



<http://researchspace.auckland.ac.nz>

ResearchSpace@Auckland

Copyright Statement

The digital copy of this thesis is protected by the Copyright Act 1994 (New Zealand).

This thesis may be consulted by you, provided you comply with the provisions of the Act and the following conditions of use:

- Any use you make of these documents or images must be for research or private study purposes only, and you may not make them available to any other person.
- Authors control the copyright of their thesis. You will recognise the author's right to be identified as the author of this thesis, and due acknowledgement will be made to the author where appropriate.
- You will obtain the author's permission before publishing any material from their thesis.

To request permissions please use the Feedback form on our webpage.

<http://researchspace.auckland.ac.nz/feedback>

General copyright and disclaimer

In addition to the above conditions, authors give their consent for the digital copy of their work to be used subject to the conditions specified on the Library Thesis Consent Form.

THE UNIVERSITY OF AUCKLAND

SCHOOL OF ENGINEERING

**The Application of Support
Vector Machines to Compression
of Digital Images**

Jonathan Robinson

A Thesis Submitted in Partial Fulfilment of the Requirements for
the Degree of

DOCTOR OF PHILOSOPHY

February 2004

Abstract

Methods exploring the application of neural networks to still image compression are detailed in both the spatial and frequency domains. In particular the sparse properties of Support Vector Machine (SVM) learning are exploited in the compression algorithms.

A classic radial basis function (RBF) neural network requires that the topology of the network be defined before training. An SVM has the property that it will choose the minimum number of training points to use as centres of the Gaussian kernel functions. It is this property that is exploited as the basis for image compression algorithms presented in this thesis.

Several novel algorithms are developed applying SVM learning to both directly model the colour surface and model transform coefficients after the surface has been transformed into the frequency domain. It is demonstrated that compression is more efficient in frequency space.

The Discrete Cosine Transform (DCT) is used to transform the colour surface into the frequency domain. A counter-intuitive result is shown where mapping the DCT coefficients to a 1-dimensional function for SVM modelling produces better results than SVM modelling of the 2-dimensional transform surface.

Results are presented in comparison to the JPEG image compression algorithm. In the frequency domain, results are superior to that of JPEG. For example, the quality of the ‘Lena’ image compressed 63:1 for JPEG is slightly worse quality than the same image compressed 192:1 with the RKi-1 algorithm presented in this thesis. Due to the commercial value of the algorithms detailed in this thesis, a patent has been filed.

Acknowledgements

I would like to thank my supervisor Assoc. Prof. Vojislav Kecman for his support and guidance during the course of this thesis. I would also like to thank my family for their love, support and encouragement. In particular, I would like to thank my wife Simone for proof-reading the text and my daughter Claudia who graciously agreed to allow her photograph to be used as a test image in this thesis.

Contents

Abstract	i
Acknowledgements	ii
1 Introduction	1
1.1 Spatial and Frequency Domains	3
1.2 Statement of the Problem	9
1.3 Previous Work	9
1.4 Support Vector Machines	11
1.5 Research Objectives and Contributions	13
1.6 Organisation of this Thesis	14
2 LOMON: LOcal MOdelling Network	16
2.1 Training	19
2.2 Results	24
2.2.1 Effect of the Gaussian Shape Parameter	28
2.3 LOMON With Multiple Gaussians	29
2.3.1 Four-Gaussian LOMON	35
2.4 Parameters Required to be Transmitted	38
2.5 Discussion	41
3 Support Vector Machine Learning	43
3.1 Introduction	43
3.2 Classification by SVM	44
3.2.1 Optimal Canonical Separating Hyperplane	46
3.2.2 The Maximal Margin	48
3.2.3 Solution by Lagrangian Multipliers	50
3.2.4 Non-separable Classification	53
3.3 Function Approximation by SVM	56
3.4 Summary	62

4	SVM Modelled Colour Surface	63
4.1	SVM Model	63
4.1.1	Implementation	64
4.1.2	Non-Negative Least Squares (NNLS)	68
4.1.3	Transformation of the QP problem to NNLS	69
4.2	Results	71
4.3	Implementation	76
4.3.1	Run Length and Huffman Encoding	79
4.3.2	Encoding the Biases	83
4.3.3	Encoding Weights	84
4.4	Discussion	87
5	LOMON combined with SVM	88
5.1	Introduction	88
5.2	The LOMON-SVM Algorithm	88
5.3	Results	92
5.4	Discussion	92
6	Compression in the Frequency Domain: The RKi-1 Algorithm	94
6.1	Introduction	94
6.2	Discrete Cosine Transform	95
6.2.1	Transformation of the DCT Matrix to 1-D	97
6.3	Combining SVM with DCT	98
6.3.1	The ‘Inversion’ bit	102
6.4	Implementation	104
6.4.1	Encoding Data for Storage	106
6.4.2	Encoding the DC Coefficients	107
6.5	Results	110
6.5.1	Performance	110
6.5.2	Effect of Reducing the Number of Inversion Bits	113
6.6	Discussion	116
7	The RKi-2 Compression Algorithm	117
7.1	2-Dimensional SVM	117
7.2	Results	120
7.3	Discussion	124
8	The RKi-3 Algorithm	126
8.1	Results	128
8.2	Discussion	134
9	Summary and Discussion	135
9.1	Further Work	136

Appendix I: Overview of the JPEG Compression Standard	139
Appendix II: Contents of the Accompanying CD	141
References	141

List of Figures

1.1	TOP: Original 512×512 ‘Lena’ image. BOTTOM: Rotated 512×512 ‘Lena’ image.	5
1.2	Rotated 512×512 ‘Lena’ image illustrating the complexity of the surface to be modeled.	6
1.3	TOP: Blocking effects seen in the JPEG compression algorithm at high compression (ratio = 27:1). BOTTOM: Smearing effects seen in wavelet compression at high compression (ratio = 114:1) .	8
1.4	Feed-forward RBF network	10
1.5	Two class classification problem	12
1.6	Example of a 1-dimensional SVM	14
2.1	The LOMON algorithm where $\epsilon = \text{max_image_error}$ is the maximum error in a pixel.	17
2.2	LOMON Strategy	18
2.3	Largest pixel error during LOMON training (No Bias)	22
2.4	LOMON trained on Lena image (no bias)	23
2.5	Largest pixel error during LOMON training (With Bias)	24
2.6	256×256 Lena image trained using LOMON	26
2.7	400×400 Claudia image trained using LOMON	27
2.8	Number of blocks versus Gaussian shape parameter for LOMON .	29
2.9	Definition of splitting in LOMON for two Gaussians	30
2.10	256×256 Lena image trained using LOMON (two Gaussians per block)	31
2.11	256×256 Lena image showing blocks generated by LOMON (two Gaussians per block)	32
2.12	Number of blocks versus Gaussian width for two Gaussians	34
2.13	LOMON results for two Gaussians per block	35
2.14	Definition of splitting in LOMON into four Gaussians	36
2.15	256×256 Lena image trained using LOMON (four Gaussians per block)	37
2.16	256×256 Lena image showing blocks generated by LOMON (four Gaussians per block)	39
2.17	Number of blocks versus Gaussian width or four Gaussian LOMON	40

3.1	Example of a 1-dimensional SVM	45
3.2	Maximal Margin	46
3.3	Optimal canonical separating hyperplane in 1-dimension	47
3.4	Maximal margin	48
3.5	Parameters used in the computation of the maximal margin	49
3.6	Non-separable binary classification	54
3.7	Non-separable binary classification with slack variables	54
3.8	The parameters used in (1-dimensional) support vector regression (from [1]).	58
4.1	Radial Basis Function network	65
4.2	256 × 256 Lena Image with no bias	72
4.3	Surface patch produced by SVM learning	73
4.4	Plot of number of support vectors versus Gaussian width	74
4.5	Plot of SNR versus compression ratio for the 400 × 400 Claudia image	75
4.6	Plot of SNR versus compression ratio for the 512 × 512 Lena image	75
4.7	256 × 256 Lena Image with bias	76
4.8	512 × 512 Lena Image with bias	77
4.9	400 × 400 Claudia image with bias	78
4.10	Huffman coding algorithm	82
4.11	Huffman encoding example	83
4.12	Huffman table data storage	84
4.13	Too few quantisation levels	86
4.14	Encoding the Weights	87
5.1	Modified LOMON Strategy	90
5.2	The LOMON algorithm where <code>max_image_error</code> is the maximum error in a pixel.	91
5.3	Lena image using LOMON combined with SVM learning	93
6.1	The Discrete Cosine Transform	97
6.2	Zig-zag mapping of DCT coefficients	98
6.3	The RKi-1 Algorithm	99
6.4	1-dimensional plot of DCT coefficients	100
6.5	DCT real values as input to the SVM	101
6.6	DCT absolute magnitude as input to the SVM	103
6.7	Input block to SVM and output block	106
6.8	XML export file	108
6.9	Plot of SNR for RKi-1 and JPEG with Lena Image	112
6.10	Plot of SNR for RKi-1 and JPEG with Claudia Image	112
6.11	Effect of reducing the number of inversion bits for the 400 × 400 Claudia image	113

6.12	Subjective comparison of RKi-1 and JPEG for the 512×512 Lena image	114
6.13	Subjective comparison of RKi-1 and JPEG for the 400×400 Claudia image	115
7.1	The RKi-2 Algorithm	119
7.2	Plot of SNR versus compression ratio for the 400×400 Claudia image	121
7.3	Plot of SNR versus compression ratio for the 512×512 Lena image	121
7.4	Subjective comparison of RKi-2 and JPEG for the 512×512 Lena image	122
7.5	Subjective comparison of RKi-2 and JPEG for the 400×400 Claudia image	123
7.6	Influence of Gaussians in RKi-1 and RKi-2	125
8.1	Shuffling DCT coefficients	127
8.2	Reduction of a ‘Super-Block’ in RKi-3	130
8.3	Subjective comparison of RKi-3 and JPEG for the 512×512 Lena image	131
8.4	Subjective comparison of RKi-3 and JPEG for the 384×384 Claudia image	132
8.5	Plot of SNR versus compression ratio for the 384×384 Claudia image	133
8.6	Plot of SNR versus compression ratio for the 512×512 Lena image	133
9.1	Plot of SnR versus compression ratio for the 512×512 Lena image for all algorithms	136
9.2	Plot of SnR versus compression ratio for the 400×400 Claudia image for all algorithms	137

List of Tables

4.1	Huffman encoding biases	85
6.1	Huffman encoding DC values	109
9.1	The Y component quantisation table	140