Edge Detection for the
Segmentation of Salmonids in
Digital Images

by
Colin Savage
B.A.Sc. The University of British Columbia 1991

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE
REQUIREMENTS FOR THE DEGREE OF MASTER OF APPLIED
SCIENCE

in
The Faculty of Graduate Studies
(Bio-Resource Engineering)

We accept this thesis as conforming
to the required standard

The University of British Columbia

April 1995
©Colin Richard Savage 1995
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of **BIO-RESOURCE ENGINEERING**

The University of British Columbia
Vancouver, Canada

Date 4/6/95
Abstract

Video footage of salmonids in sea cages was obtained from a number of areas in Johnstone and Georgia Strait. The footage represented a large number of camera orientations including camera image axis perpendicular and parallel to the water surface. Perpendicular orientations included camera image axis pointing towards the bottom and surface of the water column. Parallel orientations included camera image axis pointing towards the center and side of the sea cage. The use of a white tarp when the cameras were positioned parallel to the water surface provides separation under lower signal to noise ratios. A discussion of noise sources provides a key to understanding the difficulties applying image processing to noisy digital images. In addition, attenuation and scattering when combined with countershading of salmonids illustrates the reflectivity image degradation of salmonids in the water column when viewed from the camera position. Edge detection techniques were applied to noisy digital video images. A comparison of different edge detectors and pre-processing steps was made to determine the “best” separation technique. The Canny operator with a sigma of 2.0 proved to be the “best” edge detector. In order to smooth noisy images, a pre-processing step of Anisotropic Diffusion was applied when the signal to noise ratio was low. This work provides a basis for the automation of counting and sizing assessment of salmonids.
Contents

Abstract ii
Table of Contents iii
List of Tables v
List of Figures vi
Acknowledgements viii
1 Introduction 1
2 Objectives 6
3 Literature Review 7
  3.1 Factors Affecting Salmonid Image Quality 7
    3.1.1 Absorption and Scattering by the Hydrosphere 7
    3.1.2 Water Clarity 15
    3.1.3 Visibility Range of Salmonids 18
    3.1.4 Indices of Refraction 20
    3.1.5 Light Transmission 23
    3.1.6 Refraction and Reflection of Light at the Sea Surface 24
  3.2 Edge Detection and Smoothing 26
    3.2.1 Roberts Filter 28
    3.2.2 Nevatia-Babu 29
    3.2.3 Laplacian of Gaussian 29
    3.2.4 Canny 31
    3.2.5 Anisotropic Diffusion 35
4 Materials and Methods 43
  4.1 Field Acquisition 45
    4.1.1 Hardware Modification and Design 45
  4.2 Laboratory Acquisition 48
    4.2.1 Hardware 48
    4.2.2 Software 49
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2.3 Camera Orientation and Ambient Parameters</td>
<td>51</td>
</tr>
<tr>
<td>4.3 Determination of a Criteria for Comparative Assessment</td>
<td>52</td>
</tr>
<tr>
<td>4.4 Previous Edge Detection Findings</td>
<td>53</td>
</tr>
<tr>
<td>5 Results and Discussion</td>
<td>60</td>
</tr>
<tr>
<td>5.1 Test Images</td>
<td>60</td>
</tr>
<tr>
<td>5.2 Results of Comparative Assessment Criteria</td>
<td>60</td>
</tr>
<tr>
<td>5.2.1 Noise and the Sampled Signal</td>
<td>60</td>
</tr>
<tr>
<td>5.2.2 Nominal Noise Characteristics in the Sampled Signal</td>
<td>64</td>
</tr>
<tr>
<td>5.2.3 Nominal Edge Detection Characteristics</td>
<td>64</td>
</tr>
<tr>
<td>5.2.4 Pre-Processing Algorithms and Edge Detectors</td>
<td>68</td>
</tr>
<tr>
<td>5.2.5 Application of Pre-Processing Algorithms and Edge Detectors</td>
<td>71</td>
</tr>
<tr>
<td>5.3 Applications</td>
<td>81</td>
</tr>
<tr>
<td>6 Conclusions</td>
<td>102</td>
</tr>
<tr>
<td>6.1 Recommendations</td>
<td>105</td>
</tr>
<tr>
<td>Bibliography</td>
<td>109</td>
</tr>
<tr>
<td>Appendix A</td>
<td>115</td>
</tr>
<tr>
<td>A.1 Dome Port Lens Design</td>
<td>115</td>
</tr>
<tr>
<td>A.1.1 Methods of Calculation Simplification</td>
<td>115</td>
</tr>
</tbody>
</table>
List of Tables

5.1 Species and Strain in Test Images ....................... 61
5.2 Conditions and Related Data for Images .................. 61
5.3 Comparison of operator execution time on a 486DX/33 CPU .. 72
5.4 Object contrast against varying backgrounds. Camera angle is with respect to the water surface ......................... 72
List of Figures

3.1 Attenuation and scattering of light reaching salmonid .......... 38
3.2 Characteristic Length versus Wavelength of light ............... 39
3.3 Deviation of a ray at the air-water interface .................. 40
3.4 Dispersion of white light after refraction into a dense medium 41
3.5 Support of the non-maximum operator .......................... 42

4.1 Location of data collection sites .............................. 55
4.2 Field Acquisition of Video Footage ............................ 56
4.3 Laboratory Digital Capture of Video Footage .................... 57
4.4 Imaging Technology OFG Block Diagram ......................... 58
4.5 The constructed and modified field apparatus: A. Underwater Housings. B. Underwater Cameras in interior mount. C. Umbilical Cables with End Caps. D. Remote Control Unit. .. 59

5.1 Images tested were obtained from a variety of sites, lighting conditions, camera orientations, and salmonid species and strain 86
5.2 Edge detection with various operators with an artificially derived step edge. A. 3% ARN B. 15% ARN C. spatial brightness map at step edge with 3% ARN D. spatial brightness map at step edge with 15% ARN .................. 87
5.3 Spatial brightness maps illustrating nominal edge detection maladies .............................................. 88
5.4 Spatial brightness maps illustrating anisotropic diffusion performance for the detection of edges under average lighting condition with the camera parallel to the water surface. ............ 89
5.5 Spatial brightness maps illustrating a camera orientation which typically does not require anisotropic diffusion .................. 90
5.6 A. LoG σ = 2.0 B. Anisotropic Diffusion 3-pass LoG σ = 2.0 C. Anisotropic Diffusion 3-pass Roberts and classic thinning . 91
5.7 A1010644 A. Canny σ = 2.0 B. Canny σ = 2.0 with Anisotropic Diffusion 3 passes .......................... 92
5.8 A1402441 A. Canny σ = 2.0 B. Canny σ = 2.0 with Anisotropic Diffusion 3 passes .......................... 93
5.9 A1900611 A. Canny σ = 2.0 B. Canny σ = 2.0 with Anisotropic Diffusion 3 passes .......................... 94
<table>
<thead>
<tr>
<th>Number</th>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.10</td>
<td>A2001944</td>
<td>A. Canny $\sigma = 2.0$ B. Canny $\sigma = 2.0$ with Anisotropic Diffusion 3 passes</td>
</tr>
<tr>
<td>5.11</td>
<td>A2002357</td>
<td>A. Canny $\sigma = 2.0$ B. Canny $\sigma = 2.0$ C. Canny $\sigma = 2.0$ D. LoG $\sigma = 2.0$ with Anisotropic Diffusion 3 passes</td>
</tr>
<tr>
<td>5.12</td>
<td>A3300118</td>
<td>A. Canny $\sigma = 2.0$</td>
</tr>
<tr>
<td>5.13</td>
<td>A3302725</td>
<td>A. Canny $\sigma = 2.0$ B. Canny $\sigma = 2.0$ with Anisotropic Diffusion 3 passes</td>
</tr>
<tr>
<td>5.14</td>
<td>A3423291</td>
<td>A. Canny $\sigma = 2.0$ B. Canny $\sigma = 2.0$ with Anisotropic Diffusion 3 passes</td>
</tr>
<tr>
<td>5.15</td>
<td>A3700125</td>
<td>A. Canny $\sigma = 2.0$ B. Canny $\sigma = 2.0$ with Anisotropic Diffusion 3 passes</td>
</tr>
<tr>
<td>5.16</td>
<td></td>
<td>A novel automatic sea cage counting method</td>
</tr>
</tbody>
</table>

A.1 Relationships of dome paraxial object and image
Acknowledgements

I would like to take this opportunity to thank a number of people who were instrumental in the successful completion of this thesis.

To Dr. Royann Petrell, who wondered if I would ever finish this thesis. Thanks for all your support and suggestions. To Dr. Rabib Ward and Dr. Jim Little for their support and suggestions.

Additionally, thanks to the faculty and staff of Bio-Resource Engineering, Dr. Sie-Tan Chieng, Dr. Victor Lo, Dr. Anthony Lau, Julia Sung, Jurgen Pehlke, and Neil Jackson for their support.

Thanks to Trev for being Trev.

Thanks to my parents Ron and Donna Savage, and my sisters Kerry and Robyn Savage for their support. Thanks to Ron and Betty Henderson, Gayle and Lorne, for their kindness and support.

Finally to my wife, Karen for her everlasting support and suggestions and who will be pleased that this chapter of my life can be completed.
Chapter 1

Introduction

Justification

"It's becoming clear our fish-counting capability is not sufficiently reliable." Professor Peter Pearse 1994 [9]

Fish health, condition, size and number are important factors that impact the economy of an aquaculture enterprise. Despite the economic impact, these factors are inconsistently monitored because the means of measurement are invasive, and/or difficult and time consuming to execute.

Estimation of salmon number and size in net pen aquaculture and fisheries is currently based on systems that move and handle salmonids and/or are inherently inaccurate. Consequently inventory is either assessed infrequently or is unreliable. A reliable, non-invasive inventory method would allow for routine measurements, which could be used to solve inventory related problems of salmon farm operation and accounting. For instance, routine farm operations such as salmonid feeding, sorting and drug treatment require knowledge of salmonid number and size. Financing and loan payments require knowledge
of expected salmonid numbers. Both production and forecasting and feed contracting could be more efficiently handled, and insurance claims could be verified if changes in salmonid number and size can be measured.

Monthly inventories are ideal, but they are not practical with the use of conventional inventory methods. Piper et al. 1986 [29] discusses three conventional inventory problems: sample counts, total mass and the pilot-tank technique. Two problems exist within these techniques. Firstly, considerable stress is associated with netting activities. As a result, the stress involved will lead to lowered disease tolerance and increased mortalities. Any predator losses from herons, gulls, otters, and seals are not assessed, original stocking count could be inaccurate and the difference method does not account for escaped salmonids. These avenues of loss are typical of commercial net pen culture.

The conventional inventory techniques are unable to account for mortalities and losses due to a series of predator attacks, net breakages and inaccurate diver counts. These problems have led to the development of novel methods of inventory assessment. Videotape recording methods have been used. These methods are accurate; however, the salmonids must be manipulated over a shallow or narrow portal to be videotaped [38]. Although salmonids can be counted automatically using basic image analysis techniques [38, 13], the more common practice is the semi-automatic method which involves the laborious counting of salmonids with a hand counter from video replay. Salmon farmers have reported that the count takes 4 hours per net pen. The average
computer generated estimate is 25% less than the actual number, while the semi-automatic generated estimates varied 6.4% from the actual number of salmonids [13]. Two critical drawbacks exist in the semi-automatic method. Manipulation of salmon through the portal could induce stress [42, 35], and the counting procedure is still tedious and laborious.

Non-invasive acoustic techniques for wild fisheries inventory have been studied in farmed salmonid conditions. Acoustic techniques have been viewed as potentially the most direct and effective means of estimating the abundance of salmonids in a net pen. Unfortunately, the higher densities of salmonids associated with salmon farming (8-15 kg/m³—this represents a mass of salmonids per m³ of water) severely limits the capability of hydroacoustic techniques [4]. Random or homogeneous distribution of salmonids within the pens must be assumed. Distributional and orientation behaviour of the salmonid must be known to overcome the fundamental problem associated with the hydroacoustic techniques.

Diseases and undesirable morphological changes in appearance should also be frequently monitored so prompt remedial action can be taken. These conditions are normally only detected when salmonids manifest gross symptoms or incur a high mortality rate. By this time, the salmon farmer has no choice but to accept salmonid losses. Some salmonid diseases (i.e. sea lice, *Lepeophtheirus* sp. and physical states (spinal curvature, fin necrosis, precocious maturation) can be visually detected non-invasively, provided a vision system is in place. The early detection that a vision system would provide
could make the difference between financial loss or gain, since physical state and salmonid maladies could be recognized at an earlier stage.

Natural-light videotaping with economical underwater video cameras can be used to non-invasively capture large quantities of salmonid images inside a sea cage, thus salmonids do not need to be manipulated or removed from the water to be monitored. If the images are of good quality, they can be used to count, size and detect visually obvious diseases and undesirable conditions on salmonids. The quality of an image in natural light of the underwater environment is affected by the amount and quality of sunlight penetrating the surface of the water, sun azimuth, camera angle, light attenuation and scattering, and dome port distortion and aberrations [7]. Dome port distortion (typically radial distortion) and aberrations can be addressed through the use of good quality acrylics, polishing and correction factors for distortion [8].

This work was part of a larger project, whose goal was to develop and test economical non-invasive methods to monitor, count and size salmonids (i.e. *Oncorhynchus Tschawtywa* and *Salmo salar*) within a sea cage or tank. Currently, salmonid counting and sizing methods that have been developed are semi-automatic [23]. This work involves the use of digital image analysis techniques to develop automatic counting and sizing methods. The automatic counting procedures will involve:

- positioning underwater video cameras within a net pen
• videotaping salmonids without artificial lighting
• detecting salmonids with a computerized image analysis system
• computerized counting of salmonids from a video screen including partially occluded salmonids using image recognition techniques

The sizing procedures will be used to obtain the mass of individual salmonids and the mass distribution within the netpen. The automatic sizing procedure will involve:

• positioning a stereo pair of housed video cameras within a netpen
• videotaping salmonids without artificial lighting
• measuring individual salmonids in a prescribed number of video frames and netpen locations using a computerized sizing algorithm. These methods do not require that salmonids be moved or corralled.
Chapter 2
Objectives

This thesis addresses the separation of salmon from their backgrounds to allow feature extraction for automatic, computerized counting and biomass assessment.

The objectives of this thesis are five-fold:

- to videotape salmonid images under various lighting conditions and camera orientations.
- to analyze imaging apparatus (video cameras, videotape recorders and imaging cards) and the physical factors affecting image quality.
- to separate salmonid images from their backgrounds using different edge operators.
- to determine a basis for comparative assessment.
- to show application of results for counting and biomass assessment.
Chapter 3
Literature Review

3.1 Factors Affecting Salmonid Image Quality

3.1.1 Absorption and Scattering by the Hydrosphere

Once light, whether it is sunlight or an artificially produced light, is within the hydrosphere, it is subject to the effects of only two physical processes:

- absorption—the conversion of light into heat energy
- scattering—the change in direction of travel of light energy by foreign particulate matter.

The absorption and scattering of light by the hydrosphere affects image quality in a number of ways. The presence of particulates in the water column tends to scatter light. This decreases image quality since each of the small particulates typically is illuminated. A large density of particulates further diminishes image quality until eventually a target at some distance from a camera is invisible. Since salmon are countershaded, further complications
exist which hamper edge detection of salmonids from ambient water column conditions. The level of light in the water column drops off exponentially, therefore changes in brightness are evident from the dorsal to ventral sides of the salmonid (Figure 3.1). Combined with countershading alternatively scattering and absorbing light, the edges of the salmonid are a formidable separation task in water.

**Attenuation Properties of Water**

The attenuance of light by distilled water is well documented [12, 43, 14] (Figure 3.2). Even though the data presented reflect the properties of fresh water, when sea salts are added the selective absorption properties of the resulting mixture do not differ from those of distilled water [12, 43]. An examination of the figure yields a number of significant results. The first to consider is the extreme peak of energy transmission towards the shorter end of the visible spectrum. This occurs at a wavelength of \( \approx 470 \text{ nm} \), which is in the blue region of the visible spectrum. On each side of this peak the transmission falls off very rapidly, reaching ten-percent transmission values at wavelengths of about 570 nm in the yellow and 330 nm in the ultraviolet. The transmittance decreases very rapidly as the infrared region is entered, with characteristic length values at the wavelength of peak transmission being about 5000 times as great as that at a wavelength of 1000 nm. Beyond this, the characteristic length continues to decrease with an increase in wavelength. The characteristic length is realized by:
The absorptance is defined as the ratio of the radiant flux lost by absorption to the incident radiant flux. On the other hand, the absorption coefficient is defined as the absorptance of an infinitely thin layer divided by the layer thickness.

The absorption coefficient:

$$a = -\frac{A}{\Delta l} \quad (3.1)$$

- $A$—absorptance of material
- $a$—absorptance coefficient of material
- $\Delta l$—infinitesimal layer

The absorption coefficient of water is usually expressed in terms of a specific wavelength since water is a selective absorber.

Starting with Lambert’s law:

$$-a_{\lambda}l = \ln T_{\lambda} \quad (3.2)$$

- $a_{\lambda}$—the absorption coefficient for a particular wavelength.
- $T_{\lambda}$—the transmittance at the particular wavelength.

Returning to Lambert’s law, the absorption coefficient may vary from 0, when all the energy is transmitted ($T=1$) to infinity, when all the energy is absorbed ($T=0$).
The characteristic length is equal to the distance at which the transmittance is equal to $1/e$.

- $e$—exp function (2.7182818)

$$ T = \frac{1}{e} = \exp(-a_\lambda \mathcal{L}) $$  \hfill (3.3)

- $\mathcal{L}$—characteristic absorption length
- $a_\lambda$—absorption coefficient at specific wavelength
- $e$—exp function (2.7182818)

then

$$ a_\lambda \mathcal{L} = 1 $$  \hfill (3.4)

Finally the characteristic absorption length $\mathcal{L}$ is:

$$ \mathcal{L} = \frac{1}{a_\lambda} $$  \hfill (3.5)

Since water is a selective absorber, it becomes necessary to discuss characteristic lengths for particular wavelengths ($\mathcal{L}_\lambda$ just as in the case of absorption coefficients. Similarly, the characteristic scattering length may be defined as:

$$ \mathcal{L}_{\lambda_s} = \frac{1}{a_{\lambda_s}} $$  \hfill (3.6)
and the characteristic attenuation length may be defined as:

\[ \mathcal{L}_{\lambda_a} = \frac{1}{a_{\lambda_a}} \]  

(3.7)

As an example, if a sample of water has a characteristic attenuation length of 10 metres, this indicates that in a path length of 10 metres all but 36.7% \((1/e)\) of the energy has been lost. The corresponding attenuation coefficient of 0.1 per metre conveys little intuitive information for the transparency of the water. As a result, characteristic lengths are often used as an indicator of turbidity [15]. The transparency of the electromagnetic spectrum has maximums in three regions. The region of greatest transparency in the visible spectrum occurs at \(\approx 470\) nm (blue) [14]. Transparency occurs at high frequencies in the X-ray region. This high frequency region is not typically utilized since the individual photons have such high energy levels that the energy is dangerous to most organisms. Since salmonids must be assessed non-invasively, this wavelength is not practical. The third region is that of low frequency radio waves. This region is used for enumeration of salmonids using hydroacoustic techniques, but a problem exists in a sea cage environment. Salmonids are typically stocked at relatively high stocking densities \((5-10\ kg/m^3)\). This provides attenuation of light due to shading of salmonids from other salmonids which are present in the path of the light source.

A considerable amount of circumstantial evidence exists that suggest bacteria may contribute significantly to the attenuation of light in the water.
column. Spinrad et al. 1989 [37] observed changes in beam attenuation that correlated closely with changes in bacterial abundance during a hydroacoustic survey. Kopelevich et al. 1987 [18] examined enriched cultures of various strains of marine bacteria and determined that the scattering and absorption coefficients per unit numerical concentration vary over 3 orders of magnitude between species. Variations in the concentration of biogenous particles that absorb and scatter light occur because of physical transport by advection and diffusion within a non-homogeneous field of particle concentration and are due to local differences in the rates of particle production and loss. These local differences are principally caused by variations in the rates of microparticle production by the growth of phytoplankton cells and the loss of grazing of herbivorous zooplankton [17]. Stramski and Kiefer 1990 [41] estimate that bacteria may contribute as much as 12% to light absorption (at 390 nm) and tens of percent to light scattering (550 nm) by particulates. Another point which must be mentioned is the effect of marine bacteria on the index of refraction. Stramski and Kiefer 1990 [41] determined that the mean \( \bar{n} \) (index of refraction) of a number of marine bacteria would be approximately 1.055. This must be considered if the mean bacterial population density is determined to be of such a magnitude that the index of refraction for the column of water of interest is sufficiently affected.
Scattering

The other basic physical process involved in the loss of energy in the hydrosphere is that of scattering. This is a result of a change in direction of the light flux. There are three approaches to describing scattering of light in the natural environment.

- Assume that the scatterer acts as a reflector and that energy is reflected from the surface of individual particles.
- Assume that the particle acts as a refractor and diffractor so that the energy passing through the particle is deviated by refraction, and that the energy passing close to the particle is deviated in its path by diffraction.
- Assume that the energy impinging on a particle is absorbed by the particle and then reradiated in many different directions without a change in wavelength (electromagnetic approach).

If the particles are relatively large with respect to the wavelength of the light (particle radius > 10\(\lambda\)), the first two approaches can be used to produce a model which describes the real world reasonably well [15]. The hydrospheric environment is rarely host to particles larger than a few micrometers in size for any extended length of time. It is for this reason that the electromagnetic approach is most often used to describe scattering in the water column.

The electromagnetic theory of scattering is based on the work of Gustav Mie [44]. Mie assumed that energy impinged upon spherical conducting particles in a nonconducting medium. These particles resonated electromagnetically and reradiated energy relative to the size of the particles with respect to
the wavelength of the incident light. This relation is described by the non-
dimensional parameter:

\[
\frac{2\pi r(m^2 - 1)}{\lambda(m^2 + 1)}
\]  \hspace{1cm} (3.8)

- \( r \) — particle radius

- \( \lambda \) — wavelength of light

- \( m \) — relative index of refraction of the scatterers, with respect to the
  medium.

\[
m = \frac{\text{speed of light in medium}}{\text{speed of light in particle}}
\]  \hspace{1cm} (3.9)

When the effective scattering area coefficient (ratio of the effective scattering
cross-sectional area to the actual cross-sectional area) is plotted against the
dimensionless parameter, a number of points are evident. When the parti-
cles are small compared to the wavelength, the amount of energy scattered is
quite small but rises to a maximum as the particle size increases [15]. With
a further increase in \( r \), the value of \( K \) (\( K \) is the effective scattering-area
coefficient) decreases and oscillates about a value of 2.0 (overall effect is sim-
ilar to the ringing oscillations of the \( \text{sinc}(x) = \frac{\sin(x)}{x} \) function), becoming
asymptotic to 2.0 as the particle size becomes large. If this data were plot-
ted, a family of curves results. For very small particles, typical of mineral
suspenoids such as clay and silts, there is a great deal of scattering at the
shorter wavelengths. At the other extreme, very little scattering occurs at
3.1.2 Water Clarity

Water clarity has a critical effect on image quality. Any particulates in the water column tend to decrease water clarity. In addition, a large concentration of particulates in the water column can be detrimental to the health of salmonids. Generally, a small concentration of particulates decreases water clarity slightly, while a larger concentration tends to decrease water clarity to a larger degree.

A number of factors influence the clarity of water. Inorganic pollution is the main cause of high turbidity. This is generally composed of sand, silt, clay, and colloidal particles which have been carried out to sea by rivers and streams. Since this type of turbidity is normally the result of land erosion, a rocky coast or forested land region will tend to provide lower amounts of particulate matter carried into the water than a watershed which serves an agricultural area. In addition, if the drainage area is composed mostly of sand (large particles), there will not be as much erosion in terms of the total suspended solids since lesser amounts of clay and silt will be present.

Another factor determining the amount of turbidity in a marine area is the total amount of runoff. With increasing precipitation, more runoff occurs; since this results in a greater probability of erosion, more particulate matter will be carried off to sea.
Once the particulate matter is in suspension, it is transported by streams and eventually settles to the bottom. The rate of settling is determined by Stokes' law if the diameter of the particle is less than 0.2 mm. With a particle radius greater than \( \approx 2 \) mm, the settling rate is a function of the square root of the radius. Particles which have a radius between 0.2 mm and 2 mm have a settling rate which is proportional to a power of the particle radius between \( \frac{1}{2} \) and 2 [44].

The only difficulty with the preceding equations is the assumption that the water column which the particle descends is still. This is not the case since currents are a selected criteria of a salmon farm site. But for the purposes of most studies, the preceding equations provide a reasonable approximation to actual conditions. Since turbulence is always present, small particles are constantly re-suspended in the water column. Therefore, an area with small particulate matter is generally an area of large amounts of turbidity. In addition, areas with large currents such as those caused by tides or river streams, will result in larger amounts of suspended matter in the water column.

A number of other factors contribute to water clarity. The injection of foreign material or artificially stirring the water both tend to increase turbidity. The first case is generally considered pollution either domestic or industrial. Mechanical stirring of the water is often produced by the outfall of a hydro-electric facility.
In addition, a large contributor to turbidity is organic material (salmonid faecal material), including living organisms (phytoplankton blooms) in the water. These are typically microscopic plants and animals which exist seasonally in large numbers in certain areas. A number of factors contribute to the presence or absence of microscopic plants in the water column. Plants require both sunlight and additionally to grow, nutrients or fertilizer. Spring, summer and fall generally provide the necessary sunlight while nutrient cycles during a year return nitrates and phosphates from decaying organic material to the water column. If man supplies an additional load of nutrients, the production of phytoplankton (microscopic plants) will increase dramatically in the summer months when sunlight hours are long and upwelling is diminished. The abundance of phytoplankton translates to an increase in the turbidity.

Turbidity is also affected by wind. The wind speed, duration, and fetch (distance the wind has blown) must be considered when considering the impact of wind-driven turbidity. The greater the fetch and the greater the duration of the wind, the greater the turbulence produced.

In summary, a number of factors must be considered when assessing turbidity:

- nature of the land which borders on the area
- currents in the area
- amount of precipitation in the area
- amount of wind mixing
3.1.3 Visibility Range of Salmonids

Visibility range is simply defined as the distance at which an object can just be seen. This is defined in terms of a constant divided by the attenuation coefficient in a similar manner to the characteristic length (Section 3.1.1). In this case, the smallest contrast which may be detected \( C_T \) is set equal to \( \exp[-a_{\lambda T}V] \):

\[
C_T = \exp(-a_{\lambda T}cV)
\]  \hspace{1cm} (3.10)

where

- \( V \)—the visibility range
- \( a_{\lambda T} \) Total attenuation coefficient \( a_{\lambda a} + a_{\lambda s} \)

Then

\[
\ln \left( \frac{1}{C_T} \right) = a_{\lambda T}V
\]  \hspace{1cm} (3.11)

and
The constant $A$ varies from $\approx 6$ down to $\approx 2^{1/2}$ [44]. The visual threshold consists of a gradual transition rather than an abrupt cutoff. The threshold contrast is usually defined as the lowest contrast that can just be seen by the average eye. The visual threshold is a function of a number of parameters which are interrelated:

- Contrast
- Ambient light level
- Size of the object
- Rate of flickering

The absolute maximum-visibility range may be obtained by (assuming $C_T \approx 0.003$) [3]:

$$V = \left( \frac{1}{a_{VT}} \right) \ln \left( \frac{1}{C_T} = \frac{A}{a_{VT}} \right)$$  \hspace{1cm} (3.12)

The contrast of a scene is usually defined as the difference between the object brightness and the background brightness, divided by the background brightness.

$$C = \frac{B_O - B_B}{B_B}$$  \hspace{1cm} (3.14)

\[300\]
• $C$—Contrast of an object relative to its background

• $B_O$—Background brightness

• $B_R$—Object brightness

This relationship allows for contrast which is produced as a result of a bright object against a dark background or a dark object against a bright background. The images represented in this thesis represent both of these conditions since salmonids which are imaged from the side or top view are often the former condition and salmonids which are silhouetted against the surface are typically dark against a bright background. The sign of the contrast value indicates whether the object $+$ or the background $-$ is brighter.

### 3.1.4 Indices of Refraction

Image quality is affected by a number of conditions which are between the target and the image capturing device. A change in density of water (fresh or salt water) produces a change in the refraction index. In addition, the refraction index changes at the acrylic dome port of the underwater housing and also at the glass lens of the camera.

The index of refraction at any particular wavelength is generally proportional to the density of the material; an increase in density increases the index of refraction (A calculated value of 1.335 has been used in this thesis for salt water based on McNeill's empirical equation [23]), The deviation of a ray is
expressed by Snell's law (Figure 3.3).

\[ N \sin I = N' \sin I' \]  \hspace{1cm} (3.15)

- \( N \) — index of refraction of water
- \( N' \) — index of refraction of polymethyl acrylate
- \( I \) — incident angle (in water)
- \( I' \) — refracted angle (in polymethyl acrylate)

Generally, Snell's law indicates that a refracted ray passes closer to the normal in the denser medium as compared to a ray in the less dense medium. The refractive index of a yellow line is used to describe the optical properties of most materials. In the United States, this principal index \( N_D \) is taken at a wavelength of 589.3 nm, which is the mean wavelength between the sodium D lines of 589.6 nm and 589.0 nm.

A measure of dispersion of the material is realized by \( \Delta N \). This is indicated by the difference in the indices of refraction for C and F rays (See Figure 3.4). A measure of dispersion and its ratio with the basic refracting power of the material is often used as an indicator of the dispersive power of the material and is designated by \( \omega \). Where the angle of incidence is small, the dispersion of C and F rays \( (I'_C - I'_F) \) is proportional to:

\[ N_F - N_C \]  \hspace{1cm} (3.16)
The deviation of the D ray \((1 - \Gamma_D)\) is proportional to:

\[
N_D - 1
\]  
(3.17)

The reciprocal of the dispersive power \(V_D\) is known as the reciprocal relative dispersion, Abbe V-number or V-value:

\[
\frac{1}{\omega} = V_D = \frac{N_D - 1}{N_F - N_C}
\]  
(3.18)

The refractive index of water has been developed empirically by McNeil [23]:

\[
N_{\lambda} = 1.3247 + 3.3 \times 10^3 \lambda^{-2} - 3.2 \times 10^7 \lambda^{-4} - 2.5 \times 10^{-6}T^2 \\
+ (5 - 2 \times 10^{-2}T)(4 \times 10^{-5}S) \\
+ (1.45 \times 10^{-5}P)(1.021 - 6 \times 10^{-4}S)(1 - 4.5 \times 10^{-3}T)
\]  
(3.19)

- \(N_{\lambda}\)—refractive index of water at wavelength
- \(\lambda\)—wavelength of light, nanometers
- \(T\)—temperature of water, °C
- \(S\)—salinity of water, parts per thousand
- \(P\)—pressure of water, kg/cm²

The temperature and salinity of water have a small effect on the index of refraction. With a salinity of 25 parts per thousand and a temperature of 5°C, an \(N_{\lambda}\) of 1.339 is realized. On the other end of the scale, with a salinity of 5 parts per thousand and a temperature of 25°C, an \(N_{\lambda}\) of 1.334 is realized. The greatest impact on refraction is the wavelength. Since blue \((\approx 470 \text{ nm})\) is attenuated the least for a given depth, it is the dominant wavelength at
depths which are commonly used for video taping. The refractive index is impacted the most by wavelengths extending from $N_{400\text{nm}}$ of 1.349 to $N_{670\text{nm}}$ of 1.336. A generalization for the refractive index change may be made with respect to an increase in temperature. The refractive index change for polymethyl methacrylate (Acrylic) ranges from $-11$ to $-13 \times 10^{-5}/^\circ\text{C}$ increase [21].

3.1.5 Light Transmission

The transmission of light through acrylic (acrylic is the material selected for the dome port in the underwater camera housing) varies with the wavelength and the angle of incidence. Two grades of acrylic are available for this: UVA ultraviolet absorbing and UVT ultraviolet transmitting. When light is incident on an optical surface in air, the amount transmitted is a function of the angle of incidence. For an angle of incidence of 0 radians, the reflectance ($r$) of the surface is expressed as:

$$ r = \left( \frac{N' - N}{N' + N} \right) $$

- $N$—index for the incident medium
- $N'$—index for the refracted medium

This is due to the polarization of the incident light which increases with an increasing angle of incidence to the polarizing angle (Brewster's angle):

$$ \text{Brewster's angle} : \bar{T} = \tan^{-1} N $$

(3.21)
For acrylic, this angle is $56.15^\circ$ (0.98 radian). Beyond the polarizing angle, the transmission of light decreases rapidly with increasing incidence until at $90^\circ$ (1.57 radians) all light is reflected. At $0^\circ$ (0 radians) incidence, the reflectance of an air-acrylic interface is 3.9%. The water-acrylic interface of an underwater viewport has at 0 degrees (0 radian) incidence, a reflectance of only 0.3 percent. This indicates that multicoating the water interface of an underwater viewport mostly benefits scratch resistance, but coating the air interface would improve scratch resistance and reduce reflection losses.

3.1.6 Refraction and Reflection of Light at the Sea Surface

The scatter of light at the sea surface is due to the roughening action of the wind. Waves and swell of all wavelengths down to the shortest capillary waves combine to refract and reflect light rays into continually changing directions and additionally to focus small bundles of rays. Bubbles and foam at the sea surface tend to scatter light further. As a consequence the appearance of the sun reflected at, or refracted through, the sea surface is broken up into a large number of highlights each of which is a distorted image of the sun.

Refraction and reflection affects image quality for a number of reasons. Salmon are countershaded. The scales on a salmon tend to reflect light at certain angles. This is evident when viewing salmonids from the side. Certain areas tend to flash patches of light while others provide a lack of contrast when compared to their background. The scattering of light also tends to illu-
minimize particulates in the water column which reduces visibility of targets which are behind a cluster of illuminated particulates.

The average distribution of directions of propagation of light is broadened by scattering. In addition, the average intensity of light is altered. As an example, consider late afternoon when the sun is low on the horizon. A larger amount of sunlight enters the sea on a windy day than on a calm day since under the former condition the rays from the sun enter the sea through steeply sloped surface facets which have a greater transmissivity than facets entered on a smooth sea.

The scattering of light on the sea surface is important for marine organisms for a number of reasons. Organisms living very close to the surface are exposed to enormous fluctuations of light intensity on sunny days due to the focusing effect of surface ripples. On windy days when the typical curvature of the sea surface is large, the flickers are most prominent within the first few centimetres of the sea surface. On calm days the flicker is apparent to several metres.

The scattering of light by the sea is also important for the visibility of objects on the sea or seen through the sea surface. The horizon on a calm day is almost invisible because the reflection coefficient at incidence approaches unity. As a result at the horizon a calm sea takes on approximately the same color and brightness as the sky just above the horizon.
On average, the effect of surface roughness on reflectance is relatively small, compared to that produced by a change in the angle of incidence. The effect of the sun's altitude on reflectance by a smooth surface is described by Fresnel's equation:

\[
\rho = \frac{1}{2} \left( \frac{\tan^2(i - r)}{\tan^2(i + r)} + \frac{\sin^2(i - r)}{\sin^2(i + r)} \right)
\]  

(3.22)

The variables represent:

- \(\rho\) — the reflectance
- \(i\) — angle to the normal made by the incident light
- \(r\) — angle to the normal made by the reflected light

Assuming that the index of refraction for water is \(n = 1.335\), the reflectance works out to \(\approx 2.0\%\) at zenith position up to 40°. From this position up to a zenith with the sun at the horizon the percentage rises quickly to 100%.

### 3.2 Edge Detection and Smoothing

A problem of paramount importance in the separation of salmonids from their background is edge detection. Edges characterize object boundaries and therefore are useful for segmentation of objects in digital images. Edge points can be thought of as pixel locations of abrupt brightness change. Edge detection may be considered as a technique which generates compact descriptions which preserve most of the structural information in an image [5]. A
digitized image of a salmonid contains a great deal of redundant material before and after compression (lossy and non-lossy) which is not useful for later vision work. A method to be considered for extracting the shape of salmonids in digital images is shape from shading [10]. Salmonids may be considered smooth, but unfortunately are not of uniform reflectance. It is for this reason that shape from shading may not be applied reliably to obtain surface orientation.

An ideal edge detector should produce an edge response which is localized to a single pixel located at the midpoint of the slope. If the slope angle is 90°, the resultant edge is known as a step edge. In digital imaging systems, step edges exist only in artificially created images. Digital images, resulting from the digitization of optical images of real scenes, usually do not contain step edges because the anti-aliasing low-pass filtering prior to digitization reduces the edge slope in the digital image caused by any sudden luminance change in the scene [30].

Edge detectors can be divided into two broad classifications: gradient or first derivative operators and second derivative operators. Gradient operators respond with a broad peak at an edge location. These operators require thinning before feature extraction can be implemented. Second derivative operators respond with a zero-crossing at an edge location. The edge location can be interpolated with precision depending on the SNR. A computational advantage exists in the selection of a rotationally invariant operator such as the LoG (Laplacian of Gaussian). The LoG is separable and therefore a
row and column vector may be convolved instead of an entire mask. This decreases the number of computations from $n^2$ to $2n$.

### 3.2.1 Roberts Filter

The Roberts filter finds use in a number of machine vision applications since it lends itself to real-time processing. Given a function $f(x,y)$, the gradient of $f$ at coordinates $(x,y)$ is defined as the vector.

$$G[f(x,y)] = \left[ \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right]$$

(3.23)

In particular, a useful approximation is known as the Roberts gradient. The filter may be approximated by the following expression:

$$G[f(x,y)] = (|f(x+1,y+1) - f(x,y+1)|^2 + |f(x+1,y) - f(x,y+1)|^2)^{\frac{1}{2}}$$

(3.24)

In this case, the approximation given yields the magnitude of the gradient which is proportional to the difference in gray level between adjacent pixels. Therefore, the gradient assumes relatively large values for distinctive edges in an image, and small values where the image is smooth. The gradient would be zero in areas where the gray-values were constant.
3.2.2 Nevatia-Babu

Nevatia and Babu 1980 [26] developed gain-normalized 5x5 masks which are utilized to detect edges in 30° increments. Larger size template masks provide finer quantization of the edge orientation angle and a greater immunity to noise, but the computational requirements increase quickly. The results from this method are reliable for high SNRs, but performance decreases dramatically with decreasing SNR.

3.2.3 Laplacian of Gaussian

The use of LoG convolution masks was first suggested by Marr and Hildreth 1980 [22]. In two dimensions:

\[ \nabla^2 G(x, y) = \frac{1}{2\pi \sigma^4} \left( 2 - \frac{x^2 + y^2}{\sigma^2} \right) \cdot \exp \left[ -\frac{x^2 + y^2}{2\sigma^2} \right] \]  \hspace{1cm} (3.25)

where \( \sigma \) is the space constant of the Gaussian and \( w = 2\sqrt{2}\sigma \) is the width of the central excitatory region of the operator. Huertas and Medioni 1986 [11] determined that the width of the operator should be approximately \( 3w \). This ensures that 99.7% of the area of the 1-D Gaussian lies within \( \pm 3 \) standard deviations from the mean, giving an area close to zero for \( \nabla^2 G \).

The discrete domain version of the \( \nabla^2 G \) operator can be obtained by sampling the continuous domain impulse response function over an \( nxn \) window.
A filter size of $n = 3w$, where $w = 2\sqrt{2}\sigma$ is the width of the positive center lobe of the $\nabla^2 G$ function. The construction of the discrete $\nabla^2 G$ operator typically involves:

- scaling the operator values by some constant to match the pixel depth available on the imaging card.
- using nearest integer values for each scaled operator value
- extending the support of the filter to include all nonzero integer values.
- manipulating operator values by a small amount to ensure that the values integrate to zero.

The filter sensitivity to an intensity change with a certain width is related to $w$, the width of the central excitatory region of the operator. For example, the smaller the value of $\sigma$ (the space constant of the Gaussian), the more sensitive the filter and the more detail shown in the convolved image. The penalty is poor noise tolerance.

The detection of intensity change is shown as a peak in the first directional derivative of the intensity and a zero-crossing in the second directional derivative. Detection of intensity changes is reduced to finding the zero-crossings in the second derivative of the intensity in the direction of maximum slope at the zero-crossing. A conclusion from Canny’s 1983 work [5] indicated that one of the main disadvantage to the implementation of the LoG is that the positional error on average is $\approx 60\%$ greater than that of a directional operator of the same $\sigma$. 
3.2.4 Canny

Canny's approach to edge detection began with a traditional model of a step edge in white Gaussian noise [5]. From this he attempted to formulate effective criteria for edge detection. The detection was performed by convolving the noisy edge with a spatial function \( f(x) \) (which he was trying to determine) and marking edges at the maxima in the output of this convolution. He specified three performance criteria for good edge detection:

- A low probability should exist which would fail to mark real edge points, and a low probability of falsely marking non-edge points. Since both of these probabilities are decreasing functions of the output signal to noise ratio, this corresponds to maximizing the signal to noise ratio.

- Good localization. The points marked as edges by the operator should be as close as possible to the centre of the true edge.

- The detector should provide only one response to a single edge. This follows from the first criteria since with two nearby operators responding to the same edge one of them must be considered a false edge. The mathematical form of the first criteria did not capture this criteria, and therefore it is described explicitly.

Canny's development is based on a one-dimensional, continuous domain model of a step edge of amplitude \( h_E \) plus additive white Gaussian noise with standard deviation \( \sigma_n \). In the continuous domain, the signal to noise ratio (SNR) is specified by:

\[
SNR = \frac{h_E}{\sigma_n} \frac{S(h)}{S(h)}
\]  

\[ (3.26) \]
with:

\[ S(h) = \frac{\int_{-\infty}^{0} h(x) dx}{\int_{-\infty}^{\infty} [h(x)]^2 dx} \]  \hspace{1cm} (3.27)

The localization factor is defined as:

\[ LOC = \frac{h_E}{\sigma_n} L(h) \]  \hspace{1cm} (3.28)

with:

\[ L(h) = \frac{h'(0)}{\int_{-\infty}^{\infty} [h'(x)]^2 dx} \]  \hspace{1cm} (3.29)

where:

- \( S(h) \)—quotient of the response to the step only to the mean squared noise response
- \( SNR \)—Signal to Noise ratio
- \( h_E \)—step edge
- \( \sigma_n \)—standard deviation of white Gaussian noise
- \( h(x) \)—one-dimensional luminance
- \( h'(x) \)—the derivative of \( h(x) \)
• $L(h)$—Approximation of the standard deviation of the distance of the actual maximum from the true edge

• $w$—width of the central excitatory region of Gaussian

There should only be a single response to a true edge. The distance between peaks of the gradient when only noise is present, denoted as $x_m$ is set to some fraction $k$ of the operator width factor $w$. As a result:

$$x_m = kw$$  \hspace{1cm} (3.30)

where:

• $x_m$—distance between peaks of the gradient
• $w$—operator width
• $k$—fraction of operator width $w$

Canny combined these three criteria by maximizing the product $S(h)L(h)$ constrained by the equation for $x_m$. The Canny impulse functions tend to resemble boxcar functions for small $x_m$ while for large $x_m$ the Canny function can be closely approximated by a derivative of Gaussian impulse response function.

In two dimensions, the Gaussian is represented as:

$$G = \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right)$$  \hspace{1cm} (3.31)
The term "edge direction" as defined by Canny refers to the direction of the tangent to the contour that the edge defines in two dimensions. The image is convolved with a symmetric Gaussian, the edge direction is estimated from the gradient of the smoothed image intensity surface. The gradient is then non-maximum suppressed in that direction (Figure 3.5). This directional non-maximum suppression is equivalent to the application of the following non-linear differential predicate:

\[ G_{\pi} = \frac{\partial G}{\partial \tilde{n}} = \tilde{n} \cdot \nabla G \]  
\[ \tilde{n} = \frac{\nabla (G \ast I)}{|\nabla (G \ast I)|} \]  

where:

- \( G \) — Two-dimensional Gaussian
- \( \pi \) — Image space
- \( \tilde{n} \) — Normal vector to edge direction

The operator locates either maxima or minima, by locating the zero-crossings in the second derivative in the edge direction.
There are a number of reasons for selecting the Gaussian function as a solution for Canny's work. When a linear operator is applied to a two dimensional image, a weighted sum of a number of the input value is formed. This sum will be a difference between local averages of the different sides of the edge. The output represents a moving average of the image. The solution to this would be to use an infinite projection function, but unfortunately real edges are of finite extent. As a result, the projection function must be windowed. One problem which may evolve is if the window function is truncated and rectangular is that the filtered image will not be smooth because of the very high bandwidth of the window. The solution is to use a smooth window function. The Gaussian is a good approximation to both of these, since it has very low bandwidth for a given spatial width (The Gaussian is the unique function with minimal product of bandwidth and frequency) [5].

3.2.5 Anisotropic Diffusion

Anisotropic Diffusion is a multi-scale approach to the description of images. The central idea is to embed the original image in a family of derived images \( I(x,y,t) \) obtained by convolving the original image with a Gaussian kernel [28]. A larger value of \( t \), the scale-space parameter corresponds to images at a coarser resolution. Perona and Malik 1987 [28] pointed out a weakness in standard scale-space paradigm. In standard scale space paradigm, the true location of the boundaries at coarse scales is not directly available at the coarse scale image. The only way to determine the true location is by tracking across scale-space, to their locations in the original image, the edges
which were detected at a coarse scale.

The reason for this spatial distortion is that Gaussian blurring does not recognize the natural boundaries of objects. Perona and Malik 1987 [28] suggest the following example to illustrate. If our image of interest consisted of a treetop with the sky as background, the Gaussian blurring process would result in the green of the leaves mixed with the blue of the sky before the treetop is detected as a feature.

Based on this anomaly due to Gaussian blurring, Perona and Malik determined that three criteria were necessary to generate meaningful multi-scale descriptions of images:

- **Causality:** A scale-space representation should have the property that no spurious detail should be generated.
- **Immediate Localization:** At each resolution, the region boundaries should be sharp and coincide with semantically meaningful boundaries at that resolution.
- **Piecewise smoothing:** At all scales, intra-region smoothing should occur preferentially over inter-region smoothing.

Perona and Malik determined a simple way of modifying the scale-space paradigm to achieve the previous criteria. Applying the diffusion equation as a method of looking at scale-space, the diffusion coefficient $c$ is assumed to be constant regardless of the space location. A suitable selection of $c(x, y, t)$ enables compliance to the second and third criteria and does not sacrifice the causality criterion.
The anisotropic diffusion equation is as follows:

\[ I_t = \text{div}(c(x,y,t)\nabla I) = c(x,y,t)\Delta I + \nabla c \cdot \nabla I \quad (3.34) \]

Where \( \text{div} \) is the divergence operator, \( \nabla \) the gradient operator, and \( \Delta \), the laplacian operator with respect to the space variables. This equation reduces to the isotropic heat diffusion equation when \( I_t = c\Delta I \) if \( c(x,y,t) \) is a constant. To achieve smoothing within a region in preference to across boundaries, the conduction equation can be set to 1 in the interior of each region and 0 at the boundaries. The blurring then takes place in the region with no interaction between regions. This would allow the region boundaries to remain sharp.

Since the region boundaries at each scale are unknown, a best estimate of the location of the boundaries is calculated. If \( E(x,y,t) \) is a vector valued function defined on the image with the following properties:

- \( E(x,y,t) = 0 \) in the interior of each region.
- \( E(x,y,t) = Ke(x,y,t) \) at each edge point where \( e \) is the unit vector tangent to the edge at the point and \( K \) is the local contrast of the edge. Once an estimate of \( E(x,y,t) \) is determined, \( c(x,y,t) \) is selected to be a function \( c = g(\|E\|) \)
Figure 3.1: Attenuation and scattering of light reaching salmonid
Figure 3.2: Characteristic Length versus Wavelength of light
Figure 3.3: Deviation of a ray at the air-water interface
Figure 3.4: Dispersion of white light after refraction into a dense medium
Figure 3.5: Support of the non-maximum operator
Chapter 4

Materials and Methods

The collection of image data took place over an extended period of time. The collection of data consisted of recorded S-VHS video tape of sea cages at a number of sites which carried both Atlantic and chinook salmon. The early data collected consisted of footage which has no background added to aid in the separation of salmonids. The later footage includes the use of a white tarp which acted as a background to aid in the separation of salmonids. The project acquired the acronym NIFE (Non-Invasive Fish Enumeration).

A representative sample of typical images (Figure 5.1) employed in the NIFE project were selected. The sea cage sites were confined to the Georgia and Johnstone Strait areas of British Columbia. The sites included MEP (Marine Ecology Program), PBS (Pacific Biological Station) at Nanaimo, Burdwood and Broughton (Figure 4.1). The now defunct MEP facility provided lab and tank environments. This allowed salmonids to be video recorded and observed under controlled conditions. The lighting and camera orientation were easily changed. Equipment was calibrated and tested for field use. The
PBS facilities provided controlled conditions for use of a sea cage. A number of small netpens (5 m x 5 m x 6 m) were available to test a number of camera orientations. These facilities provided some of the first depth tests (6 m) for the underwater housings. The B.C. Packers Broughton site provided footage of Atlantic salmon *in situ*. In addition, the first sea cage typical depth tests (25 m) were performed on the housings. Burdwood provided footage which included the use of the white tarp to increase separation of salmonids from their background. These images represent different camera orientations and lighting conditions. Camera orientations included:

- image axis perpendicular to the water surface, camera pointed toward the surface.
- image axis perpendicular to the water surface, camera pointed down from the surface.
- image axis parallel to the water surface, camera pointed toward the center of the sea cage.
- image axis parallel to the water surface, camera pointed toward the outside of the sea cage.

Natural lighting ranged from sunlit conditions to overcast. Sunny clear days provided the greatest depth penetration of natural lighting. The use of supplemental lighting was not employed since the project was designed to be non-invasive.
4.1 Field Acquisition

4.1.1 Hardware Modification and Design

System hardware (Figure 4.2) used to capture salmonid images included Panasonic WV-BD400 black and white cameras with a minimum sensitivity of 0.5 lux. The image sensor was a 1.7 cm interline CCD (Charge Coupled Device) with 768 (Horizontal) and 493 (Vertical) pixels. The camera was electronically shuttered to prevent blurring of moving targets in a video frame. A Cosmicar 4.8 mm wide angle lens was used with an in-air diagonal field of view of 95°. With the extended depth of field of the lens, an object positioned 0.4 m to infinity from the camera was in focus. The underwater housing design was made in conjunction with International Hardsuits of North Vancouver, B.C. Design specifications were provided which determined the material, and dimensions which would ensure a large safety factor and minimize long term stress failure in the aluminum.

The internal camera mount which holds the camera in place in the housing was secured using o-rings which relied on friction to prevent movement. The use of an index pin in the lens hood was used to align the housing with the internal camera mount. The aluminum underwater housing was anodized to increase the long-term durability of the housing. The end cap used an o-ring to provide the rear seal for the housing while a locking collar secured the end cap in place. An adjustable threaded rod was used to secure the position of the internal camera mount. This provided reliable positioning of the critical rear dome port surface to lens front element distance. The dome port was
machined from polymethyl methacrylate and was designed to provide a diagonal field of view in-water which is approximately equivalent to its in-air value. The dome port provides a 76° viewing angle in both the horizontal and vertical plane. A right regular pyramid is the resulting view volume. The height of the view volume, as measured perpendicular to the base of the pyramid, was considered to be the visibility (m) within the water column as defined by a Secchi disk reading (A circular disk which is lowered in the water column and viewed from the water surface. The reading is taken when the Secchi disk disappears from view). The front lens hood was threaded to allow the replacement of different dome ports if required and provide a means for complete maintenance of the cylindrical housing body. The front seal consisted of an o-ring which was recessed in the threaded area of the lens hood. Two large stainless steel hose clamps were used which secured an external mounting bracket to the cylindrical housing. This allowed fastening a mounting bracket without drilling the housing. The external mounting bracket used an index pin to eliminate torque rotation of the external camera mount which was caused by the large umbilical cable.

The umbilical cable was assembled using a fish tape to thread the various control wires through the hose. A 1" braid reinforced hose was used which had a large wall thickness. The large wall thickness provided an additional factor of safety which minimized the possibility of leakage due to umbilical cable failure. The underwater unit was connected to the surface via a 22.5 m umbilical cable which provided, 120VAC, video signal, genlock (for synchronization of stereo cameras) and a control cable which allows the shutter
speed, aperture, AGC (Automatic Gain Control) and field/frame modes to be adjusted from the surface. The remote adjustment of these parameters greatly increased the efficiency of field capture of image data.

The remote control was designed from “scratch”. The unit connected to each umbilical control cable via a 15-pin twist on connector. The case was high impact injection molded plastic with an aluminum face. This ensured high reliability and low corrosion rates in the harsh salt water environment. The cameras power was verified on the surface by an LED (Light Emitting Diode). Thumbwheel resistors were used for the ALC, level and shutter control, while toggle switches were used for AGC control, frame/field mode and moisture detection. In order to accommodate for remote control, the Panasonic WV-BD400 was extensively modified to accept a control cable and a number of PCB (Printed Circuit Boards) were modified to accept the control features provided by the umbilical control cable.

A Panasonic AG-1960 S-VHS video cassette recorder accepted an RS-170 video signal from the underwater camera unit. The S-VHS VCR permitted approximately 400 lines of horizontal resolution thus allowing the recognition of finer detail and smaller salmonids. A Panasonic WJ-FS10 Digital Frame Switcher allowed multiple camera inputs to be recorded to one video tape.
4.2 Laboratory Acquisition

The procedure followed in the laboratory (Figure 4.3) consisted of grabbing suitable video frames from the footage obtained using an Imaging Technology OFG (Overlay Frame Grabber). The JVC BRS-822U was controlled by RS-232 to allow selection of specific frames for analysis using either LTC (Longitudinal Time Code) or VITC (Vertical Interval Time Code). The RS-170 signal out of the BRS-822U was applied to one of the inputs of the Imaging Technology OFG. The resulting image was transferred to OFG RAM and a pre-processing and/or edge detection was applied.

4.2.1 Hardware

Computing hardware consisted of an Intel™ 486DX/33 MHz CPU (Central Processing Unit) with 8 MB RAM (Random Access Memory). Direct programming of DMA (Direct Memory Access) hardware allowed for fast memory management.

The images were played back using a JVC BRS-822U S-VHS editing VCR. This VCR has an RS-232 interface for control, an infinite window TBC (Time Base Corrector), and SMPTE (Society of Motion Picture and Television Engineers) VITC and LTC. LTC allowed time code to be striped when the tape was returned from the field. Images could be selected and noted to almost frame accuracy (+/- 1 frame) using LTC. The RS-232 interface allowed software to be written which controlled the transport of tape and the positioning
of frames. For instance, an index point could be noted and using a search function in the BRS-822U, a specific frame could be re-seeked at a later date. This was not possible with the Panasonic AG-1960 since no LTC or RS-232 control capabilities were available with this unit.

Images were captured with an Imaging Technology OGF (Figure 4.4). This card has four video inputs, an input LUT (Look Up Table), 16-triplet output LUTs, and 1024x512 pixel frame memory. The input LUT allowed the incoming pixel data to be transformed before they are stored in frame RAM. The output LUTs allow image data transformations to be viewed without altering frame RAM. The OGF is unique since 12 bits are available for manipulation. The lower eight bits are used for image data (providing $2^8$ or 256 discrete gray-level values), and the upper 4 bits are used for graphic overlay. When dynamic LUT mode is enabled, the value in the upper 4 bits (0-15 range in decimal) of each pixel location determines which of the 16 LUT transformations is selected. The underlying 8-bit pixel value determines the LUT entry. In static LUT mode, the currently selected LUT is used to transform the entire displayed image.

4.2.2 Software

The Imaging Technology Vision-Plus AT subroutine library was used as the basic building block for the manipulation of OGF internal registers and pixel block transfers. The code for this application was written in Borland C++ 3.1 with some external calls to Borland Turbo Assembler 3.0. Assembler was
utilized to optimize routines which were CPU intensive. The Imaging Technology library was re-written for 80486 protected mode and utilized a Phar Lap 286/DOS Extender. The main reason for protected mode programming was the large size of images which were transformed in RAM. Arrays of this size are not supported in real-mode programming for the 80x86 processors.

The software developed for this project consisted of a protected mode DOS (Disk Operating System) program which operated on two monitors. One monitor provided the output from the OFG card and menuing, while the second monitor provided system status updates and a visual directory of captured images which could be selected with mouse control. The software developed included a GUI (Graphical User Interface) for user control, 2-D graphics, image processing routines for edge detection, thinning, field interpolation, image re-sizing and manipulation, remapping using code developed by Naiberg [25, 40]. Software to control the JVC BRS-822U VCR was developed to allow the software selection of frames. An image directory was developed and implemented which displayed a small thumbnail of each of the images. This system allowed the visual selection of images for processing.

A set of graphics file format filters were developed. These included code for encoding and decoding .gif, .jpg and .pcx file formats. The .gif format was specified by Compuserve™. The .pcx file format was documented by Zsoft Corporation. .jpg file format is outline by JPEG (Joint Photographic Experts Group). Routines to store brightness topographical data in a number of file formats were developed. These included dbase III, .csv, .wq1, .wk1 and .txt. Dbase III and .wq1 file formats were documented by Borland, .wk1 was doc-
umented by Lotus Corporation, .csv is a comma and carriage return/line feed (CR/LF) file format for use with CorelChart (Corel Corporation) and Surfer.

4.2.3 Camera Orientation and Ambient Parameters

Fish were video taped in tanks and sea cages. Video recording of 50-500 g chinook salmon, (approximately 8-20 kg/m³) were taken in 3.3 m diameter tanks at the Marine Ecology Program, West Vancouver Fisheries and Oceans Laboratories. The video camera was positioned at the bottom of the tank oriented toward the centre standpipe and tilted upward between 5-10°. A number of different types of screens were put over the tank so that the background was relatively brighter than the foreground. This procedure effectively backlit the salmonids within the tank, thereby producing salmonid silhouettes in medial or side, and anterior salmonid views.

Video recordings of salmonids in several sizes of sea cages were taken from an aquaculture research site at the Pacific Biological Station, Fisheries and Oceans Canada, Nanaimo, B.C. Salmonids within these sea cages were of various age-classes and stocking densities. Two species, chinook and Atlantic salmon (Salmo salar) were present in different cages in densities ranging from 4 to 8 kg/m³. The largest sea cage used was 7.5 m x 7.5 m x 6 m deep. Secchi disk readings have ranged from an average of 6 m in the summer to 11.5 m in the winter. Using different camera angles, salmonids were recorded in a variety of medial, ventral, and anterior views. With the cameras orientated towards the water surface, ventral views of silhouetted or backlit salmonids
PBS facilities provided controlled conditions for use of a sea cage. A number of small netpens (5 m x 5 m x 6 m) were available to test a number of camera orientations. These facilities provided some of the first depth tests (6 m) for the underwater housings. The B.C. Packers Broughton site provided footage of Atlantic salmon in situ. In addition, the first sea cage typical depth tests (25 m) were performed on the housings. Burdwood provided footage which included the use of the white tarp to increase separation of salmonids from their background. These images represent different camera orientations and lighting conditions. Camera orientations included:

- image axis perpendicular to the water surface camera towards the surface.
- image axis perpendicular to the water surface camera down from the surface.
- image axis parallel to the water surface camera towards center of sea cage.
- image axis parallel to the water surface camera towards the outside of sea cage.

Natural lighting ranged from sunlit conditions to overcast. Sunny clear days provided the greatest depth penetration of natural lighting. The use of supplemental lighting was not employed since the project was designed to be non-invasive.
were recorded.

Recordings were also made at two salmon farming sites in the Port McNeill area of Northern Vancouver Island. The sites are known as Broughton Island and Burdwood. At Burdwood, a white 20 foot x 20 foot tarp was placed behind the salmonids parallel to the netting (Figure 5.16). The tarp was an attempt to provide a consistent background upon which edge detection could be facilitated with some consistency.

4.3 Determination of a Criteria for Comparative Assessment

A criteria was established to compare the efficiency of the various edge operations. The criteria were qualitative and consisted of comparing the original image with an overlaid edge map. The edge map provided an indication of the edges seen by the human eye and those detected by the edge detection algorithm. This provided a method which could be compared amongst numerous operations performed on the same image. The best edge operators should be able to handle a significant amount of noise and still detect edges consistently. The images which were tested were always noisy and therefore constantly challenging even the best edge operators. The ability to separate salmonids under a wide variety of lighting conditions is crucial to the final determination of a suitable edge operator. A number of operators were employed during this study, all having merit in the final determination of the
"best" edge detection. In cases where poor edge detection was encountered, spatial brightness maps were plotted to allow a visual sense of the difficulty the edge detector was encountering with the provided data.

Edge operators used in this thesis included the Canny, LoG (Laplacian of Gaussian) and Roberts. In addition, a Sobel filter was used in a comparative assessment of noise at a step edge. The Sobel filter was not used in the treatment of results since its performance is similar to the Roberts filter. A number of pre-processing steps were used to smooth noisy video images and de-interlace video images. Pre-processing included Anisotropic Diffusion and Field Interpolation.

4.4 Previous Edge Detection Findings

During work for an undergraduate thesis [33], selected algorithms [32, 1, 16, 36] were applied with marginal success. The main difficulty was the level of noise and lack of contrast in the captured images [34]. Compass Gradient Masks [32] were used which indicated a slope direction of maximum response. These masks were 3x3 and tended to detect a large number of "false" edges due to varying background contrast, countershading, and noise which was added to the digitized image as a result of VHS signal decoding and a host of other maladies which were described in section 5.2. Another difficulty with this method was the number of calculations required per pixel. At the time, the imaging station used had a 386SX/16 MHz processor. Typical
processing times were 30 minutes for a 512x512/8-bit image. The implementa-
tion of these eight compass masks appears to have been excessive since an algorithm employing only two of the eight masks has been outlined by Nadler 1990 [24]. A number of contrast enhancement techniques were em-
ployed to increase the contrast in local areas of the images captured. One method by Beghdadi and Le Negrate [1992] [1], involved contrast enhance-
ment based on local edge detection. This method tended to increase local contrast, but a large computational penalty was evident. A typical operation on a 512x512/8-bit image took approximately 60 minutes to complete on a 386SX/16 Mhz processor. Another disadvantage of these algorithms is that as the sampled signal increased in contrast so did the noise in the image. An image typically ended up with a large number of artifacts in the image due to this process. An edge detector which provides one-pixel wide edge detection [36] was also employed. This operator first took a raw edge detection from either a Sobel [6] or Roberts 1965 [31] edge detection and then performed a thinning operation on the broad peaks of the edge detection. The thinning operation checked either side of a central pixel on a line-by-line basis and set the pixels on either side to zero if they were less than the central pixel. This operation was performed in both the x and y directions. The problem with this approach was the detected images were sensitive to noise and therefore tended to detect a large number of false edges. The conclusions from this work indicated that the only feasible method should include some form of image smoothing. This technique was first realized in the application of the LoG operator and subsequently with the Canny operator.
Figure 4.1: Location of data collection sites
Figure 4.2: Field Acquisition of Video Footage
Figure 4.3: Laboratory Digital Capture of Video Footage
Figure 4.4: Imaging Technology OFG Block Diagram
Figure 4.5: The constructed and modified field apparatus: A. Underwater Housings. B. Underwater Cameras in interior mount. C. Umbilical Cables with End Caps. D. Remote Control Unit.
Chapter 5

Results and Discussion

5.1 Test Images

Nine test images (Figure 5.1) were selected which represented different camera orientations (Table 5.1 and 5.2), lighting, and background setup. This provided images with the camera image axis oriented parallel (towards inside and outside of the sea cage) and perpendicular (towards the surface and bottom of the sea cage) with respect to the water surface.

5.2 Results of Comparitive Assessment Criteria

5.2.1 Noise and the Sampled Signal

Noise is a result of non-optimal conditions for video capture and the result of real world phenomenon. These include the length of umbilical cable, the video cables proximity to noisy sources at the extender boxes, corrosion by cyclic salts due a salt water environment, noisy power supply sources, noise
<table>
<thead>
<tr>
<th>Image Code</th>
<th>Species</th>
<th>Strain</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1010644</td>
<td>Chinook</td>
<td>NA</td>
</tr>
<tr>
<td>A1402441</td>
<td>Chinook</td>
<td>NA</td>
</tr>
<tr>
<td>A1900611</td>
<td>Atlantic</td>
<td>Horax</td>
</tr>
<tr>
<td>A2001944</td>
<td>Atlantic</td>
<td>Horax</td>
</tr>
<tr>
<td>A2002357</td>
<td>Atlantic</td>
<td>Horax</td>
</tr>
<tr>
<td>A3301118</td>
<td>Atlantic</td>
<td>Mowi</td>
</tr>
<tr>
<td>A3302725</td>
<td>Atlantic</td>
<td>McConnell</td>
</tr>
<tr>
<td>A3423291</td>
<td>Atlantic</td>
<td>Mowi</td>
</tr>
<tr>
<td>A3700125</td>
<td>Atlantic</td>
<td>McConnell</td>
</tr>
</tbody>
</table>

Table 5.1: Species and Strain in Test Images

<table>
<thead>
<tr>
<th>Image Code</th>
<th>Location</th>
<th>Lighting</th>
<th>Camera L</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1010644</td>
<td>PBS DFO Nanaimo</td>
<td>Sunny</td>
<td>Towards surface</td>
</tr>
<tr>
<td>A1402441</td>
<td>PBS DFO Nanaimo</td>
<td>Sunny</td>
<td>Towards center</td>
</tr>
<tr>
<td>A1900611</td>
<td>BCP Broughton</td>
<td>Overcast</td>
<td>Towards center</td>
</tr>
<tr>
<td>A2001944</td>
<td>BCP Broughton</td>
<td>Overcast</td>
<td>Towards center</td>
</tr>
<tr>
<td>A2002357</td>
<td>BCP Broughton</td>
<td>Overcast</td>
<td>Towards center</td>
</tr>
<tr>
<td>A3301118</td>
<td>BCP Burdwood</td>
<td>Sunny</td>
<td>Tarp towards side</td>
</tr>
<tr>
<td>A3302725</td>
<td>BCP Burdwood</td>
<td>Sunny</td>
<td>Tarp towards side</td>
</tr>
<tr>
<td>A3423291</td>
<td>BCP Burdwood</td>
<td>Sunny</td>
<td>Tarp towards side</td>
</tr>
<tr>
<td>A3700125</td>
<td>BCP Burdwood</td>
<td>Sunny</td>
<td>Toward side</td>
</tr>
</tbody>
</table>

Table 5.2: Conditions and Related Data for Images
added during recording of the video tape, noise added during the decoding of the S-VHS signal, noise introduced in the cable between the output of the S-VHS VCR and the image processing board and noise introduced during digitization of the sampled image.

The umbilical cables were typically 20 metres in length. Belden 9259 is an RG59/U 75Ω cable. This cable was used for all video and genlock signals. Expected losses for this length of cable is 3.0-3.5 dBs. This amount seems small when looking at the dynamic range of the camera (60 dB), but becomes critical at the typically low illumination seen in the sea cages.

The power supply available on the salmon farm site is typically of generated power. This power source is very noisy and a number of safeguards had to be employed to prevent destruction of video and computer equipment. Nevertheless, this power supply does not provide a rock solid AC voltage and line frequency. This tends to "wreak havoc" with video equipment which relies on a stable power source.

Noise added during the recording of video tape is unavoidable. Electronics can minimize added noise, but cannot eliminate it. The main difficulty with recording of VHS and S-VHS is in the recording method. VHS uses a technique of encoding the analog signal on the video tape. Both the chroma and luminance are mixed together and then decoded again during playback. This method alone tended to add considerable noise to the decoded signal. S-VHS was developed to decrease the amount of noise due to decoding of the signal.
The chroma and luminance are combined during recording and separated during playback. This tends to produce an image with much higher quality. Unfortunately, the RS-170 signal still suffers during the encoding/decoding process. Problems with noise added during cabling can be minimized by insuring that the length of cable is as short as possible.

There are a number of disadvantages of the RS-170 composite video signal [2]. Consecutive lines are transmitted in different fields (interlacing). Adjacent rows are imaged 1/60 second apart. Long term stabilities of horizontal synchronization and level detection are required. No scale in x is explicitly transmitted, this leads to an uncertainty in the origin and scale. The synchronization signals, which themselves carry a certain amount of imprecision, must be separated from video information at the receiver, thereby introducing additional uncertainties from sensing elements. With the state of technology, at the time of equipment purchases for the NIFE project, an RS-170 system was the most cost effective. Digital cameras were very expensive and few imaging boards supported their output.

A number of problems occur during the digitization of video using an image processing board. A frame grabber using true PLL (Phase Locked Loop) synchronization, such as the OFG frame grabber from Imaging Technology use a fixed number of clock cycles per line (in our case 640 for the OFG with 512 Pels, leaving 128 cycles for the horizontal blanking interval. PLL circuits are not perfect and there is always some line jitter (0.333-0.500 Pels for an OFG640). Most of the line jitter occurs after the vertical blanking interval,
where line sync is usually lost due to either missing or falsely interpreted serrated horizontal sync pulses [19].

5.2.2 Nominal Noise Characteristics in the Sampled Signal

Figure 5.2 shows the response of four edge operators: Roberts, Sobel, LoG, and Canny. By creating a step edge with a height of 15 (8-bit image) and then applying 3% additive random noise (ARN), all four operators correctly detected the location of the step edge. By increasing the ARN to 15%, the performance of the operators at the step edge is greatly diminished. The Roberts and Sobel detected a great deal of background noise as valid edges and had a difficult time distinguishing the step edge. The LoG filter tended to detect the step edge but the ARN, tended to produce a large number of discontinuities in the detected edge segment. The Canny operator with the addition of non-maximum suppression tended to accurately follow the contours of the step edge.

5.2.3 Nominal Edge Detection Characteristics

The criteria developed for comparing the success of the edge operators employed was applied to image A1402441. This image was chosen since it typifies some of the characteristics and difficulties of separating salmonids in digital images. The difficulties include noise in the sampled image and lack of contrast of the salmonid relative to its background due to countershading.
The effect of countershading is evident at every area investigated. The tail, head, ventral, and dorsal regions all exhibited the same difficulties with separation due to countershading. In addition, particulates in the water column tended to decrease the contrast of the salmonid. The decrease in contrast was due to less light being available to illuminate the salmonid. The image A1402441 was acquired on June 6, 1992, a sunny day at the aquaculture study facilities at Pacific Biological Station, Department of Fisheries and Oceans, Nanaimo, B.C. The tank combined both normal and albino chinook salmon (*Oncorhynchus tshawytywa*). Figure 5.3 provides an 8-bit gray scale image with spatial brightness maps showing different features.

**Ventral Edge (A)**

The ventral edge in image A1402441 is prevalent with respect to the background. This is not always the case since salmonids in deeper schooling regimes have less contrast between their ventral edge and the ambient background luminance. An edge of this magnitude and simplicity is always detected accurately with the Canny operator when the camera image axis is oriented parallel to the water surface.

**Tip of Nose (B)**

The tip of nose shows the difficulty in resolving an accurate outline of the head region. A lack of contrast exists between the head area and the ambi-
ent background luminance. The peak indicated in the luminance map is the result of a spectral reflectance off a non-edge region. This provides a false edge detection which is difficult to resolve.

**Texture Feature Due to Spotting Above the Lateral Line (C)**

The texture feature due to spotting provides a challenging problem with edge detection. This is due to the localized spots showing up as edge information. The smoothing provided by the Gaussian portion of the Canny operator suppresses this somewhat but it is entirely dependent on the camera to salmonid distance since at close range the texture takes a larger number of pixels than the same texture at a greater distance. With more intelligence in the system using stereo imaging, an estimate of distance could be used as gauge for the selection of an appropriate $\sigma$ for the Canny operator. Little 1993 [20] suggested a method of texture exclusion using multi-resolution detection.

**Ambient Background Noise (D)**

The ambient background noise provides a clue to the level of false detections to be expected. It was determined that a $\sigma$ of 1.0 detected the ambient background noise as valid edges whereas $\sigma$'s larger than 1.0 tended to suppress the detection of these false edges.
Texture Feature and Dorsal Edge (E)

The texture pattern due to the spotting above the lateral line provides a formidable separation task. Due to biological diversity, the pattern is different on each salmonid. One possibility is the use of multi-scale edge detection to filter out the spotting. This technique was demonstrated by Little 1993 ([20]) and may provide a technique for separation. This difficulty is mostly evident in the case of chinook salmon which exhibit spotting on the dorsal region. Atlantic salmon exhibit some spotting, but very little when compared to chinook salmon.

Dorsal Side Caudal Fin Edge (F)

The caudal fin separation is dependent on the orientation of the salmonid with respect to the surface and therefore light source (sun). In this case, the ambient background brightness provides a contrast between the caudal fin and background. If the camera were oriented towards the bottom on a slight angle from horizontal, the background and caudal fin are typically of the same brightness, which provides a very difficult if not impossible task for separation.

Mid caudal fin edge (G)

The mid caudal fin provides an indication of the changing contrast between the caudal fin and background. The difficulty in separation is evident from
the lack of distinct contours in the spatial brightness map.

5.2.4 Pre-Processing Algorithms and Edge Detectors

Edge detection operators employed included the Roberts, LoG and Canny Edge Detectors. Smoothing was provided by Anisotropic Diffusion while de-interlacing was provided by field averaging. The Roberts Filter would be optimal computationally since it requires the least number of calculations per pixel. In addition, the hardware lends itself to the efficient implementation of this operator. Unfortunately it is very sensitive to noise since it is a first derivative operator. In addition, a great deal of pre and post processing is required to produce one-pixel wide edge detection.

Anisotropic Diffusion

The use of the Anisotropic Diffusion operator stemmed from a need to smooth noisy images. Figure 5.4 (A2002357) illustrates the smoothing that is provided as a result of multiple iterations of Anisotropic Diffusion. This tended to decrease the number of false detections in the ambient background without sacrificing the location and number of edges detected. Figure 5.5 (A1010644) typifies lighting conditions with the camera oriented normally to the water surface. Under these conditions, a high signal to noise ratio exists and provides an excellent edge without having to apply the anisotropic diffusion operator. Most operators will detect these edges and provide accurate edge detection. The difficulty occurs at the edges of the sea cage where con-
Contrast decreases rapidly and objects and background become indistinguishable. These areas can produce good edge detection when the Canny operator is applied. Other operators would require local contrast enhancement which would increase the computational load. The ambient background noise ranges in brightness approximately 15 (8-bit image), when compared to the edges which are typically a height of 200.

Canny

The Canny was selected as the edge operator of choice for a number of reasons. This operator was able to consistently separate objects from their backgrounds in a large number of scenarios. The edge detections were shift invariant, and of sufficient magnitude that thresholding techniques could be employed to minimize false edge detections. The use of non-maximum suppression minimized the number of false edge detections and provided connectivity on others which were in low SNR. A selection of operator width was done experimentally by trying \( \sigma \)'s from 1.0 to 5.0. A \( \sigma \) of 1.0 typically detected a large percentage of false edges in the background of the image. The most remarkable change was an increase in \( \sigma \) to 2.0. A large portion of the false edges in the background were suppressed leaving mostly valid edge detections. A \( \sigma \) of 3.0 provided more smoothing, but tended to over smooth features of the salmonid which would be used for classification. \( \sigma \) widths of 4.0 and 5.0 provided no advantage in detection, were typically very long in processing and due to an error in the selection of coefficients for the
mask produced a double edge during detection of valid edges. This operator also proved to be the fastest. This criteria is realized if a comparison is made between using Anisotropic Diffusion (AD), Roberts Filter (RF), and a classic thinning (CT) [27] algorithm and using the Canny operator (A preprocessing de-interlacing (DI) step was applied to all images). Based on a 12100 pixel sample, the combination of DI, AD, RF, and CT took a total of 31.87 seconds, while the DI and Canny algorithm using a $\sigma$ of 2.0 took only 11.22 seconds. The Canny algorithm provided one-pixel wide edge detection and a robustness to noise in the image (Table 5.3).

**Laplacian of Gaussian**

The Laplacian of Gaussian (LoG) was employed as it showed the most promise in dealing with noisy video images. The smoothing effect of the Gaussian function and the rotational invariance of the Laplacian combined to produce more consistent results. But as the images degraded in quality, so did the detection of valid edges of the salmonids. One of Canny’s conclusions in his work was that on average, the positional error of the LoG is about 60% greater than that of a directional operator of the same $\sigma$.

**Roberts**

The Roberts filter was used as a first attempt on images which tended to have a high signal to noise ratio. This filter approximates the first derivative
and as such produces a broad peak which must be thinned and is sensitive to noise in the captured image. As an alternative to the Canny operator, a series of steps were used for edge detection based on the Roberts filter [31]. The image was first smoothed with 3 passes of Anisotropic diffusion [28], it was then edge detected with the Roberts filter, a global threshold was selected for the image and finally the classic thinning algorithm [27] was employed to provide a finished edge detected image (Figure 5.6). The results tended to be very close to the detections provided by Canny's operator with the exception of performance differences due to the use of a local thresholding strategy. The local thresholding strategy tended to provide better overall performance from shadow detail to areas of extreme highlight where as the global thresholding scheme tended to only show detection which had a minimum global strength. The preceding scheme would not be recommended since the computational penalty is large (Table 5.3).

The nine images selected for analysis are summarized by contrast at four landmarks on the salmon including the anterior, posterior, dorsal, and ventral views (Figure 5.4).

5.2.5 Application of Pre-Processing Algorithms and Edge Detectors on Test Images

A1010644

The image displayed was taken on September 16, 1991 at the aquaculture study facility, Pacific Biological Station, Department of Fisheries and Oceans,
<table>
<thead>
<tr>
<th>operation</th>
<th>number of pixels</th>
<th>execution time (s)</th>
<th>µs/pixel</th>
</tr>
</thead>
<tbody>
<tr>
<td>De-interlace</td>
<td>12100</td>
<td>0.70</td>
<td>57.9</td>
</tr>
<tr>
<td>Roberts</td>
<td>12100</td>
<td>0.44</td>
<td>36</td>
</tr>
<tr>
<td>Classic Thin</td>
<td>12100</td>
<td>28.75</td>
<td>2380</td>
</tr>
<tr>
<td>Canny σ 1.0</td>
<td>12100</td>
<td>7.96</td>
<td>658</td>
</tr>
<tr>
<td>Canny σ 2.0</td>
<td>12100</td>
<td>10.52</td>
<td>869</td>
</tr>
<tr>
<td>Canny σ 3.0</td>
<td>12100</td>
<td>13.77</td>
<td>1140</td>
</tr>
<tr>
<td>Canny σ 4.0</td>
<td>12100</td>
<td>14.72</td>
<td>1220</td>
</tr>
<tr>
<td>LoG σ 2.0</td>
<td>12100</td>
<td>8.06</td>
<td>666</td>
</tr>
<tr>
<td>Binary Threshold</td>
<td>12100</td>
<td>2.30</td>
<td>190</td>
</tr>
<tr>
<td>Nevatia Babu</td>
<td>12100</td>
<td>18.79</td>
<td>1550</td>
</tr>
<tr>
<td>Anisotropic Diff 3 pass</td>
<td>12100</td>
<td>1.98</td>
<td>164</td>
</tr>
</tbody>
</table>

Table 5.3: Comparison of operator execution time on a 486DX/33 CPU

<table>
<thead>
<tr>
<th>Image</th>
<th>Anterior</th>
<th>Posterior</th>
<th>Dorsal</th>
<th>Ventral</th>
<th>Cam. ℓ</th>
<th>Tarp</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1010644</td>
<td>-0.62</td>
<td>-0.46</td>
<td>+0.36</td>
<td>-0.62</td>
<td>T</td>
<td>No</td>
</tr>
<tr>
<td>A1402441</td>
<td>-0.63</td>
<td>-0.23</td>
<td>+0.12</td>
<td>-0.48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A1900611</td>
<td>-0.89</td>
<td>-0.18</td>
<td>+0.57</td>
<td>+0.60</td>
<td>T</td>
<td>No</td>
</tr>
<tr>
<td>A2001944</td>
<td>+0.43</td>
<td>+0.55</td>
<td>+0.95</td>
<td>+0.47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A2002357</td>
<td>+0.58</td>
<td>-0.21</td>
<td>+0.13</td>
<td>-0.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A3301118</td>
<td>-0.22</td>
<td>-0.21</td>
<td>-0.22</td>
<td>-0.22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A3302725</td>
<td>-0.55</td>
<td>-0.39</td>
<td>-0.54</td>
<td>-0.66</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A3423291</td>
<td>-0.44</td>
<td>-0.52</td>
<td>-0.59</td>
<td>-0.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A3700125</td>
<td>-0.31</td>
<td>-0.26</td>
<td>-0.59</td>
<td>-0.21</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5.4: Object contrast against varying backgrounds. Camera angle is with respect to the water surface.
Nanaimo, B.C. The weather conditions were sunny and clear. This provided a Secchi disk reading of 3.5 metres. The column in the left portion of the image is a column which was used to hold the cameras rigidly. The cameras are oriented with their image planes parallel to the water surface at a depth of approximately 6 metres. This orientation has been considered for counting assessment under lower stocking densities (≤ 4 kg/m³). The ALC (Automatic Level Control) averaged the ambient light falling on the image plane and maximized the number of salmonids which were detectable. A number of salmonids closer to the surface would not be visible due to attenuation in the water column and the exponential increase in brightness which the cameras dynamic range could not accommodate completely. A 3-iteration anisotropic diffusion was performed on this image. This tended to help minimize the number of iso-contours of brightness from being detected as valid edges.

Since the signal to noise ratio in an image of this camera orientation is high, edges are typically very well defined (Figure 5.7). Detections were obtained with all operators selected with minor differences. The top left and right corners of the image both proved difficult to separate. The brightness separation between the background and object was so small that valid objects could not be separated from their backgrounds. The top left and right corners of this image are typical of the results obtained when sea cage netting becomes the background. This scenario with the left and right top corners of the image being netting is unusual since sea cage size is typically too large for both corners of the net to be visible in one frame.
A1402441

A1402441 (Figure 5.8) represents footage of chinook salmon under a controlled aquaculture setting at the Pacific Biological Station, Department of Fisheries and Oceans, Nanaimo (Figure 4.1). A portion of the sample were albinos. This image is typical of the problems with separating salmonids when the camera is oriented parallel to the water surface. The mid-ventral section of the salmonid typically separates well while the mid-dorsal section separates poorly due to texture and countershading. The tail and head are typically unresolved due to poor contrast or “flashing” on the operculum area of the head. The implementation of a white tarp under these conditions would likely provide good separation of the entire salmonid due to increased contrast of the salmonid with respect to its background. Another difficulty which typically hampers repeat success in separation is the abundance of occlusion due to high stocking densities. This can typically provide a large number of overlapping salmonids. A technique which would likely diminish false detection is the use of super segments for the extraction of features. This technique is described in Naiberg's 1994 work [25].

A1900611

A1900611 (Figure 5.9) was obtained on December 3, 1992 at B.C. Packers, Broughton Island, Port McNeill, B.C. during overcast conditions (Figure 4.1). The camera was oriented with the image axis perpendicular to the water surface and the camera pointing downwards. The netting for this sea
cage has been pulled up to approximately 3 metres depth to allow salmonid sizing. This type of separation is difficult since the netting provides a texture which may be expressed as an edge detection. Selecting the appropriate $\sigma$ for the Canny operator ensured that the netting was not detected as valid edges. One advantage to this orientation was the lack of countershading which must be overcome to completely separate the salmonids. It is likely that this orientation would not be useful for biomass determination since a lack of data is evident with respect to the girth of the salmonids. This orientation could be improved by providing a white tarp at the bottom of the sea cage. This would likely provide excellent separations. The smoothing provided by the use of a 3-pass Anisotropic Diffusion tended to delete valid edges in low SNR areas of the image.

A2001944

A2001944 (Figure 5.10) was obtained at B.C. Packers, Broughton Island, Port McNeill, B.C. (Figure 4.1), typifies the edge detection result from a salmonid which is mid-frame. The dramatic decrease in contrast is evident from the dorsal to ventral features of the salmonid. Separation is evident from the background on the dorsal side of the fish from the nose to the base of the caudal tail. A feature which is typically separated, but which did not in this case is the dorsal fin (A weak detection is noted on the upper side of the dorsal fin). The main difficulties with this image occurs in the ventral region of this salmonid. The decrease in contrast provides a low contrast which
evaluates to a spotty edge detection. The location of the salmonid in the frame is evidently critical judging from the edge detection of the salmonids below the salmonid of interest which show little separation from the background. On the other hand, the same problem exists with salmonids above the salmonid of interest. The two Atlantics appearing in the upper right are strongly separated from their background indicating that if these were centered in the frame above the salmonid of interest, they would also be well separated. This camera orientation is typical of the one used for salmonid sizing. These results should be compared with A3423291 (Figure 5.14) and A3302725 (Figure 5.13). These camera orientations are similar to A2001944 (Figure 5.10) but with the addition of a white tarp. The difference in separation is remarkable. A side effect that was first noted in A1900611 (Figure 5.9) was evident in this image. The use of a 3-pass Anisotropic Diffusion tended to break connectivity at various points in low SNR regions of the image.

A2002357

A2002357 (Figure 5.11) was obtained at B.C. Packers, Broughton Island, Port McNeill, B.C. (Figure 4.1), provides an insight into the difficulty separating a typically countershaded salmon from its varying background brightness. The main reason that the separations were difficult was due to changes in spectral reflectivity on different scales of the salmonid and additionally countershading provided an effective camouflage against ambient water column lighting.
conditions. Image A shows a detection with a small $\sigma (\sigma = 1.0)$. The iso-contours of brightness in the background are detected with a large frequency. The largest problem with this $\sigma$ is the difficulty following contours of valid edges on borderline objects (borderline since their brightness does not vary a great deal from their background). This is most evident in the difference between Figure 5.11A and Figure 5.11B. Figure 5.11B uses a $\sigma$ of 2.0. This provides separation and edge contour following of borderline edge detection which were lost in the smaller $\sigma$. In particular, the faint outline of the salmon at the bottom of the frame is quite evident in figure 5.11B but is suppressed in figure 5.11A. The increased computation time for larger $\sigma$ (3.0-4.0) (Figure 5.11C) and the additional smoothing provided proved unsatisfactory for the accurate separation of salmon from their backgrounds (Table 5.3 is provided to indicate the difference in relative computational times). It is interesting to note that if this object were centered and at a slightly greater distance from the camera, countershading as detected edges would likely be suppressed and the mouth area would be more accurately described by the edge detection. Figure 5.11D illustrates the use of the LoG operator with a $\sigma$ of 2.0 and the incorporation of 3-pass Anisotropic Diffusion. The lack of non-max suppression tends to detect numerous background features as valid edge information. An interesting result is the use of the Roberts operator in conjunction with 3-passes of Anisotropic Diffusion and Pavlidis thinning (Figure 5.6). The major drawback is the huge computational load required to complete this operation (Table 5.3). Once again selection of the “best” operator leads to the Canny operator. The employment of 3-pass Anisotropic Diffusion helped to minimize the number of false edge detection due to iso-contours of bright-
ness in the background.

A3300118

A3300118 (Figure 5.12) provides an insight to the difference in detection ability using the white tarp as a background. The left side of this image uses the natural background in an attempt to separate salmon. The existence of countershading and light attenuation in the water column suppresses most detectable edges. The tarp on the right hand side of the image provides a backdrop which separates the salmon so well that it is likely that some counting algorithms could be developed assuming the stocking densities and therefore occlusion was minimal. The ambient background iso-contours of brightness are detected as a low value of edge strength. This would allow their deletion from the sample by simple thresholding. The decrease in contrast from the top of the frame to the bottom is typical of attenuation of light in the water column with increasing depth. By providing levelling in the camera housing a portion of the bottom of the frame where contrast is too low to provide adequate separation of the salmonids from their background could be determined to be out of the sampling window. In addition this same strategy could be employed to determine a region in the top of the frame which is out of the sampling window. Since the camera housing dome port to tarp distance could be estimated, a sampling view volume could be determined from known angle of views in-water for the dome port in use ($\angle = 76.12^\circ$ on horizontal). This technique would be coupled with behaviour
on schooling for strains of chinook and Atlantic salmon. This could provide a technique to determine a sample count within a sea cage.

A3302725

A3302725 (Figure 5.13) was obtained on a sunny day at B.C. Packers Burdwood site. This image represents use of a white tarp as a background to the salmon and is typical of the optimal results due to position of the camera with its image axis parallel to the water surface and the direction of the camera pointing towards the edge of the sea cage. The largest object in the frame shows the result of countershading providing false detections of valid edges. In addition, the anterior end of the salmon shows detail in the mouth area of the salmon. Objects which are further away from the camera tend to provide a smooth edge detection around this region. The salmonid at the top of the image shows the transition between this smoothed mouth region and a completely separated mouth region. The key appears to be to keep the salmon greater than a minimum threshold of distance to ensure the suppression of countershading as valid edge information and to ensure a smoothed detection in the mouth region of the salmon.

A3423291

A3423291 (Figure 5.14) was obtained on a sunny day at B.C. Packers Burdwood site using a white tarp. The salmon were typically well separated. Iso-contours of brightness were detected in the background. The majority of
these may be suppressed by thresholding the detection level for a valid edge on the output of the Canny or Canny/Anisotropic Diffusion operation. A noteworthy result is the object contrast obtained with the use of the tarp. Table 5.4 provides a guide to the sign and numeric value of the object contrast relative to its background. The use of the tarp provided a negative value for all areas of the salmonid. This indicated that the object brightness was less than the background brightness. In effect, the tarp ensures that the salmonid will be well contrasted relative to its background. This simple implementation of a tarp to aid in separation greatly reduces the complexity of the pre-processing and edge detection required to separate the salmonid from its background. To insure accurate edge detection, a $\sigma$ of 2.0 was found to be best for overall separation. Larger values of $\sigma$ tended to "over smooth" the detected image deleting object information which was important to the characterization of the salmon. A $\sigma$ of less than 2.0 tended to detect an over-abundance of background iso-contours of brightness. A fine threshold exists between a fish which is accurately detected and another whose countershading expresses "false" edges typically between the lateral line and dorsal surface of the salmon. This phenomenon was common to both chinook salmon and Atlantic salmon. The two salmon towards the bottom of the frame are typical of those which are a sufficient distance from the camera to benefit from attenuation in the water column. The attenuation tended to suppress the detection of "false" edges in the lateral line to dorsal region of the salmon. This effect is coupled with the salmon being towards the bottom half of the image. A disadvantage to be noted of the tarp method when applied at close range for salmonid sizing is due to the uneven nature of the tarps surface in
the water column; a number of tarp contours are detected as valid edges. This would not be a hindrance with the use of a feature extraction method such as that described by Stein and Medioni's 1992 [39]. This technique as applied to salmonids for salmonid sizing is outlined in Naiberg's 1994 work [25].

A3700125

A3700125 (Figure 5.15) is a mid range salmonid sizing solution with the use of the netting for a background. In this case most of the salmonids are well separated from their background providing a basis for feature extraction. For comparative measures a LoG operation with a $\sigma$ of 2.0 was also examined. It performed well with salmonids in the upper half of the video frame. It fails in the lower half where contrast decreases. The Canny operator with a $\sigma$ of 1.0 illustrates the large number of false detections which occur in the background of the image. An increase in operator width illustrates a large decrease in false detections of the background. To further eliminate false detections, a 3-pass Anisotropic Diffusion was applied which further decreased these false background detections. It is likely that this step would not be required when implementing Stein and Medioni's 1993 work [39].

5.3 Applications

The ability to separate salmonids from their background is reliant on the orientation of the image axis with respect to the water surface. When salmon
are videotaped with the image axis parallel to the water surface, the most consistent separation occurs. Typically the top and bottom quarters of the image provide poor images for separation. The center half consistently provides images which are separated from ventral to dorsal sides. An optimal distance for imaging is evident from the observed footage obtained. This distance would have to be determined empirically and is therefore out of the scope of this thesis work. Salmon which have the benefit of attenuation which is occurring between the camera CCD (charge coupled device) and the object are typically silhouetted at a threshold distance from the camera. The threshold distance could be determined empirically. Salmon which are closer than this distance typically have a number of false detections due to countershading. On the other hand, salmon which are between the threshold distance and the distance of perceptible object detection are silhouetted against the background (assuming the implementation of a white tarp) without the distraction of countershading to impede the detection of valid edges.

A camera which is oriented with the image axis perpendicular to the water surface tends to provide images with a consistent background. The rear illumination provided by the sun provides silhouetted images. The main criteria is to match the dynamic range (The Panasonic WV-BD400's dynamic range is 60dB) with the presented signal. One method of achieving this is to provide a reference object at the surface of the water (or at the minimum observed depth of salmon). The camera may then be lowered until the object is just visible as a result of light attenuation in the water column. This ensures that the targeted sample will be completely described by the dynamic range of
the sensor available.

The most reliable method of insuring that the salmon are the correct distance away from the camera is videotaping with the image axis parallel to the water surface, with the camera pointing towards the outside of the sea cage. The correct distance is defined as the distance where countershading expression is suppressed (camera to salmonid distance). A discussion of determining the correct distance may be found in the Recommendations section. The absolute distance between the dome port of the camera housing and the white background tarp may be estimated and maintained. In addition, a minimum threshold distance is evident. This occurs when the attenuation in the water column suppresses the expression of countershading as valid edge information. The mouth area of the salmon tends to be smoothed as an added benefit which ensures an accurate outline of the salmon.

Image A3301118 provides some insight to a counting technique which could be employed in a sea cage environment. This technique uses a white tarp which is placed at the edge of the sea cage and a camera which has its image axis normal to the plane formed by the tarp. By placing a sampling window over the image area to eliminate salmonids at the top of the frame which are at the top of the dynamic range of the camera and at the bottom of the frame which are at the bottom of the dynamic range of the camera, a view volume emerges which may be estimated from the camera housing dome port to tarp distance and the angle of view provided by the dome port and camera lens in use. This view volume combined with schooling behaviour could provide a
The use of a white tarp has provided a consistent method of total salmonid separation from their background. This may be confirmed by checking the value of the contrast for all salmonids which have the white tarp as their background. The negative value indicates that in all cases the background image is brighter than the salmonid. One difficulty encountered is the "flashing" which occurs due to the countershading of the salmonid. This provides a false edge detection slightly above the lateral line. This may be ignored if the feature extraction technique outlined in Naiberg's 1994 thesis work is employed[25, 40].

The recording of a noise minimized signal can be accomplished using a number of methods. The simplest "fix" is the use of a "betacam" VCR. This type of VCR records the chroma and luminance on two separate heads. The advantage is that the chroma/luminance signal does not require encoding and decoding. Unfortunately, these VCRs are not meant to be used in a controlled environment, rather than an unpredictable humid environment such as a salmon farm site. A direction to be considered for the future is the use of digital cameras and recorders. This method virtually eliminates the problem with noise. Unfortunately, digital VCRs and digital cameras are extraordinarily expensive and do not fare well in a harsh environment.

The selection of the Canny operator was made after extensive testing of
other operators. The Canny operator offered the best compromise between computational load, localization of edges, singular response to edges, and suppression of false edges which were a result of background noise. A $\sigma$ of 2.0 was selected based on an assessment of the best balance between noise suppression and edge smoothing.

This work could find use as previously mentioned in counting applications and as a preliminary step towards automating a stereo system based on edge detection of stereo pairs. Prior knowledge of the camera geometry would allow a prediction to be made of the location of the second image of the stereo pair. This would minimize the search area between stereo pairs for valid salmonid edge information. The application of Naiberg’s [25] work to this thesis work would allow the automatic extraction of valid salmonids for salmonid sizing. The system would analyze the edges detected and determine which edges represented valid anterior and posterior data.
Figure 5.1: Images tested were obtained from a variety of sites, lighting conditions, camera orientations, and salmonid species and strain.
Figure 5.2: Edge detection with various operators with an artificially derived step edge. A. 3% ARN B. 15% ARN C. spatial brightness map at step edge with 3% ARN D. spatial brightness map at step edge with 15% ARN
Figure 5.3: Spatial brightness maps illustrating nominal edge detection maladies
Figure 5.4: Spatial brightness maps illustrating anisotropic diffusion performance for the detection of edges under average lighting condition with the camera parallel to the water surface.
Figure 5.5: Spatial brightness maps illustrating a camera orientation which typically does not require anisotropic diffusion.
Figure 5.6: A. LoG $\sigma = 2.0$ B. Anisotropic Diffusion 3-pass LoG $\sigma = 2.0$ C. Anisotropic Diffusion 3-pass Roberts and classic thinning
Figure 5.7: A1010644 A. Canny $\sigma = 2.0$ B. Canny $\sigma = 2.0$ with Anisotropic Diffusion 3 passes
Figure 5.8: A1402441 A. Canny $\sigma = 2.0$ B. Canny $\sigma = 2.0$ with Anisotropic Diffusion 3 passes
Figure 5.9: A1900611 A. Canny $\sigma = 2.0$ B. Canny $\sigma = 2.0$ with Anisotropic Diffusion 3 passes
Figure 5.10: A2001944 A. Canny $\sigma = 2.0$ B. Canny $\sigma = 2.0$ with Anisotropic Diffusion 3 passes
Figure 5.11: A2002357 A. Canny $\sigma = 2.0$ B. Canny $\sigma = 2.0$ C. Canny $\sigma = 2.0$
D. LoG $\sigma = 2.0$ with Anisotropic Diffusion 3 passes
Figure 5.12: A3300118 A. Canny $\sigma = 2.0$
Figure 5.13: A3302725 A. Canny $\sigma = 2.0$ B. Canny $\sigma = 2.0$ with Anisotropic Diffusion 3 passes
Figure 5.14: A3423291 A. Canny $\sigma = 2.0$ B. Canny $\sigma = 2.0$ with Anisotropic Diffusion 3 passes
Figure 5.15: A3700125 A. Canny $\sigma = 2.0$ B. Canny $\sigma = 2.0$ with Anisotropic Diffusion 3 passes
Figure 5.16: A novel automatic sea cage counting method
Chapter 6

Conclusions

This thesis addressed five objectives which aid in the separation of salmonids from their backgrounds. This will allow feature extraction to be carried out and could be applied to an automatic, computerized counting and biomass system.

Salmonids were video taped under a wide variety of conditions and camera orientations (Figure 5.1). The camera image axis was oriented parallel with respect to the water surface. This orientation allowed for video taping with the camera pointing inwards towards the center of the sea cage and pointing outwards towards the edge of the net pen. The camera pointing outwards allowed salmonids to be placed against sea cage netting as a background and also against a white tarp. The camera image axis perpendicular with respect to the water surface pointing downwards allowed separation of salmonids from an anterior view. This may be useful for counting, but is not very useful for salmonid sizing since little is revealed about the girth of the salmonid. The camera image axis perpendicular with respect to the water surface and
pointing upwards from the bottom of the sea cage allowed simplified edge detection under high SNR. This orientation may be useful for counting, but since little is revealed about the girth of the salmonid, this orientation may not reveal enough data to allow for salmonid sizing.

A qualitative analysis was provided which discussed imaging apparatus (video cameras, videotape recorders and imaging cards) and the physical factors which affect image quality. Noise is added to images at every stage. In the field, noise is added at the CCD (Charge Coupled Device), video camera electronics, umbilical cabling, video cassette recorder, and during S-VHS tape encoding (encoding chrominance and luminance). In the lab, noise is added during decoding of the S-VHS signal (decoding chrominance and luminance), in the VCR electronics, in the cabling from the VCR to the imaging board, and before the image is digitized on the imaging board. A number of physical factors affect image quality including turbidity (organic and non-organic), attenuation and scattering in the water column and ambient lighting conditions (sunny or overcast) which are related to the refraction and reflection of light at the sea surface.

A number of edge detectors and pre-processing algorithms were employed. The Roberts filter when combined with Anisotropic Diffusion and Thinning provided good separation under a variety of conditions (Figure 5.6). Unfortunately, the processing time involved in applying these algorithms proved to be a hindrance. The LoG was applied to a number of images with marginal success (Figure 5.6). Under higher SNR, the LoG successfully separated
salmonids from their background, but as the SNR decreased, the LoG operator detected a larger portion of false edges. The Canny operator with a $\sigma$ of 2.0 provided consistent separation of salmonids under a wide variety of lighting conditions and camera orientations. Smaller values of $\sigma$ such as 1.0 tended to detect a large number of false edges which were often iso-contours of brightness in the background of images (Figure 5.11). Larger values of $\sigma$ such as 3.0 tended to excessively smooth the edges detected (Figure 5.11). This would tend to hide specific attributes which would aid in identification. Anisotropic diffusion was useful when images had low SNR (Figure 5.4). Typically, three passes of Anisotropic Diffusion were applied before edge detection was performed. This tended to decrease the number of false detections, in particular this tended to suppress the iso-contours of brightness which were detected in the background.

Table 5.4 provided an insight to the standardization of salmonid separation which could be obtained by placing a white tarp between the salmonids and the sea cage netting. The white tarp ensures that the object contrast is negative for all attributes of the salmonid. A3301118, A3302725, and A3423291 illustrate the benefits of implementing a white tarp. In particular a simplified and contiguous edge is detected around all salmonids in the video frame.

A basis for comparative assessment was discussed which provided a qualitative approach to determine the success or failure of a given edge detection. The edge detection was laid over top of the original image and a comparison was made between the edges in the image and the edges detected by
the operator. In particular, a number of key areas on the salmonid were examined. These included edges detected on the anterior, posterior, dorsal and ventral regions. By comparing a number of operators by this method, a comparison could determine how well or poorly the different edge detectors had performed in these key areas.

A technique was demonstrated which allows counting under low (<4 kg/m$^3$) stocking density and therefore lower incidence of occlusion (Figure 5.16). By combining the results of this thesis with the work of Naiberg [25], a step towards the automation of salmonid sizing and counting can be considered.

6.1 Recommendations

To facilitate the future success of edge detection of salmonids, a number of treatments should be considered. The threshold distance which attenuates countershading expression should be derived empirically. Observations of numerous images indicated that countershading expression was suppressed due to attenuation and scattering in the water column. The presence of particulates in the water, the ambient lighting conditions (overcast or sunny) and the camera orientation and camera to subject distance are all factors which are likely to influence the suppression of countershading expression. The ambient lighting conditions could be examined quantitatively using an underwater quantum photo meter. These values would be recorded at the time video tape was obtained. The camera to subject distance could be estimated by
attaching a graduated rod to the camera platform which extended parallel to the image axis. Salmonids distance could be estimated with this and using stereo methods. Images which demonstrate countershading expression suppression could be manually selected from the video tape and data of the subject to camera distance, ambient lighting conditions, secchi disk reading (horizontal and vertical) would be used to determine an empirical relationship of the form:

\[ D_s = f(S_h, S_v, L_{PPFD}, G_s) \]  \hspace{1cm} (6.1)

where:

- \( D_s \) — distance that countershading expression is suppressed (m)
- \( S_h \) — horizontal Secchi disk reading (m)
- \( S_v \) — vertical Secchi disk reading (m)
- \( L_{PPFD} \) — Photosynthetic Photon Flux Density (\( \mu E s^{-1} m^{-2} \))
- \( G_s \) — Genus and Species of salmonid of interest

There are many ways to find connectivity (none are perfect for all images). One way to arrive at a good technique is to learn it by mixing elements of a variety of methods. Genetic Algorithms (GA), as an aid to connectivity and feature extraction could provide additional connectivity which is currently not detected. GAs are finding use in a large number of applications
which previously employed specific algorithm solutions. A GA is a class of stochastic search algorithms. They were developed by John Holland at the University of Michigan in the mid-70's [9]. GAs behave much like biological genetics. GAs encode information into strings, just as living organisms encode characteristics into strands of DNA (Deoxyribonucleic Acid). A string in GA is analogous to a chromosome in biology. A population of strings competes and those strings that are the fittest procreate, the rest eventually die off, childless. As with biological parents, two strings combine and contribute part of their characteristics to create their offspring, the new individual. The new string joins the population and the fight to produce the next generations.

One of the GA's attributes is its ability to evaluate many possible solutions simultaneously. This ability, known as implicit parallelism is the key to GA's power [6]. Implicit parallelism results from simultaneous evaluation of the numerous building blocks that comprise the string. Each string may contain millions of building blocks, and the GA assesses all of them simultaneously each time the fitness of the string is calculated. In effect, the algorithm selects for patterns inside the string that exhibit high worth, and passes these building blocks on the next generation. This selection process allows GAs to perform well where traditional algorithms do not, such as problems with huge search spaces (such as images).

GAs have the quality of robustness. While special-case algorithms may find more optimal solutions to specific problems, GAs perform well over a large number of problem categories. This robustness results because GAs usually
apply their search against a large number of points rather than a single point. Another contribution to their robustness is that GAs use the string's fitness to direct the search. This enables them to operate well on search spaces that have gaps, jumps, or noise.

The images captured and separated in this thesis were based on single frame processing. The implementation of temporal processing would allow salmonids to be separated under higher stocking densities since temporal processing would indicate image changes which were due to changes in occlusion of salmonids. In addition, temporal processing would allow tracking of salmonids through the view volume of the cameras. This would allow multiple salmonid sizing estimations for averaging and could be used for counting salmonids which pass through the view volume of the cameras.

Another problem is the recorded quality of images captured in the field. The use of higher quality tape recording methods such as separation of the chrominance and luminance during recording and surveillance cameras with orders of magnitude greater sensitivity would provide higher SNR. This would ensure a larger percentage of valid edges are expressed in an edge detection. Perhaps the best method would be the use of digital cameras and imaging boards. These would eliminate most sources of noise when combined with a white tarp as a background. As a result, highly reliable edge detection would likely evolve by implementing these methods.
Bibliography


Appendix A

A.1 Dome Port Lens Design

A.1.1 Methods of Calculation Simplification

A hemispherical viewport was selected since it forms a half of a spherical pressure vessel. When the supporting edge is properly seated there is a minimum of distortion of the dome surfaces, therefore it becomes a very practical optical element in a pressure environment. Optical rays which are confined to small enough angles that the sine and tangent can be considered equal are said to be in the paraxial region, an infinitely narrow bundle on the optical axis. The first order formulae which are employed produce image relationships for the perfect optical system. Obviously, there is always a departure from the first-order layout. If the object's space refractive index is 4/3 (a reasonable value for salt water), then the second focal length \( f' \) is negative and equal to three time the dome radius. A virtual image appears to be located at a distance of \( 3r \) toward the object from the dome apex (principal image distance \( d' \)). A number of conventions are established to allow paraxial analysis of the dome port:
• Light travels from left to right.

• All radii are positive when the center of curvature is to the right of the surface and negative when it is the left.

• All object distances are positive when they are measured to the left of a reference point and negative when they are measured to the left of a reference point.

• All image distances are positive when they are measured to the right of a reference point and negative when they are measured to the left of a reference point.

• The first and second focal lengths are positive for a converging lens system and negative for a diverging system.

• All object and image dimensions are positive when measured upward from the optical axis and negative when measured downward.

The lens system can be described by the location of six points along the optical axis:

• $F_1$ and $F_2$—First and second focal points.

• $P_1$ and $P_2$—First and second principal points.

• $N_1$ and $N_2$—First and second nodal points.
The general formula for a thick lens in first-order optics can be applied to the dome port:

\[
\frac{n}{f} = \frac{n'}{f'} = \frac{n' - n}{r_1} + \frac{n'' - n'}{r_2} - \left[ \frac{n' - n}{r_1} \times \frac{n'' - n'}{r_2} \times \frac{r_1 - r_2}{n'} \right] \quad (A.1)
\]

where:

- \( f \) — First focal length
- \( f' \) — Second focal length
- \( n \) — Refractive index for object space
- \( n' \) — Refractive index for the dome port
- \( n'' \) — Refractive index for image space
- \( r_1 \) — Dome outer radius
- \( r_2 \) — Dome inner radius

First-order object and image nodal distances \( (D \text{ and } d) \) are specified relative to the nodal points.

\[
d = \frac{nDf}{nD - f} = \frac{Ndf'}{D - f'} \quad (A.2)
\]
\[ D = \frac{df}{n(d-f)} = \frac{df'}{d-nf'} \]  \hspace{1cm} (A.3)

The sign convention produces positive number of objects that are a distance \( D \) to the left of the dome (Figure A.1). If the images are observed from the center of the dome, then there is no distortion and only a minimum of aberration, since all rays from the object are at or near normal to the dome’s refracting surface.

\[ f_1 = \frac{n_2}{n_1} f \]  \hspace{1cm} (A.4)

Substitution of the values pertinent to NIFE requirements into the thick lens equation yields the following results:

- \( n=1.355 \) — Refraction index for object space (salt water)
- \( n'=1.491 \) — Refraction index for the dome port (acrylic)
- \( n''=1.000 \) — Refraction index for image space (air)
- \( r_1=50.8 \) mm — Radius of outer dome port
- \( r_2=38.1 \) mm — Radius of inner dome port
- \( f=-136.65 \) mm — First focal length
- \( f'=-91.65 \) mm — Second focal length
The effective focal length of the dome port and 4.8 mm auto-iris Cosmicar lens was determined by:

\[ M_L = \frac{f_L}{-3.8r_c + f_L} \]  \hspace{1cm} (A.5)

where:

- \( M_L \)—Magnification
- \( f_L \)—Focal length of the camera lens
- \( r_c \)—Inner radius of the dome port

substituting:

\[ M_L = \frac{4.8 \text{ mm}}{-3.8(38.1 \text{ mm}) + 4.8 \text{ mm}} = -3.429^{-2} \]  \hspace{1cm} (A.6)

The effective focal length may then be calculated:

\[ f = \frac{1}{n}(1 - M_L)f_L \]  \hspace{1cm} (A.7)

where:

- \( n \)—Refractive index of object space (salt water)
• $M_L$—Magnification

• $f_L$—Focal length of the camera lens

substituting:

$$f = \frac{1}{1.355} (1 - (-3.429^{-2})) 4.8 = 3.7\,mm \quad (A.8)$$

The effective focal length of a 4.8 mm lens with the specified dome port is 3.7 mm.
Figure A.1: Relationships of dome paraxial object and image