Multiple Prediction Combination and Confidence Measures for Marine Object Detection

by

Michael Horton, BComp (Hons)

Submitted in fulfilment of the requirements for the Degree of Doctor of Philosophy

University of Tasmania
February 2009
Declaration of Originality

I, Michael Horton, do hereby declare that this thesis contains no material that has been accepted for the award of any other degree or diploma in any tertiary institution, except by way of background information and duly acknowledged in the thesis. To the best of my knowledge and belief it contains no material previously published by another person, except where due reference is made in the text of the thesis, nor does the thesis contain any material that infringes copyright.

Signed:  

Date:  
Authority of Access

This thesis may be made available for loan and limited copying in accordance with the Copyright Act 1968.

Signed:..............................

Date:.................................
Statement of Co-authorship

The publications of the work undertaken as part of this thesis are the following:


Mr. Michael Horton (60%) is the primary author. He conducted the research and prepared the material for publication.

Dr. Mike Cameron-Jones (30%) of the School of Computing and Information Systems, University of Tasmania, suggested the ‘both balanced subsets’ algorithm and provided general guidance and editing advice as supervisor.

Dr. Raymond Williams (10%) of the School of Computing and Information Systems, University of Tasmania, provided general guidance and editing advice as supervisor.


Mr. Michael Horton (80%) is the primary author. He conducted the research and prepared the material for publication.

Dr. Mike Cameron-Jones (10%) and Dr. Raymond Williams (10%) of the School of Computing and Information Systems, University of Tasmania, both provided general guidance and editing advice as supervisors.
We the undersigned agree with the above stated proportion of work undertaken for each of the above published or submitted manuscripts contributing to this thesis.

Signed:..........................  Signed:..........................
Date:..............................  Date:..............................
Dr Mike Cameron-Jones   Dr Raymond Williams
Supervisor               Supervisor
School of Computing and  School of Computing and
Information Systems       Information Systems
University of Tasmania   University of Tasmania
Abstract

This thesis considers two problems in classification – a field within artificial intelligence. One is the general problem of classifier learning, for which a meta-classification technique called ‘virtual attribute subsetting’ is developed and tested. The other is object detection, with emphasis on marine creature detection, using the ‘Haar Classifier Cascade’ method (Viola & Jones, 2001b).

Haar Classifier Cascades are built from a feature set of simple rectangular patterns of relative light and dark. Adaptive boosting selects those features that best tell the difference between objects and non-objects. In this thesis, a new cascade confidence measure is proposed, equivalent to the boosting ‘margin’; it uses information about how well the cascade features match the image region being classified. Tests on the common application of face detection show that this confidence measure improves detection accuracy. Virtual attribute subsetting is also used to modify the cascade; it further improves accuracy at the expense of classification time.

In addition, Haar Classifier Cascades are trained to detect two types of marine animal (fish and seahorses). This requires object detection across a wide range of orientations, so approaches using both image and cascade rotation are compared. Results show that image rotation is more accurate than cascade rotation, and that cascades trained to detect objects over a range of angles should have their training images randomly perturbed over a similar (but not always equal) range. Confidence-based detections are also made and show themselves to be more accurate than binary detection. The confidence-based results sum the confidences from similar detections and show that confidence measurements from multiple Haar Classifier Cascades may be combined effectively.

Seahorse detection poses an additional problem: seahorses are too flexible to be found by single cascades in any orientation. To solve this, separate seahorse head and body detectors are trained and their detections matched to create whole seahorse detections. Both designed and learnt matching cost formulae are created and two matching algorithms are implemented to link together head and body detections given a cost measurement. The best of the resulting whole seahorse detectors is more accurate
than either of the component part detectors.

The confidence measurement and virtual attribute subsetting algorithms make no use of domain knowledge, so should improve the accuracy of most other Haar Classifier Cascades. They also have the unusual property of being applicable to already learnt classifiers.
Acknowledgements

The salmon images used in this research were provided by AQ1 Systems Pty. Ltd. The seahorse images are from a film taken by Rob Fearn with permission from the School of Aquaculture, University of Tasmania. This research was also supported by a Tasmanian Graduate Research Scholarship.

This thesis would not exist without my supervisors, Mike Cameron-Jones and Raymond Williams, and their willingness to meet every week to ask why I hadn’t implemented my experiments yet. When finally ready those experiments were run in the School of Computing labs maintained by Christian McGee and Andrew Spilling, and on the cluster assembled by Brad Goldsmith and Alistair Atkinson.

I was also fortunate to have two parents, Brian and Jan Horton, who had written theses of their own, and a cat, Perdita Nitt, who could be relied upon not to ask awkward questions about my progress. Finally, I must thank Margaret Hoban of the Launceston Youth & Community Orchestra and Matthews Tyson of the St Cecilia Chamber Orchestra. While their rehearsals were hardly an anchor on the reality outside my office, they did provide a counterweight.
# Contents

Abstract vi

Acknowledgements viii

Contents ix

List of Tables xiv

List of Figures xvi

1 Introduction 1

1.1 Classifier learning ......................................................... 1

1.2 Object detection ............................................................. 2

1.2.1 Detecting rotated objects ............................................... 2

1.2.2 Detection with confidence measures ................................. 2

1.2.3 Detected segment matching ........................................... 3

1.3 Outline .............................................................................. 3

2 Literature review 5

2.1 Machine learning .............................................................. 5

2.2 Introduction to classifier learning ....................................... 5

2.3 Multiple classifier learning ............................................... 8

2.3.1 Bagging ........................................................................ 8

2.3.2 Attribute subsetting ...................................................... 10

2.3.3 Stacking ....................................................................... 10

2.3.4 Boosting ...................................................................... 11

2.3.4.1 Boosting the margin ................................................ 12

2.3.5 Implementation ............................................................ 12

2.4 Introduction to computer vision ........................................ 12

2.4.1 Object detection ............................................................ 13

2.4.1.1 Marine creature detection ..................................... 13

2.5 Image features ............................................................... 14
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.6</td>
<td>Haar Classifier Cascades</td>
<td>14</td>
</tr>
<tr>
<td>2.6.1</td>
<td>Rectangle calculation</td>
<td>15</td>
</tr>
<tr>
<td>2.6.2</td>
<td>Features</td>
<td>18</td>
</tr>
<tr>
<td>2.6.3</td>
<td>Stage training samples</td>
<td>20</td>
</tr>
<tr>
<td>2.6.4</td>
<td>Cascade stage boosting</td>
<td>20</td>
</tr>
<tr>
<td>2.6.5</td>
<td>Region testing</td>
<td>21</td>
</tr>
<tr>
<td>2.6.6</td>
<td>Merging neighbouring detections</td>
<td>23</td>
</tr>
<tr>
<td>2.6.7</td>
<td>Rotated objects</td>
<td>23</td>
</tr>
<tr>
<td>2.6.8</td>
<td>Implementation</td>
<td>24</td>
</tr>
<tr>
<td>2.6.9</td>
<td>Extensions</td>
<td>25</td>
</tr>
<tr>
<td>2.7</td>
<td>Classifier evaluation</td>
<td>25</td>
</tr>
<tr>
<td>2.7.1</td>
<td>Cross-validation</td>
<td>25</td>
</tr>
<tr>
<td>2.7.2</td>
<td>Matching detected objects to annotated objects</td>
<td>26</td>
</tr>
<tr>
<td>2.7.3</td>
<td>ROC curves</td>
<td>26</td>
</tr>
<tr>
<td>2.7.3.1</td>
<td>Combining ROC curves</td>
<td>28</td>
</tr>
<tr>
<td>2.8</td>
<td>Conclusions</td>
<td>29</td>
</tr>
<tr>
<td>3</td>
<td>Virtual attribute subsetting</td>
<td>30</td>
</tr>
<tr>
<td>3.1</td>
<td>Unknown attributes in classification</td>
<td>30</td>
</tr>
<tr>
<td>3.2</td>
<td>The virtual attribute subsetting algorithm</td>
<td>31</td>
</tr>
<tr>
<td>3.2.1</td>
<td>Subset choice</td>
<td>32</td>
</tr>
<tr>
<td>3.2.1.1</td>
<td>Random subsets</td>
<td>33</td>
</tr>
<tr>
<td>3.2.1.2</td>
<td>Classifier balanced subsets</td>
<td>33</td>
</tr>
<tr>
<td>3.2.1.3</td>
<td>Attribute balanced subsets</td>
<td>34</td>
</tr>
<tr>
<td>3.2.1.4</td>
<td>Both balanced subsets</td>
<td>34</td>
</tr>
<tr>
<td>3.2.2</td>
<td>Base classifiers</td>
<td>36</td>
</tr>
<tr>
<td>3.2.3</td>
<td>Combining predictions</td>
<td>36</td>
</tr>
<tr>
<td>3.3</td>
<td>Method</td>
<td>36</td>
</tr>
<tr>
<td>3.4</td>
<td>Results</td>
<td>37</td>
</tr>
<tr>
<td>3.4.1</td>
<td>Naïve Bayes</td>
<td>37</td>
</tr>
<tr>
<td>3.4.2</td>
<td>Decision trees</td>
<td>38</td>
</tr>
<tr>
<td>3.4.3</td>
<td>Rule learning</td>
<td>39</td>
</tr>
<tr>
<td>3.4.4</td>
<td>Training time</td>
<td>40</td>
</tr>
</tbody>
</table>
3.4.5 Classifier size ................................................. 40
3.4.6 Accuracy tables ................................................. 40
3.5 Conclusions ......................................................... 44

4 Datasets for object detection ........................................ 45
  4.1 Faces .......................................................... 46
    4.1.1 Cascade selection ......................................... 48
  4.2 Fish ........................................................... 49
    4.2.1 Required and optional annotations ....................... 49
  4.3 Seahorses ..................................................... 51
    4.3.1 Matching and merging seahorse segment detections ... 54
  4.4 Conclusions .................................................... 55

5 Rotated object detection ............................................ 56
  5.1 Cascade training and testing implementation ................ 56
  5.2 Cascade training data ......................................... 57
    5.2.1 Fish ....................................................... 57
    5.2.2 Seahorses .................................................. 62
  5.3 Cascade training settings .................................... 65
  5.4 Results ......................................................... 66
    5.4.1 Cascade features ........................................ 67
    5.4.2 Angle ranges for fish detection ......................... 71
    5.4.3 Angle ranges for seahorse segment detection .......... 74
    5.4.4 Seahorse segment comparison ............................ 79
    5.4.5 Angle steps ............................................... 80
    5.4.6 Rotated images against rotated cascades .............. 82
  5.5 Conclusions .................................................... 83

6 Confidence measures for object detection ......................... 84
  6.1 Uses ........................................................ 85
    6.1.1 Hill-climbing ............................................. 85
    6.1.2 Confidence mapping ....................................... 85
      6.1.2.1 Merging confidences ................................ 86
      6.1.2.2 Multiple cascades .................................. 86
6.2 Stage variations ........................................... 87
  6.2.1 Stage failure tolerance ................................. 87
  6.2.2 Virtual attribute subsetting ............................ 87
  6.2.3 Computation costs ..................................... 88
6.3 Cascade selection ......................................... 88
6.4 Results ..................................................... 88
  6.4.1 Confidence mapping merge comparison ................. 88
  6.4.2 Normalising confidence maps from multiple cascades .. 90
  6.4.3 Stage failure tolerance ................................ 92
  6.4.4 Virtual attribute subsetting ............................ 95
  6.4.5 Method comparison .................................... 97
  6.4.6 Angle ranges .......................................... 100
  6.4.7 Hill-climbing steps .................................... 101
  6.4.8 Classification time .................................... 102
6.5 Conclusions ............................................... 104

7 Detected segment matching ................................ 105
  7.1 Cascades .................................................. 105
  7.2 Matching formulae ........................................ 106
    7.2.1 Designed matching .................................. 107
    7.2.2 Learnt matching ...................................... 108
  7.3 Matching algorithms ..................................... 109
    7.3.1 Greedy .............................................. 109
    7.3.2 Closest match ....................................... 109
  7.4 Evaluation method ....................................... 109
  7.5 Results .................................................. 110
  7.6 Conclusions ............................................... 113

8 Conclusions and further work .............................. 114
  8.1 Rotated object detection ................................. 114
    8.1.1 Training angle ranges ................................ 115
  8.2 Confidence measures ..................................... 115
  8.3 Seahorse segment matching ............................... 116
8.4 Virtual attribute subsetting ........................................ 116
8.5 Recommendations .................................................. 117
8.6 Further work ......................................................... 118

References ................................. 119

A Face detection tables ...................... 126
B Confidence-based ROC curves .......... 128
C Example images with detections ...... 136
List of Tables

2.1 Example classifier learning applications ........................................... 6
2.2 Example classifier training data (Quinlan, 1986) ............................... 7
2.3 Counts of regions tested within a $640 \times 480$ pixel image by a $20 \times 20$
unit cascade under OpenCV defaults .................................................. 22

3.1 Examples of subsets created by different algorithms with $a = 4$, $s = 5$
and $p = 0.7$ ......................................................................................... 33
3.2 Wins/draws/losses for standard/virtual attribute subsetting with varying
subset choice algorithms compared with a single Naive Bayesian
classifier .............................................................................................. 37
3.3 Wins/draws/losses for standard/virtual attribute subsetting with varying
proportion compared with a single Naive Bayesian classifier ............. 37
3.4 Wins/draws/losses for standard/virtual attribute subsetting with varying
subset choice algorithms compared with a single J4.8 classifier ........ 38
3.5 Wins/draws/losses for standard/virtual attribute subsetting with varying
proportion compared with a single J4.8 classifier ............................. 38
3.6 Wins/draws/losses for standard/virtual attribute subsetting with varying
subset choice algorithms compared with a single PART classifier .... 39
3.7 Wins/draws/losses for standard/virtual attribute subsetting with varying
proportion compared with a single PART classifier .......................... 39
3.8 Sizes of classifiers trained on the entire dataset and under standard attribute
subsetting ............................................................................................. 41
3.9 Percentage accuracy over the 31 datasets for J4.8 and standard/virtual
attribute subsetting using J4.8 base classifiers .................................... 42
3.10 Percentage accuracy over the 31 datasets for PART and standard/virtual
attribute subsetting using PART base classifiers ................................ 43

5.1 Fish detection cascade window sizes ................................................. 59
5.2 List of cascade training settings ......................................................... 65
5.3 Rotated cascade feature counts ......................................................... 68
5.4 Rotated cascade area per feature ........................................ 69

6.1 Confidence returned by rotated fish cascades on training images .... 90
6.2 ROC curve figure numbers for binary detections and their corresponding
    hill-climbing curves, including the selected ‘best’ angles .............. 100
6.3 ROC curve figure numbers for binary detections and their corresponding
    confidence mapping curves, including the selected ‘best’ angles ....... 100
6.4 Summary of hill-climbing steps carried out during object detection ... 102
6.5 Time in seconds taken to classify the face dataset with different methods 103

A.1 Face detection true and false positive counts for binary detection, binary
detection followed by hill-climbing and confidence mapping ............ 127
List of Figures

2.1 Example decision tree (Quinlan, 1986) .................................................. 7
2.2 Bootstrap sampling example: 3 bootstrap samples created from a 5-instance dataset .............................................................. 9
2.3 Attribute subsetting example: 3 attribute subsets created from a 4-attribute dataset ................................................................. 11
2.4 Haar Classifier Cascade classification process ....................................... 15
2.5 Integral image illustration ...................................................................... 17
2.6 Tilted integral image illustration ............................................................. 17
2.7 Haar Classifier Cascade features (Viola & Jones, 2001b; Lienhart & Maydt, 2002) ................................................................. 19
2.8 Haar Wavelet features (Papageorgiou et al., 1998) ................................. 19
2.9 Example face with features from stage 1 of a face detection cascade .... 19
2.10 Face detections before and after merging neighbouring detections ..... 23
2.11 Example Receiver Operating Characteristic (ROC) curves ................. 28
2.12 Example of combining points from multiple ROC curves to create a single ‘best’ ROC curve ......................................................... 29

3.1 Virtual attribute subsetting example: 3 copies of a 4-attribute instance created with attribute values erased ........................................ 32
3.2 Visualisation of steps in both balanced attribute subsets algorithm ...... 35

4.1 Example face images ............................................................................ 47
4.2 Example face images, annotated with testing positives ...................... 47
4.3 ROC curves for face detection by the OpenCV detectors .................... 48
4.4 Example fish images ........................................................................... 50
4.5 Example fish image, annotated with testing positives ....................... 50
4.6 Example seahorse images ................................................................... 52
4.7 Seahorse head and body segment lines ............................................... 52
4.8 Example seahorse images, annotated with testing positives ............. 53
4.9 Values measured to compare a seahorse segment annotation $A$ against a seahorse segment detection $D$ ........................................ 54

5.1 Example fish images, annotated with training positives .................. 58

5.2 Positive sample orientations for 7 cascades fixed on angles from $-45^\circ$ to $+45^\circ$ with random angle ranges; darker areas show where random ranges overlap ........................................ 59

5.3 Positive fish training regions at different orientations .................... 60

5.4 First 10 negative fish training regions before flipping, mirroring and rotation ......................................................... 60

5.5 Creating a fish negative training sample from the first negative training region ................................................................. 61

5.6 First 10 negative fish training regions after flipping, mirroring and rotation 61

5.7 Example seahorse images with segments annotated for training ...... 63

5.8 First 8 seahorse heads extracted from example images and forced to $0^\circ$ orientation ............................................................... 63

5.9 First 8 seahorse bodies extracted from example images and forced to $90^\circ$ orientation ............................................................... 63

5.10 Example seahorse images with heads blanked out....................... 64

5.11 Example seahorse images with heads blanked out after flipping, mirroring and rotation ................................................................. 64

5.12 Graph of rotated cascade feature counts .................................. 70

5.13 Graph of rotated cascade feature areas ................................... 70

5.14 ROC curves for fish detection on rotated images, varying the cascade random angle range ........................................ 72

5.15 ROC curves for fish detection by rotated cascades, varying the cascade random angle range ........................................ 73

5.16 ROC curves for seahorse head detection on rotated images, varying the cascade random angle range ........................................ 75

5.17 ROC curves for seahorse body detection on rotated images, varying the cascade random angle range ........................................ 76

5.18 ROC curves for seahorse head detection by rotated cascades, varying the cascade random angle range ........................................ 77
5.19 ROC curves for seahorse body detection by rotated cascades, varying the cascade random angle range ...................................... 78
5.20 ROC curves for seahorse head detection compared with seahorse body detection ................................................................. 79
5.21 ROC curves for fish detection with varying angle steps .................. 80
5.22 ROC curves for seahorse segment detection on rotated images with varying angle steps ......................................................... 81
5.23 ROC curves for seahorse segment detection by rotated cascades with varying angle steps ............................................................ 81
5.24 ROC curves comparing rotated images with rotated cascades ........... 82

6.1 Haar Classifier Cascade confidence measurement process .............. 85
6.2 ROC curves for face detection by confidence mapping, varying local maximum usage ............................................................... 89
6.3 ROC curves for fish detection by confidence mapping, varying local maximum usage ............................................................... 89
6.4 Graph of rotated fish cascade confidences ........................................ 90
6.5 ROC Curves for fish detection by rotated cascade confidence mapping, varying normalisation .................................................... 91
6.6 ROC curves for face detection by confidence mapping, varying the number of stage failures permitted ................................. 92
6.7 ROC curves for fish detection by confidence mapping, varying the number of stage failures permitted ................................. 93
6.8 ROC curves for seahorse segment detection by confidence mapping, varying the number of stage failures permitted ..................... 94
6.9 ROC curves for face detection using confidence mapping and virtual attribute subsetting .................................................. 95
6.10 ROC curves for fish detection using confidence mapping and virtual attribute subsetting .................................................. 96
6.11 ROC curves for face detection using binary detection, binary detection followed by hill-climbing, and confidence mapping ..................... 97
6.12 ROC curves for fish detection using binary detection, binary detection followed by hill-climbing, and confidence mapping ..................... 98