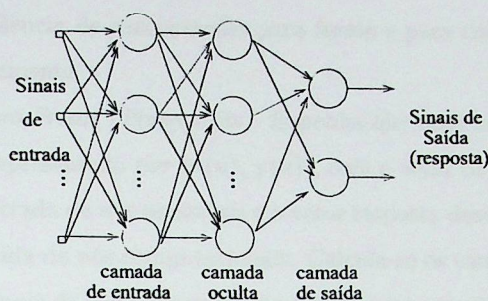


Figura 4.7: Rede perceptron de múltiplas camadas

camadas oculta. A camada de saída pode possuir mais de dois neurônios, tudo depende da configuração desejada da rede para a aplicação em um determinado problema.

A rede perceptron de múltiplas camadas empregada aqui é reconfigurável em sua camada de saída. Isso servirá para a adição de novos padrões de nadadeiras de golfinhos para serem analisadas, caso surja um novo padrão.

4.4.1 Algoritmo de retropropagação

O algoritmo de retropropagação utilizado em uma rede formada por perceptrons de múltiplas camadas funciona no modelo “com professor”, ou seja, trata-se de uma rede supervisionada.

A atualização seqüencial dos pesos é o método preferido para a implementação em tempo de execução (*on-line*) do algoritmo de retropropagação. Para este modo de operação, o algoritmo percorre os dados da amostra de treinamento $\{x(n), d(n)\}$ de $n = 1$ a N como segue:

1. *Inicialização.* Assumindo que nenhuma informação prévia esteja disponível, retira-se os pesos sinápticos e limiares de uma distribuição cuja média é zero e cuja variância é escolhida para que o desvio padrão dos campos locais induzidos dos neurônios se encontre na transição entre as partes lineares e saturada da função de ativação sigmóide.
2. *Apresentação dos Exemplos de Treinamento.* Apresenta-se uma época de exemplos de treinamento à rede. Para cada exemplo do conjunto, ordenado de alguma forma,

realiza-se a seqüência de computações para frente e para trás descritas nos pontos 3 e 4, respectivamente.

3. *Computação para Frente (Propagação)*. Suponha que num exemplo de treinamento da época seja representado por $(\mathbf{x}(n), \mathbf{y}(n))$, com o vetor de entrada $\mathbf{x}(n)$ aplicado à camada de entrada de nós sensoriais e o vetor resposta desejada $\mathbf{d}(n)$ apresentado à camada de saída de nós computacionais. Calcula-se os campos locais induzidos e os sinais funcionais da rede prosseguindo para frente através da rede, camada por camada. O campo local induzido $v_j^{(l)}(n)$ para o neurônio j na camada l é

$$v_j^{(l)}(n) = \sum_{i=0}^{m_0} w_{ji}^{(l)}(n) y_i^{l-1}(n) \quad (4.10)$$

onde $y_i^{l-1}(n)$ é o sinal (função) de saída do neurônio i na camada anterior $l-1$, na iteração n , e $w_{ji}^{(l)}(n)$ é o peso sináptico do neurônio j da camada l , que é alimentado pelo neurônio i da camada $l-1$. Para $i=0$, temos $y_0^{l-1}(n) = +1$ e $w_{j0}^{(l)}(n) = b_j^{(l)}(n)$ é o bias aplicado ao neurônio j na camada l . Assumindo-se o uso de uma função sigmóide, o sinal de saída do neurônio j na camada l é

$$y_j^{(l)} = \varphi_j(v_j(n)) \quad (4.11)$$

Se o neurônio j está na primeira camada oculta (i.e., $l=1$), faz-se

$$y_j^{(0)}(n) = x_j(n) \quad (4.12)$$

onde $x_j(n)$ é o j -ésimo elemento do vetor de entrada $\mathbf{x}(n)$. Se o neurônio j está na camada de saída (i.e. $l=L$, onde L é denominado a *profundidade* da rede), faz-se

$$y_j^{(L)} = o_j(n) \quad (4.13)$$

Calcula-se o sinal de erro

$$e_j(n) = d_j(n) - o_j(n) \quad (4.14)$$

onde $d_j(n)$ é o j -ésimo elemento do vetor resposta desejada $\mathbf{d}(n)$.

4. *Computação para trás (Retropropagação)*. Calcula-se os δ (i.e., gradientes locais) da rede, definidos por

$$\delta_j^{(l)}(n) = \begin{cases} e_j^{(L)}(n) \varphi_j'(v_j^{(L)}(n)) & \text{para a camada de saída} \\ \varphi_j'(v_j^{(l)}(n))' \sum_k \varphi_k^{l+1}(n) w_{kj}^{l+1}(n) & \text{para a camada oculta} \end{cases}$$

onde o apóstrofe em $\varphi'_j(\cdot)$ representa a diferenciação em relação ao argumento. Ajusta-se os pesos sinápticos na camada l de acordo com a regra delta generalizada:

$$w_{ji}^{(l)}(n+1) = w_{ji}^{(l)}(n) + \alpha[w_{ji}^{(l)}(n-1)] + \eta\varphi_j^{(l)}(n)y_i^{l-1}(n) \quad (4.15)$$

onde η é o parâmetro da taxa de aprendizagem e α é a constante de momento.

5. *Iteração.* Itera-se as computações para frente e para trás dos pontos 3 e 4, apresentando novas épocas de exemplos de treinamento para a rede, até que seja satisfeito o critério de parada.

A ordem de apresentação dos exemplos de treinamento deve ser aleatória, de época para época. Os parâmetros de momento e a taxa de aprendizagem tipicamente são ajustados (e normalmente reduzidos) quando o número de iterações de treinamento aumenta.

O algoritmo da BPNN encontra-se correntemente bem definido e pode ser encontrado em vários trabalhos [RHW86, Wer88, RB93].

4.5 Identificação e reconhecimento

Para identificação dos animais usamos um perceptron em multi-camadas treinado com um algoritmo do tipo “Backpropagation” (BPNN) [RB93]. A BPNN é uma rede totalmente conectada, com ativação avante, isto é, da camada de entrada para a última camada. O conhecimento adquirido fica codificado nas sinapses (ligações ou pesos), em cada conexão entre as unidades da rede (nós ou neurônios). Os pesos são ajustados por um processo de treinamento que envolve basicamente duas fases. Na primeira delas, dada um vetor na camada de entrada, o sistema calcula a ativação avante, usando os pesos correntes. Na segunda fase, o erro na última camada é calculado como a diferença entre a saída atual da rede e a saída desejada e este erro é propagado para trás, para as camadas intermediárias visando corrigir os pesos destas camadas segundo uma função de ajuste. Após treinada, ativação obtida na última camada representa uma resposta a um determinado vetor de valores entrado na rede.

A Figura 4.8 mostra a estrutura da rede usada neste trabalho. Ela possui uma entrada para cada característica determinada acima (total de 14). O número de nós da camada intermediária muda de acordo com o número corrente de golfinhos conhecidos. O número

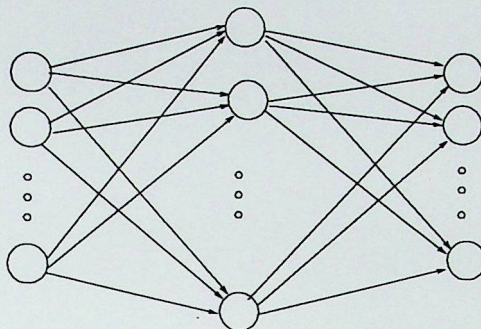


Figura 4.8: Topologia da BPNN usada.

de nós na camada de saída pode ser mudado dinamicamente (e automaticamente) pelo sistema se mais golfinhos ainda não conhecidos forem detectados, na Baía dos Golfinhos por exemplo. Introduzimos um mecanismo de auto-crescimento, ou seja, definimos experimentalmente um limiar que informa se uma representação é nova; se a ativação para um determinado padrão encontra-se abaixo deste limiar, um módulo de aprendizado supervisionado é automaticamente acionado, inserindo o novo padrão na memória de longo termo do sistema, incrementa (cria novos nós ou neurônios) e re-treina a BPNN. Ou seja, a rede aprende forma automática uma representações para o novo padrão e as insere nas memórias do sistema. Para determinar se um golfinho ainda não é conhecido, usamos uma função que considera uma ponderação entre os erros mínimo e máximo da rede, dados pela fase de treinamento. O número de nós na camada intermediária também muda. Empiricamente, determinamos que 1,5 vezes o número de nós da camada de saída produz bons resultados. A Equação 4.16 é usada para a ativação avante (determinar a melhor identificação). A Equação 4.17 é usada para o treinamento da rede.

$$o_i = (1 + e^{-\sum_{i=0}^A \omega_{ij} x_i})^{-1} \quad (4.16)$$

$$\Delta \omega_{ij}(t+1) = \epsilon \delta_j o_i + \alpha \Delta \omega_{ij}(t), \quad (4.17)$$

onde o_i foi definido acima e

$$\delta_j = \begin{cases} o_j(1 - o_j)(y_j - o_j), & \forall j \in \text{output} \\ o_j(1 - o_j) \sum_{k=1}^B \delta_k \omega_{jk}, & \forall j \in \text{other} \end{cases}$$

Capítulo 5

Sistema para Identificação de Golfinhos

5.1 Introdução

Neste capítulo é apresentado o sistema de identificação de golfinhos desenvolvido. Todo o sistema foi elaborado visando uma maior facilidade na interação entre usuário e máquina. Desde que os usuários deste sistema não possuíam conhecimento técnico, procurou-se facilitar ao máximo a interação para que os usuários não se sentissem desconfortáveis ao utilizar o sistema. Assim, a interface gráfica foi desenvolvida de modo a ser intuitiva e transparente ao usuário. Assim, a interface gráfica foi desenvolvida de modo a ser intuitiva e transparente ao usuário.

Este software foi produzido no ambiente Linux usando o compilador GCC, sendo o ambiente de programação integrado KDevelop 2.3.3 para o desenvolvimento dos programas. O ambiente gráfico é baseado no Qt 4.0. Para a interface gráfica foi utilizado o Qt Designer, que permite a criação de interfaces gráficas de usuário de forma visual. O sistema foi desenvolvido em C++ e o compilador GCC foi utilizado para a compilação do código fonte. O sistema foi desenvolvido em C++ e o compilador GCC foi utilizado para a compilação do código fonte.

Montaram-se um exemplo de como utilizar o sistema para a identificação de golfinhos, mostrando as interfaces gráficas e os resultados obtidos. O sistema foi desenvolvido em C++ e o compilador GCC foi utilizado para a compilação do código fonte.

Capítulo 5

Sistema para Identificação de Golfinhos

5.1 Introdução

Neste capítulo é apresentado o sistema de identificação de golfinhos desenvolvido. Todo o sistema foi elaborado visando uma maior facilidade na interação entre usuário e máquina. Desde que os usuários deste sistema são potencialmente alunos da área biológica, procurou-se facilitar ao máximo a interface para que as técnicas usadas fossem totalmente transparentes ao usuário. Assim, a interface baseia-se basicamente em um ambiente de interação através de janelas e manipulações por mouse.

Esse software foi produzido em plataforma Linux usando RedHat versão 8. Foi utilizado o ambiente de programação integrado KDevelop 2.1.3 para o desenvolvimento dos programas. O ambiente gráfico é baseado em Qt 3.0. Este ambiente gráfico tem a enorme vantagem de poder ser facilmente convertido para o ambiente Microsoft Windows, caso seja esse o sistema operacional desejado. Toda a programação foi desenvolvida usando a linguagem orientada a objetos C++. O compilador usado foi o *g++* da versão atual do sistema RedHat.

Mostraremos um exemplo dos passos necessários para identificação de golfinhos, apresentando as interfaces envolvidas. Neste capítulo também são apresentados os dados resultantes da extração de características de 8 imagens distintas de golfinhos.

5.2 Aquisição de Imagens

A aquisição de imagens para o sistema baseia-se em fotos registradas por biólogos com câmeras analógicas. Em seguida, é feita a digitalização da imagem. Logicamente, há a possibilidade de que essas imagens possam ser registradas com câmeras digitais, o que irá facilitar o trabalho.

As imagens obtidas caracterizam-se por grandes áreas de mar e a presença do golfinho geralmente no centro do frame da foto. Note-se que registros de fotos que apareçam mais de um animal irá prejudicar o bom andamento do sistema, caso esses animais apareçam juntos, além de fotos de animais muito distantes, ou seja, fotos em que o animal (essencialmente sua nadadeira) aparece muito pequena no frame da foto.

O processo de digitalização das imagens analógicas deve ser feito com resolução alta para que se preserve ao máximo a imagem obtida. Resoluções em torno de 1000X1000 píxeis são adequadas.

5.3 Pré-processamento

Devido a grande diferenciação em termos de brilho e luminosidade em que podem se encontrar as imagens digitais a serem analisadas, são necessárias sofisticadas ferramentas e o intuito de se obter uma melhor imagem possível para posterior segmentação. Dentro deste princípio, optou-se, neste trabalho, pelo uso da transformada KLT, seguida de um auto-contraste. A aplicação desta transformada juntamente com o auto-contraste mostrou-se bastante eficaz.

5.3.1 Visualização da imagem

O sistema está habilitado para a leitura de imagens do tipo JPEG, PNG ou GIF. A figura 5.1 mostra a janela principal do sistema e uma janela de diálogo para a abertura do arquivo de imagem a ser processado.

Como as imagens são normalmente de dimensões muito elevadas devido a resolução, então é necessária uma redução dessa imagem apenas para visualização. A figura 5.2 apresenta uma imagem aberta e uma numeração para identificação das partes da janela

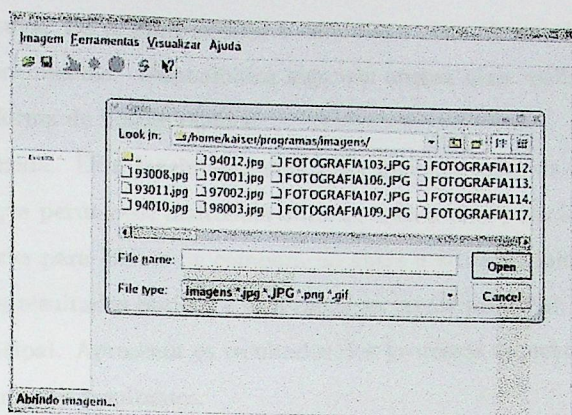


Figura 5.1: Janela principal do sistema

do sistema, conforme descritas em seguida.

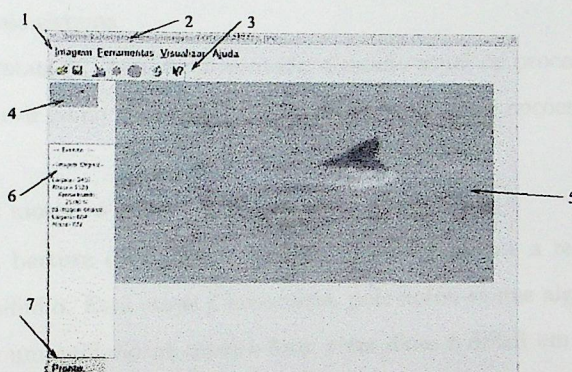


Figura 5.2: Descrição da janela principal do sistema

1. Barra de Menus. Contém as opções: *Imagem* serve para abrir, fechar ou salvar a imagem atual ou sair do sistema; *Ferramentas* contém as opções: Selecionar nadadeira, Pré-processamento e Classificar; *Visualizar* contém as opções: Imagem original delimitada, Escala de cinzas, Imagem pré-processada, Imagem paramétrica X Processada e Fendas detectadas; *Ajuda* que contém as opções Como funciona? e Sobre.

2. Barra de títulos. Serve para mostrar o caminho e nome do arquivo aberto.
3. Barra de ferramentas. Disponibiliza algumas opções mais utilizadas da Barra de Menus em forma de botões rápidos.
4. Janela reduzida. Essa janela apresenta uma imagem reduzida da imagem principal inicial que permanece inalterada durante toda a manipulação da imagem atual aberta. Serve para efeitos de comparação entre a situação inicial apresentada ao sistema e os resultados obtidos e mostrados na janela principal.
5. Janela principal. Apresenta os resultados dos processos sequenciais do sistema em forma de imagens resultantes.
6. Descrição de processos. Esta janela de textos apresenta alguns atributos das imagens manipuladas pelo sistema ao longo do processamento. Nela também podem ser vistos alguns atributos do golfinho na imagem, como o ponto de pico da nadadeira e o número de padrões de golfinhos atualmente cadastrados no arquivo de dados mantido pelo sistema.
7. Barra de "status". Aqui é apresentado o estado atual do processamento realizado, além de servir como uma ajuda rápida sobre os botões e opções do sistema.

A figura 5.3 mostra o passo a passo do uso da ferramenta.

Logo após a abertura da imagem desejada, o usuário define a região que delimita a nadadeira do golfinho. Esta etapa é necessária, pois notou-se que algumas imagens apresentam mais de um golfinho na mesma foto, além disso é difícil um método automático para extração apenas da nadadeira do golfinho, pois a textura desta é a mesma da do corpo do animal. O usuário tem a opção de retornar à imagem anterior para nova delimitação, caso a delimitação feita tenha sido inadequada. Essa região de delimitação não precisa ser minuciosamente precisa, mas apenas para destacar a região da nadadeira do animal, separando-a de outros objetos da imagem. Essa opção é habilitada através da escolha da opção *Ferramentas* → *Delimitar nadadeira* ou o botão correspondente na Barra de Ferramentas.

O próximo passo é selecionar os três pontos da nadadeira: (1) ponto da base mais próximo da cabeça do animal, (2) ponto do pico da nadadeira, (3) ponto da base mais próximo da cauda do animal. Entre os pontos (1) e (3) é gerada automaticamente uma

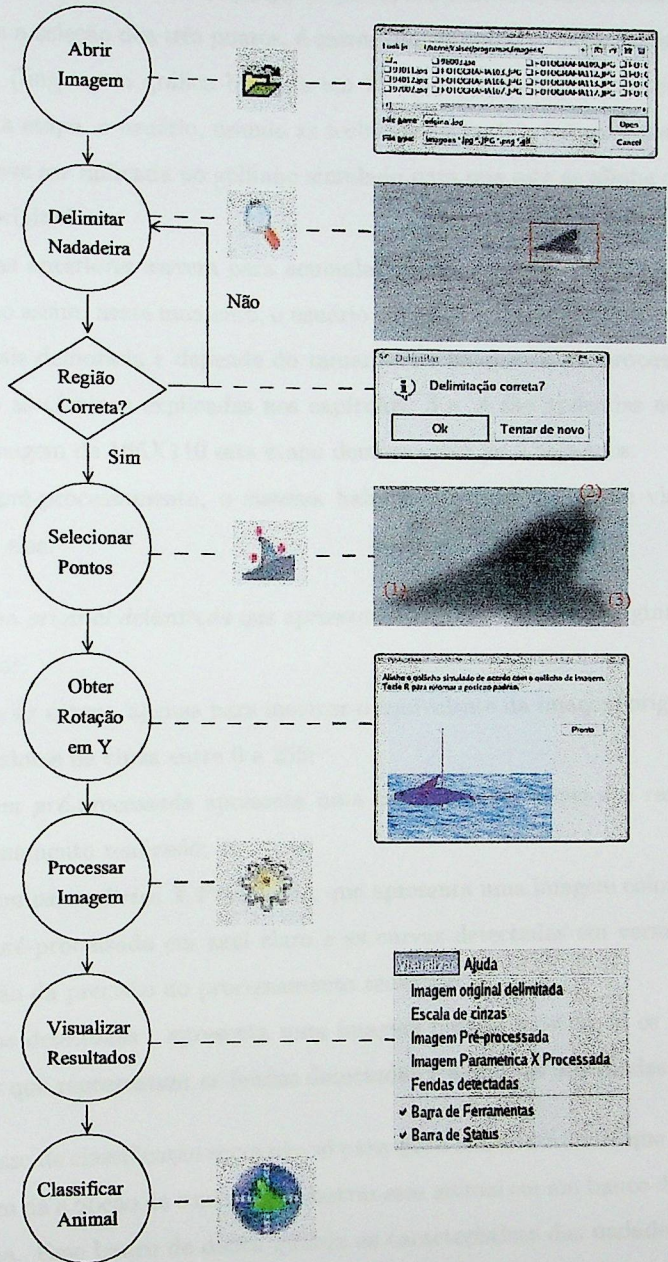


Figura 5.3: Fluxograma para processamento de uma imagem

reta para facilitar a escolha destes, veja a figura para melhor entendimento.

Logo após a seleção dos três pontos, é carregado automaticamente um simulador feito em OpenGL (linguagem gráfica baseada em C) que mostra uma cena de um golfinho virtual. Nesta etapa, o usuário, usando as teclas de “+” e “-”, informa qual a rotação no eixo y que deve ser aplicada no golfinho simulado para que este se alinhe com o golfinho da imagem original.

As etapas anteriores servem para acumular variáveis necessárias à fase de processamento. Sendo assim, neste momento, o usuário clica no botão de pré-processamento. Esta etapa é a mais demorada e depende do tamanho da imagem a ser processada. Praticamente todas as técnicas explicadas nos capítulos 3 e 4 são aplicadas neste momento. Para uma imagem de 166X110 esta etapa demora cerca de 3 segundos.

Após o pré-processamento, o sistema habilita algumas opções de visualização disponíveis são elas:

- *Imagem original delimitada* que apresenta a região da imagem original delimita pelo usuário;
- *Escala de cinzas*, apenas para mostrar o equivalente da imagem original delimitada com valores de cinza entre 0 e 255;
- *Imagem pré-processada* apresenta uma imagem binária que é o resultado do pré-processamento realizado;
- *Imagem paramétrica X Processada*, que apresenta uma imagem colorida com a imagem pré-processada em azul claro e as curvas detectadas em vermelho para comparação da precisão do processamento realizado;
- *Fendas detectadas*, apresenta uma imagem binária mostrando os aglomerados de píxeis que representam as fendas detectadas e as curvas detectadas.

O processo de classificação serve não só para que o sistema informe que animal é aquele, mas também dá a opção de o usuário cadastrar esse animal em um banco de dados mantido pelo sistema. Esse banco de dados guarda as características das nadadeiras dos animais para que posteriormente esses arquivo seja lido pelo sistema para comparação sistemática de cada padrão mantido com a imagem atualmente processada. Ao clicar neste botão de classificação, o sistema informa (na caixa de mensagens disponível na parte esquerda do

software) quantos padrões de nadadeiras já existem cadastradas no arquivo de dados.

O processo de classificação é explicado no capítulo seguinte.

Capítulo 6

Experimentos e Resultados

O sistema de identificação de golfinhos utiliza diversos algoritmos para fazer o processamento de imagens. Neste capítulo, mostra-se um exemplo de um golfinho já pesquisado pelos biólogos entre as praias de Pipa e Tabatinga no litoral sul do Rio Grande do Norte. Neste exemplo, mostra-se as diversas imagens resultantes da aplicação desses algoritmos empregados.

É também descrito aqui os resultados da comparação de 8 imagens distintas de golfinhos, ressaltando as curvas e as fendas detectadas.

6.1 Exemplo de processamento

Como já descrito aqui, é natural que as imagens utilizadas nesse programa apresentem uma alta resolução. Essa é uma exigência do software para que a forma e as fendas presentes na nadadeira do animal sejam detectadas. Sendo assim, ao abrir a imagem a ser processada pelo sistema a janela principal é carregada com uma imagem de no máximo 640X640 píxeis. A figura 6.1 mostra a imagem já delimitada que foi usada neste exemplo.

Nas seções seguintes são mostrados os resultados da aplicação das diversas técnicas empregadas.



Figura 6.1: Imagem original já delimitada

6.1.1 Transformada KLT

Antes da aplicação da transformada KLT a imagem original foi decomposta em três imagens, as quais são: imagem R (Red), imagem G (Green) e imagem B (Blue). Obtidas as imagens R, G e B passou-se à aplicação da transformada KLT a cada uma delas. Esta transformação consiste em encontrar uma matriz A que opera em cada píxel da imagem de forma a gerar uma imagem transformada. Enfatiza-se que a transformada KLT, além de descorrelacionar as informações contidas nos píxeis enaltecendo a melhor informação possível presente na imagem, faz uma mudança da imagem para um novo espaço de cores. O processo da transformada é composto dos seguintes passos:

- Cálculo da matriz de correlação;
- Aplicação do método de Jacobi para encontrar os autovalores e autovetores;
- Aplicação da transformada propriamente dita, a qual irá proceder a mudança do espaço de cores atual para um novo espaço de cores totalmente descorrelacionado, fornecendo para isto os autovetores e a matriz de correlação, equação 3.3;
- Aplicação de um ajuste na imagem para que os níveis de cinza nos planos RGB se encontrem no intervalo de 0 a 255.

A figura 6.2 mostra o resultado da aplicação do processo KLT na imagem original. Note que esta imagem resultante ficou difícil de ser analisada segundo critérios do olho humano, mas o processo foi aplicado corretamente.

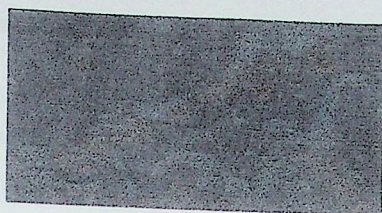


Figura 6.2: Resultado da aplicação da KLT na imagem

6.1.2 Comparação com o processo do gradiente

A decisão de se usar a sequência de processos de KLT, autocontraste, autosegmentação no pré-processamento das imagens deu-se principalmente devido a qualidade da imagem após esses processos. As imagens resultantes destes processos, seguidos de uma limiarização, apresentam contornos suaves. Isso é muito importante quando essa imagem é submetida a processos de morfologia matemática. Para efeitos de comparação apresenta-se na figura 6.3 como fica o esqueleto das imagens após a aplicação do pré-processamento feito com o gradiente e com a KLT. Note que o esqueleto gerado da imagem após o processo de gradiente prejudicaria fortemente os processos seguintes.

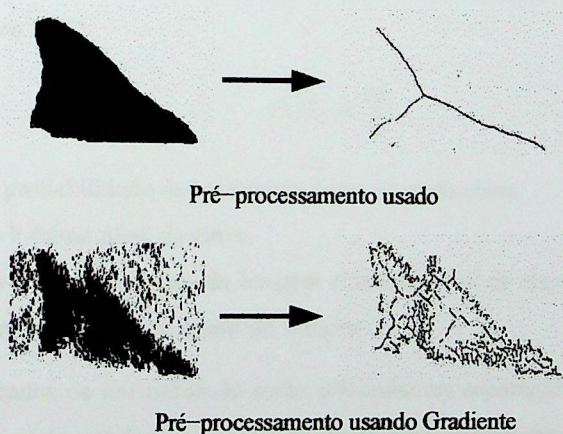


Figura 6.3: Comparação: pré-processamento com KLT e com gradiente

6.1.3 Auto-contraste

O auto-contraste visa melhorar a qualidade da imagem sob os critérios subjetivos do olho humano. O processo de auto-contraste consiste em efetuar uma espécie de esticamento do histograma, sem que isso afete a forma do mesmo. Na figura 6.4 temos uma demonstração do que ocorre na imagem ao realizarmos um auto-contraste, após a aplicação da KLT.

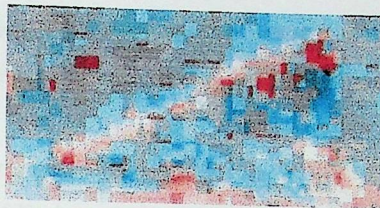


Figura 6.4: Aplicação do auto-contraste, após a KLT

- Inicialmente montou-se o histograma dos planos RGB, fazendo uma contagem dos píxeis em seus devidos níveis e planos;
- Em seguida, realizou-se a normalização dos níveis do histograma de cada uma das componentes RGB.

$$P(rk) = \frac{nk}{n} \quad (6.1)$$

- P é a probabilidade de ocorrência deste nível de cinza;
 - rk é o k -ésimo nível de cinza;
 - nk é o número de píxeis da imagem com esse nível de cinza;
 - n é o número total de píxeis da imagem.
- Estes resultados da normalização serão utilizados na separação dos níveis. O que estiver abaixo de um limiar inferior n_1 ficam com valores iguais a n_1 e todos os valores superiores a n_2 ficam com valores iguais a n_2 .
 - Por último, foi aplicada uma transformação linear T_L nos níveis intermediários de forma que, ao final, o histograma tivesse um esticamento e que os níveis se encon-

trassem no intervalo de 0 a 255.

$$T_L = \left(\frac{\text{pixel}_{\text{atual}} - \text{limiar}_{\text{inferior}}}{\text{limiar}_{\text{superior}} - \text{limiar}_{\text{inferior}}} \right) \times 255 \quad (6.2)$$

6.2 Segmentação por textura

A segmentação por texturas na imagem foi realizada através da auto-segmentação. Um sistema de quantização vetorial baseado em uma rede neural competitiva foi utilizado para fazer a classificação dos píxeis da imagem. Para cada píxel escolheu-se um descritor. Neste trabalho, escolhemos a média dos píxeis R, G e B para representar a textura de cada píxel.

Em uma janela de tamanho $W \times W$, centrada no píxel do qual se deseja extrair os atributos, calculou-se a média para cada componente R, G e B, obtida através do cálculo do primeiro momento, construindo, desta forma, um vetor composto de 3 médias. Estes blocos seguem a seguinte regra de construção $[(2 \times W + 1) \times (2 \times W + 1)]$. O que acontece na verdade é o seguinte: ao atribuir-se um valor W igual a 2 para o bloco tem-se na verdade um bloco abrangendo 25 posições de píxeis. O valor W corresponde ao número de píxeis iniciais para formação do bloco na regra explicada acima.

A segmentação do golfinho foi feita baseada na textura da imagem. Para identificar cada píxel utilizou-se um sistema de quantização vetorial baseado em uma rede neural competitiva conforme foi mostrado no capítulo 4. A idéia é que cada píxel tenha um vetor de atributos que descreva de alguma forma a textura a que este píxel está presente. Os atributos escolhido foram as médias RGB das janelas $W \times W$. Estes atributos formam vetores. Cada vetor tem três médias, sendo uma para cada componente RGB.

No treinamento da rede não foram utilizados os atributos de todos píxeis da imagem. É feita uma escolha aleatória de P píxeis da imagem para formar os vetores que treinarão a rede neural. O cálculo da média dos blocos $W \times W$, formam um espaço de atributos de dimensão 3 em que cada vetor neste espaço tem coordenadas m_r , m_g e m_b . Podemos entender o cálculo da média de um píxel qualquer como sendo a transposição do píxel do espaço R, G e B de cores para o espaço de atributos.

A entrada da rede neural competitiva será os X vetores com as médias ou atributos e sua saída será 2 classes que representarão as texturas de fundo e objeto da imagem.

Após escolhidos os 2 centros (ou vetores representantes de cada classe ou textura) todos os píxeis são mapeados para o espaço de atributos e classificados em uma das 2 classes geradas.

A classificação é realizada utilizando o classificador de distância mínima, já descrito anteriormente. O qual, após escolhidos os 2 centros (vetores representantes de cada textura ou classe), irá fazer o mapeamento de todos os píxeis da imagem para o espaço de atributos e classificação em uma das duas texturas. A segmentação segue os seguintes passos:

- Seleção aleatória dos P píxeis;
- Formação das janelas $W \times W$;
- Elaboração dos diferentes histogramas RGB de cada janela;
- Cálculo das médias m_r , m_g e m_b e formação dos vetores de média;
- Entrega dos vetores contendo as médias para a rede competitiva:
 - Treinamento da rede competitiva e cálculo dos centros, conforme explicado no algoritmo de aprendizagem competitiva, seção 4.3.1.
- Obtidos os 2 centros, todos os píxeis são mapeados para o espaço de atributos e classificados em uma das duas classes (texturas). O classificador utilizado é o classificador de distâncias mínimas, detalhado na seção 4.3.2.

6.3 Limiarização ou Binarização

Aplicação de uma limiarização é necessária, pois após a segmentação da imagem, necessita-se aplicar operadores morfológicos no intuito de extrair as bordas e o esqueleto dos objetos. Neste bloco, cada objeto da imagem é trabalhado separadamente. A limiarização é feita como a seguir:

- Inicialmente faz-se uma subtração entre o maior e menor valor dos píxeis da imagem e guarda-se os resultados em uma variável C .
- Em seguida, aplica-se uma regra simples para encontrar-se um valor de limiar:

$$limiar = C \times p + menorvalor \quad (6.3)$$

Onde : p é um parâmetro percentual para o cálculo do limiar. No nosso caso, usamos $p = 0.7$, ou seja, 70% da escala de cores presente na imagem.

- Após se encontrar o valor de limiar faz-se a binarização da imagem, colocando 1 para valores maiores que o limiar e 0 para os demais valores.

A figura 6.5(a) mostra o resultado da aplicação do processo de limiarização. O lado esquerdo da figura mostra a imagem original e o lado direito o resultado da limiarização após as etapas de pré-processamento anteriormente descritas. Note, nesta figura, que

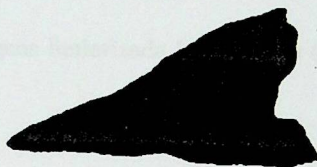


Figura 6.5: Imagem limiarizada

houve uma real separação do golfinho e a área de mar presente na imagem.

6.4 Cortar a imagem

Esta é uma etapa do processamento que não pode ser desconsiderada. Qualquer que seja o processo realizado para extrair a nadadeira de forma totalmente automática, dificilmente irá fazer um corte apenas da nadadeira do animal, pois o corpo do animal tem textura muito semelhante. Portanto é necessário obter do usuário do sistema aqui implementado os pontos que delimitam a base da nadadeira. De posse desses pontos é calculada a angulação da reta que passa entre esses dois pontos e, logo em seguida, feito o corte adequado para separação da nadadeira do animal. A figura 6.6(b) exhibe a imagem cortada ao lado direito da imagem limiarizada 6.6(a).

6.5 Reflexão da imagem

Esta etapa do processamento visa padronizar o sentido da nadadeira. Conforme comentado anteriormente adotamos o sentido em que a cabeça do golfinho aponta para o lado direito da imagem. Assim, imagens que tem um sentido contrário a esse devem ser imediatamente refletidas. A figura 6.6(c) demonstra que esta imagem necessitou de uma reflexão.

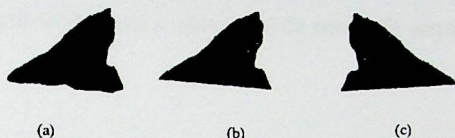


Figura 6.6: Imagens limiarizada (a), cortada (b) e refletida (c)

6.6 Rotulação

Após o processo de corte da nadadeira, temos uma imagem binária, mas que pode ainda possuir outros objetos além da nadadeira. Sendo assim, é necessária uma rotulação dos objetos presentes na imagem e, em seguida, a extração apenas da nadadeira. Essa extração é feita simplesmente calculando qual é o maior objeto presente na imagem. Por fim, apenas o maior objeto é deixado na imagem.

Nesse exemplo houve apenas um objeto rotulado, portanto a imagem resultante é a mesma a da figura 6.6(b).

6.7 Operadores Morfológicos

A morfologia matemática é vista como um caminho para descrever ou analisar a forma de objetos, onde algumas operações matemáticas em conjunto de pixels podem ser usadas para ressaltar detalhes específicos das formas, permitindo que sejam contadas ou reconhecidas. Neste trabalho, pretende-se utilizar dois operadores muito importantes da morfologia matemática como método auxiliar na obtenção de vários descritores de forma,

são eles: a esqueletização e extração de bordas. Para a determinação dos elementos estruturantes utilizados neste trabalho nos baseamos nos estudos teóricos e testes práticos realizados. Foi obtido bons resultados com os elementos estruturantes “quadrado para extração de borda e “cruz” para esqueletização. Um problema detectado com a utilização de elementos estruturantes “quadrados” na esqueletização está no fato de não garantirem total conectividade dos pontos num esqueleto resultante. Os trechos de descontinuidades, quando gerados, são pouco perceptíveis. Apesar de o elemento estruturante “cruz” também não garantir esta conectividade tem a vantagem de produzir esqueletos consideravelmente finos.

6.7.1 Esqueletização

O esqueleto dos objetos foi obtido fazendo-se sucessivas erosões e aberturas, através da equação 3.20 que demonstra que $S(A)$ é o esqueleto de A , o qual pode ser adquirido pelo subconjunto de esqueletos obtidos da equação 3.21. A erosão ocorre até que o último passo seja antes que o conjunto A seja erodido a um conjunto vazio. O esqueleto não é feito simplesmente com a erosão, pois temos que ter a capacidade de detectar quando a erosão de um píxel levaria o objeto a um desconcerto ou conjunto vazio. O elemento estruturante utilizado foi:

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (6.4)$$

Na figura 6.7(a) apresenta-se um exemplo da esqueletização realizada:

O elemento estruturante citado foi aplicado sucessivas vezes até que se chegou a representação do objeto da imagem (nadadeira dorsal) como se fosse um grafo conectado, como visto na figura.

Com essa informação e a informação seguinte (fronteira do objeto) podemos empregar algoritmos que avaliem a forma e peculiaridades da nadadeira dorsal do golfinho. Hillman e colaboradores [HKA⁺98] destacam a localização do pico da nadadeira dorsal como um descritor muito importante na distinção entre animais. Assim, pretende-se, a partir da informação do esqueleto, obter uma localização do pico da nadadeira. Essa informação

(esqueleto do objeto) certamente não servirá apenas para obtenção deste descritor, mas para outros também.

6.7.2 Extração de bordas

A extração de borda dos objetos da imagem foi obtida fazendo uso da equação 3.19, a qual é descrita como a aplicação da erosão de A por B seguida da diferença de conjuntos entre A e a sua erosão. O elemento estruturante utilizado foi:

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad (6.5)$$

Na figura 6.7 apresenta-se um exemplo da extração de fronteira ou borda realizada.

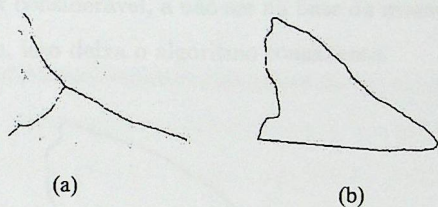


Figura 6.7: Resultados da esqueletização (a) e extração de borada (b)

Na figura apresentada temos a borda da nadadeira de um golfinho, onde os píxeis se apresentam contínuos ao longo da fronteira. Com esse resultado poderemos aplicar um algoritmo que faça a seqüenciação dos píxeis, colocando-os em uma estrutura de dados. A partir daí, podemos aplicar diversos algoritmos para extração de características.

6.8 Extração do Pico da Nadadeira

Nesta etapa é usado o algoritmo de extração de pico da nadadeira descrito em 3.5.2. Esse valor do pico da nadadeira servirá como ponto médio entre as duas curvas, ou seja, as duas curvas iniciarão neste ponto com $\lambda = 0$ e terminarão na base da nadadeira com $\lambda = 1$.

6.9 Sequenciação de Pontos

Após a extração do pico da nadadeira do animal, procede-se uma seqüenciação dos pontos externos da fronteira da imagem, mas sem considerar a base da nadadeira. É através desta seqüência de pontos colocados em dois vetores A e B que pode-se determinar os polinômios cúbicos que regem cada curva. O algoritmo de seqüenciação de pontos é descrito na seção 3.5.2. A figura 6.8 mostra o resultado do algoritmo de sequenciação de pontos. No que na região da maior fenda da nadadeira, houve uma descontinuidade dos pontos. Isso ocorre porque neste ponto é feita uma pesquisa de pontos ao longo do eixo x , quando é fixado um ponto em y . A pesquisa pára quando é encontrado o primeiro ponto, deixando os demais sem serem sequenciados. Isso não acarretou um problema na geração de curvas, além do que, diante das várias imagens testadas, é muito difícil encontrar uma secção reta ao longo de x considerável, a não ser na base da mesma, mas como a base não é incluída neste processo, isso deixa o algoritmo consistente.

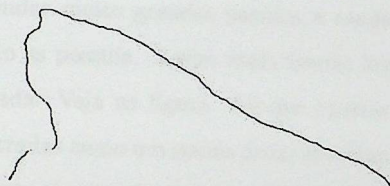


Figura 6.8: Seqüenciação de pontos da fronteira da imagem

6.10 Obtenção das curvas paramétricas

Após o processo de seqüenciação de pontos da nadadeira podemos obter as duas curvas cúbicas paramétricas que descrevem a nadadeira através do processo explicado na seção 3.5.2. O resultado deste processo, aplicado à imagem deste exemplo, resulta na figura 6.9 apresentada a seguir. Note que a nadadeira deste golfinho está com uma fenda tão acentuada que a própria forma de sua nadadeira mudou. Da mesma forma os polinômios tentam se ajustar o máximo possível da forma original.

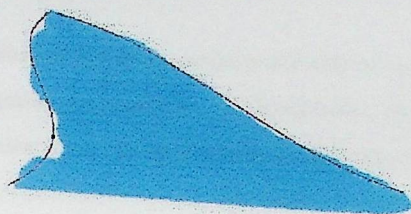


Figura 6.9: Representação das curvas polinomiais encontradas

6.11 Detecção de fendas

Em alguns casos, o processo de detecção de fendas fica sujeito a erros devido a angulação da imagem original passada, que pode esconder essas fendas. Outro fator que prejudica a detecção correta das fendas é a presença de fendas muito grandes na nadadeira. Nadadeiras com fendas muito grandes passam a mudar a sua forma original de quando o animal ainda não as possuía. Assim essas fendas levam o polinômio de curvas a uma forma mais acentuada. Veja na figura 6.9 que apenas 1 fenda foi detectada (as fendas detectadas são mostradas como um ponto preto em cima da curva paramétrica que está em vermelho), por causa da grande curvatura que o polinômio teve que fazer para conseguir se ajustar à forma nadadeira. Em imagens como a mostrada na figura 6.10 as fendas são detectadas com grande precisão.

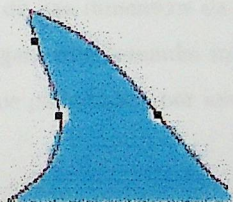


Figura 6.10: Detecção de fendas

6.12 Normalização

O processo de normalização dos dados gerados até aqui se faz necessário, pois queremos comparar um mesmo animal em condições de iluminação, escala ou rotação diferentes, assim como essa mesma comparações entre animais distintos para a classificação correta. Os passos usados para normalização são os seguintes:

1. Rotacionar polinômios em torno do eixo y da angulação determinada pelo usuário na interface;
2. Rotacionar os polinômios em torno do eixo z de forma que a base da nadadeira fique alinhada com o eixo x;
3. Transladar polinômios para o ponto (0,0,0) do eixo de coordenadas;
4. Fazer nova translação de forma que a base do polinômio fique em cima do eixo x. Isso servirá para o processo de mudança de eixos. Na imagem os valores de y crescem “para baixo”, enquanto que no eixo padrão normalmente adotado, os valores de y crescem “para cima”;
5. Proceder a mudança de eixos;
6. Fazer a normalização em escala de forma que o maior valor em x seja “1”. Com isso, o maior valor de y fica dividido pelo valor aplicado a x, não necessariamente, e dificilmente, o valor máximo de y será exatamente “1”.

Após todo esse processo os polinômios normalizados são guardados pelo sistema, caso o usuário deseje. A figura 6.11 apresenta o resultado da normalização aplicado nos polinômios que formam as curvas. Essa imagem aparece pequena, pois é apresentada multiplicada por um valor que dê nas dimensões da imagem original, já que seria impossível apresentarmos uma imagem com tamanho máximo igual a “1” na tela. Foram plotados apenas alguns pontos dos polinômios, por isso nota-se uma figura pontilhada.

6.13 Descritores de Curvatura

Os descritores de curvatura são o raio médio de curvatura e o índice de discrepância que é calculado pela diferença entre a distância de cada um dos n pontos de uma curva paramétrica ao ponto central (x_c, y_c) e o valor do raio médio. Esses descritores são obtidos

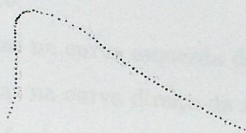


Figura 6.11: Resultado do processo de normalização

após a normalização dos polinômios. A figura 6.12 demonstra o resultado da obtenção apenas do raio de curvatura em uma imagem ainda não normalizada, apenas para demonstrar o lugar do centro e o raio de curvatura.

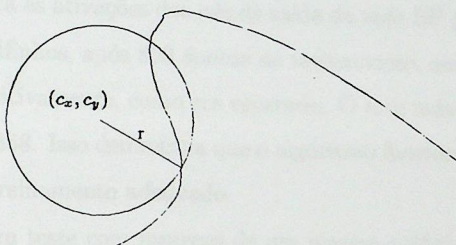


Figura 6.12: Apresentando o raio de curvatura

6.14 Classificação

Para o processo de classificação dos animais usamos um perceptron em multi-camadas, treinado com um algoritmo do tipo "Backpropagation Neural Network" (BPNN) [RB93].

Nos experimentos realizados, os resultados foram aceitáveis. Basicamente, algumas instâncias de animais foram escolhidas e calculadas as características para eles. Estas características foram usadas no classificador para uma fase inicial de treinamento (BPNN). A seguir, algumas outras imagens das mesmas instâncias de golfinhos foram escolhidas e apresentadas ao sistema.

As características da nadadeira submetida à rede para treinamento e classificação são

as seguintes:

1. Posição x do pico;
2. Posição y do pico;
3. Número de fendas na curva esquerda da nadadeira;
4. Número de fendas na curva direita da nadadeira;
5. Localização das fendas presentes na curva esquerda, relativo a λ ;
6. Localização das fendas presentes na curva direita, relativo a λ ;
7. Raio de curvatura da curva esquerda;
8. Raio de curvatura da curva direita;
9. Índice de discrepância da curva esquerda;
10. Índice de discrepância da curva direita.

A Tabela 6.1 mostra as ativações dos nós de saída da rede BP para vetores de entrada correspondentes a 8 golfinhos, após 500 épocas de treinamento, em pose ideal. Todos eles foram identificados positivamente, como era esperado. O erro máximo ficou em 0.076041. O erro médio em 0.015358. Isso demonstra que o algoritmo funciona da maneira desejada, convergindo para um treinamento adequado.

Foi feito também um teste com imagens de um mesmo golfinho. Para este caso submetemos a rede a um treinamento de 10000 épocas. A figura 6.13 mostra duas imagens distintas de um mesmo golfinho. A tabela 6.2 mostra o resultado da saída da rede comparada com a resposta desejada. O erro máximo ficou em 0.747619. O erro médio em 0.025822.

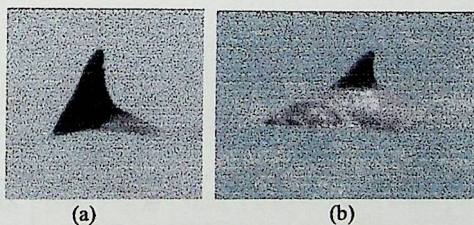


Figura 6.13: Imagens distintas de um mesmo golfinho.

Escolhemos esse caso para mostrar que imagens muito ruidosas prejudicarão fortemente a identificação do animal desejado. Neste caso, na imagem (a) da figura 6.13 foi

Rede	0.95	0.04	0.00	0.00	0.01	0.00	0.02	0.00
Desej	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Rede	0.03	0.92	0.04	0.00	0.01	0.00	0.00	0.01
Desej	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00
Rede	0.00	0.04	0.94	0.03	0.00	0.00	0.02	0.03
Desej	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00
Rede	0.00	0.00	0.02	0.97	0.00	0.01	0.00	0.02
Desej	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00
Rede	0.01	0.01	0.00	0.00	0.97	0.01	0.00	0.01
Desej	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00
Rede	0.01	0.00	0.00	0.00	0.01	0.98	0.00	0.02
Desej	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00
Rede	0.03	0.00	0.03	0.00	0.01	0.00	0.97	0.00
Desej	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00
Rede	0.00	0.04	0.03	0.02	0.01	0.03	0.00	0.96
Desej	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00

Tabela 6.1: Ativação dada pela última camada da rede e ativação ideal.

Rede	0.00	0.00	0.25	0.64	0.00	0.00	0.00	0.04
Desej	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00

Tabela 6.2: Identificação errada devido ao ruído.

identificada uma fenda na nadadeira do golfinho, já na imagem (b) esta fenda não pôde ser identificada, por isso a rede não conseguiu classificar corretamente esse animal. Para os outros casos a identificação procedeu corretamente.

A ativação para golfinhos presentes na memória, mas cuja imagem era diferente da imagem com que ele foi inserido inicialmente na memória foi de mais de 0,9 em todos os caso testados (testamos 8 casos assim). ou seja, pudemos verificar que as rotações foram bem suportadas pelo classificador. Quanto ao tempo de treinamento, o gráfico mostrado na Figura 6.14 mostra o tempo em segundos versus número de golfinhos na rede. Pode-se notar que para 21 golfinhos a rede gasta cerca de 22 segundos. o gráfico mostrado na Figura 6.15 mostra o número de épocas versus número de golfinhos (nós na última camada).

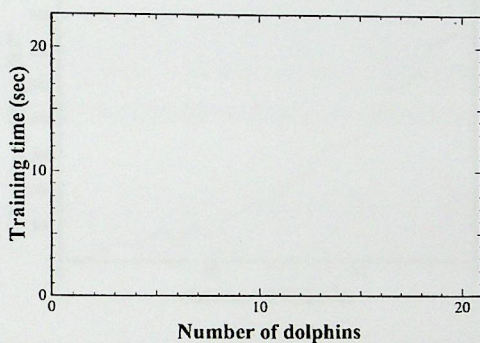


Figura 6.14: Tempo versus número de golfinhos.

Notamos que os vetores eram compostos com 16 entradas e que a rede possui 22 neurônios na camada intermediária. Os gráficos mostram uma função aparentemente exponencial. Isto não é um problema, uma vez que o número de golfinhos em uma dada família não deve ultrapassar de 50. Ainda, numa situação prática, o sistema é treinado off-line e os pesos salvos para serem usados na ocasião em que for necessário. Isto certamente não interfere na performance do sistema.

Capítulo 7

Discussões, conclusões e perspectivas

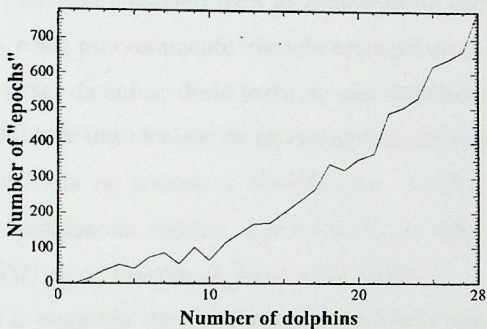


Figura 6.15: Épocas versus número de golfinhos.

Capítulo 7

Discussões, conclusões e perspectivas

Desenvolvemos um sistema completo para identificação de golfinhos, que vai desde a submissão da imagem, o seu processamento visando extração de características e sua posterior identificação. Apesar da ênfase deste texto ser nos classificadores, convém ressaltar a importância e generalidade das técnicas de processamento de imagens usada. O sistema procedeu da forma esperada no tocante a identificação. Outras alternativas de classificadores deverão ser testadas em seguida, por exemplo, as redes bayesianas, os mapas auto-organizáveis (SOM) ou as funções de base radial (RBF).

Vale ressaltar que o conjunto de características utilizado pode ser melhorado. Um bom conjunto deve, sem dúvida, considerar a textura. Para o uso de textura, deve ser definida uma maneira de se contornar os efeitos da água na iluminação da imagem. Apesar disso, destacamos a importância das características apresentadas, com bons resultados. Há algumas restrições a serem observadas na obtenção das características. Por exemplo, caso ocorra uma rotação em y próximo de 90 graus, as fendas e contornos podem não ser visíveis.

A rede BP usada mostrou-se eficiente no tocante a inclusão de novos indivíduos na memória de longo termo, devido ao seu mecanismo de auto-crescimento. Ainda, alguma estratégia pode ser desenvolvida para que o sistema consiga aprender qual classe de características pode ser melhor empregada no problema em questão. Desta forma, um conjunto menor de características pode ser empregado, diminuindo o tempo de processamento da rede. Algo baseado em inferência estocástica poderá ser tentado em trabalhos futuros.

Bibliografia

- [ASHW61] H. N. Ashwell, N. Kenderdine, G. Hillman, and R. Waring. Evaluation of computer models for dolphin photo identification. *IBR*, 1961.
- [BRI85] L.O. Brown, D.P. DeLong, and C.K. Ray. Status of studies on local marine mammals. *Marine Mammal Science*, 1:1-24, 1985.
- [DIN86] R. D. Dinstein. *Digital Image Processing*. Prentice Hall Inc, New Jersey, 1986.
- [GRO86] R.M. Cross Jr. and L. de F. Costa. Pattern recognition: Pattern Recognition, 79-130, 1986.
- [GRO87] R.M. Cross Jr. and L. de F. Costa. Application and assessment of nonlinear bending energy for morphometric characterization of aerial video. *Rev. Sci. Instrum.*, 58(5):2175-2186, 1987.
- [ISH84] M.R. Douglas, G.D. Arnold, and D. J. Dough. Observations between inshore and offshore coastal dolphins in the western Long Island Sound. *Journal of Mammalogy*, 65(3):275-281, 1984.
- [RSH80] R. H. Dufrenoy, G. M. Shotts, and R.W. Miller. A technique for the photographic identification and recording of dorsal fins of the bottlenose dolphin (*Tursiops truncatus*). Technical report, International Whaling Commission (Japan - Special Issue 12), 1980.
- [DWH99] R.H. Dufrenoy and R.W. Miller. Comparing identification techniques and value of bottlenose dolphins (*Tursiops truncatus*) and other coastal cetaceans. *Marine Mammal Science*, 15(2):245-256, 1999.

Bibliografia

- [AKHW01] B. N. Araabi, N. Kehtarnavaz, G. Hillman, and B. Wursig. Evaluation of invariant models for dolphin photo-identification. *IEEE*, 2001.
- [BDR85] L.G. Barnes, D.P. Dooming, and C.E. Ray. Status of studies on fossil marine mammals. *Marine Mammal Science*, 1:1–53, 1985.
- [Cas96] K. R. Castleman. *Digital Image Processing*. Prentice Hall Inc, New Jersey, 1996.
- [CJC96] R.M. Cesar Jr. and L. da F. Costa. Pattern recognition. *Pattern Recognition*, 29:1559, 1996.
- [CJC97] R.M. Cesar Jr. and L. da F. Costa. Application and assessment of multiscale bending energy for morphometric characterization of neural cells. *Rev. Sci. Instrum.*, 68(5):2177–2186, 1997.
- [DSH84] M-E. Douglas, G.D. Schnell, and D.J. Hough. Differentiation between inshore and offshore spotted dolphins in the eastern tropicalpacific ocean. *Journal of Mammalogy*, 65(3):375–387, 1984.
- [DSW90] R. H. Defran, G. M. Shultz, and D.W. Weller. A technique for the photographic identification and cataloging of dorsal fins of the bottlenose dolphin (*tursiops truncatus*). Technical report, International Whaling Commission (Report - Special issue 12), 1990.
- [DW99] R.H. Defran and D.W. Weller. Occurence, distribution, site fidelity and school size of bottlenose dolphins (*Tursiops truncatus*) off san diego, california. *Marine Mammal Science*, 15(2):366–380, 1999.

- [Fac00] J. Facon. *Processamento e análise de imagens*, 2000.
- [Flo99] P.A.C. Flores. Preliminary results of photoidentification study of marine tucuxi, *Sotalia fluviatilis* in southern brazil. *Marine Mammal Science*, 15(3):840–847, 1999.
- [GTE97] IBAMA GTEMA. *Mamíferos Aquáticos do Brasil: Plano de Ação*. IBAMA, Brasília, 1997.
- [GWG92] Rafael C. GONZALES, RICHARD E. (Contributor) WOODS, and Ralph C. GONZALES. *Digital Image Processing*. Addison-Wesley Publication Company, 1992.
- [Hay99] S. Haykin. *Neural Networks a Comprehensive Foundation*. Prentice Hall, 1999.
- [Het93] Bia Hetzel. *Baleias, Botos e golfinhos: guia de identificação para o Brasil*. Nova Fronteira, Rio de Janeiro, 1993.
- [Hil95] M. Hildebrand. *Análise da estrutura dos vertebrados*. Atheneu Editora São Paulo, São Paulo, 1995.
- [HKA⁺98] G. R. Hillman, Tagare H. and Elder K., Drobyshevski A., Weller D., and Wursig B. Shape descriptors computed from photographs of dolphins dorsal fins for use as database indices. In *Proc. of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, volume 20, 1998.
- [JAI89] Anil K. JAIN. *Fundamentals of Digital Image Processing*. Prentice-Hall, October 1989.
- [Lin00] Luciana de Oliveira Link. Ocorrência, uso do habitat e fidelidade ao local do boto cinza. Master's thesis, Universidade Federal do Rio Grande do Norte, 2000.
- [Mar90] R.A. Martin. *Whales and Dolphins*. Salamandra press, London, 1990.

- [Mar01] A.M. Martins. Sistema inteligente para classificação e contagem de células utilizando redes neurais e morfologia matemática. Master's thesis, UFRN - Universidade Federal do Rio Grande do Norte, 2001.
- [MCTW00] J. Mann, R.C. Connor, P.L. Tyack, and H. Whitehead. *Cetaceans societies - field studies of dolphins and whales*. The University Chicago Press, Chicago, 2000.
- [Orr86] R.T. Orr. *Biologia dos Vertebrados*. Livraria Roca LTDA, São Paulo, 5 edition, 1986.
- [Par97] J.R. Parker. *Algorithms for Image Processing and Computer Vision*. Wiley Computer Publishing, 1997.
- [PM92] F.C.W. Pinedo, M.C. anda Rosas and M. Marmontel. *Cetáceos e Pinípedes do Brasil*. UNEP/FUA, Manaus, AM, 1992.
- [Pra89] W.K. Pratt. *Digital Image Processing*. Addison-Wesley, 1989.
- [RB93] M. Riedmiller and H. Braun. A direct adaptive method for faster back-propagation learning: The rprop algorithm. In *Proc. of the International Conference on Neural Networks (ICNN'93)*, pages 123–134. IEEE Computer Society Press, 1993.
- [RCA⁺00] R.A.O Rabelo, L.M.B Claudino, A.A Araujo, L Perroton, and S.J.F Guimarães. Segmentação e classificação automática de imagens segundo informação textural, 2000.
- [RH99] K.A. Rossbach and D.L. Herzing. Inshore an offshore bottlenose dolphin *Tursiops truncatus* communities distinguished by association patterns near grand bahama island, bahamas. *Canadian Journal of Zoology*, 77(4):581–592, 1999.
- [RHW86] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. *Learning internal representations by error propagation*. In D. E. Rumelhart and J. L. McClelland, editors, *Parallel Distributed Processing: Explorations in the microstructure*

of cognition, volume 1: Foundations. The MIT Press, Cambridge, Massachusetts, 1986.

- [RJ96] M.M. Rollo Jr. Variações diurnas e sazonais nos padrões de atividade do boto *Tursiops truncatus*, (montagu, 1821) (cetacea, delphinidae), na região de laguna, estado de santa catarina. Master's thesis, Universidade de São Paulo, São Paulo, 1996.
- [San99] M.C.O. Santos. Novas informações sobre cetáceos no litoral sul de são paulo e norte do paran  com base em estudos sobre encalhes e na aplica  o da t cnica de foto-identifica  o individual de *Sotalia fluviatilis* (cetacea, delphinidae). Master's thesis, Universidade de S o Paulo - Instituto de Bioci ncias, S o Paulo, 1999.
- [Sch90] J.M.M. Schmiegelow. Estudo sobre cet ceos odontocetos encontrados em praias da regi o entre igua e (sp) e enseada de paranagu  (pr) (24  42' s - 25  28' s), com especial refer ncia a *Sotalia fluviatilis* (gervai, 1853) (delphinidae). Master's thesis, Instituto Oceanogr fico, Universidade de S o Paulo, S o Paulo, 1990.
- [SQMB98] A.P. Silva, E.L. Queiroz, G. Monc rvo, and Y.S. Braga. Distribuic o e ocorr ncia de *Sotalia fluviatis* (cetacea, delphinidae) na enseada de todos os santos (bts), bahia, brasil - durante o per odo de janeiro a mar o de 1998, 1998.
- [VM84] F.R.D. Velasco and N.D.A. Mascarenhas. Processamento digital de imagens, 1984. IV Escola de Computa  o.
- [W79] B. W rsig. Dolphins. *Scientif American*, 240(3):108-119, 1979.
- [Wer88] P. Werbos. Backpropagation: Past and future. *IEEE International Conference on Neural Networks*, pages 343-353, 1988.
- [Wil98] B. Wilson. *Dolphins*. Colin Baxter Photography, Grantown-on-Spey - Scotland, 1998.

- [WJ90] B. Würsig and T.A. Jefferson. Methods of photoidentification for small cetaceans. Technical report, Report of international whaling commission, special issue, 1990. In: P.S. Hammond, S.A. Mizroch and G.P. Donovan (Eds) Individual recognition of cetaceans: use of photoidentification and other techniques to estimative parameters.
- [WTH97] B. Wilson, P.M. Thompson, and P.S. Hammond. Habitat use by bottlenose dolphins: seasonal distribution and stratified movement patterns in moray firth. *Journal of applied ecology*, 34:1365–1374, 1997.
- [WW77] B. Würsig and M. Würsig. The photographic determination of group size, composition, and stability of coastal porpoises (*Tursiops truncatus*). *Science*, 198, pages 755–756, 1977.

IMPRESSOS

Universidade Federal do Rio G. do Norte
BIBLIOTECA CENTRAL
"ZILA MAMEDE"



L187038

