Wavelet Domain Image Resolution
Enhancement Methods

Alptekin Temizel

Submitted for the Degree of
Doctor of Philosophy

Centre for Vision, Speech and Signal Processing
School of Electronics and Physical Sciences
University of Surrey
Guildford, Surrey GU2 7XH, UK

January 2006

© Alptekin Temizel 2006
Summary

Resolution enhancement of pictorial data is desirable in many applications such as monitoring, surveillance, medical imaging and remote sensing when images at desirable resolution levels are not available. It is a classic signal interpolation problem and several conventional approaches such as zero-order interpolation (sample-and-hold), bilinear and spline interpolation are widely used. However, undesirable levels of smoothing across salient edges in the higher resolution images obtained using these conventional methods resulted in a search for more effective algorithms. Recently several efforts in the field have utilised wavelet-domain methodologies with the intention of overcoming some of the problems associated with conventional treatment.

In this thesis, we propose three wavelet domain image resolution enhancement algorithms.

The first proposed algorithm is based on the estimation of detail wavelet coefficients at high resolution scales by exploiting the wavelet coefficient correlation in a local neighbourhood sense. The unknown detail coefficients are estimated by employing linear least-squares regression parameters of which are obtained from the quad-tree decomposition of the low-resolution image.

The second algorithm starts with an initial high-resolution approximation to the original image obtained by means of zero-padding in the wavelet domain. This is further processed using the cycle spinning methodology which reduces ringing. Linear regression using a minimal training set of high-resolution originals is finally employed to rectify the degraded edges.

For the third algorithm we propose a directional variant of the cycle spinning methodology with an aim of reducing the over-smoothing of salient image features as well as offering a reduction in computational complexity. In particular we take into account local edge orientation information derived from a wavelet decomposition of the available low resolution image to influence key parameters of the cycle spinning algorithm.

Our results show that the proposed methods are considerably superior to conventional image interpolation techniques, both in objective and subjective terms, while also comparing favourably with competing methods.
Acknowledgments

Above all else, I would like to express gratitude to my wife Tugba for her constant support and understanding throughout the ups and downs of this challenging journey. She was always there when the light at the end of tunnel seemed too far away with her support and also to share and to intensify the cheerful mood when things went well. This work would never have been achieved without her constant encouragement.

I wish to express my sincere thanks to my supervisor Dr. Theo Vlachos for patiently putting up with my endless distractions, constant drifts and coming up with numerous strange ideas. He has been a great supervisor helping me with not only his technical guidance but also with his wisdom and much needed encouragement and morale support.

I also would like to thank to my collaborative supervisor Dr. Bilgay Akhan and Gerard O’Driscoll from Visioprime for facilitating this work and their understanding.

Last but not least, I would like to thank my parents and my sister whose hearts beat next to mine all the time even though they are far away.
## Contents

Summary ............................................................................................................................................................... i
Acknowledgments ........................................................................................................................................... iii
Contents ............................................................................................................................................................... v
List of Figures .................................................................................................................................................. ix
List of Tables .................................................................................................................................................. xv
Glossary of Terms ....................................................................................................................................... xvii

**CHAPTER 1** ....................................................................................................................................................... 1

Introduction ......................................................................................................................................................... 1

1.1 Image Resolution Enhancement in the Literature .................................................................................. 2
1.2 Assumptions ........................................................................................................................................... 4
1.3 Contributions of the Thesis ................................................................................................................. 5
1.4 Publications Resulting from the Thesis ........................................................................................... 6
1.5 Structure of the Thesis ......................................................................................................................... 6

**CHAPTER 2** ...................................................................................................................................................... 9

Image Resolution Enhancement ................................................................................................................. 9

2.1 Introduction ............................................................................................................................................ 9
2.2 Image Acquisition Process .................................................................................................................. 9
2.3 Linear Interpolation Methods ........................................................................................................... 11

2.3.1 Signal processing approach to interpolation .............................................................................. 11
2.3.2 Interpolation in signal domain .................................................................................................. 14
2.4 Advanced Methods ............................................................................................................................. 18

2.4.1 Discontinuity adaptive techniques .......................................................................................... 18
2.4.2 Regularisation based approaches ............................................................................................ 22
2.4.3 Fractal and Triangulation Based Methods ............................................................................ 23
2.4.4 Discrete Cosine Transform (DCT) based methods ............................................................. 24
2.4.5 Wavelet Based Techniques ....................................................................................................... 25
2.5 Conclusions .......................................................................................................................................... 25

**CHAPTER 3** .................................................................................................................................................... 27
Contents

Wavelet Domain Image Resolution Enhancement ................................................................. 27
  3.1 Introduction ...................................................................................................................... 27
  3.2 Wavelet Transform ........................................................................................................... 27
    3.2.1 Continuous Wavelet Transform (CWT) ...................................................................... 27
    3.2.2 Discrete Wavelet Transform (DWT) ........................................................................... 28
    3.2.3 Signal Processing Approach to DWT .......................................................................... 29
    3.2.4 2-D Discrete Wavelet Transform .............................................................................. 33
  3.3 Wavelet Based Image Resolution Enhancement Methods ........................................ 36
    3.3.1 Background ................................................................................................................... 36
  3.4 Wavelet Domain Image Resolution Enhancement Methods ........................................ 43
    3.4.1 Methods Using the Wavelet Transform Extrema in Coarser Subbands .................... 43
    3.4.2 Methods Using the Hidden Markov Tree (HMT) modelling of wavelet coefficients ... 46
    3.4.3 Methods Using Multiresolutional Basis Fitting Reconstruction (MBFR) .................... 49
    3.4.4 Others .......................................................................................................................... 50
    3.4.5 Wavelet Superresolution .............................................................................................. 51
  3.5 Conclusions ...................................................................................................................... 51

CHAPTER 4 ...................................................................................................................................... 53
Wavelet Domain Image Resolution Enhancement with Coefficient Estimation .................... 53
  4.1 Introduction ...................................................................................................................... 53
  4.2 Wavelet-domain resolution enhancement ...................................................................... 53
    4.2.1 Discussion ....................................................................................................................... 56
  4.3 Experimental results ........................................................................................................ 57
  4.4 Frequency Domain Analysis ........................................................................................... 67
  4.5 Conclusions ...................................................................................................................... 70

CHAPTER 5 ...................................................................................................................................... 71
Wavelet Domain Image Resolution Enhancement Using Cycle Spinning and Edge Rectification
  5.1 Introduction ...................................................................................................................... 71
  5.1.1 Cycle Spinning (CS) ..................................................................................................... 71
  5.2 Resolution Enhancement Using Cycle Spinning ............................................................ 72
    5.2.1 Wavelet Domain Zero Padding (WZP) ........................................................................ 72
    5.2.2 Adaptation of Cycle Spinning to Image Resolution Enhancement ............................ 73
    5.2.3 Discussion ...................................................................................................................... 73
  5.3 Post Processing Using Edge Rectification ....................................................................... 79
    5.3.1 Edge Model .................................................................................................................... 79
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.3.2 Post Processing of Wavelet Compressed Images</td>
<td>82</td>
</tr>
<tr>
<td>5.3.3 Edge Rectification</td>
<td>83</td>
</tr>
<tr>
<td>5.4 Experimental Results</td>
<td>90</td>
</tr>
<tr>
<td>5.5 Conclusions</td>
<td>90</td>
</tr>
<tr>
<td>CHAPTER 6</td>
<td>91</td>
</tr>
<tr>
<td>Wavelet Domain Image Resolution Enhancement with Directional Cycle-Spinning</td>
<td>91</td>
</tr>
<tr>
<td>6.1 Introduction</td>
<td>91</td>
</tr>
<tr>
<td>6.2 Resolution Enhancement Using Directional Cycle Spinning</td>
<td>91</td>
</tr>
<tr>
<td>6.3 Algorithm Description</td>
<td>94</td>
</tr>
<tr>
<td>6.4 Experimental Results</td>
<td>96</td>
</tr>
<tr>
<td>6.5 Discussion</td>
<td>96</td>
</tr>
<tr>
<td>6.6 Conclusions</td>
<td>102</td>
</tr>
<tr>
<td>CHAPTER 7</td>
<td>103</td>
</tr>
<tr>
<td>Comparative Assessment</td>
<td>103</td>
</tr>
<tr>
<td>7.1 Introduction</td>
<td>103</td>
</tr>
<tr>
<td>7.2 Assessment Process</td>
<td>103</td>
</tr>
<tr>
<td>7.3 Experimental Results</td>
<td>111</td>
</tr>
<tr>
<td>7.4 Conclusions</td>
<td>111</td>
</tr>
<tr>
<td>CHAPTER 8</td>
<td>113</td>
</tr>
<tr>
<td>Conclusions</td>
<td>113</td>
</tr>
<tr>
<td>8.1 Summary</td>
<td>113</td>
</tr>
<tr>
<td>8.2 Future Work</td>
<td>118</td>
</tr>
<tr>
<td>Appendix A - Least Squares Estimation</td>
<td>121</td>
</tr>
<tr>
<td>Appendix B - Subjective Evaluation Criteria</td>
<td>123</td>
</tr>
<tr>
<td>Appendix C - Test Images</td>
<td>125</td>
</tr>
<tr>
<td>Bibliography</td>
<td>127</td>
</tr>
</tbody>
</table>
List of Figures

Figure 1: Courtesy of “The Daily Telegraph” newspaper, published on Thursday, September 16, 2004, No. 46427 ..................................................... 3
Figure 2: 64x64 and 32x32 Lena images and corresponding grids ........................................................ 10
Figure 3: Sampling rate converter ........................................................................................................ 11
Figure 4: Illustration of sampling rate conversion in the frequency domain ........................................ 12
Figure 5: Ideal low-pass filter, recovering the interpolated signal ....................................................... 13
Figure 6: Illustration of bilinear interpolation ..................................................................................... 15
Figure 7: (a) Zero-order hold, (b) linear and (c) cubic interpolation functions and their corresponding frequency responses ...................................................................................................... 16
Figure 8: Images enlarged by 4x (128x128 to 512x512) using (a) zero-order hold, (b) linear and (c) cubic interpolation functions. The images on the left are magnified portions of these enlarged images and 8x amplified difference images where mid-grey represents zero error are shown on the right .............................................................................................................................. 17
Figure 9: Wavelet Decomposition (analysis) ................................................................................... 29
Figure 10: Wavelet Reconstruction (synthesis) .................................................................................... 30
Figure 11: Frequency responses of Haar filter pair ............................................................................. 32
Figure 12: Frequency responses of Daubechies 9/7 filter pair ............................................................ 32
Figure 13: Illustration of 2-D wavelet decomposition .......................................................................... 34
Figure 14: Further levels of 2-D wavelet decomposition ...................................................................... 35
Figure 15: Persistency of wavelet coefficient magnitudes through different scales of wavelet decomposition. The coefficient magnitudes at corresponding spatial locations decay at a similar rate from lower resolution scales to higher resolution scales ........................................... 37
Figure 16: Histogram of horizontal detail subband wavelet coefficients of Lena image ..................... 37
Figure 17: Wavelet Domain Zero Padding (WZP): the available image (X) is assumed to be the low resolution subband of the target high resolution image. The unknown subband elements (detail coefficients) are filled with 0s and resolution enhancement is achieved by inverse wavelet transform ...................................................................................................................................... 39
Figure 18: Illustration of Wavelet Domain Zero Padding (WZP) with a real image ......................... 40
Figure 19: Original Lena image (top left) and an extract from this image (top right). (a)Reconstructed image using WZP -34.45 dB- and extract. (b) Reconstructed image using coefficients generated from their parents with bilinear interpolation -30.44 dB- and extract. 41
Figure 20: A summary of wavelet domain image resolution enhancement methods. ...................... 42
Figure 21: Image resolution enhancement using wavelet transform extrema extrapolation ............43
Figure 22: Illustration of extrapolation using wavelet coefficient extrema evolution among the scales using data from vertical detail subbands at different scales of Lena image. The solid line shows the fitted line, crosses are the actual data points. The unknown coefficient at scale 4 is estimated with extrapolation of the fitted line. ............................................................44
Figure 23: Original Lena image (top left) and an extract from this image (top right). (a) Extract from reconstructed image using WZP and amplified error. (b) Extract from reconstructed image using original wavelet coefficient magnitudes while coefficients signs are copied from their parents and amplified error images. .............................................................................45
Figure 24: Illustration of two-state, zero-mean Gaussian mixture model with probability density functions (pdf). In these plots $w$ is the observed coefficient value and $S$ is the corresponding hidden state variable. $S=1$ is the low-variance Gaussian corresponding the lower valued coefficients (top left) and $S=2$ is the high-variance Gaussian corresponding to higher valued coefficients representing singularities (top right). The combined pdf which models the distribution of all wavelet coefficients is shown at the below plot .............................................48
Figure 25: Illustration of state transitions......................................................................................49
Figure 26: The estimator coefficients are calculated by the top sub-system where Least-Squares Regression is denoted by LSR. High-resolution image estimate is generated by the below sub-system where H-Estimator and V-Estimator linearly combines the weighted coefficients in the horizontal and vertical direction respectively, $W^T$ denotes the inverse wavelet transform and $HH'$ is all zeros.............................................................................................................56
Figure 27: Top-left quadrant of original Zone Plate image (top-left), difference image using bilinear interpolation for resolution enhancement (top-right), difference image using 0s in the “unknown” subbands (bottom-left), difference image using predicted coefficient values in the “unknown” subbands (bottom-right).............................................................................58
Figure 28: Extracts from original (top) and 4x reconstructed Lena images: (a) reconstructed and residual images using bilinear interpolation, (b) Reconstructed and residual images using the proposed method (residual images have been amplified for inspection purposes) ..................59
Figure 29: Extracts from original (top) and 4x reconstructed Peppers images: (a) reconstructed and residual images using bilinear interpolation, (b) Reconstructed and residual images using the proposed method (residual images have been amplified for inspection purposes) ..........60
Figure 30: Extracts from original (top) and 4x reconstructed Baboon images: (a) reconstructed and residual images using bilinear interpolation, (b) Reconstructed and residual images using the proposed method (residual images have been amplified for inspection purposes) ...............61
Figure 31: Original Snoopy image: (a) reconstructed and residual images using bilinear interpolation, (b) Reconstructed and residual images using the proposed method (residual images have been amplified for inspection purposes) ........................................62
List of Figures

Figure 32: Illustrations of the added detail to Lena image. (a) original coefficients of the detail subbands for 2x enlargement (top left), (b) predicted coefficients for 2x enlargement (top right), (c) original coefficients of the detail subbands for 4x enlargement (bottom left), (d) predicted coefficients for 4x enlargement (bottom right). 63

Figure 33: Absolute reconstruction error using various techniques for a cross-section taken from the highlighted part of Baboon image. The cross-section is plotted in the top plot where s(x) is the pixel amplitude at position x. The below plot shows the absolute reconstruction error |Err(x)| at position x using various techniques. 64

Figure 34: Absolute reconstruction error using various techniques for an edge. The above plot shows cross section of an edge where s(x) is the amplitude of a pixel at position x. The below plot shows the reconstruction error |Err(x)| using various techniques. 65

Figure 35: Illustration of the estimation process using filtering operations. The aim is to find an estimator $F(z)$ which estimates $HL_q$ given $LL_q$. Assuming next level subbands has also have similar characteristics, the $F(z)$ can be derived using $HL_i$ and $LL_i$ as they are both known, hence allowing calculation of an estimator. 68

Figure 36: Frequency response of the filter $F(z)$ obtained by combining the high pass filter $H(z)$ and $W(z)$ which is the filter created using least square weights obtained using 512x512 Lena image. 69

Figure 37: Frequency analysis using 512x512 Lena image (a) the original subbands $LL$, $HL$, $LH$ and $HH$ (b) estimated HL and LH subbands, $HL''$ and $LH''$ (including the intermediary estimation stages $HL'$ and $LH'$). 69

Figure 38: Block diagram of the proposed method. 74

Figure 39: PSNR(dB) vs. shift amount (k) for Lena image. 75

Figure 40: Various stages of the reconstruction where cycle spinning is applied in the neighbourhood denoted by the shaded areas in the boxes. 76

Figure 41: Extracts from original (top) and reconstructed Lena images: (a) reconstructed and residual images using bilinear interpolation, (b) reconstructed and residual images using the proposed method (residual images have been amplified for inspection purposes). 77

Figure 42: Extracts from original (top) and 4x reconstructed Peppers images: (a) reconstructed and residual images using bilinear interpolation, (b) reconstructed and residual images using the proposed method (residual images have been amplified for inspection purposes). 78

Figure 43: Simplified Edge Model. 81

Figure 44: An edge and its modelled version using the described edge model. 82

Figure 45: Edge Cluster Map. 84

Figure 46: Block diagram of the proposed method. 85

Figure 47: PSNR(dB) as a function of shift for Lena (From 256x256 to 512x512). 85
List of Figures

Figure 48: Extracts from original (top) and 4x reconstructed Lena images: (a) reconstructed and residual images using bilinear interpolation, (b) reconstructed and residual images using the proposed method (residual images have been amplified for inspection purposes) .................. 86

Figure 49: Extracts from original (top) and 4x reconstructed Peppers images: (a) reconstructed and residual images using bilinear interpolation, (b) reconstructed and residual images using the proposed method (residual images have been amplified for inspection purposes) .......... 87

Figure 50: Extracts from original (top) and 4x reconstructed Baboon images: (a) reconstructed and residual images using bilinear interpolation, (b) reconstructed and residual images using the proposed method (residual images have been amplified for inspection purposes) ............. 88

Figure 51: Original Snoopy image (top) and (a) reconstructed and residual images using bilinear interpolation, (b) reconstructed and residual images using the proposed method (residual images have been amplified for inspection purposes) ..................................................................... 89

Figure 52: PSNR (dB) values for the different quadrants of the synthetic image using different cycle spinning directions ........................................................................................................................92

Figure 53: Shifts used for different directional CS types .......................................................................................................................... 93

Figure 54: Cross section of a vertical edge extracted from the top-right quadrant of the synthetic image in Figure 52 and its reconstructed versions using horizontal and vertical CS (top) and absolute error (\(|\text{Err}(x)|\)) of these reconstructions (bottom) .................................................................................. 93

Figure 55: Lena image partitioned into 8x8 blocks and corresponding wavelet subbands with 4x4 blocks................................................................................................................................................................. 95

Figure 56: Extracts from original (top) and 4x reconstructed Lena images: (a) reconstructed and residual images using bilinear interpolation, (b) reconstructed and residual images using the proposed method (residual images have been amplified for inspection purposes) .................. 97

Figure 57: Extracts from original (top) and 4x reconstructed Peppers images: (a) reconstructed and residual images using bilinear interpolation, (b) reconstructed and residual images using the proposed method (residual images have been amplified for inspection purposes) .......... 98

Figure 58: Extracts from original (top) and 4x reconstructed Baboon images: (a) reconstructed and residual images using bilinear interpolation, (b) reconstructed and residual images using the proposed method (residual images have been amplified for inspection purposes) ............. 99

Figure 59: Original Snoopy image (top) and 4x reconstructed images: (a) reconstructed and residual images using bilinear interpolation, (b) reconstructed and residual images using the proposed method (residual images have been amplified for inspection purposes). ................. 100

Figure 60: Cross section of an edge and its reconstructed versions using various methods (top) and absolute error of the reconstructed edges (bottom) ............................................................................ 101

Figure 61: Extract from original Lena image (top) and corresponding images enlarged using (a) Genuine Fractals (b) NEDI. 2x enlargements are presented on the left and 4x on the right.106
Figure 62: Extract from original *Elaine* image (top) and corresponding images enlarged using (a) Genuine Fractals (b) NEDI. 2x enlargements are presented on the left and 4x on the right. 107

Figure 63: Extract from original *Peppers* image (top) and corresponding images enlarged using (a) Genuine Fractals (b) NEDI. 2x enlargements are presented on the left and 4x on the right. 108

Figure 64: Extract from original *Baboon* image (top) and corresponding images enlarged using (a) Genuine Fractals (b) NEDI. 2x enlargements are presented on the left and 4x on the right. 109

Figure 65: Original *Snoopy* image (top) and corresponding images enlarged using (a) Genuine Fractals (b) NEDI. 2x enlargements are presented on the left and 4x on the right. 110

Figure 66: 4x reconstructed images using nearest neighbour interpolation (top) and proposed wavelet domain coefficient prediction method explained in chapter 4 (bottom). Image courtesy of “The Daily Telegraph” newspaper, published on Thursday, September 16, 2004, No. 46427. 115

Figure 67: 4x reconstructed images using proposed cycle-spinning and edge reconstruction method as explained in Chapter 5 (top) and directional cycle-spinning method explained in chapter 6 (bottom). Image courtesy of “The Daily Telegraph” newspaper, published on Thursday, September 16, 2004, No. 46427. 116

Figure 68: Illustration of 8x (left) and 16x (right) enlargement using (a) bilinear interpolation, (b) coefficient prediction, (c) CS and ER, (d) Directional CS. 117

Figure 69: Illustration of a step edge (left) and a line edge (right). 118
List of Tables

Table 1: Haar filter coefficients. ................................................................. 31
Table 2: Daubechies 9/7 filter coefficients. .................................................. 31
Table 3: Correlation values between $HL_i$ and $HL'_j$ subbands for Lena ........ 55
Table 4: Correct coefficient sign percentages for various techniques. The calculations are done for the coefficients whose magnitude exceeds the threshold $\theta$. The threshold $\theta$ is found by constraining the percentage of coefficients exceeding the $\theta$ to the values in the second column. ......................................................................................................................... 66
Table 5: PSNR(dB) performance as a function of maximum shift amount for Lena, Elaine, Baboon and Peppers images (2x resolution enhancement: from 128x128 to 256x256) .... 75
Table 6: Estimation of horizontal edge width parameter $w$ for a variety of images and reconstruction techniques ................................................................. 83
Table 7: PSNR(dB) as a function of shift for Lena (From 256x256 to 512x512) ........ 85
Table 8: PSNR (dB) values for 2x enlargement (256x256 to 512x512) .................. 105
Table 9: PSNR (dB) values for 4x enlargement (128x128 to 512x512) ................. 105
# Glossary of Terms

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Term</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCD</td>
<td>Charge Coupled Device</td>
</tr>
<tr>
<td>CCTV</td>
<td>Close Circuit Television</td>
</tr>
<tr>
<td>CS</td>
<td>Cycle-Spinning</td>
</tr>
<tr>
<td>CWT</td>
<td>Continuous Wavelet Transform</td>
</tr>
<tr>
<td>Db9/7</td>
<td>Daubechies 9/7 Wavelet Filter</td>
</tr>
<tr>
<td>DCS</td>
<td>Diagonal Cycle Spinning</td>
</tr>
<tr>
<td>DFT</td>
<td>Discrete Fourier Transform</td>
</tr>
<tr>
<td>DWT</td>
<td>Discrete Wavelet Transform</td>
</tr>
<tr>
<td>EM</td>
<td>Expectation Maximisation</td>
</tr>
<tr>
<td>ER</td>
<td>Edge Rectification</td>
</tr>
<tr>
<td>HCS</td>
<td>Horizontal Cycle Spinning</td>
</tr>
<tr>
<td>HHₙ</td>
<td>nth level diagonal detail wavelet subband</td>
</tr>
<tr>
<td>HLₙ</td>
<td>nth level vertical detail wavelet subband</td>
</tr>
<tr>
<td>HMRF</td>
<td>Huber-Markov Random Field</td>
</tr>
<tr>
<td>HMT</td>
<td>Hidden Markov Tree</td>
</tr>
<tr>
<td>HP</td>
<td>High-Pass</td>
</tr>
<tr>
<td>HR</td>
<td>High Resolution</td>
</tr>
<tr>
<td>IDWT</td>
<td>Inverse Discrete Wavelet Transform</td>
</tr>
<tr>
<td>LHₙ</td>
<td>nth level horizontal detail wavelet subband</td>
</tr>
<tr>
<td>LLₙ</td>
<td>nth level low-pass wavelet subband</td>
</tr>
<tr>
<td>LP</td>
<td>Low-Pass</td>
</tr>
<tr>
<td>LR</td>
<td>Low Resolution</td>
</tr>
<tr>
<td>MAP</td>
<td>Maximum-a-posteriori</td>
</tr>
<tr>
<td>MBFR</td>
<td>Multiresolutional Basis Fitting Reconstruction</td>
</tr>
<tr>
<td>MSE</td>
<td>Mean Square Error</td>
</tr>
<tr>
<td>NCS</td>
<td>Non-directional Cycle Spinning</td>
</tr>
<tr>
<td>NEDI</td>
<td>New Edge Directed Interpolation</td>
</tr>
<tr>
<td>PSNR</td>
<td>Peak Signal-to-Noise Ratio</td>
</tr>
<tr>
<td>QMF</td>
<td>Quadrature Mirror Filters</td>
</tr>
<tr>
<td>RMSE</td>
<td>Root Mean Square Error</td>
</tr>
<tr>
<td>VCS</td>
<td>Vertical Cycle Spinning</td>
</tr>
<tr>
<td>WZP</td>
<td>Wavelet Domain Zero Padding</td>
</tr>
<tr>
<td>X</td>
<td>Low resolution input image</td>
</tr>
</tbody>
</table>
Glossary of Terms

Y  High resolution ground-truth image
Y'  Reconstructed high resolution image
CHAPTER 1

Introduction

With the advancement of the digital technology, digital images have become an indispensable part of our daily lives. The ease of use, low-cost and convenience of digital image acquisition, manipulation and reproduction are a few reasons which made the digital images and associated equipment so popular. Nowadays, an average person in developed world is highly likely to own a digital TV receiver, a camera phone on which he/she can take or view digital pictures anytime, a personal computer with internet connectivity facilitating access to virtually unlimited number of images from all around the world and potentially a photo printer to reproduce these images on printed medium. Nowadays, many people carry -almost anytime- an imaging device in the form of camera-phones. An interesting application of the ubiquitous-ness of imaging equipment is ease of obtaining pictures for unexpected incidents hence allowing amateurs to capture events as they happen. Witnesses of events can now take pictures long before professional reporters and photographers arrive at the scene. Recently, it has been usual to see news headlines bearing images captured with camera-phones -albeit the inferior resolution of the captured images-.

Besides their broad penetration into our daily lives, digital images are also widespread in professional applications such as medical imaging, surveillance, recording for evidential purposes and satellite imaging to give a few examples. Digital images and availability of hardware and software dealing with digital images facilitated many things that hadn’t been possible or had been simply impractical previously. Some benefits of digital images can be named as easy retrieval, storage, transmission and copying; allowing both automatic and manual analysis, compression, processing and enhancement for particular purposes.

Although digital images are abundant, their resolution might not always be sufficient and it is frequently needed to “enlarge” these digital images at the best possible quality. “Enlargement” might be required for a number of reasons:

- The capture device might not be capable of capturing high resolution images. Fewer number of CCDs in an imaging equipment results in lower resolution images. The imaging devices on most camera-phones can be given as an example to this. The resolution of the images might be prohibitive when viewing or printing these images. In such cases, enlargement before displaying or printing these images is desired.

- The part of the image which is of concern might only be a small part in the overall image and hence represented by insufficient number of pixels. An example to this is a face image captured by a security camera at distance. In this case, even though the captured image is at high resolution, when the face image is extracted, it might be a low resolution image itself.
Then, in a similar fashion to the previous case enlargement might be desirable before displaying or printing.

- Although the source image is sufficiently high resolution, due to restrictions in the bandwidth or processing power available it might be transferred or displayed in a progressive fashion (i.e. initially displayed as a low-resolution image and getting better in time as more data is received and/or decompressed). Progressive JPEG images are generally transmitted in this fashion over the internet, allowing users to see a general sketch of the images first while they are refined when new data is received. These initial sketches could be made more useful with enlargement saving network bandwidth and/or time.

Figure 1 is an example of an image captured -potentially- from a PAL video recording and enlarged for printing purposes. The bottom image in this figure is an enlarged version (using zero-order hold) of the marked area in the top image. Although an image is available picturing the incident, the resolution is clearly insufficient to provide sufficient data for further analysis.

The image resolution enhancement is an ill posed problem with the implication that it is not possible to recover the missing data perfectly. A better reconstructed enlarged image could be defined as the reconstruction which best approximates the unavailable high resolution image. The digital image acquisition process (or the down-sampling of a digital image) results in loss of high spatial frequency to avoid aliasing. Hence, the resolution enhancement problem could be stated as “to estimate the missing higher spatial frequency information using the available low-frequency information”. Recovery of high-frequency information also implies the recovery of edges. Edges convey significant information about the images and generally are the most important features in an image.

1.1 Image Resolution Enhancement in the Literature

In the literature, a number of terms are used to describe the process of generating an enlarged image from a single available low-resolution image. The following list summarises the terms that we encountered in the literature:

- Enlargement
- Expansion
- Resizing
- Magnification
- Zooming
- Up-sampling
- Interpolation
Chapter 1. Introduction

- Resolution enhancement
- Super-resolution

The terms enlargement, expansion, resizing, magnification and zooming have a general scope and might also be used for optical and analogue techniques achieving an enlarged view for an image.

Up-sampling is usually defined as increasing the sampling rate of a signal by inserting zero-valued samples between the original samples and as such, no estimation is implied for these positions.

Interpolation is defined as (Free On-Line Dictionary of Computing, 2005):

Figure 1: Courtesy of “The Daily Telegraph” newspaper, published on Thursday, September 16, 2004, No. 46427.
A mathematical procedure which estimates values of a function at positions between listed or given values. Interpolation works by fitting a "curve" (i.e. a function) to two or more given points and then applying this function to the required input.

Although "interpolation" is by far the most commonly used term for the process, we refrain from using the term as it implies generating missing samples by fitting a curve to the available neighbouring samples and hence more restrictive in the method used to enlarge the image.

We prefer to use the term "image resolution enhancement" when describing the proposed algorithms in this thesis. The proposed algorithms generate high-frequency information based on the available low-resolution image and result in generation of an enlarged higher resolution image, hence "enhancing" the resolution.

It is also interesting to note that, although the term "super-resolution" is normally used in the case of availability of multiple images, in a few instances, it is also used to describe schemes utilising only a single image. In the context of this thesis, we use the term "super-resolution" when a method uses more than one input image to generate an enlarged result and refrain from using the term when there is only a single input image.

1.2 Assumptions

There are a number of assumptions we made constraining the reviewed literature and presenting solutions to the problem in hand. The assumptions as well as the underlying reasons and possible ways to overcome these assumptions if so required are as follows:

- *Only a single input image is available.* This is a realistic assumption for digital imaging equipment capturing still images only. It is also valid for video when super-resolution methods are difficult to apply due to various restrictions. Such potential restrictions are explained below.

- *Enlargement factor is a power of 2 (quadrupling the original pixel density at each increment).* Different factors of enlargement could be achieved by combining the algorithms with a conventional interpolation scheme such as bilinear interpolation. Enlargement by a factor of $k$ where $k$ is not a power of 2 is achieved as follows: (i) Iterate the algorithm $n$ times where $n$ is the smallest integer satisfying the condition $k < 2^n$, (ii) Down-sample the image by a factor of $\frac{k}{2^n}$ using a conventional technique. For example, to realise 3x enlargement, first 4x enlarged image is obtained with the proposed algorithms followed by down-sampling with a factor of 0.75x using bilinear interpolation.
• *Both the input and output images are greyscale.* The resolution enhancement of colour images is also possible by applying the same algorithm to each colour component separately. For example, red, green and blue components of an RGB composite image could be processed independently to generate a resolution enhanced colour image.

It might be argued that super-resolution image reconstruction is a better alternative to resolution enhancement from a single source image when multiple images are available. Although super-resolution has distinctive benefits in various scenarios such as when the input images are translated versions -particularly if the translation is controlled- and there is little change in the scene (such as no moving objects and non-rigid motion), especially in the case of the object in interest going through a non-rigid motion which is typical in real life scenarios such as security video captures, super-resolution image reconstruction loses its advantages and more robust results could be obtained using a single image.

### 1.3 Contributions of the Thesis

The main contributions of this thesis could be summarised as presenting a comprehensive review of the wavelet domain image resolution enhancement methods available and proposing a number of methods for the problem.

More specifically the contributions could be listed as follows:

• Presenting a review of wavelet domain image resolution enhancement methods, as well as a comparison of these methods. This is important in that most of the research on wavelet domain image resolution enhancement methods seem to have been conducted independently and infrequently cite other relevant work. As a consequence, the information regarding the relative performance comparisons of wavelet domain methods is not easily available.

• Bringing attention to the basic wavelet domain image resolution enhancement method which we coin as wavelet-domain-zero-padding (WZP).

• Introducing a wavelet domain resolution enhancement technique utilising the local neighbourhood information.

• Proposing a resolution enhancement method by adapting the cycle spinning methodology to the resolution enhancement problem.

• Suggesting an edge rectification technique to improve the results obtained with the cycle spinning based method.

• Introducing the concept of directional cycle spinning with application to image resolution enhancement.
1.4 Publications Resulting from the Thesis

During the course of this PhD thesis, the following papers have been published or accepted for publication in journals:


The following paper has been presented in a conference:


1.5 Structure of the Thesis

In chapter 2, we introduce the image resolution enhancement problem and the conventional interpolation methods. Then we present the more advanced methods that appeared in the literature and/or used in the commercial applications. This chapter aims to give a broader view to the problem and proposed solutions before focusing particularly on the wavelet transform based methods.

Chapter 3 starts with a brief introduction to the continuous and discrete wavelet transform leading to presentation of 2-D discrete wavelet transform which is suitable for imaging applications. Then we introduce the basic methods utilising wavelet transform for the resolution enhancement and the problems associated with these basic approaches. In this section we also bring the wavelet-domain-zero padding (WZP) approach into light. A summary of the wavelet domain methods are then presented starting from the foremost method proposed by Chang et al. (1995) to the current state-of-the-art methods.

In chapter 4 we propose a wavelet domain image resolution algorithm which works by estimating the unavailable high-frequency coefficients. We then present the experimental results and provide a frequency domain analysis to the proposed algorithm.

In chapter 5, first we introduce cycle spinning. We then propose an image resolution enhancement algorithm by adapting cycle spinning to the problem in hand. Following this, we show that the
results could further be improved upon by the proposed edge rectification method. The chapter is finished with the presentation of experimental results followed by concluding remarks.

In chapter 6 we propose an extension to the cycle spinning method introduced in the previous chapter. More particularly, we show that, by benefiting from the multi dimensionality of image signals, cycle spinning could be selectively done in different directions to prevent over-smoothing and hence obtain better results. A subjective evaluation of the results is presented before concluding this chapter.

Chapter 7 presents the comparative assessment of all the proposed methods as well as a number of competing methods. The results are presented in the form of Peak-Signal-to-Noise Ratio (PSNR) measure. This chapter aims to present subjective results to complement to the subjective results presented in individual chapters.

Chapter 8 concludes this thesis by summarising the work.
CHAPTER 2

Image Resolution Enhancement

2.1 Introduction

Resolution enhancement of pictorial data is desirable in many applications such as monitoring, surveillance, medical imaging and remote sensing when images at desirable resolution levels are not available. It is a classic signal interpolation problem and several conventional approaches such as zero-order interpolation (sample-and-hold), bilinear and spline interpolation are widely used. However, undesirable levels of smoothing across salient edges in the higher resolution images obtained using these conventional methods resulted in a search for more effective algorithms. This interest in generation of higher quality images resulted in proposal of several approaches.

In this chapter, we discuss the image acquisition process and the resolution implications. Then we present the signal processing approach to the interpolation problem and discuss the conventional methods. Next, we classify the advanced interpolation methods into groups in relation to their core approaches to the image resolution enhancement problem and present an overview of the proposed methods in the literature. We conclude the chapter with a critical summary of the related literature.

2.2 Image Acquisition Process

The most common digital image acquisition technology in modern imaging is Charge Coupled Device (CCD). A CCD sensor is comprised of a rectangular grid of electron-collecting regions laid over thin silicon wafer. Each element of the grid records the amount of light energy reaching on its surface. The number of electrons recorded at the cell located at row \( m \) and column \( n \) of a CCD can be modelled as:

\[
I(m, n) = \int_{\lambda} E(p, \lambda) R(\lambda) q(\lambda) d\lambda
\]

where \( I(m, n) \) is the electron collection time and the integral is calculated over the spatial domain \( S(m, n) \) of each CCD grid element and the range of wavelengths to which the CCD has a nonzero response, \( E \) is the power per unit area and unit wavelength (irradiance) arriving at the point \( p \), \( R \) is the spatial response of the site, and \( q \) is the quantum efficiency of the device (i.e. the number of electrons generated per unit of incident light energy) (Forsyth and Ponce, 2003). This accumulated charge over the integration time (exposure time) is then converted into a proportional voltage measure by the output amplifier and subsequently converted into a digital value.
Digitisation most commonly generates an 8-bit number while 12-bit resolution has also been commonplace in professional imaging equipment recently.

This process is demonstrated for 64x64 and 32x32 element CCDs in top and bottom images of Figure 2 respectively. Each square in these grids (pixel) corresponds to a CCD element, which produces a single quantity proportional to the light intensity corresponding to its spatial position. As can be seen from these images, fewer CCD cells result in a lower resolution image which is blockier.

Taking this model into account, a similar process could also be proposed for acquisition of an LR (Low Resolution) image from an HR (High Resolution) image. The pixels of the LR image could be generated from the integration of pixel values of the HR image which spatially correspond to the same location followed by normalisation. This process also corresponds to first smoothing of the image using low-pass filtering followed by discarding of samples.
2.3 Linear Interpolation Methods

2.3.1 Signal processing approach to interpolation

Theoretically, interpolation of a digital signal involves generation of corresponding continuous time signal followed by resampling at a higher rate. However this is not a feasible option in practice because of the non-ideal nature of analog reconstruction filter and digital to analogue and analogue to digital converters. Hence an all discrete-operation is desirable for the sampling rate conversion.

Let \( x[m] \) is the available discrete low sampling rate signal whose sampling rate is sought to be enhanced by a factor of \( \frac{L}{N} \) where \( L,N \in \mathbb{Z} \), \( L > N \). Denoting the underlying continuous signal as \( x_c(m) \), the low sampling rate discrete signal could be written as \( x[m] = x_c(mT) \) where \( T \) is the sampling rate determined by the low-resolution acquisition process. Similarly the HR image could be written as \( y[m] = x\left[\frac{mN}{L}\right] = x_c(mT') \) where \( T' = \frac{TN}{L} \) and \( m = 0, \pm\frac{L}{N}, \pm 2\frac{L}{N}, \ldots \).

A system for obtaining \( y[m] \) from \( x[m] \) is shown in Figure 3 (Oppenheim and Schafer, 1989). In this figure, the sampling rate expander generates an expanded signal with missing data points having a value of 0:

\[
x_e[m] = \begin{cases} x[\frac{mN}{L}], & m = 0, \pm\frac{L}{N}, \pm 2\frac{L}{N}, \ldots \\ 0, & \text{otherwise} \end{cases}
\]

This expanded signal is then fed as input to the lowpass filter, which has a cut-off frequency of \( \frac{\pi N}{L} \) and a gain of \( \frac{L}{N} \). This lowpass filter reconstructs the missing data points.

\[ \text{Figure 3: Sampling rate converter} \]
This process could be better illustrated in frequency domain. Equation (2.2) could equivalently be written using dirac-delta operator as follows:

\[ x_e[m] = \sum_{k=-\infty}^{\infty} x[k] \delta[m - \frac{kL}{N}] \]  

(2.3)

The Fourier transform of which is:

\[ X_e[e^{j\omega}] = \sum_{n=-\infty}^{\infty} \left( \sum_{k=-\infty}^{\infty} x[k] \delta[m - \frac{kL}{N}] \right) e^{-j\omega n} \]

\[ = \sum_{k=-\infty}^{\infty} x[k] e^{-j\omega \frac{k}{N}} \]  

(2.4)

Implying that the expander produces a frequency scaled version of its input (scaled by a factor of \( \frac{L}{N} \)). This effect is illustrated in Figure 4.

An inspection of this figure reveals that, if all the replicates of the frequency scaled versions except the ones at multiples of \( 2\pi \) are removed and the amplitude is converted from \( \frac{1}{T} \) to \( \frac{1}{T'} \), interpolated signal would be obtained. This could be achieved using a lowpass filter with a gain of
\( \frac{L}{N} \) and a cut-off frequency of \( \frac{\pi N}{L} \). The ideal lowpass filter and resulting interpolated signal are illustrated in Figure 5. The impulse response of this filter corresponds to equation (2.5) which is the well-known sinc signal, requiring infinite number of samples prohibiting practical implementations.

\[
h[m] = \frac{\sin \left( \frac{\pi mN}{L} \right)}{\pi mN}
\]  \hspace{1cm} (2.5)

In practice, it is not possible to implement ideal lowpass filters. Instead, approximations are utilised or simpler interpolation procedures, such as zero-order hold and linear interpolation, are applied.

It should also be noted that the above analysis makes the assumption that the low sampling rate discrete signal \( x[m] \) was obtained without aliasing.

![Figure 5: Ideal low-pass filter, recovering the interpolated signal.](image)

Zero-order hold and linear interpolation could be achieved using the above procedure and using a filter with impulse response as given in equations (2.6) and (2.7) respectively.
Chapter 2. Image Resolution Enhancement

\[ h_{2k}[m] = \begin{cases} 1, & |m| \leq \frac{N}{L} \\ 0, & \text{otherwise} \end{cases} \quad (2.6) \]

\[ h_{2l}[m] = \begin{cases} 1 - |m| & \frac{N}{L}, \quad |m| \leq \frac{N}{L} \\ 0, & \text{otherwise} \end{cases} \quad (2.7) \]

2.3.2 Interpolation in signal domain

An alternative way to visualise interpolation is to imagine it as a mapping from a coarser grid to a finer grid. If the mapping is done so that the value in the finer grid is determined by the nearest corresponding value in the coarser grid, the process is called nearest neighbour interpolation or zero-order hold interpolation. Although this method is simple to implement, it has the drawback of producing blocky images with jagged edges. Better results could be obtained by assuming some degree of smoothness in the image and using bilinear or bicubic interpolation methods.

Bilinear interpolation uses the four nearest neighbours. Figure 6 illustrates the bilinear interpolation, where \(x_1, x_2, x_3\) and \(x_4\) are the known pixel values corresponding to low-resolution grid while \(y_1, y_2, \ldots, y_6\) are the pixel values at the high-resolution grid (same factor of interpolation is assumed in both directions). Effectively, bilinear interpolation is achieved in a separable way, i.e. first in horizontal direction and followed by in vertical direction or vice versa.

Ignoring the boundary conditions and hence calculation of boundary pixel values, the interpolated pixels are obtained as follows:

\[ y_1 = x_1 \]
\[ y_2 = ax_1 + (1 - a)x_2 \]
\[ y_3 = x_2 \]
\[ y_4 = ax_2 + (1 - a)x_3 \]
\[ y_5 = x_3 \]
\[ y_6 = ax_3 + (1 - a)x_4 \]
\[ y_7 = x_4 \]
\[ y_8 = ax_4 + (1 - a)x_5 \]
\[ y_9 = x_5 \]

and

\[ y_5 = ay_2 + (1 - a)y_6 = ay_4 + (1 - a)y_5 = a^2x_1 + a(1 - a)x_2 + a(1 - a)x_3 + (1 - a)^2x_4 \]
Figure 6: Illustration of bilinear interpolation

Although bilinear interpolation produces good results at smooth parts of the image, the results are not as good at discontinuities. The inherent assumption of continuity results in over-smoothing of the salient image features.

For better results than the linear interpolation, one can resort to cubic interpolation which was introduced by Keys (1981). Keys introduced a set of functions which are made of piecewise cubic polynomials. Cubic interpolant functions are defined as follows depending on parameter $c$:

$$h_{cub}(x) = \begin{cases} 
(c+2)x^3 - (c+3)x^2 + 1, & 0 \leq |x| < 1 \\
-5c x^2 + 8cx + 4c, & 1 \leq |x| < 2 \\
0, & 2 \leq |x| 
\end{cases}$$

(2.9)

$c$ is reported to be $-0.5$ in (Keys, 1981).

The interpolating functions for nearest-neighbour, linear and cubic and their frequency responses are illustrated in Figure 7. Images obtained by interpolation utilising these functions are illustrated in Figure 8 for an enlargement factor of 4 (128x128 to 512x512).

Although cubic interpolation results in better results than nearest-neighbour and bilinear, it tends to overshoot sharp discontinuities, producing a ringing effect at edges (Schultz and Stevenson, 1994).

Interpolation using a higher number of neighbouring pixels is also possible by using higher order interpolants. Hou and Andrews (1978) investigated the suitability of "B-splines" for interpolation. "B-splines" is a family of functions with fixed degree area piecewise polynomials. Nth degree B-spline is a piecewise polynomial of degree $n$ and denoted as $\beta^n$, where $n \in \mathbb{N}$. Interpolation with $\beta^0$ converges to nearest-neighbour interpolation and with $\beta^1$ converges to linear.
Figure 7: (a) Zero-order hold, (b) linear and (c) cubic interpolation functions and their corresponding frequency responses.
Figure 8: Images enlarged by 4x (128x128 to 512x512) using (a) zero-order hold, (b) linear and (c) cubic interpolation functions. The images on the left are magnified portions of these enlarged images and 8x amplified difference images where mid-grey represents zero error are shown on the right.
interpolation. Interpolation using $\beta^3$ results in a similar result to the cubic interpolation of Keys (1981). As a general rule, the higher the degree of B-spline, the closer to sinc it is and the better are the results (Thevenaz et al., 2000). However this comes with the cost of more computational complexity.

Bilinear and bicubic interpolation techniques are the standard image resolution enhancement techniques, which are commonly found in popular computer imaging applications such as Adobe® PhotoShop® (Adobe, 2005) and PaintShop™ Pro® (Corel, 2005).

For a more detailed discussion of B-splines and their application to image interpolation, readers are referred to (Unser et al., 1995), (Unser, 1999) and for a comprehensive review of interpolation methods to (Meijering, 2002). The latest developments in linear interpolation techniques could be found in (Blu et al., 2004) where it is demonstrated that the linear interpolation performance can be increased to the levels that of cubic interpolation by applying a simple shift.

### 2.4 Advanced Methods

Despite the simplicity and popularity of the linear interpolation techniques described in the preceding section, images generated using these techniques are not of sufficient quality as they do not take the characteristics of the data that is being interpolated into account. Prominently, the linear techniques fail to preserve the salient image features such as edges and tend to smoothen them out. Therefore more advanced techniques are required for more satisfactory reconstructions.

In this section, we classify these techniques into the following groups and present with an emphasis to the latest developments:

- Discontinuity adaptive techniques
- Regularisation based approaches
- Fractal based methods
- Discrete Cosine Transform (DCT) based methods
- Wavelet based techniques

#### 2.4.1 Discontinuity adaptive techniques

Discontinuities or more specifically edges are the most important features in an image and most of the visual information is deducted from the edge information which leads naturally to image resolution enhancement methods which exploits this information.
Chapter 2. Image Resolution Enhancement

This subset of techniques is comprised of methods that use standard interpolation techniques such as bilinear interpolation in the smooth regions of the image while in the neighbourhood of an edge resort to various nonlinear techniques to preserve the sharpness and prevent over-smoothness of the edges.

The earliest example of this class of methods is the one proposed by Martinez and Lim (1989). Their method is designed to generate a frame from a single video field in a 2 to 1 interlaced sampling set-up such as PAL and NTSC television signals. Hence the interpolation is carried out over the vertical direction in vertically under-sampled fields to estimate the missing lines and no processing is done in the horizontal direction. A constant luminance model, which assumes pixel intensities remain constant along the path of motion, is used. First the line shift, i.e. the spatial shift between the adjacent lines are calculated using a group of 10 pixels (5 pixels in each adjacent line) and then the interpolation is done along the direction of the estimated line shift. Subsequently, Ayazifar and Lim (1992) suggested a generalisation of this line shift model through a concentric circular shift model in which small segments of concentric arcs of an image are related to adjacent arcs on either side of them by an angular shift. If there is a sufficient fit, the estimation is done using the concentric circular shift model, if not, then the algorithm falls back to Martinez and Lim's line shift model to estimate the pixels. By adapting to the contours as well as the linear edges, artefacts are reduced. Both these algorithms were designed for generation of frames from vertically under-sampled fields and operate only in vertical direction whereas horizontal resolution is not increased. While these methods produce good quality results in the regions where the contours can be estimated correctly, the contour estimation could be adversely affected by the lack of lines. Due to under-sampling in the vertical direction, in the high spatial frequency regions, incorrect estimations result in objectionable artefacts in the resolution enhanced image.

Jensen and Anastassiou (1995) used edge fitting within 3x3 overlapping windows to generate edges in the high resolution image. The proposed algorithm assumes that edges partition the pixels around them into two regions with nearly constant intensity values. The edge parameters are estimated with the proposed edge fitting operator and using these parameters as templates and the aforementioned assumption of constant intensity regions, the edges on higher resolution grid are constructed. However, edge detection using 3x3 regions might not be robust in many cases.

In (Thumhofer and Mitra, 1996), each pixel is categorised into three sets (constant, oriented, irregular) and then a different interpolation algorithm is applied to each different set. Also a quadratic Volterra filter is applied as a post-processing operation to improve the sharpness and the contrast of the enlarged image. A more recent pixel classification based method is proposed in (Atkins et al., 2001). This approach differs from the earlier work of Thumhofer and Mitra in that, the classification and interpolation filter parameters are generated by training where maximum
Chapter 2. Image Resolution Enhancement

likelihood (ML) estimation is used. The training is done using a number of ground-truth high-resolution images and low-resolution images obtained from these images by filtering followed by down-sampling. Once obtained, these parameters can then be used for images not in the training set. To generate the enhanced resolution image, the low-resolution image pixels are classified and the appropriate interpolation operator obtained at training is then applied which acts as a projection operator converting the low-resolution pixels into high-resolution pixel estimates.

Bayrakeri and Mersereau (1995) propose a directional interpolation scheme where the effect of interpolation in a particular direction is relative to the directional derivative magnitude in that direction. A similar approach is proposed in (Battiato et al., 2002) where a weighted interpolation is carried out controlled by the gradient. More specifically, the edge directions are determined using the gradient and the pixels are estimated as average of known pixels at the direction of the gradient (which implies the maximal directional derivative). The technique in (Arandiga et al., 2003) interpolates only from one side of the data (for either side of the edge) when close to an edge to preserve sharpness of the edges in the high resolution image to be generated.

Another gradient directed method was proposed by Jiang and Moloney (2002). In this method, the problem is formulated as a partial differential equation with both smoothness and orientation constraints. By solving the partial differential equation, the conversion rules from low-resolution grid to high-resolution grid is found and subsequently used to obtain the high-resolution images.

An algorithm which requires training is advocated in (Carrato and Tenze, 2000). In this method, first the desired outputs for various edge patterns are identified and an interpolator is derived using these desired responses with an optimisation process. Subsequently, this derived interpolator is used to obtain the enlarged images.

Morse and Schwartzwald (1998) propose an iterative approach where the initial image which is generated using bilinear interpolation is refined by iteratively reconstructing the isophotes (intensity level curves) using constrained smoothing. The method aims to cure the jagged edges while maintaining the intensity levels and the isophote topology and ordering. They further extended their work in (Morse and Schwartzwald, 2001).

A subclass of discontinuity adaptive techniques is the edge directed interpolation. Edge directed interpolation techniques try to detect the edges and then interpolate along these avoiding interpolation across edges -which is undesirable as this tends to smooth-out the edges and diminish the image details-. Allebach and Wong (1996) propose a technique where sub-pixel edge estimation using an approximation of the Laplacian-of-Gaussian filter is performed. Then a variation of bilinear interpolation which prevents interpolation across edges is used.

New Edge Directed Interpolation (NEDI) of Li and Orchard (2001) is based on the edge directed interpolation proposed by Allebach and Wong (1996). In NEDI algorithm, the local covariance
matrix is used for estimating the missing pixels. Covariances are calculated in a square window in the low-resolution image. The covariance gives a measure of relative flow of differential characteristics by which the interpolator could be modified so that its support is along the edges only. Effectively, as opposed to the bilinear interpolation which averages the neighbouring pixels to estimate the unknown pixel, the covariance of each neighbour - calculated individually for each pixel- is used as a weight. To reduce the computational complexity, covariance based adaptive interpolation is only done at edge regions, while bilinear interpolation is used in the smooth regions.

Biancardi et al. (2002) used bicubic interpolation to identify the position of the zero-crossings to locate the edges and then used polynomial interpolation of a degree proportional to the magnification factor. The polynomial is constrained to produce a specific value for the magnitude of the gradient at the exact location of the edge identified. This value is calculated to interpolate the values of magnitude of the gradient of the neighbouring samples.

Hwang and Lee (2004) suggested applying inverse gradient to the bilinear and bicubic interpolation. Inverse gradient helps adapting to the edge direction and prevent over-smoothing. In the context of this work, the authors derive two interpolation algorithms taking bilinear and bicubic interpolation as bases. In both cases, the gradient in horizontal, vertical and diagonal directions are calculated and the corresponding weights for the bilinear or bicubic interpolation are divided by the magnitude of this gradient. Hence, a much smoother interpolation is achieved across edges, preventing the edge blurring. The proposed algorithm also allows for arbitrary magnification factors other than multiplies of two as it is a direct extension to the bilinear and bicubic interpolation techniques.

Kim et al. (2003) proposed obtaining each high-resolution image pixel by using adaptive weighting followed by summation of 4 low frequency image pixels (windows). In essence the algorithm is a combination of zero-order hold interpolation and bilinear interpolation. The algorithm aims to utilise the nearest zero-order hold interpolation at the edge locations to benefit from the sharp edges produced by this method as opposed to blurring resulting from averaging of bilinear interpolation.

In (Wang and Ward, 2003) it is advocated that one shortcoming of edge-directed methods is not differentiating between the edge and ridge pixels. As the ridge lines are usually thinner than edge lines, ridge reconstructions generally appear twisted or broken. The authors suggest a method to detect ridges and apply the directional interpolation to these ridges as well as the edges to preserve the ridge characteristics.

In (Chuah and Leou, 2001), the initial high-resolution image estimate is partitioned into \( M \times M \) non-overlapping blocks where \( M \) is the expansion ratio in a single direction and all the pixels are
classified as non-edge or edge pixels which are further classified according to their orientations. Then a filter incrementing or decrementing the pixel value by one followed by a 2-D edge sensitive filter is applied in an iterative fashion until the convergence criteria is achieved. The former filter aims to increase the high frequency content by shifting the pixels away from the mean value of the block they belong to. Then the proposed 2-D edge sensitive filter is applied to further sharpen the edges. Finally to reduce the artefacts due to block based processing, a de-blocking filter is applied.

Recently, optimal recovery has also been applied to image resolution enhancement problems. In this approach, image resolution enhancement is viewed as a problem of estimating missing samples of an image. Then, this problem is examined using the theory of optimal recovery by modelling the image as belonging to an ellipsoidal signal class $K$ as follows (Muresan and Parks, 2001a; 2001b):

$$K = \{x \in \mathbb{R}^n : x^T Q x \leq \varepsilon \}$$  \hspace{1cm} (2.10)

$x$ is a subset of the image containing the pixel that is to be estimated and $Q$ is derived adaptively from the neighbourhood using a set of training vectors constructed from the low resolution image. Missing pixels are then reconstructed using optimal recovery. Note that when $Q$ is constant, this method could be made equivalent to bilinear or bicubic interpolation. A similar approach is proposed in (Muresan and Parks, 2004).

A commercial image resolution enhancement product utilising edge adaptive interpolation is pxl Smartscale™ from Extensis™ (A Celartem™ Company) (2005).

### 2.4.2 Regularisation based approaches

These techniques make assumptions about the image characteristics and make use of a priori image constraints. The initial high-resolution image estimate which is obtained by a conventional method such as bilinear or bicubic interpolation from the low-resolution image is iteratively refined until a convergence criterion is satisfied. The high-resolution image to be constructed is expected to conform to the a priori image model.

A popular regularisation based work is Schultz and Stevenson’s Bayesian approach for the image resolution enhancement (1994). In this work, the authors place the problem into a statistical framework using maximum a posteriori (MAP) estimation. A convex Huber-Markov Random Field (HMRF) function model is assumed for the image; this model helps preserve discontinuities by penalising the discontinuities to a lesser extent compared to other commonly used models such as Gauss-Markov random field. Then, the problem is formulated as an MAP estimation problem maximising the log likelihood which is converted into a minimisation problem using Bayes’ rule.
Chapter 2. Image Resolution Enhancement

The resolution enhanced image is generated by minimising this function by MAP estimation. Disadvantages of this method are that it requires three parameters that need to be tuned for different images for the best results and the deficiency of HMRF in modelling sharp edges. This method has also been adapted by the authors to super-resolution problems where multiple low-resolution frames are available (Schultz and Stevenson, 1996).

The method proposed by Ratakonda and Ahuja (1998) uses Projection Onto Convex Sets (POCS) methodology to find the intersection of the convex sets which complies with the following rules: (i) the Discrete Fourier Transform (DFT) of the low-resolution image should be the same as the low-frequency part of the resolution enhanced image and (ii) the pixel values for the non-edge locations are allowed to vary more than the ones at edge locations. Although POCS is a well-established methodology in super-resolution image reconstruction, in the availability of multiple images it isn't a commonly used technique in the context of image resolution enhancement from a single image as the results suffer from blotching artefacts.

In a very recent work, Aly and Dubois (2005) suggested utilisation of total-variation regulariser as well as a new observation model. The total-variation regulariser is favoured as it preserves the sharpness of the edges (Strong and Chan, 2003) and it results in edges with the least oscillation in the isophotes resulting in smooth contours avoiding the jagged and oscillatory solutions.

2.4.3 Fractal and Triangulation Based Methods

Fractals offer an alternative means to achieve image resolution enhancement. Fractals are used to exploit the self-similarity between low-resolution and high-resolution images. Based on this property, higher resolution images are reconstructed using well-defined contractive mappings. The word fractal was coined by Mandelbrot (1977) to describe self-similar objects which had no clear dimension. Fractal image coding is based on the theory of iterated contractive transformations by Barnsley (1988) and first proposed by Jacquin (1992). In this technique, images are coded by exploiting the self-similarity property. Hence the reconstructed images can be of any size regardless of the initial coded image sizes. Alkhansari et al. (1997) and Honda et al. (1999) proposed image enlargement algorithms based on this principle. However, in practice, due to usage of sub-optimal contractive mappings, the enlarged images are far from satisfactory (Chung et al., 2003) and perform fractionally better if not inferior to bilinear interpolation (Wohlberg and Jager, 1999). In addition, being a block based processing method, resolution enhanced images suffer from blocking artefacts. Chung et al. (2003) proposed an enhancement layer which uses the error information from the low-resolution image, interpolates the error and subtracts from the resolution enhanced image. The method also involves application of the

---

1 An object is said to be self-similar if it looks "roughly" the same on any scale (Mathworld, 2005).
algorithm twice using blocks with different centre points which are subsequently averaged to generate the resulting image. This processing is aimed at reducing the blocking artefacts by averaging them out.

Triangulation is a popular method in geometric modelling and recently it has found some use in image resolution enhancement applications. As the standard interpolation methods are applied along the image coordinate axes, the artefacts appear along those axes as well. By utilising data-dependent triangulation (DDT) (Dyn et al., 1990), these artefacts could be moved off the sampling axes, reducing the visual impact of these errors (Yu et al., 2001). The data dependent triangulation works by matching the edges in the image. This idea has been further enhanced by Su and Willis (2004) so that it doesn’t require iterations and any cost function.

A commercially available image resolution algorithm utilising the fractal technique is “Genuine Fractals algorithm” from LizardTech (LizardTech, 2004). Availability of a trial version of this tool facilitates evaluation of this algorithm with the test images and enables usage as a benchmark.

2.4.4 Discrete Cosine Transform (DCT) based methods

Discrete Cosine Transform (DCT) based methods exploit the idea of estimation of the missing high-frequency information in the DCT domain by using the existing information followed by application of inverse DCT to obtain the resolution enhanced image. Other techniques have also been proposed to increase the computational efficiency of spatial domain techniques by carrying out the necessary operations in the transform domain when the transform domain data is already available.

An approach to restore the high-frequency components in the DCT domain was proposed in (Shinbori and Takagi, 1994). The authors use Gerchberg-Papoulis (GP) iterative algorithm with two constraints: correct information in a pass-band is known and the spatial extent of an image is infinite.

In the method proposed by Dugad and Ahuja (2001), the image is divided into 4 x 4 non-overlapping blocks. Then DCT is performed on each block. The resulting 4 x 4 DCT blocks are then placed in the low-frequency quadrant of an 8 x 8 block and the remaining coefficients are set to zero. The enlarged image is then obtained by performing inverse DCT on each 8x8 block. This algorithm has been further improved upon by Mukherjee and Mitra (2002). While the high-frequency coefficients are still set to zero, the low-frequency coefficients are multiplied by a factor depending on their frequency content when copied. Inherently these methods are limited to enlargement factors which are power of two. To eliminate this limitation, Park et al. (2003) (see also (Park and Park, 2004)) suggested modifications using the multiplication-convolution property.
of DCT (Martucci, 1994) to the method in (Mukherjee and Mitra, 2002). They also proposed a new filter with more suitable characteristics. A common property of these algorithms is that they all involve block based processing which inevitably result in blocking artefacts.

Martucci (1995) proposed that the equivalent image resolution enhancement process whereby the high resolution image is obtained by sampling rate expansion followed by low-pass filtering as explained in section 2.3.1 could be done in the discrete cosine transform domain. The proposed technique allows carrying out the interpolation using the methods which could be represented as a Finite Impulse Response (FIR) filter in the spatial domain -such as bilinear or bicubic interpolation- by making use of the convolution-multiplication property of DCT.

An alternative DCT domain image resolution enhancement approach has been proposed by Hong et al. (1996). The authors proposed extraction of the edge information directly from the DCT coefficients using some edge templates and finding the best match to these templates. In non-edge regions, bilinear interpolation is applied while in edge regions a one dimensional directional interpolation is applied according to the direction of the edge. Finally, ellipsoidal Gaussian lowpass filters are applied to eliminate the blocking artefacts.

2.4.5 Wavelet Based Techniques

Wavelet based techniques use a model whereby the available LR image is assumed to correspond to the low-frequency subband of HR image that is sought to be estimated. In this framework, resolution enhancement is achieved by means of estimating the unavailable virtual high-frequency subbands followed by inverse wavelet transform operation.

This subset of techniques exploits the correlation between the coarser subband parent wavelet coefficients and finer subband child coefficients (persistency) as well as the non-Gaussianity of wavelet coefficients in some cases.

As this thesis deals, in particular, with the wavelet based techniques, these are examined in more detail in the following chapter.

2.5 Conclusions

This chapter doesn’t aim to provide a review of all the image resolution enlargement algorithms. As the literature about the image resolution enhancement is quite rich and there are numerous published materials on the subject, this would be a bold undertaking. Hence, in this chapter, only the fundamental and distinct methods which form the basis for many methods in the literature as well as the most recent developments in the field are presented. Among these methods, the discontinuity adaptive methods are the most popular while there has been some interest in the
regularisation based methods recently. On the other hand, most of the standard imaging tools still utilise bilinear or bicubic (or nearest neighbour when the speed is of utmost importance) interpolation as standard image resolution enhancement algorithms.
CHAPTER 3

Wavelet Domain Image Resolution Enhancement

3.1 Introduction

As mentioned in the previous chapter, the high resolution images obtained with conventional approaches are not satisfactory for various purposes which led way to proposal of new methods for image resolution enhancement. Among these methods, wavelet domain techniques, which are relatively recent, have been shown to be effective in overcoming some of the problems associated with conventional treatment.

In this chapter, we first introduce the continuous and discrete wavelet transform before discussing the utilisation of wavelet domain for image resolution enhancement and the problems associated with this particular set of methods. After discussing the basic techniques for wavelet domain resolution enhancement, we present the more advanced methods in the literature. We conclude this chapter by discussing the pattern of research conducted on the subject and the potential problems that can be encountered when implementing and evaluating these algorithms.

3.2 Wavelet Transform

3.2.1 Continuous Wavelet Transform (CWT)

The Continuous Wavelet Transform (CWT) of a 1-D signal \( x(t) \) is defined as (Mallat, 1998):

\[
CWT_x^\psi (\tau, s) = \frac{1}{\sqrt{|s|}} \int_{-\infty}^{\infty} x(t)\psi(t-\tau) dt
\]  
(3.1)

The transformed signal is a function of two variables, \( \tau \) and \( s \) which are the translation and scale parameters respectively. \( \psi(t) \) is the transforming signal which is called the mother wavelet.

CWT is a reversible transform provided that the equation (3.2) called admissibility criterion is satisfied. This implies that \( \psi(0) = 0 \) and \( \psi(u) \rightarrow 0 \) fast enough to make \( C_\psi < \infty \).

\[
C_\psi = \int_{-\infty}^{\infty} \frac{||\psi(u)||^2}{|u|} du < \infty
\]  
(3.2)

The inverse CWT is defined as:
Chapter 3. Wavelet Domain Image Resolution Enhancement

\[
    x(t) = \frac{1}{C^2} \int \int \text{CWT} \left( \psi \left( \frac{t - \tau}{s} \right) \right) \psi \left( \frac{\tau}{s} \right) d\tau ds
\]  

(3.3)

Essentially, wavelet transforms look for similarities between the mother wavelet and the original signal. This is analogous to conventional transform techniques such as Fourier and Cosine Transforms which capture similarities between a fixed basis function (namely a sine or cosine) and the original signal. Unlike these conventional transform techniques, the wavelet transform enables using basis functions which are not infinite in time. A vast number of wavelets could be used for wavelet transform which allows selection of a wavelet for a specific application. In equation (3.1), the translation parameter \( \tau \) acts as a shifting operator moving the mother wavelet on the signal that is being transformed and the scaling parameter \( s \) dilates or expands the mother wavelet depending on the scale. High scales, where the mother wavelet is expanded, correspond to global information about the signal while the lower scales give finer detail. Unlike the transform techniques such as Fourier Transform, the wavelet transform provides a way to analyse the signal in different time and frequency resolutions. Two major advantages of the wavelet transform over these are as follows:

- The transform domain still carries time information (or spatial location information in images) as well as the frequency information.
- Wavelet transform inherently provides a means for multiresolution analysis; different wavelet scales present information belonging to different frequency subbands.

Alongside these aforementioned advantages, CWT has two major weaknesses for practical purposes: it is highly redundant and not suitable for discrete processing making way for Discrete Wavelet Transform (DWT).

3.2.2 Discrete Wavelet Transform (DWT)

A discrete signal \( x[m] \) could be decomposed using discrete wavelet transform (Mallat, 1998):

\[
    x[m] = \sum_k a_{J,k} \phi \left( \frac{m}{2^j} - k \right) + \sum_{j=1}^J \sum_k d_{j,k} \psi \left( \frac{m}{2^j} - k \right)
\]  

(3.4)

Where \( \phi(m) \) and \( \psi(m) \) are scaling and detail wavelet functions (i.e. the basis functions) respectively, \( J \) denotes the number of levels of the transform. \( a_{J,k} \) and \( d_{j,k} \) represent, respectively, approximation and detail coefficients which are found as follows:

\[
    a_{J,k} = \sum_m x[m] \phi \left( \frac{m}{2^j} - k \right)
\]  

(3.5)
Chapter 3. Wavelet Domain Image Resolution Enhancement

\[ d_{j,k} = \sum_{m} x[m] \psi \left( \frac{m}{2^j} - k \right) \]  (3.6)

Similar to CWT, DWT works by analysing the similarities between the signal and dilated and shifted versions of the detail wavelet function. An approximation of the signal is obtained using shifted versions of the scaling wavelet function. Note that although the DWT is dependent on a scaling function as well as a wavelet function, whereas CWT is only dependent on a wavelet function. This is due to starting scale \( j \) of CWT being \(-\infty\), hence eliminating scaling function dependency (Gonzales and Woods, 2002)

3.2.3 Signal Processing Approach to DWT

In practice, wavelet transforms could be implemented by filtering and down-sampling the signal (Mallat, 1998). Filtering is basically convolution of the signal with the impulse response of the filter. Letting \( g[m] \) be the impulse response of a filter, this is written as:

\[ x[m] * g[m] = \sum_{k=\infty}^{\infty} x[k] g[m-k] \]  (3.7)

Letting \( L(z) \) and \( H(z) \) form an analysis filter bank while \( \hat{L}(z) \) and \( \hat{H}(z) \) form a synthesis filter bank where \( L(z) \) and \( \hat{L}(z) \) are low-pass filters and \( H(z) \) and \( \hat{H}(z) \) are high-pass filters, the discrete wavelet decomposition is shown in Figure 9 and corresponding reconstruction in Figure 10.

After passing the signal through a low-pass filter (or a high-pass filter) which removes half of the frequency content, the half of the samples become redundant according to Nyquist’s rule and can be discarded without any loss of information. After this process, the approximation coefficients contain the frequency content in \([0, \pi/2]\) and the detail coefficients the remaining \([\pi/2, \pi]\). Further levels of wavelet transform is done by repetition of the same process on the approximation part, hence the next level divides the frequency content of \([0, \pi/2]\) into \([0, \pi/4]\) and \([\pi/4, \pi/2]\) and so on.

![Figure 9: Wavelet Decomposition (analysis).](image-url)
The group of low- and high-pass filters which allows perfect reconstruction are called 'Quadrature Mirror Filters' (QMF).

Noting that for a discrete-time signal \(x[m]\) its downsampled by two version \(x_d[m]\) could be written in z-domain as:

\[
X_d(z) = \frac{1}{2} [X(z^{1/2}) + X(-z^{1/2})] \tag{3.8}
\]

and expansion of \(x[m]\) by two \(x_s[m]\) as:

\[
X_s(z) = X(z^2) \tag{3.9}
\]

and considering the analysis procedure followed by the synthesis, the output of the entire system could be written as:

\[
\hat{x}(z) = \frac{1}{2} [L(z)\hat{L}(z) + H(z)\hat{H}(z)]X(z) + \frac{1}{2} [L(z)\hat{L}(-z) + H(z)\hat{H}(-z)]X(-z) \tag{3.10}
\]

Recognising that the right hand side of this equation represents aliasing (as it contains \(-z\) dependence), for perfect reconstruction, it should be eliminated. This could be achieved by imposing the following condition:

\[
L(z)\hat{L}(-z) + H(z)\hat{H}(-z) = 0 \tag{3.11}
\]

where \(\hat{L}(z)\) and \(\hat{H}(z)\) are the perfect reconstruction filters. Equivalently:

\[
\hat{L}(z) = H(-z) \quad \text{and} \quad \hat{H}(z) = -L(-z) \tag{3.12}
\]

and using the inverse z transform of \(X(-z)\) as following:

\[
X(-z) \leftrightarrow (-1)^n x[m] \tag{3.13}
\]

The reconstruction filters could be written as:
Chapter 3. Wavelet Domain Image Resolution Enhancement

\[ \hat{h}[m] = (-1)^n h[m] \]

and

\[ \hat{h}_0[m] = (-1)^{n+1} h[m] \]  \hspace{1cm} (3.14)

Equations (3.13) and (3.14) imply that when the decomposition filter pair is known, then the low-pass reconstruction filter could be generated by inverting the sign of every other coefficient of the high-pass decomposition filter. Similarly, the high-pass reconstruction filter could be generated by inverting the sign of every other coefficient of the low-pass decomposition filter.

Decomposition wavelet filter coefficients for the Haar wavelet are given in Table 1 and for Daubechies 9/7 in Table 2. Frequency responses of these filter pairs are given in the Figure 11 and Figure 12 respectively (Mallat, 1998).

<table>
<thead>
<tr>
<th>Low-Pass Coefficients</th>
<th>High-Pass Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.70710678118654</td>
</tr>
<tr>
<td>0</td>
<td>0.70710678118654</td>
</tr>
</tbody>
</table>

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.02674875741081</td>
</tr>
<tr>
<td>0</td>
<td>0.02674875741081</td>
</tr>
</tbody>
</table>

Table 1: Haar filter coefficients.

<table>
<thead>
<tr>
<th>Low-Pass Coefficients</th>
<th>High-Pass Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>-0.01686411844287</td>
</tr>
<tr>
<td>3</td>
<td>0.04563588155713</td>
</tr>
<tr>
<td>2</td>
<td>-0.02877176311425</td>
</tr>
<tr>
<td>1</td>
<td>-0.29563588155713</td>
</tr>
<tr>
<td>0</td>
<td>0.55754352622850</td>
</tr>
<tr>
<td>-1</td>
<td>0.04563588155713</td>
</tr>
<tr>
<td>-2</td>
<td>-0.29563588155713</td>
</tr>
<tr>
<td>-3</td>
<td>-0.02877176311425</td>
</tr>
<tr>
<td>-4</td>
<td>0.04563588155713</td>
</tr>
</tbody>
</table>

Table 2: Daubechies 9/7 filter coefficients.
Chapter 3. Wavelet Domain Image Resolution Enhancement

Figure 11: Frequency responses of Haar filter pair.

Figure 12: Frequency responses of Daubechies 9/7 filter pair.
3.2.4 2-D Discrete Wavelet Transform

The 1-D wavelet transform can easily be extended to 2-D. In 2-D, the scaling function \( \phi(m,n) \) scales in both directions and three detail wavelet functions are required: horizontal \( \psi^H(m,n) \), vertical \( \psi^V(m,n) \) and diagonal \( \psi^D(m,n) \). In a separable fashion these could be written as:

\[
\begin{align*}
\phi(m,n) &= \phi(m)\phi(n) \\
\psi^H(m,n) &= \psi(m)\phi(n) \\
\psi^V(m,n) &= \phi(m)\psi(n) \\
\psi^D(m,n) &= \psi(m)\psi(n)
\end{align*}
\]  

(3.15) (3.16) (3.17) (3.18)

Hence the wavelet transform in 2-D could be written in a similar fashion:

\[
x[m,n] = \sum_{k,l} a_{j,k,l} \phi\left(\frac{m}{2^j} - k, \frac{n}{2^j} - l\right) + \sum_{j=1}^J \sum_{k,l} d^H_{j,k,l} \psi^H\left(\frac{m}{2^j} - k, \frac{n}{2^j} - l\right) + \sum_{j=1}^J \sum_{k,l} d^V_{j,k,l} \psi^V\left(\frac{m}{2^j} - k, \frac{n}{2^j} - l\right) + \sum_{j=1}^J \sum_{k,l} d^D_{j,k,l} \psi^D\left(\frac{m}{2^j} - k, \frac{n}{2^j} - l\right)
\]  

(3.19)

The 2-D wavelet transform can be implemented in a separable way, row-wise processing followed by column-wise or vice versa. 2-D wavelet transform of an image is illustrated in the Figure 13. Further levels of wavelet transform could be achieved by using the approximation subband (LL) as input into the same system and repeating the procedure. This is shown in Figure 14. The original image can still be reconstructed exactly by iterating the inverse transform as many times as the forward transform levels.

Implementation of the discrete wavelet transform using filtering operations results in the problem of total number of input samples being not equal to the total number of output coefficients. With a filter size of \( L \), when an input image of size \( N \times N \) is filtered, the output size is \( N+L-1 \) and after the decimation, this result in wavelet subbands having a size of \( \frac{N+L}{2} \times \frac{N+L}{2} \). Thus \( N \) original input samples result in a total of \( N+L \) wavelet coefficients after one level of transform. More levels of transform makes the problem worse, since each further level results in addition of \( L \) more samples. This presents a significant problem in wavelet domain image resolution enhancement algorithm implementation as it is generally required that there is a mapping amongst the scales and coefficients corresponding to the same spatial location in different scales could be matched.
MATERIAL REDACTED AT REQUEST OF UNIVERSITY
A simple way to eliminate coefficient expansion is to use circular convolution (where the second sequence is circularly time reversed and circularly shifted as opposed to linear convolution where the time reversal and shifting is linear (Oppenheim and Schafer, 1989)). While solving the coefficient expansion problem, circular convolution might lead to the introduction of artefacts if there is a large difference in the magnitude of beginning and end coefficients. Yet, the artefacts in the neighbourhood of an image boundary can be eliminated by performing a symmetric periodic extension before the convolution. This symmetric extension guarantees continuity across replicas of the input and eliminates the large wavelet coefficients caused by border discontinuities. Normally symmetric extension doubles the number of input samples. This is not a problem initially since half the samples are redundant because of symmetry. However when the input is filtered and downsampled, it results in outputs that are not necessarily symmetric periodic. As a result, half the coefficients cannot be eliminated by symmetry and there is a doubling of the number of coefficients required to represent the input. Thus, this case is even worse than the linear convolution case where the number of coefficients only increased by the filter length $L$ and
is as problematic when the coefficients corresponding to the same spatial location across scales are required to be obtained. Fortunately, symmetry can be preserved across scales of the wavelet transform by imposing an additional constraint on the wavelet filters: they must be either symmetric or antisymmetric (linear phase). For this special case, periodic symmetric inputs give periodic symmetric outputs and the result is no coefficient expansion.

In the remaining of this thesis, when wavelet transform is mentioned it should be taken as discrete wavelet transform and as more specifically the data that is being dealt with are images, the discrete wavelet transform is assumed to be 2-D.

### 3.3 Wavelet Based Image Resolution Enhancement Methods

Image resolution enhancement is a classic signal interpolation problem and conventional approaches, such as zero-order hold, cause severe pixelation impairments while bilinear and spline interpolation invariably result in undesirable levels of smoothing across salient edges. Recently several efforts in the field have utilised wavelet-domain methodologies with the intention of overcoming some of the problems associated with conventional treatment. In this section, we present a review of the wavelet domain image resolution enhancement techniques.

#### 3.3.1 Background

It is important to have some knowledge about the characteristics of wavelet coefficients among the scales before appreciation of the wavelet domain methods used in image resolution enhancement problems. The wavelet coefficients of natural images have two important properties: persistency and non-Gaussianity:

- **Persistency** refers to the observation that the magnitudes of wavelet coefficients corresponding to the same spatial location tend to propagate from the lower resolution scales through to the higher resolution scales (Figure 15) (Simoncelli, 1999; Crouse et al., 1998). This property is exploited in state-of-the-art wavelet based image compression algorithms such as Embedded Zerotree Wavelet Coder (Shapiro, 1993) and Set Partitioning in Hierarchical Trees (SPIHT) coder (Said and Pearlman, 1996).

- Empirically, the statistics of wavelet coefficients of natural images are highly non-Gaussian. A typical wavelet coefficient histogram has higher density around zero and heavy-tailed than the Gaussian (Simoncelli, 1999; Crouse et al., 1998). As an example, the histogram of horizontal detail subband coefficients of *Lena* image is given in Figure 16.

A common feature of wavelet domain image resolution enhancement algorithms is the assumption that the low-resolution (LR) image to be enhanced is the low-pass filtered subband of a high-
resolution (HR) image which has been subjected to a decimated wavelet transform. A trivial approach would be to reconstruct an approximation to the HR image by filling the unknown, so-called ‘detail’ subbands (normally containing high-pass spatial frequency information) with zeros followed by the application of the inverse discrete wavelet transform (IDWT). Using a given LR image $X$ of size $m \times n$, the unknown HR image $Y$ is reconstructed by using zero padding of high-frequency subbands (i.e. setting all elements of these subbands to zeros) followed by inverse wavelet transform:

$$Y' = W^{-1} \begin{bmatrix} X & 0_{m,n} \\ 0_{m,n} & 0_{m,n} \end{bmatrix}$$  \hspace{1cm} (3.20)

Figure 15: Persistency of wavelet coefficient magnitudes through different scales of wavelet decomposition. The coefficient magnitudes at corresponding spatial locations decay at a similar rate from lower resolution scales to higher resolution scales.

Figure 16: Histogram of horizontal detail subband wavelet coefficients of Lena image.
where $0_{m,n}$ is an all-zero sub-matrix of dimensions $m \times n$ and $W^T$ is the inverse discrete wavelet transform. The underlying assumption is that $X$ can be approximated by the low-order wavelet coefficients of $Y$. This implies a simplified image formation process whereby the point spread function is associated with the low-pass wavelet filter kernel used for the above transformation. We call this method Wavelet Domain Zero Padding (WZP) and summarise in Figure 17 and demonstrate on an image in Figure 18.

In this method, the interpolative reconstruction is achieved by the synthesis wavelet filter pairs and, as a consequence, the selection of a mother wavelet which better models the regularity of natural images yields better results. For example, the well-established in the wavelet literature Daubechies 9/7 filters are expected to generate better results than Haar wavelets (Chapter 7). On the other hand it should also be noted that increasing the wavelet support too much results in oversmoothness. Our experiments suggest that Daubechies 9/7 filters achieve a good balance for natural images (Chapter 7).

It is interesting to note that while this approach is capable of outperforming bilinear interpolation, it has never appeared in the literature probably due to its simplicity.

A natural extension to this method would be to copy the missing high-frequency subband coefficients from the next available lower-frequency subband. This could be done by using a scheme such as bilinear interpolation followed by proper scaling of magnitude coefficients to account for filter gains. Although this method still produces improved results compared to bilinear interpolation, the results are of inferior quality compared to the WZP method, which prevents copying of parent coefficients to be a useful method for image enlargement purposes. The result of this method is demonstrated in Figure 19 where a part of the image is enlarged to demonstrate the artefacts.

More sophisticated methods have attempted to estimate the unknown detail wavelet coefficients in an effort to improve the sharpness of the reconstructed images. These methods could be classified according to the underlying technique used to estimate the detail wavelet coefficients as follows:

- Methods Using the Wavelet Transform Extrema in Coarser Subbands
- Methods Using the Hidden Markov Tree (HMT) modelling of wavelet coefficients
- Methods Using Multiresolutional Basis Fitting Reconstruction (MBFR)
- Others

These methods are summarised in Figure 20 showing the relations between base methods and the methods derived. These methods are discussed in more detail in the following section.
Figure 17: Wavelet Domain Zero Padding (WZP): the available image (X) is assumed to be the low resolution subband of the target high resolution image. The unknown subband elements (detail coefficients) are filled with 0s and resolution enhancement is achieved by inverse wavelet transform.
MATERIAL REDACTED AT REQUEST OF UNIVERSITY
Figure 20: A summary of wavelet domain image resolution enhancement methods.
3.4 Wavelet Domain Image Resolution Enhancement Methods

3.4.1 Methods Using the Wavelet Transform Extrema in Coarser Subbands

These methods attempt to exploit the persistency of wavelet coefficients among the scales. The estimation of high-frequency subband coefficients is carried out by examining the evolution of wavelet transform extrema from coarser to finer subbands and extrapolating the coefficient magnitude decay.

This idea was first proposed by Chang et al. (1995) motivated by the fact that wavelet transform modulus maxima capture the sharp variations of a signal and evolution of coefficient magnitudes across the scales could be characterised by Lipschitz regularity.

The wavelet transform modulus maxima (the local maxima of the wavelet coefficients at a given scale) indicate the singularities of the input signal. Let \( w_{s,l} \) represent the wavelet coefficient of function \( f(x) \) at scale \( s \) and position \( l \) and \( 0 < \alpha < 1 \). If there exists a constant \( C > 0 \) such that for all \( x \in [a,b] \) in the support of the wavelet, the wavelet coefficients satisfy

\[
 w_{s,l} \leq C 2^{-\alpha x}
\]

then \( f(x) \) is said to be uniformly Lipschitz \( \alpha \) over \([a,b]\) (Mallat and Hwang, 1992). The Lipschitz regularity is computed by finding \( \alpha \) such that \( C 2^{-\alpha x} \) approximates the decay of the wavelet coefficients over a range of scales. For each coefficient to be estimated, an exponential fit is performed using its parents. If the fit is found to be close enough to the exponential, then detail coefficient is predicted by extrapolating the decay of wavelet coefficient magnitudes from coarser to finer subbands. Otherwise, prediction is not performed. Block diagram of the process is presented in Figure 21. Figure 22 illustrates the extrapolation process where the coefficient at 4\(^{th}\) scale is estimated by extrapolation of exponential fitting of its parents.

![Figure 21: Image resolution enhancement using wavelet transform extrema extrapolation](image-url)
As the Lipschitz property doesn’t generally hold outside the maxima points, Chang et al. (1995) proposed that the initial estimates could be further improved. The authors identified three constraints that the high-frequency coefficients needed to obey and proposed reconstruction by alternating projections into these sets. They utilise undecimated wavelet transform during reconstruction and the initial estimate of the approximation subband is generated from the original signal using a standard interpolation technique such as bilinear interpolation.

These three constraints are:

(i) The magnitudes of estimated coefficients should be in the range of the wavelet transform.

(ii) The downsampled version of the approximation subband should be equal to the original low-resolution image.

(iii) The local extrema of the estimated detail subband should correspond to the sharp variations in the input image.

A similar but less computationally expensive approach is advocated by Carey et al. (1999). In this work, the initial estimate using the magnitude decay is assumed to be sufficiently accurate and the computational complexity is reduced by restricting the prediction exclusively to extrema points hence avoiding the iterative approach of alternating projections proposed in (Chang et al., 1995).
Nicolier and Truchetet (2000) proposed a method whereby the parent and child relationship among the scales is learned using a set of images. During the learning process, 3 coarser scales are employed to calculate a pattern matrix which minimises the estimation error. Once this pattern matrix is calculated, it is then used to estimate the unknown detail subband.

In these methods, only the coefficients with significant magnitudes are successfully estimated as the evolution of the wavelet coefficients among the scales is found to be difficult to model for other coefficients. Significant magnitude coefficients correspond to salient image discontinuities and consequently only the portrayal of edges in the enhanced resolution image can be targeted while moderate activity detail escapes treatment. Furthermore, due to the fact that wavelet filters have support which spans a number of neighbouring coefficients, edge reconstruction is also based on contributions from such neighbourhoods. As methods based on extrema evolution only target locations of coefficients with significant magnitudes, such neighbourhoods will inevitably provide incomplete information ultimately affecting the quality of edge reconstruction.

Performance is also affected by the fact that the signs of estimated coefficients are replicated directly from ‘parent’ coefficients (in a quad-tree hierarchical decomposition sense) without any attempt being made to estimate the actual signs. This is contradictory to the commonly accepted fact that there is very low correlation between the signs of parent coefficients and their descendants. In a coding context for example, the signs of descendants were generally assumed to be random (Shapiro, 1995; Said, 1996). As a result, the signs of the coefficients estimated using extrema evolution techniques cannot be relied upon. Nevertheless, the accurate sign information plays an important role in reconstructed image quality. This is demonstrated in Figure 23 where the Lena image is reconstructed from its wavelet coefficients where the high-frequency wavelet coefficient signs are copied from their parents whereas the magnitudes of all the coefficients are kept intact. Although the coefficient magnitudes are unchanged, calculations reveal a significant reduction of PSNR to 31.73 dB and the reconstructed images have objectionable artefacts.

3.4.2 Methods Using the Hidden Markov Tree (HMT) modelling of wavelet coefficients

Hidden Markov Tree (HMT) based methods make use of both the persistency and non-Gaussianity properties of wavelet coefficients. These methods model the unknown wavelet coefficients as belonging to mixed Gaussian distributions (states) which are symmetrical around the zero mean. The motivation being that the coefficient distributions which are heavy-tailed and have high density around zero could be modelled by a mixture of Gaussian distributions. Even though it is possible to use an arbitrary number of Gaussian distributions, a two state model where one Gaussian is used to model the coefficients around zero and one for the higher-magnitude coefficients, which constitutes the singularities is generally adopted. This two state model is
illustrated in Figure 24. Each coefficient is assumed to fall into one of these distributions. The
HMT model is trained using Expectation Maximisation (EM) which finds the state transition
parameters which are most likely to result in the coefficients in the observation set by iteration
until a specified convergence error is achieved. In training, the coefficients in the same type of
subband (i.e. HL, LH and HH) are tied together so that a single state transition parameter is
calculated for that type of subband. This is done to prevent over-fitting to the training image.

The HMT models are used to find out the most probable state for the coefficient to be estimated
(i.e. to which distribution it belongs to). The posterior state is found using state-transition
information from lower-resolution scales and then the coefficient estimates are randomly
generated using this distribution. Letting \( p(S^{s-1}_i = m) \) denote the probability of the state of the coefficient \( i \) at scale \( s-1 \) being \( m \), \( p(S^{s}_i = m) \) which is the probability of the state of the coefficient \( i \) at scale \( s \), that is to be estimated, can be written as (for two state model \( m = 1 \) or 2):

\[
p(S^s_i = 1) = p(S^s_i = 1 | S^{s-1}_i = 1)p(S^{s-1}_i = 1) + p(S^s_i = 1 | S^{s-1}_i = 2)p(S^{s-1}_i = 2) \quad (3.22)
\]

\[
p(S^s_i = 2) = p(S^s_i = 2 | S^{s-1}_i = 1)p(S^{s-1}_i = 1) + p(S^s_i = 2 | S^{s-1}_i = 2)p(S^{s-1}_i = 2) \quad (3.23)
\]

Figure 25 shows an illustration of the state transitions.

The seminal work for this type of approach is the work by Crouse et al. (1998) who applied the
technique to signal denoising and related applications. This technique was applied to image
enlargement by Kinebuchi et al. (2001).

An extended version of this approach is presented in (Zhao et al., 2003). The HMT based method
has also been further developed so that it does not require any training data set (Woo et al., 2004a
and 2004b). In this method, state transition parameters are obtained from the low-resolution image
in hand. This is achieved by using the coarser subbands of the image which are obtained by
further applications of wavelet transform. Despite the need to calculate the state transition
parameters for each image to be enlarged, the algorithm generates improved results as the
parameters fit better to the particular image.

In the HMT methodology, once the states of the coefficients are estimated, coefficient magnitudes
are assigned randomly using the Gaussian distribution which is associated with the state. As
Gaussian distributions are symmetrical around zero, coefficients generated randomly using these
distributions have an equal chance of having assigned a negative or a positive sign. This doesn’t
represent a problem when the algorithm is used for denoising purposes, as in denoising
applications the coefficient signs in the finest scale already exist. However, this is not ideal in
resolution enhancement problems where correct sign estimation has an important effect on
resulting image quality as demonstrated in Figure 23.
Figure 24: Illustration of two-state, zero-mean Gaussian mixture model with probability density functions (pdf). In these plots $w$ is the observed coefficient value and $S$ is the corresponding hidden state variable. $S=1$ is the low-variance Gaussian corresponding the lower valued coefficients (top left) and $S=2$ is the high-variance Gaussian corresponding to higher valued coefficients representing singularities (top right). The combined pdf which models the distribution of all wavelet coefficients is shown at the below plot.
3.4.3 Methods Using Multiresolutional Basis Fitting Reconstruction (MBFR)

In (Nguyen, 2000; Nguyen and Milanfar, 2000), a wavelet-based superresolution method was presented based on the Multiresolutional Basis Fitting Reconstruction (MBFR) technique in (Ford and Etter, 1998). The algorithm exploits the interlaced sampling structure in the LR data in the existence of multiple LR images. Finally, a similar approach was proposed in (Mitevski and Bogdanov, 2001) on the basis of the availability of a single LR image. This method is based on the estimation of the desired signal \( y = [f(0), f(1), f(2), \ldots, f(M-1)] \) from an under-sampled available signal \( x \) with \( P \) available samples \( \{P \leq M\} \), \( \{f(t_k), f(t_{1}), f(t_{2}), \ldots, f(t_{P-1})\}, \{t_k \in \{0, 1, \ldots, M-1\}, k=0, 1, \ldots, P-1\} \). It is assumed that \( x \) can be obtained from \( y \) using a sparse \( P \times M \) matrix with values 0 and 1. In vector notation:

\[
x = G^x_j c_j + \sum_{j=1}^J H^y_j d_j
\]  

(3.24)

where \( G^x_j \) is the matrix of scaling function shifts at time indexes \( t_k \) and \( H^y_j \) is the wavelet function shifts at level \( j \). Then with the usual assumption that the available signal \( x \) is an approximation of the low-frequency subband of \( y \), is approximated as:

\[
x \approx G^x_j c_j
\]  

(3.25)

and solving this equation in a least-squares sense, the estimates of the low-frequency scaling function coefficients \( \hat{c}_j \) are obtained. Using these coefficients, the low frequency subband estimate is obtained by \( LL_0 = G_j \hat{c}_j \) where \( G_j \) is the scaling function shift matrix at integer shifts \( \{0, 1, \ldots, M-1\} \). Then, at each available sample, the difference error signal \( e_0 \) is calculated as:
Chapter 3. Wavelet Domain Image Resolution Enhancement

\[ \mathbf{e}_0 = \mathbf{x} - \mathbf{L}\mathbf{L}_0 \mid_{r=n+1,\ldots,r-1} = \mathbf{H}_j^x \mathbf{d}_j \]  

(3.26)

By solving this equation similar to the equation (3.25), the first refinement (low frequency subband and next higher frequency subband) is obtained:

\[ \mathbf{x}_1 \approx \mathbf{G}_j^x \hat{\mathbf{e}}_j + \mathbf{H}_j^x \hat{\mathbf{d}}_j \]  

(3.25)

This process could be iterated until all the desired subbands are constructed.

The basis of this approach, MBFR technique, was designed to take advantage of the non-uniform sampling of a signal using sections with higher sampling rates to interpolate higher frequencies locally. However availability of only a single LR image, with implication that the sampling is uniform, prohibits taking full advantage of this scheme. In addition, the MBFR technique doesn’t allow making use of persistence of wavelet coefficients between scales.

3.4.4 Others

In (Muresan and Parks, 2000a), authors propose two wavelet based image interpolation algorithms. One of the techniques starts with generating a high-resolution image using a standard interpolation method such as zero-order hold interpolation followed by wavelet decomposition of the interpolated image. Then the end result is produced by discarding the low magnitude coefficients and applying inverse wavelet transform. However, this method is highly dependent on the initial interpolation technique. The following operation is, in effect, constrains the results of the initial interpolation algorithm to locations of wavelet extrema (i.e. discontinuities in the image) while neutralising the effect of interpolation in smoother sections of the image. The other technique proposed by the authors is explained in more detail in (Muresan and Parks, 2000b). The authors propose using optimal recovery approach for estimation of high-frequency subband wavelet coefficients. The optimal recovery is applied to select the parent coefficient in the support of wavelet which is the best approximation. However this method suffers from similar complications as the methods using the wavelet transform extrema in coarser subbands described in section 3.4.1; such as inaccurate sign information due to sign information being copied from coarser scale and non-optimal edge reconstruction due to operating only at the extrema points.

In (Huang and Chang, 1999), the authors propose utilising Multi-Layer Perceptron (MLP) type neural networks for the estimation of wavelet subbands. Weights of the MLP predictor is generated by the back-propagation algorithm by training using a sample image set. Then the predictor whose values were obtained in the previous step is used to estimate the missing high-frequency subband coefficients. Unfortunately the paper fails to provide sufficient information for implementation of the proposed method; particularly configuration parameters of the neural network such as the training algorithm, number of epochs, number of hidden layers etc.
Xu et al. (2001), suggest a method utilising fractals to estimate missing coefficients. This method works by exploiting the self-similarity in the detail subbands of wavelet transformed image. Authors propose calculation of the fractal index from the change in the variance of wavelet coefficient values among scales. Then the detail subband coefficients are estimated using fractal interpolation method.

### 3.4.5 Wavelet Superresolution

Although the term superresolution is sometimes used for image resolution enhancement problems from a single low-resolution image, mostly it is used for problems where a sequence of low-resolution images (each with minor variations in their characteristics, hence carrying unique information) is available for the same scene. In the context of this thesis, since we deal with enlargement using a single image, these methods are not examined in detail. For a review of methods reconstructing high-resolution images from multiple low-resolution images, or as widely referred to as “superresolution” methods using wavelet domain techniques, the reader is referred to (Chan et al., 2003), (Bose and Lertrattanapanich, 2001) and (Nguyen, 2000; Nguyen and Milanfar, 2000).

### 3.5 Conclusions

Advances in the wavelet domain image resolution enhancement have the characteristic that most of the research work is conducted independently and citing of other wavelet domain methods are infrequent. This hinders the attempts of making an immediate assessment as the methods are only compared against standard techniques such as bilinear or bicubic interpolation*, if any objective result is provided at all. It is hoped that the comparative assessment provided in chapter 7 serves to fill this gap by providing objective results in the form of Peak Signal to Noise Ratio (PSNR) for various techniques. Nevertheless, there were some obstacles to achieve this aim which are worth mentioning:

(i) Reports of inconsistent PSNR values:

Even for standard algorithms such as bilinear interpolation, the PSNR values vary significantly between different articles in the literature. The deviation in the PSNR values reported in different articles reaches to fairly significant values. For example, PSNR value for the original 512x512 Lena image obtained with bilinear interpolation in (Huang and Chang, 1999) is reported to be 35.77 dB while the same procedure on the same image is given in (Woo et al., 2004a and 2004b) as 30.29 dB, in (Carey et al., 1999) 27.3 dB and

---

*In the reviewed papers, the only exception to this was (Woo et al., 2004a and 2004b) who compared their results to HMT based technique in (Kinebuchi et al., 2001).
results obtained using bicubic spline algorithm—which is expected to produce higher PSNR values—in (Zhao S. et al., 2003) 22.3 dB. Some of the difference in the PSNR values could be attained to floating point precision differences in numeric data, possible use of different down-sampling methods in generating a low-resolution image and dissimilarities in the original image—potentially a product of using a different source obtaining the test images. However, even taking these into account, it is difficult to explain the significant difference in the resulting PSNR values.

Consequently, the PSNR values presented for a particular algorithm in a paper cannot be compared against the PSNR values for a different algorithm in another paper.

(ii) Some papers fail to present objective results such as MSE and PSNR values:

Objective results are not always given in a paper and presentation of the results could simply be illustration of obtained images preventing direct comparison of the methods.

(iii) Difficulty in obtaining ready-to-run versions of some algorithms and implementation problems:

Combination of the difficulties mentioned in (i) and (ii) forces one to conduct independent experiments to be able to come up with an objective comparison of proposed methods in the literature. However this is also complicated for the reasons that it isn’t always possible to acquire the exact parameters to replicate the original author’s experiments, if not explained in sufficient detail in the paper.

An interesting point to note is that, although capable of producing results comparable to state-of-the-art wavelet domain image resolution enhancement techniques, the WZP method is not mentioned and not compared against the techniques proposed in the literature.
CHAPTER 4

Wavelet Domain Image Resolution Enhancement with Coefficient Estimation

4.1 Introduction

As already discussed in the preceding chapters, wavelet-domain resolution enhancement is commonly achieved by regarding the available image (i.e. the image whose resolution we seek to enhance) as the low-resolution (LR) version (i.e. a low-pass wavelet-filtered image both horizontally and vertically) of the unknown high-resolution (HR) image. Then by zero-filling the unknown detail subbands (i.e. using zeros in place of the unknown high-order wavelet coefficients) and combining these with the LR image, an enhanced resolution image can be obtained by taking the inverse wavelet transform. Even though the results obtained by zero padding of unknown detail subbands generates good results, if reliable estimation of the unknown coefficients could be carried out, high spatial frequency image features could be further enhanced.

The algorithms proposed in the literature exploits the persistency and non-Gaussianity properties of wavelet transform to estimate this high-frequency detail information. In this chapter, we show that, the estimation results can be improved by exploiting correlation among neighbouring coefficients as well as persistency which exploits the correlation across the subbands. We first explain the proposed method and the motivation. Then we present the experimental results on a number of images. After a frequency domain discussion of the proposed method, we conclude the chapter with a summary and the demonstration of results obtained by implementation of the method.

4.2 Wavelet-domain resolution enhancement

Let \( L(z) \) and \( H(z) \) represent respectively the low-pass (LP) and high-pass (HP) filters constituting an analysis/synthesis filter pair for the discrete wavelet transform in the z-domain and let \( Y \) be the unknown high-resolution image we seek to reconstruct. We use a notation in which the direction of filtering (i.e. row/column) is shown explicitly as a subscript while \((\downarrow 2)\) denotes decimation by a factor of 2 in both directions. Using this notation we can write that:

\[
LL_{\downarrow 2}(z) = (\downarrow 2) L_{col}(z)L_{row}(z) Y(z)
\] (4.1)

and

\[
HL_{\downarrow 2}(z) = (\downarrow 2) L_{col}(z)H_{row}(z) Y(z)
\] (4.2)
where $LL_0$ and $HL_0$ are respectively low-pass subband and vertical detail subband of wavelet coefficients. We assume that $LL_0$ is the available LR image $X$ whose resolution we seek to enhance and consequently our task is to estimate the elements of detail coefficients contained in detail subbands such as those in $HL_0$.

In the literature, the estimation is generally done using the next level detail subbands. Our approach is based on the assumption that a high-pass filtered undecimated version of the available LR image ($LL_0$):

$$H_{L0}'(z) = H_{U} z) L_{L0}(z)$$ (4.3)

is sufficiently correlated with $HL_0$ horizontally to provide a basis for the estimation of the latter. $H_{L0}'$ is obtained by row-wise high-pass filtering of $LL_0$, while the next level HL subband ($HL_1$) would be obtained by column-wise low-pass filtering and decimation in both directions following this. As the purpose of the estimation is to reconstruct the high-frequency details, we don't apply low-pass filtering and decimation to prevent loss of data and keep high-frequency information as much as possible. Our estimation process is based on the following properties:

- The low-pass and high-pass wavelet filters have sufficiently large coefficient magnitudes not only at the centre but also in the vicinity of the centre (Table 2). Hence, the coefficients obtained as a result of filtering are affected by the neighbours of their parents as well as their immediate parents.
- Local stationarity in the most images allow using the neighbouring parents to have more robust results. This correlation between neighbouring coefficients is also utilised in image coding applications such as (Shapiro, 1993) and (Said and Pearlman, 1996) to increase coding efficiency.

Exploiting the fact that the next-level subbands would have similar characteristics, i.e. the relation between $LL_1$ and $HL_1$ is expected to be similar to the relation between $LL_0$ and $HL_0$, we test our assumption that there is considerable levels of correlation between $HL_0'$ and $HL_0$ by applying undecimated wavelet filtering on the available image ($LL_0$) to obtain low- and high-pass filtered versions $LL_1$ and $HL_1$:

$$LL_1(z) = L_{col}(z) L_{row}(z) L_{L0}(z)$$ (4.4)

and

$$HL_1(z) = L_{col}(z) H_{row}(z) L_{L0}(z)$$ (4.5)

Subsequently, in accordance with (4.3), we obtain

$$HL_1'(z) = H_{row}(z) LL_1(z)$$ (4.6)

and demonstrate that substantial levels of horizontal correlation exist between $HL_1$ and $HL_1'$. Measured correlation values between $HL_1$ and $HL_1'$ are given in Table 3 where range of correlation is [-1,1]. This table shows the correlation between a coefficient located at $(m,n)$ in $HL_1$.
and coefficients in the neighbourhood of \( (m,n) \) in \( \text{HL}' \). The highest correlation values are observed at locations \( (m-1,n), (m,n), (m+1,n) \) and \( (m+2,n) \). These locations define the horizontal neighbourhood that will be used in the remainder of this work. Calculations on different test images confirm that the highest correlation values occur at the same locations.

<table>
<thead>
<tr>
<th></th>
<th>( m-2 )</th>
<th>( m-1 )</th>
<th>( m )</th>
<th>( m+1 )</th>
<th>( m+2 )</th>
<th>( m+3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n-2 )</td>
<td>0.002</td>
<td>-0.045</td>
<td>0.045</td>
<td>0.007</td>
<td>-0.011</td>
<td>0.015</td>
</tr>
<tr>
<td>( n-1 )</td>
<td>0.053</td>
<td>-0.148</td>
<td>0.034</td>
<td>0.181</td>
<td>-0.161</td>
<td>0.008</td>
</tr>
<tr>
<td>( n )</td>
<td>0.092</td>
<td>-0.327</td>
<td>0.206</td>
<td>0.208</td>
<td>-0.324</td>
<td>0.122</td>
</tr>
<tr>
<td>( n+1 )</td>
<td>0.015</td>
<td>-0.165</td>
<td>0.185</td>
<td>0.028</td>
<td>-0.137</td>
<td>0.064</td>
</tr>
<tr>
<td>( n+2 )</td>
<td>0.034</td>
<td>-0.005</td>
<td>0.018</td>
<td>0.034</td>
<td>-0.041</td>
<td>0.006</td>
</tr>
</tbody>
</table>

Table 3: Correlation values between \( \text{HL} \) and \( \text{HL}' \) subbands for Lena

The above observations have motivated an estimation approach which is based on linear least-squares regression. By denoting the coefficient at position \( (m,n) \) in \( \text{HL} \) as \( y_{m,n} \) and the coefficient at position \( (m,n) \) in \( \text{HL}' \) as \( x_{m,n} \) an estimate of \( y_{m,n} \) is obtained according to the following:

\[
y_{m,n} = a_0 + a_1 x_{m-1,n} + a_2 x_{m,n} + a_3 x_{m+1,n} + a_4 x_{m+2,n}
\] (4.7)

Applying (4.7) to all coefficients of \( \text{HL} \) in the specified neighbourhood, we can solve the over-determined problem by linear least-squares regression to obtain the weights \( a_0, a_1, a_2, a_3 \) and \( a_4 \) that minimise the approximation error. These weights are then subsequently used to obtain an estimate of the unknown \( \text{HL}_0 \) coefficients as a linear combination of neighbouring \( \text{HL}_0' \) coefficients which were found by (4.3). As the derivation of the weights are done by using the next-level subbands of the image that is sought to be enlarged, image dependency of the weights are essentially taken into account.

A similar treatment is possible for the \( \text{LH}_0 \) subband. In this case, as expected, the highest correlation values are found in the vertical neighbourhood and the same weights as for \( \text{HL}_0 \) could be used in vertical direction. Estimation of \( \text{HH}_0 \) coefficients was omitted as its contribution in subjective terms was found to be negligible. This is not inconsistent with human visual perception since the \( \text{HH}_0 \) subband is generated by high-pass filtering the image horizontally and vertically. Such filtering promotes the retention of the spatially diagonal image activity and it is widely accepted that the human visual system is relatively less sensitive to diagonal detail. This is an established consideration in image coding systems which invariably tolerate higher levels of quality loss in diagonal detail (Shapiro, 1995).

After coefficient estimation is carried out as above, the high-resolution image is obtained by applying the inverse wavelet transform.
The overall algorithm is described in Figure 26 where the estimator coefficients are calculated by the sub-system at the top. The below sub-system utilises these coefficients to generate the resulting high-resolution image estimate.

The baseline algorithm described above generates HR images which are 2x the original LR image size. Further factors of enlargement can be achieved by applying the same algorithm iteratively exploiting the hierarchy of the quad-tree wavelet decomposition.

4.2.1 Discussion

It could be argued that the operation described by (4.3) is self-contradictory i.e. it amounts to high-pass filtering image data which are essentially low-pass frequency in nature. Nevertheless it should be taken into account that all filters involved are non-ideal and consequently there will be frequency components that will survive this operation. Such components correspond to the aliased part of the image spectrum. In that sense our method seeks to exploit correlation between aliased frequency components in low-frequency subband images on the one hand and frequency components which are unknown and reside in high-frequency subband images i.e. the wavelet
coefficients which we endeavour to estimate. Our measurements have established that such a
degree of correlation exists. A similar case can be argued for the operation described by (4.6).

4.3 Experimental results

In our experiments we have used four well-known test images: Lena, Elaine, Baboon and
Peppers, all having a size of 512x512 pixels. We have filtered the HR images using a Gaussian
kernel and then downsampled by factors of 2 and 4 respectively for 2x and 4x resolution
enhancement and used those as the available LR images.

Figure 27 shows results obtained using a quadrant of a Zone Plate test image. Zone plates are
obtained by appropriate modulation of sinusoids so that a continuous variation of spatial detail in
cycles per picture width/height is achieved. They are often used to measure directly the frequency
response of processing algorithms and, as such, they are particularly useful for assessing
resolution enhancement methods. We generated the 512x512 zone plate image using the
following formula:

\[ x(m,n) = 127.5 + 127.5 \cos \left( \frac{\pi}{2} \left( \frac{(m-256)^2 + (n-256)^2}{512} \right) \right) \]  

(4.8)

We show the differences between the original HR zone plate (used as ground truth) and
reconstructed images obtained by applying three competing algorithms on an LR version of the
original. In the difference images, mid-grey tones represent zero error while darker and lighter
tones negative and positive errors respectively. Our results clearly indicate that the proposed
technique safeguards the integrity of a wider range of spatial frequencies. As the proposed
algorithm generates the unknown frequency components using the known lower frequency
components, constant-varying frequency content has a boosting effect on the algorithm
performance.

Figure 28, Figure 29, Figure 30 and Figure 31 show comparisons with bilinear interpolation for 4x
reconstructions of extracts of Lena, Peppers, Baboon and Snoopy images respectively. It can be
seen that the proposed method offers considerable improvements. An appreciation of the added
detail contributed by our method is shown in Figure 32. This is a comparison between the actual
detail contained in \( H_L_0 \) on the one hand and the estimated detail using our method on the other.
Such added detail images were obtained by setting to zero all the elements of the \( L_L_0 \) subband,
keeping either the original or the estimated information in all remaining subbands (other than \( L_L_0 \))
and finally applying the inverse wavelet transform. In this way, the added high frequency
information can be visualised as grey level variations on a mid-grey canvas.
Figure 34: Absolute reconstruction error using various techniques for an edge. The above plot shows cross section of an edge where $s(x)$ is the amplitude of a pixel at position $x$. The below plot shows the reconstruction error $|\text{Err}(x)|$ using various techniques.
Table 4: Correct coefficient sign percentages for various techniques. The calculations are done for the coefficients whose magnitude exceeds the threshold $\theta$. The threshold $\theta$ is found by constraining the percentage of coefficients exceeding the $\theta$ to the values in the second column.

Table 4 provides a comparison in terms of coefficient sign agreement achieved by the estimation process. This is expressed as a percentage of correctly estimated signs relative to the total number of coefficients. Results are additionally classified according to coefficient magnitude. Only the coefficients exceeding a pre-specified magnitude threshold $\theta$ are contributing to the results. This threshold is adjusted to achieve the percentage of contributing coefficients shown in the second column. It should be noted that the method proposed by Carey et al. (1999) only estimates coefficients at extrema points while the other coefficients are assigned zero values. As a consequence, when all coefficients are taken into account, the coefficient percentage drops significantly. To be able to carry out a fair comparison with the other methods, two different percentage values are shown for the Regularity Preserving Image Interpolation of Carey et al.
(1999); the first is relative to the total number of coefficients while the second (in brackets) is relative to the subset of estimated coefficients. These results also confirm the observation that there is little correlation among signs of co-sited coefficients across scales as discussed in chapter 3. However, it is interesting to note that there is higher correlation as the coefficient magnitude increases. As expected, the HMM based estimation (Kinebuchi et al., 2001), due to the inherent randomness of estimated signs, achieves just under 50% agreement, increasing slightly with coefficient magnitude. Our results demonstrate that the exploitation of correlation in the proposed neighbourhood improves significantly coefficient sign agreement and for images such as Elaine, up to 93.36% agreement was achieved.

4.4 Frequency Domain Analysis

In the proposed algorithm, we first high-pass filter the $LL_j$ so that the high frequency part is retained (containing the frequency content adjacent to $HL_j$). The result is called $HL_j'$, which is the first estimate of $HL_j$. Then we employ linear least-squares regression to obtain the optimal estimator of $HL_j$ coefficients as weighted combinations of neighbouring $HL_j'$ coefficients. This is done to learn the "reshaping" needed to generate an approximation of $HL_j$ using $HL_j'$. Subsequently this process is used to estimate $HL_0$ from $LL_0$ by generating an initial estimate $HL_0'$ followed by application of the weights obtained from the procedure explained above.

These weights could also be considered as a filter. The intercept (which is represented with $a_0$ in Equation 4.7) could be added to the signal as a DC shift, however this could also be ignored as experimental results reveal that the shift is marginal and discarding of the shift doesn’t affect the results considerably. Naming this filter as $W(z)$, filter for direct estimation of $HL_0$ from $LL_0$ (or equivalently estimation of unknown $HL_0$ from the readily available $LL_0$) can be written as:

$$F(z) = H(z)W(z)$$ (4.9)

where both filters are applied row-wise.

Using the property that multiplication in z-domain is equivalent to convolution of the filter coefficients in the time domain, the process can be reduced to a single filtering operation:

$$f[n] = h[n]*w[n]$$ (4.10)

The overall process is illustrated in Figure 35.

Analysis of frequency response of the direct estimation filter (illustrated in Figure 36) reveals that it acts as a band-pass filter with a peak around $0.7\pi$. In effect, this filter extracts the higher frequencies of the $LL_j$ and shapes it to emphasise the frequencies between the lowest frequency that $HL_j$ would have (which is with “ideal” filters $0.5\pi$) and maximum frequency $\pi$. This way, the
Figure 35: Illustration of the estimation process using filtering operations. The aim is to find an estimator $F(z)$ which estimates $HL_0$ given $LL_0$. Assuming next level subbands have also similar characteristics, the $F(z)$ can be derived using $HL_1$ and $LL_1$ as they are both known, hence allowing calculation of an estimator.
Figure 36: Frequency response of the filter $F(z)$ obtained by combining the high pass filter $H(z)$ and $W(z)$ which is the filter created using least square weights obtained using 512x512 Lena image.

Figure 37: Frequency analysis using 512x512 Lena image (a) the original subbands $LL$, $HL$, $LH$ and $HH$ (b) estimated HL and LH subbands, $HL''$ and $LH''$ (including the intermediary estimation stages $HL'$ and $LH'$)
coarser "features" in the lower frequency subbands are extended to higher frequency by learning this conversion using observations from existing LL to HL conversion from the available subbands.

Frequency domain analyses of the estimated subbands (including the intermediary estimation stages) and real subbands (obtained using a HR ground-truth image, which was assumed to be unavailable for the estimation purposes) are presented in Figure 37. These figures were obtained by taking DFT of the subbands. Comparison of the estimated subbands shows the generation of coefficients using the leaked information in the LL subband. It can also be observed that the estimated content density is the highest around midpoint between 0.5\(\pi\) and \(\pi\) while decaying sharply towards either side. On the other hand the original subband data has fairly dense content between 0.5\(\pi\) and \(\pi\). While the estimation algorithm generates some previously unavailable higher frequency content, the highest end of the frequency content is mostly not estimated.

4.5 Conclusions

We have presented a wavelet-domain image resolution enhancement algorithm which operates in a quad-tree decomposition framework and exploits wavelet coefficient correlation in a local neighbourhood sense. The proposed method employs linear least-squares regression to estimate the unknown detail coefficients. Our results show that our method outperforms conventional image resolution enhancement methods such as bilinear interpolation, for a wide range of standard test images. The availability of fast wavelet transform and filtering implementations makes our scheme attractive for real-time and on-line applications. The objective evaluation (in terms of PSNR) of the proposed algorithm is presented in chapter 7.
CHAPTER 5

Wavelet Domain Image Resolution Enhancement Using Cycle Spinning and Edge Rectification

5.1 Introduction

Wavelet Domain Zero Padding (WZP) approach which has been described in chapter 3 is a powerful alternative to conventional interpolation techniques such as bilinear and bicubic interpolation. However this technique suffers from impairments commonly associated with conventional resolution enhancement approaches due to unavailability of higher spatial frequencies.

In this chapter a new method is proposed to reduce the problems associated with WZP method. An initial high-resolution (HR) approximation to the original image is obtained by means of WZP. Then this image is further processed using the cycle-spinning methodology which reduces ringing.

A critical element of the algorithm is the adoption of a simplified edge profile suitable for the description of edge degradations such as blurring due to loss of resolution. The degraded edges are classified into groups according to the parameters calculated using the edge model. Then linear regression using a minimal training set of high-resolution originals is finally employed to rectify the degraded edges by using a separate set of parameters for each of these groups.

In this chapter, first the cycle spinning methodology which is typically used in the context of wavelet based denoising applications is explained. Then a resolution enhancement technique which utilises WZP and cycle spinning is proposed. It is then shown that the results of this technique could further be improved upon by rectification of the edges in the reconstructed high-resolution image.

5.1.1 Cycle Spinning (CS)

The decimated wavelet transform is not shift-invariant and as a result, distortion of wavelet coefficients, due to quantisation of coefficients in compression applications or non-exact estimation of high-frequency coefficients in resolution enhancement applications, introduces cyclostationarity into the image which manifests itself as ringing in the neighbourhood of discontinuities (Coifman and Donoho, 1995). Cycle-spinning has been shown to be an effective method against ringing when used for de-noising purposes in the wavelet domain (Coifman and Donoho, 1995) and also for reducing ringing and increasing the perceptual quality of compressed
images. Cycle-spinning method works by averaging-out the inherent translation dependency of wavelet transform.

Translation dependency implies that, if the input signal is shifted, denoised and then shifted back, the result would be different from the estimate obtained from denoising without shifting with ringing artefacts occurring in different places. Cycle spinning utilises the periodic time-invariance of the wavelet transform by reiterating the denoising algorithm on the shifted signals and averaging-out the artefacts.

In (Nosratinia, 2001) and (Nosratinia, 2003), it was shown that cycle-spinning applied as a post-processing operation after decompression results in significant improvements in the framework of JPEG and JPEG2000 image compression respectively. Cycle-spinning has also been proven to be an effective tool to improve the results subsequent to wavelet denoising. Chen and Suter (2004) showed that cycle-spinning when used in the context of established wavelet denoising methods such as HMT method (Crouse et al., 1998) successfully improved the results.

In the following section, we show that cycle-spinning could also be adapted for the resolution enhancement problem.

5.2 Resolution Enhancement Using Cycle Spinning

We propose an algorithm which consists of two steps. In the first step, an initial HR approximation is generated using wavelet domain zero padding. This approximation commonly exhibits ringing artefacts. To reduce these artefacts, we propose a variant of the cycle spinning methodology which is applied as a second step.

5.2.1 Wavelet Domain Zero Padding (WZP)

We generate an initial approximation to the unknown HR image using wavelet-domain zero padding (WZP). Using a given LR image $X$ of size $m \times n$, the initial estimate of unknown HR image $Y$ is reconstructed by using zero padding of high-frequency subbands (i.e. setting all elements of these subbands to zeros) followed by inverse wavelet transform:

$$Y_0^I = W^{-1} \begin{bmatrix} X & 0_{m,n} \\ 0_{m,n} & 0_{m,n} \end{bmatrix}$$

(5.1)

where $0_{m,n}$ is an all-zero sub-matrix of dimensions $m \times n$ and $W^{-1}$ is the inverse discrete wavelet transform. This process is explained in more detail in chapter 3.
The HR approximation $Y_0'$ obtained as above commonly exhibits smoothing and ringing artefacts. The ringing emerges at the vicinity of discontinuities as alternating undershoots and overshoots of the intensity level.

### 5.2.2 Adaptation of Cycle Spinning to Image Resolution Enhancement

For the purpose of image resolution enhancement, we adapt the cycle-spinning methodology to operate in the wavelet domain as follows. First, a number of LR images $X'_{i,j}$ are generated from $Y_0'$ by spatial shifting, wavelet transforming and discarding the high frequency (HF) coefficients:

$$X'_{i,j} = DWS_{ij} Y_0'$$

where $D$ represents discarding of HF coefficients, $W$ denotes wavelet transform and $S_{ij}$ is an operator applying horizontal and vertical shifts of $(i,j)$ in the range $i,j \in \{-k,-k+1,\ldots,k-1,k\}$. Then, equation (5.1) is applied to all $X'_{i,j}$ yielding $N Y'_{i,j}$ images, where $N = (2k+1)(2k+1)$.

Finally, these intermediate HR images are re-aligned and averaged to give the final HR reconstructed image:

$$Y' = \frac{1}{N} \sum_{i=-k}^{k} \sum_{j=-k}^{k} S_{ij}^{-1} Y'_{i,j}$$

where $S_{ij}^{-1}$ is the inverse of the shifting operator $S_{ij}$. This process is summarised in Figure 38 as the CS block and subsequently referred to as CS. In this figure DWT denotes the discrete wavelet transform and notation for spatial shifting is in z-domain.

### 5.2.3 Discussion

In Table 5, we tabulate PSNR values as a function of the maximum shift $k$ for the Lena test image (Also plotted in Figure 39). It can be seen that the results are highly dependent on the number of shifted images used and $k=2$ yields the best performance. Using a smaller neighbourhood suffers from not having sufficient data while a larger neighbourhood is liable to crosstalk from spatially uncorrelated image features. The PSNR figures also reveal the noise cancelling behaviour of cycle spinning. When $k$ is an odd-number, the results are significantly lower than the cases when $k$ is even.

Figure 40 illustrates the intermediary images from the various stages of the algorithm. The shaded areas in the boxes correspond to the coordinates for which cycle-spinning is applied. These images illustrate the progressive improvement in the reconstructed images and reveal that vertical edges are improved with the horizontal cycle-spinning while horizontal edges are with the vertical. This observation is fundamental to the directionally adaptive approach which is introduced in the following chapter.
Chapter 5. Wavelet Domain Image Res. Enhancement with Cycle Spinning and Edge Rectification

Figure 38: Block diagram of the proposed method
Enlargement by a factor of 4x can be achieved by two alternative methods: either (i) generating the initial image at 4x resolution by iterating WZP twice and then applying cycle-spinning on this 4x enlarged image once, or (ii) iterating twice the baseline method described in this chapter, i.e. 2x enlargement by WZP followed by cycle-spinning and another 2x enlargement of the output in the same manner. Experiments show that the former method gives better results, as well as being computationally less demanding. Both methods can be further iterated in a similar fashion to achieve higher factors of enlargement. Experimental results for Lena and Peppers are shown in Figure 41 and Figure 42 respectively, together with reconstruction error images. Quantitative comparisons are provided in chapter 7.

<table>
<thead>
<tr>
<th>Shift/Image</th>
<th>Lena</th>
<th>Elaine</th>
<th>Baboon</th>
<th>Peppers</th>
</tr>
</thead>
<tbody>
<tr>
<td>(-1,-1) – (1,1)</td>
<td>30.76</td>
<td>34.39</td>
<td>24.77</td>
<td>31.97</td>
</tr>
<tr>
<td>(-2,-2) – (2,2)</td>
<td>31.37</td>
<td>34.85</td>
<td>25.10</td>
<td>32.48</td>
</tr>
<tr>
<td>(-3,-3) – (3,3)</td>
<td>31.04</td>
<td>34.61</td>
<td>24.91</td>
<td>32.21</td>
</tr>
<tr>
<td>(-4,-4) – (4,4)</td>
<td>31.31</td>
<td>34.80</td>
<td>25.04</td>
<td>32.43</td>
</tr>
<tr>
<td>(-5,-5) – (5,5)</td>
<td>31.11</td>
<td>34.66</td>
<td>25.02</td>
<td>32.26</td>
</tr>
<tr>
<td>(-6,-6) – (6,6)</td>
<td>31.28</td>
<td>34.71</td>
<td>25.03</td>
<td>32.40</td>
</tr>
</tbody>
</table>

Table 5: PSNR(dB) performance as a function of maximum shift amount for Lena, Elaine, Baboon and Peppers images (2x resolution enhancement: from 128x128 to 256x256)

Figure 39: PSNR(dB) vs. shift amount (k) for Lena image.
MATERIAL REDACTED AT REQUEST OF UNIVERSITY
5.3 Post Processing Using Edge Rectification

As already stated it is assumed that the LR image to be enlarged is the LL (low-pass) subband of a quad-tree wavelet decomposition of the unknown HR image. LL subbands are obtained by variable-separable (first horizontal, then vertical) low-pass wavelet filtering which results in loss of high frequency data. The unavailability of high-frequency spatial information normally residing in the unknown subbands results invariably in blurring and ringing of salient image features such as sharp edges. Nevertheless, low-pass filtered versions of edges actually survive within the LL subband. While these are typically smoothed and widened versions of the original edges it is worth noting that, for simple edge profiles, their notional centre of symmetry (i.e. the midpoint between the minimum and the maximum intensity value) has not moved from its original location. The availability of most of the edges in the low-resolution image motivates an approach where degraded edges are attempted to be rectified. The rectification is done using linear regression using a minimal training set of high-resolution originals. The proposed technique segments the edges according to their characteristics and employs a different set of regression parameters for each edge group.

5.3.1 Edge Model

To describe edge evolution as it undergoes low-pass filtering we adopt the model proposed in (Beek and Van, 1995). According to this model, image edges can be approximated by Gaussian-smoothed step functions. In one dimension (i.e. along an image scan line) an edge \( s(x) \) is approximated by:

\[
s(x) = s(x; b, c, w) = h(x; b, c) \ast g(x; w)
\]

where \( h(x; b, c) \) is a step function and \( g(x; w) \) is a Gaussian as given in equations (5.5) and (5.6) respectively. Letting \( u(x) \) represent the unit step function;

\[
h(x; b, c) = b + cu(x)
\]

\[
g(x; w) = \frac{1}{w\sqrt{2\pi}} e^{-\frac{x^2}{2w^2}}
\]

\[
h(x) \ast g(x) = \int_{0}^{\infty} h(t) g(x - t) dt
\]

Using (5.4)-(5.6) and with convolution defined as in (5.7), an edge could be represented as

\[
s(x; b, c, w) = b + \frac{c}{2} \left( 1 + \text{erf} \left( \frac{x}{w\sqrt{2}} \right) \right)
\]
where \( \text{erf}(x) \) is defined as in (5.9).

\[
\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt
\]

(5.9)

In this model, \( c \) corresponds to the edge contrast, \( b \) is the edge minimum and \( * \) denotes convolution. Parameter \( w \) describes the width of the edge (i.e. distance from the midpoint) and is implicitly related to the variance of the Gaussian smoothing function. The above profile is illustrated in Figure 43 for edge parameters \( b=20, c=120 \) and \( w=1.5 \).

Edge detection could be implemented similar to Canny Edge Detector by convolving the image with the first derivative of Gaussian to localise the edges.

Using the following properties (Oppenheim and Schafer, 1989)

(i) \[
\frac{d}{dx}(h(x) * g(x)) = \frac{d}{dx} h(x) * g(x) = h(x) * \frac{d}{dx} g(x)
\]

(ii) Convolution of Gaussians with variances \( \sigma_1 \) and \( \sigma_2 \) is a Gaussian with variance \( \sqrt{\sigma_1^2 + \sigma_2^2} \)

the output of edge detection is given as follows:

\[
d(x; w, c, \sigma) = s(x; b, c, w) * g'(x; \sigma) = c \cdot g(x; \sqrt{w^2 + \sigma^2})
\]

(5.10)
where $\sigma$ is usually set to 1 for natural images (Beek and Van, 1995). Taking the discretisation into account, the detected edge centre has some error, letting $s_0(x)$ is a signal, an edge at position $x=0$, the true edge centre is at $x=x_0$, with $|x_0|<0.5$.

Using the edge representation in (5.8) as an equation, the remaining three equations to be able to find a solution to the parameters $w$, $b$, $c$ and $x_0$ are obtained by sampling the edge detector output at positions $+1, 0$ and $-1$ (Equations (5.11)-(5.13)).

\[
\begin{align*}
    d_0 &= d(0,c,w,\sigma) = \frac{c}{\sqrt{2\pi(w^2+\sigma^2)}} \exp \left( \frac{-x_0^2}{2(w^2+\sigma^2)} \right) \\
    d_1 &= d(1,c,w,\sigma) = \frac{c}{\sqrt{2\pi(w^2+\sigma^2)}} \exp \left( \frac{-1-x_0^2}{2(w^2+\sigma^2)} \right) \\
    d_{-1} &= d(-1,c,w,\sigma) = \frac{c}{\sqrt{2\pi(w^2+\sigma^2)}} \exp \left( \frac{(-1-x_0)^2}{2(w^2+\sigma^2)} \right)
\end{align*}
\]

Solving (5.8) and (5.11)-(5.13), edge parameters are obtained as follows:

\[
\begin{align*}
    w &= \frac{1}{\ln \left( \frac{d_0d_{-1}}{d_1d_{-1}} \right)^{-\sigma^2}} \\
    x_0 &= \frac{\ln \left( \frac{d_{-1}}{d_1} \right)}{2\ln \left( \frac{d_{-1}}{d_0} \right)} \\
    c &= d_0 \sqrt{\frac{2\pi}{\ln \left( \frac{d_0d_{-1}}{d_1d_{-1}} \right)^2}} \left( \frac{d_1}{d_{-1}} \right) \\
    b &= s_0(x_0) - \frac{c}{2}
\end{align*}
\]

An edge extracted from a real image and the plot of the corresponding model is illustrated in Figure 44. In this figure, the approximate edge parameters were calculated as follows: $b = 42$, $c = 64$ and $w = 1.6813$; hence the edge is described with the following equation:

\[
s(x) = 42 + \frac{64}{2} \left( 1 + \text{erf} \left( \frac{x}{1.6813\sqrt{2}} \right) \right).
\]

Using the above model, smoothing due to low-pass filtering is observed as an increase in parameter $w$, while the other parameters remain relatively unaffected (assuming for simplicity
wavelet filters of unity gain). If \( s_d(x) = s(x; b_0, c_0, w_0) \) is an edge in the unavailable HR image, the corresponding surviving edge in the available image would be represented by \( s_l(x) = s(x; b_0, c_0, w_l) + q_s(x) \) where \( w_l = \lambda w_0 \) with an edge widening factor of \( \lambda \) and \( q_s(x) \) is a term accounting for other residual degradations such as ringing. A critical component of the proposed edge rectification method is the estimation of edge width \( w \). We account for edge distortion (primarily smoothing) by establishing a correspondence between available LR image data and a training set of HR image data using linear regression.

![Figure 44: An edge and its modelled version using the described edge model.](image)

Furthermore, our experiments reveal that the application of cycle-spinning widens the edges further. Table 6 shows the mean value and standard deviation of \( w \) for a range of images obtained using various techniques. In this table Db.9/7 denotes the well-known Daubechies 9/7 discrete wavelet filter-bank which was used throughout our experiments.

### 5.3.2 Post Processing of Wavelet Compressed Images

In the framework of this thesis the methods for post-processing of wavelet compressed images, although not targeting the image resolution enhancement applications, are also of interest. The lossy wavelet compression inherently results in some impairment in coefficient information. As the degradation model adopted in this work assumes loss of all the finest level wavelet

82
coefficients, it could also be thought of as an extreme case of wavelet compression. Hence it is expected that the artefacts in wavelet compressed images have similar characteristics to the artefacts wavelet resolution enhanced images.

Although the literature is rich in post-processing methods designed to battle with the artefacts in images compressed with Discrete Cosine Transform (DCT) based methods -which are mainly intended against the blocking artefacts- (Shen and Kuo, 1998), there are only a few post-processing methods for wavelet compressed images.

A method to improve the perceptual quality of wavelet compressed images has been proposed in (Fan and Cham, 2000). In this work, edges are represented using the edge model devised in (Beek and Van, 1995). The edges are then reconstructed by changing the edge parameters and making them more similar to the ideal edge model. As the inherited edge model is an artificial one, the algorithm results in non-natural looking edges.

<table>
<thead>
<tr>
<th>Image/Method</th>
<th>Lena</th>
<th>Elaine</th>
<th>Baboon</th>
<th>Peppers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original</td>
<td>1.3086</td>
<td>1.5197</td>
<td>1.0206</td>
<td>1.4627</td>
</tr>
<tr>
<td>Bilinear</td>
<td>±0.3863</td>
<td>±0.2843</td>
<td>±0.2924</td>
<td>±0.3110</td>
</tr>
<tr>
<td>WZP (Db.9/7)</td>
<td>1.4672</td>
<td>1.7326</td>
<td>1.2373</td>
<td>1.6970</td>
</tr>
<tr>
<td>WZP and CS</td>
<td>±0.4085</td>
<td>±0.3477</td>
<td>±0.3627</td>
<td>±0.3422</td>
</tr>
<tr>
<td>WZP,CS and ER</td>
<td>1.3635</td>
<td>1.6284</td>
<td>1.1157</td>
<td>1.5743</td>
</tr>
<tr>
<td></td>
<td>±0.402</td>
<td>±0.3217</td>
<td>±0.3355</td>
<td>±0.3302</td>
</tr>
<tr>
<td></td>
<td>1.3867</td>
<td>1.6683</td>
<td>1.1373</td>
<td>1.6174</td>
</tr>
<tr>
<td></td>
<td>±0.3979</td>
<td>±0.3158</td>
<td>±0.3301</td>
<td>±0.3316</td>
</tr>
<tr>
<td></td>
<td>1.3325</td>
<td>1.6058</td>
<td>1.1564</td>
<td>1.5514</td>
</tr>
<tr>
<td></td>
<td>±0.3774</td>
<td>±0.3037</td>
<td>±0.3250</td>
<td>±0.3170</td>
</tr>
</tbody>
</table>

Table 6: Estimation of horizontal edge width parameter $w$ for a variety of images and reconstruction techniques.

5.3.3 Edge Rectification

It would be reasonable to assume that a method of readjusting (i.e. reducing) the edge width while maintaining the edge profile otherwise would improve the quality of the reconstructed image. We propose an alternative method which estimates edge parameters directly from a training set of HR image data and further refines them using local neighbourhood information.

First we detect the edges in $Y$, the HR image used as a ground truth training set, using a Canny Edge Detector. Then we calculate the edge parameters using (5.14)-(5.17) and cluster the edges. Edges are clustered into 9 groups according to their $w$ and $c$ parameters. These groups are equally separated in between maximum and minimum values of $w$ and $c$ as illustrated in Figure 45. The following algorithm is applied to each cluster independently.
The edges in $\hat{Y}$, the output of the previous step, are degraded versions of the corresponding edges in $Y$. Linear regression is employed to estimate pixel values in the neighbourhood of the centre of symmetry of an edge. In particular we express edge pixel values in $Y$ as a weighted linear mix of neighbouring pixels in $\hat{Y}$. By denoting the coefficient at position $(m,n)$ in $Y$ as $y_{m,n}$ and the coefficient at position $(m,n)$ in $\hat{Y}$ as $\hat{y}_{m,n}$, the estimates of $y_{m,l,n}$, $y_{m+n}$, and $y_{m+1,n}$ are obtained according to the following:

\begin{align}
\hat{y}_{m,n} &= a_0 + a_1 \hat{y}_{m-2,n} + a_2 \hat{y}_{m-1,n} + a_3 \hat{y}_{m,n} \\
y_{m,n} &= a_4 + a_5 \hat{y}_{m-1,n} + a_6 \hat{y}_{m,n} + a_7 \hat{y}_{m+1,n} \\
y_{m+1,n} &= a_8 + a_9 \hat{y}_{m,n} + a_{10} \hat{y}_{m+1,n} + a_{11} \hat{y}_{m+2,n}
\end{align}

(5.18) (5.19) (5.20)

Using (5.18)-(5.20) for all detected edge points falling into the current cluster, three over-determined equations could be constructed. These equations are solved by linear least squares regression to find the estimator weights sets $\{a_0, a_1, a_2, a_3\}$, $\{a_4, a_5, a_6, a_7\}$, and $\{a_8, a_9, a_{10}, a_{11}\}$ minimising the error. These estimated parameters are then used to rectify the edges in $\hat{Y}$. By repeating this process for each cluster, a separate estimator sets for each cluster is obtained.

It is interesting to note that estimated parameters obtained from a very small training set, even a single image, work well for a wide variety of natural images. The algorithm works in a variable-separable way, i.e. it is first applied horizontally followed by vertical application or vice versa. As horizontal and vertical edges in natural images have the potential of different degrees of sharpness, for example due to non-isotropic sensors, it was found beneficial to calculate different estimator weights for each direction. A simplified block diagram for the overall algorithm is shown in Figure 46. This edge rectification process is subsequently referred to as ER in the remainder of this chapter.
Further levels of enlargement, such as a factor of 4, can be achieved by first generating the 4x enlarged image directly with WZP method and then applying the cycle-spinning method followed by enhancement of edges on this resulting image. Iterative application of the algorithm is also possible but it was observed that this method generates inferior quality images in terms of calculated PSNR values as well as being computationally more expensive.

![Block diagram of the proposed method](image)

**Figure 46: Block diagram of the proposed method**

<table>
<thead>
<tr>
<th>Shift Method</th>
<th>WZP and CS</th>
<th>WZP, CS and ER</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>34.45</td>
<td>34.73</td>
</tr>
<tr>
<td>(-1,-1) - (1,1)</td>
<td>34.43</td>
<td>34.82</td>
</tr>
<tr>
<td>(-2,-2) - (2,2)</td>
<td>34.93</td>
<td>35.23</td>
</tr>
<tr>
<td>(-3,-3) - (3,3)</td>
<td>34.68</td>
<td>35.15</td>
</tr>
<tr>
<td>(-4,-4) - (4,4)</td>
<td>34.89</td>
<td>35.27</td>
</tr>
<tr>
<td>(-5,-5) - (5,5)</td>
<td>34.73</td>
<td>35.20</td>
</tr>
<tr>
<td>(-6,-6) - (6,6)</td>
<td>34.87</td>
<td>35.26</td>
</tr>
</tbody>
</table>

**Table 7: PSNR(dB) as a function of shift for Lena (From 256x256 to 512x512)**

![PSNR graph](image)

**Figure 47: PSNR(dB) as a function of shift for Lena (From 256x256 to 512x512)**
5.4 Experimental Results

We have experimented with a number of well-known test images including Lena, Elaine, Baboon and Peppers. HR versions of these images (512x512) were used as ground truth for performance evaluation purposes and also as a training set for linear regression.

Our experiments have revealed that the choice of parameter $k$ (the maximum shift to be applied in the cycle-spinning) has a significant effect on the algorithm efficiency. In Table 7, PSNR values as a function of the parameter $k$ for the 512x512 Lena test image are tabulated, and a plot of PSNR vs. $k$ is presented in Figure 47. As can be seen, choosing $k = 4$ results in the best performance, while results for $k = 2$ are not far off. Using a smaller neighbourhood suffers from not having sufficient data while a larger neighbourhood is liable to crosstalk from spatially uncorrelated image features. However it should be noted that the choice of the optimum $k$ parameter is also potentially dependent on image size. For example, for a 256x256 version of Lena, a value of $k = 2$ gives the best results. Our results show that $k = 2$ is a realistic trade-off between computational complexity and quality for a range of image sizes.

Figure 48, Figure 49, Figure 50 and Figure 51 show subjective comparisons with bilinear interpolation for 4x reconstructions of Lena, Peppers, Baboon and Snoopy respectively. Amplified residual images are also shown for better appreciation of reconstructed details. Overall our results confirm that the portrayal of salient image features such as edges and contours is consistently improved while no perceivable artefacts are introduced. Lena and Peppers images benefit from the proposed algorithm more while Baboon image exhibits marginal improvement. This is due to relative lack of salient image features in this image compared to the others.

5.5 Conclusions

An image resolution enhancement algorithm operating in the wavelet domain was presented. The main elements of this algorithm were zero-padding of high-frequency wavelet subbands, cycle spinning to reduce ringing arising from zero-padding and finally edge rectification to alleviate blurring due to the unavailability of high spatial frequency information. The proposed algorithm was based on the adoption of a simplified profile model and the parameterisation and rectification of blurring using linear regression. The objective evaluation (in terms of PSNR) of the proposed algorithm is presented in chapter 7.
CHAPTER 6

Wavelet Domain Image Resolution Enhancement with Directional Cycle-Spinning

6.1 Introduction

Since its introduction, the basic cycle-spinning algorithm has hardly evolved apart from an iterative variant proposed in a denoising context in (Fletcher et al., 2002). This involves iterated applications of a processing chain consisting of forward translation, wavelet denoising and backward translation. Although this scheme was shown to improve upon the basic cycle spinning in wavelet denoising problems, it doesn’t lend itself to wavelet domain resolution enhancement. The resolution enhanced image, which is the output of the previous iteration, has to be downsampled before being fed as input into the next, nullifying the effects of the preceding iteration.

In this chapter we introduce the concept of directional cycle-spinning. In particular we take into account local edge orientation information derived from a wavelet decomposition of the available low resolution image to influence key parameters of the cycle spinning algorithm.

6.2 Resolution Enhancement Using Directional Cycle Spinning

In conventional cycle spinning as described in the previous chapter, all possible shifts within a range which defines a local neighbourhood are used. Nevertheless, a closer inspection of artefacts in a typical HR image generated using WZP reveals that ringing artefacts not only occur in the vicinity of strong edges but, more importantly, that they are strongly correlated with the orientation of those edges. In particular, for an edge of a given orientation, ringing is more pronounced in the normal direction. This observation suggests that cycle spinning should be predominantly applied across edges (i.e. vertically in the vicinity of a horizontal edge and vice versa) avoiding other orientations and thus preventing unnecessary smoothing of image detail.

To test this conjecture, we generated a synthetic image shown in Figure 52 and used for testing. The image is a standard greyscale image of size 256x256 and with a dynamic range of [0,255]. The top left quadrant of this image is smoothly varying, while the other three quadrants contain respectively strong vertical, horizontal and diagonal edges.

We consider this synthetic image as the ground truth HR image we seek to reconstruct. Then we generate an LR version of it by low-pass filtering HR followed by down-sampling. Finally we obtain approximations to the original HR image using WZP followed by cycle spinning applied at...
different orientations: non-directional (NCS), horizontal (HCS), vertical (VCS) or diagonal (DCS). Figure 53 indicates the patterns used for cycle spinning (shift locations shown as shaded elements) with the no shift location occupying (0,0). For example HR reconstructions using HCS and VCS are obtained respectively using (6.1) and (6.2):

$$C_h = \frac{\sum_{i=-k}^{k} \hat{Y}_{i,0}}{2k + 1}$$  \hspace{1cm} (6.1)

$$C_v = \frac{\sum_{j=-k}^{k} \hat{Y}_{0,j}}{2k + 1}$$  \hspace{1cm} (6.2)

where $\hat{Y}_{i,j}$ are the intermediate HR images corresponding to a shift of $(i,j)$ and $k$ is the maximum shift amount as explained in the previous chapter.

PSNR values of the HR reconstruction for each quadrant and for each cycle-spinning orientation are shown in Figure 52 while the best value for a particular quadrant is shown in boldface. The plot at the top of Figure 54 illustrates reconstruction of a vertical edge (taken from top-right quadrant of the synthetic image in Figure 52) using horizontal and vertical cycle-spinning, the plot at the bottom shows the absolute error of these reconstructions. Our results confirm that the best reconstruction always occurs by cycle spinning in the direction of the normal to dominant edges while in the absence of such edges non-directional cycle spinning is preferable.

![Figure 52: PSNR (dB) values for the different quadrants of the synthetic image using different cycle spinning directions.](image-url)
Chapter 6. Wavelet Domain Image Res. Enhancement with Directional Cycle-Spinning

Figure 53: Shifts used for different directional CS types

Figure 54: Cross section of a vertical edge extracted from the top-right quadrant of the synthetic image in Figure 52 and its reconstructed versions using horizontal and vertical CS (top) and absolute error ($|\text{Err}(x)|$) of these reconstructions (bottom).
6.3 Algorithm Description

The observation that cycle spinning normal to dominant edges produces better results motivates an approach where image directional activity is used to influence the orientation in which cycle spinning is applied. Wavelet decomposition, which is already available as it plays a central role in the WZP algorithm, provides a suitable setting to measure the directional activity. When a single-level wavelet decomposition of an image is obtained, higher valued coefficients in the high-frequency horizontal and vertical subbands correspond to strong horizontal and vertical edge information respectively. Hence sum of absolute values of the coefficients in these subbands gives information regarding the directional activity. These aforementioned subbands provide a measure of directional activity in an image-wide global sense. Nevertheless, in natural images, various parts of the image exhibit different characteristics.

An important property of wavelet decomposition is its ability to portray spatial and frequency information simultaneously. This property could be exploited to measure the local directional activity in different parts of the image. A trivial approach is to partition the image into non-overlapping blocks and process these blocks individually. Figure 55 (top) shows 256x256 Lena image partitioned into 8x8 pixel non-overlapping blocks. The first level wavelet decomposition of the image is shown in Figure 55 (bottom). In these subbands, corresponding 4x4 blocks represent the same spatial area as 8x8 pixel blocks in the original image. In these figures, the extracted blocks are examples of areas where horizontal and vertical activity are dominant. As can be seen from the corresponding wavelet subband blocks, the wavelet coefficients in relevant high-frequency subbands contain large number of high-valued coefficients (lighter tones represent higher-valued coefficients while darker tones represent lower-valued coefficients).

The proposed algorithm is as follows: Let $LH_0$ and $HL_0$ denote respectively the high-frequency horizontal and vertical detail subbands of $X$ and $w_{h}^{r,s}$ and $w_{v}^{r,s}$ denote the wavelet coefficients at position $(r,s)$ of $LH_0$ and $HL_0$ respectively. We partition $X$ into non-overlapping blocks and estimate horizontal and vertical activity measures for the $p$th block as:

$$A_p^h = \sum_{r,s} |w_{h}^{r,s}|$$  \hspace{1cm} (6.3)

$$A_p^v = \sum_{r,s} |w_{v}^{r,s}|$$  \hspace{1cm} (6.4)

where summation is over all wavelet coefficients contained in block $p$. Finally, block $p$ in the HR image is reconstructed by using the horizontal and vertical activity values as weights to the corresponding reconstructions generated using (6.3) and (6.4):

$$\hat{y}_p = \frac{A_p^h C_{p}^h + A_p^v C_{p}^v}{A_p^h + A_p^v}$$  \hspace{1cm} (6.5)
MATERIAL REDACTED AT REQUEST OF UNIVERSITY
6.4 Experimental Results

The proposed method has been tested on a number of well-known test images including Lena, Elaine, Baboon, and Peppers. An HR version of these images (512x512) was used as ground truth for performance evaluation purposes. The wavelet transform was implemented using the well-known Daubechies 9/7 filters. Our experiments showed that although the algorithm is relatively insensitive to the size of selected block size, 8x8 pixels blocks generate the best results.

Figure 56, Figure 57, Figure 58 and Figure 59 respectively show the enlarged images and amplified residual images for subjective comparisons for 4x reconstructions of Lena, Peppers, Baboon and Snoopy images. It can be seen that directional cycle spinning adapts better to edge orientation and avoids jagged edge (staircase) artefacts as well as ringing. An objective evaluation (in terms of PSNR) of the proposed algorithm is presented in chapter 7.

The top image in Figure 60 shows cross section of an edge (taken from the synthetic image in Figure 52) and corresponding reconstructed versions using various methods. The bottom image in this figure illustrates the absolute error of these reconstructions. These plots reveal that the edges reconstructed using the proposed method is closer to the real one.

Overall our results show that the directional cycle-spinning outperforms the competing methods for a variety of images and offers modest but consistent improvements over baseline cycle-spinning. Additionally, the proposed scheme offers a reduction in computational complexity relative to conventional cycle spinning because only shifts in the normal to salient edges are required, involving $2k$ WZP calculations ($k$ for horizontal direction and $k$ for vertical direction) compared with a total of $k^2$ in conventional cycle spinning.

6.5 Discussion

It might be argued that a number of extensions could be made to the proposed directional cycle-spinning algorithm. These potential extensions, as well as the reasons for excluding these are as follows:

- When both horizontal and vertical activity is relatively low, instead of utilizing HCS and VCS, the corresponding block from the WZP could be used. Activity level could be measured by summing up the horizontal and vertical activity: $C^h + C^v$, and low activity decision could be taken if this sum is below a predefined threshold. However experiments show that the improvement in the results are only marginal implying that combination of HCS and VCS does a good job for smooth regions which also eliminates the problem of having to select a threshold for "low activity" which is likely to be different for images having different characteristics.
The proposed algorithm exploits the horizontal and vertical directions and doesn’t do any processing in diagonal direction. Diagonal cycle-spinning could also be incorporated into the algorithm as well as the horizontal and vertical in a similar fashion, i.e. by measuring the diagonal activity using $HH_0$—diagonal detail subband—and using this measure as a weight to the diagonal component. However, the experiments reveal that incorporating the diagonal component has little effect on the overall performance in terms of PSNR. This could be

Figure 60: Cross section of an edge and its reconstructed versions using various methods (top) and absolute error of the reconstructed edges (bottom).
attributed to the widely accepted fact that the human visual system is relatively less sensitive to diagonal detail. This is an established consideration in image coding systems which invariably tolerate higher levels of quality loss in diagonal detail.

- It has been found empirically that, optimal results are obtained with 8x8 block size. The blocks used to calculate the directional activity are non-overlapping blocks corresponding to the same spatial location with the blocks used in reconstruction. On the other hand, it might be argued that overlapping blocks which span a larger spatial area than the non-overlapping blocks used in construction and having the same centre with the overlapping blocks could be used to avoid blocking artefacts. Again, the experiments reveal only a slight increase in the performance when a 50% larger overlapping blocks (12x12) are used to calculate the directional activity while the same size blocks (8x8) are used in reconstruction.

6.6 Conclusions

A directional variant of the cycle spinning methodology was considered for image resolution enhancement in the wavelet domain. The proposed method estimates local edge orientation from a wavelet decomposition of the available low resolution image and uses this information to influence cycle spinning parameters. Our experimental results confirm that the proposed method outperforms competing methods for the benchmark test images offering modest but consistent improvements.
CHAPTER 7

Comparative Assessment

7.1 Introduction

In the context of this thesis, a number of wavelet domain algorithms have been proposed for image resolution enhancement problems. These algorithms depending on the method used in the resolution enhancement process are as follows:

- Coefficient estimation (Chapter 4)
- Cycle spinning (Chapter 5)
- Cycle spinning followed by edge rectification (Chapter 5)
- Directional cycle spinning (Chapter 6)

The chapters in which these algorithms are introduced demonstrate examples for subjective evaluation. In this chapter we provide comparisons in terms of PSNR with a variety of techniques.

7.2 Assessment Process

We compared the proposed algorithms with a variety of wavelet domain methods as well as other established methods. We used Matlab® 6.1 (MathWorks, 2001) for implementation of the algorithms. For implementation of basic wavelet routines, we used Wavelab 802 which has been developed by researchers in Stanford University, Department of Statistics (1999). The methods we evaluated are as follows:

- Non-wavelet based methods:
  - Bilinear Interpolation: This is the conventional interpolation method widely utilised in practical applications. Bilinear interpolation algorithm is commonly found in commercial software and hardware implementations. Also it is used as a benchmark method in overriding majority of technical papers. We utilised the implementation provided in Image Processing Toolbox of Matlab® for evaluation purposes.
  - Bicubic Interpolation: Bicubic interpolation produces improved results compared to bilinear interpolation with the increased computational overhead. We utilised the bicubic interpolation implementation provided in Image Processing Toolbox of Matlab®.
Chapter 7. Comparative Assessment

- New Edge Directed Interpolation (NEDI): We used this algorithm as a representative method for discontinuity adaptive techniques. Authors kindly provided Matlab® source code for this algorithm, which is available to download freely (Li and Orchard, 2001b).

- Genuine Fractals: We used this algorithm as it is a state-of-the-art fractal based method, to represent this set of algorithms. A limited evaluation version of this fractal based method was available in the related website (LizardTech, 2004).

- Wavelet domain methods:
  - WZP using Haar wavelets: WZP is the most basic wavelet domain algorithm. WZP has been implemented as described in chapter 3. The filter coefficients for the Haar filter pair are given in Table 1.
  - WZP using Daubechies 9/7 wavelets: The filter coefficients for the Daubechies 9/7 filter pair are given in Table 2.
  - Regularity Preserving Image Interpolation: We implemented this method as described in (Carey et al., 1999).
  - Image Interpolation Using Wavelet-Based Hidden Markov Trees: We implemented this method as described in (Kinebuchi et al., 2001).

As a subjective evaluation measure, we used Peak Signal-to-Noise Ratio (PSNR) criteria. PSNR is defined in Appendix B.

The test images used are shown in Appendix C, characteristics of these images can be briefly explained as follows:

- Lena: A typical head-shoulder image. The background has fairly strong straight edges.

- Elaine: The foreground is a head-shoulder image similar to Lena while the hair contains more edges rather than finer details. The background of this image is not in focus and varies smoothly, lacking strong edges.

- Baboon: An image with high amount of fine details. Although this image has abundance of high-frequency information, most are not in the form of edges.

- Peppers: A natural image containing a number of objects with distinct boundaries.

- Snoopy: A hand-sketched cartoon image which mainly consists of lines with a white background.
### Table 8: PSNR (dB) values for 2x enlargement (256x256 to 512x512)

<table>
<thead>
<tr>
<th>Image/Method</th>
<th>Lena</th>
<th>Elaine</th>
<th>Baboon</th>
<th>Peppers</th>
<th>Snoopy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bilinear</td>
<td>30.13</td>
<td>30.60</td>
<td>22.85</td>
<td>30.01</td>
<td>18.84</td>
</tr>
<tr>
<td>Bicubic</td>
<td>31.34</td>
<td>31.17</td>
<td>22.98</td>
<td>30.28</td>
<td>20.33</td>
</tr>
<tr>
<td>Genuine Fractals</td>
<td>33.65</td>
<td>32.71</td>
<td>23.78</td>
<td>33.79</td>
<td>24.12</td>
</tr>
<tr>
<td>NEDI</td>
<td>34.10</td>
<td>32.89</td>
<td>23.87</td>
<td>33.54</td>
<td>23.88</td>
</tr>
<tr>
<td>WZP (Haar)</td>
<td>31.46</td>
<td>31.71</td>
<td>23.61</td>
<td>31.45</td>
<td>21.35</td>
</tr>
<tr>
<td>WZP (Db. 9/7)</td>
<td>34.45</td>
<td>33.26</td>
<td>24.22</td>
<td>33.94</td>
<td>24.63</td>
</tr>
<tr>
<td>HMM</td>
<td>34.52</td>
<td>33.31</td>
<td>24.24</td>
<td>34.04</td>
<td>24.70</td>
</tr>
<tr>
<td>Regularity Preserving</td>
<td>34.48</td>
<td>33.29</td>
<td>24.24</td>
<td>34.03</td>
<td>24.69</td>
</tr>
<tr>
<td><strong>Coefficient Estimation</strong></td>
<td><strong>35.39</strong></td>
<td><strong>33.40</strong></td>
<td><strong>24.52</strong></td>
<td><strong>34.46</strong></td>
<td><strong>25.59</strong></td>
</tr>
<tr>
<td>WZP and CS</td>
<td>34.93</td>
<td>33.56</td>
<td>24.28</td>
<td>34.32</td>
<td>23.63</td>
</tr>
<tr>
<td>WZP, CS and ER</td>
<td>35.23</td>
<td>33.69</td>
<td>24.43</td>
<td>34.63</td>
<td>23.81</td>
</tr>
<tr>
<td>WZP and Directional CS</td>
<td>35.09</td>
<td>33.73</td>
<td>24.37</td>
<td>34.50</td>
<td>23.90</td>
</tr>
</tbody>
</table>

### Table 9: PSNR (dB) values for 4x enlargement (128x128 to 512x512)

<table>
<thead>
<tr>
<th>Image/Method</th>
<th>Lena</th>
<th>Elaine</th>
<th>Baboon</th>
<th>Peppers</th>
<th>Snoopy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bilinear</td>
<td>24.06</td>
<td>25.38</td>
<td>20.43</td>
<td>24.37</td>
<td>16.45</td>
</tr>
<tr>
<td>Bicubic</td>
<td>26.76</td>
<td>28.93</td>
<td>21.02</td>
<td>26.86</td>
<td>17.16</td>
</tr>
<tr>
<td>Genuine Fractals</td>
<td>28.01</td>
<td>29.58</td>
<td>21.10</td>
<td>28.55</td>
<td>17.75</td>
</tr>
<tr>
<td>NEDI</td>
<td>28.81</td>
<td>29.97</td>
<td>21.18</td>
<td>28.52</td>
<td>17.49</td>
</tr>
<tr>
<td>WZP (Haar)</td>
<td>26.67</td>
<td>28.06</td>
<td>21.11</td>
<td>26.89</td>
<td>16.40</td>
</tr>
<tr>
<td>WZP (Db. 9/7)</td>
<td>28.84</td>
<td>30.44</td>
<td>21.47</td>
<td>29.57</td>
<td>18.76</td>
</tr>
<tr>
<td>HMM</td>
<td>28.86</td>
<td>30.46</td>
<td>21.47</td>
<td>29.58</td>
<td>18.81</td>
</tr>
<tr>
<td>Regularity Preserving</td>
<td>28.81</td>
<td>30.42</td>
<td>21.47</td>
<td>29.57</td>
<td>18.78</td>
</tr>
<tr>
<td><strong>Coefficient Estimation</strong></td>
<td><strong>29.44</strong></td>
<td><strong>30.92</strong></td>
<td><strong>21.59</strong></td>
<td><strong>30.16</strong></td>
<td><strong>19.26</strong></td>
</tr>
<tr>
<td>WZP and CS</td>
<td>29.27</td>
<td>30.78</td>
<td>21.54</td>
<td>29.87</td>
<td>16.49</td>
</tr>
<tr>
<td>WZP, CS and ER</td>
<td>29.36</td>
<td>30.89</td>
<td>21.56</td>
<td>30.05</td>
<td>16.51</td>
</tr>
<tr>
<td>WZP and Directional CS</td>
<td>29.55</td>
<td>30.98</td>
<td>21.67</td>
<td>30.14</td>
<td>16.87</td>
</tr>
</tbody>
</table>
MATERIAL REDACTED AT REQUEST OF UNIVERSITY
Chapter 7. Comparative Assessment

7.3 Experimental Results

In our experiments we have used aforementioned test images: Lena, Elaine, Baboon, Peppers and Snoopy. Experimental results obtained by simulating the algorithms are shown in Tables 8 and 9 where PSNR performance is tabulated for 2x and 4x enlargement factors respectively.

As expected, the best results are obtained for Lena and Peppers images for all the proposed algorithms. This is due to the relatively more salient edge content in these images. Only marginal improvements could be obtained for the Baboon image, as this image contains large amount of isolated high-frequency content rather than more regular high-frequency content such as edges, the proposed algorithms do not work particularly well. PSNR improvements for Elaine image are somewhat lower than Lena and Peppers due to the blurry background in this image.

Subjective evaluations are presented in the chapters where the algorithms are introduced.

Figure 61, Figure 62, Figure 63 and Figure 64 illustrate 2x and 4x enlarged images using Genuine Fractals (LizardTech, 2004) and NEDI (Li and Orchard, 2001b). The NEDI algorithm first finds the edges and then interpolate along these. However, inaccuracy in the edge detection might result in interpolation of non-edge areas (where intensity changes occur) or of neighbourhood of edges, intensifying undesired features of an image. The vicinities of nose and eyes in the enlarged image in Figure 61 and the area where the stem and pepper meet in Figure 63 exhibit such artefacts. Although images enlarged with Genuine Fractals exhibit less artefact, the images lack detail and appear coarser. A closer inspection of pixel intensity values also reveal that they vary somewhat less in a neighbourhood relative to other techniques. Baboon image presented in Figure 64 remains a challenging image to enlarge for these particular algorithms as well.

7.4 Conclusions

Our results have shown that the proposed methods outperform conventional image interpolation approaches for the images tested, both in objective and subjective terms, while they also compare favourably with state-of-the-art methods operating in the wavelet domain, utilising fractals and discontinuity adaptive techniques.

It should be noted that the PSNR results taken as the only measure for the reconstruction quality is not sufficient for the most cases as the constructed images show different characteristics.

In the selection of a particular algorithm, the application area and image characteristics should be considered. Cycle spinning based methods produce smoother images which are more pleasant to eye, on the other hand, the coefficient estimation based method preserves the finer details better. The coefficient estimation approach is suitable for real-time operations as it is computationally less demanding than cycle spinning based methods.
CHAPTER 8

Conclusions

8.1 Summary

In the context of this thesis, we first introduced the image enlargement problem and scenarios when it might be required. We then presented the conventional methods used for the problem and gave an overview of both the classic and contemporary approaches suggested in the literature. After giving a brief introduction into the wavelet transform, we showed that wavelets could be utilised to achieve resolution enhancement and wavelet domain methods have potential to achieve noteworthy results. We formulated the wavelet domain approaches by relating the LR image in hand and unavailable HR image. We have done it by using the commonly used model where LR image is assumed to be normalised version of first level low-frequency subband of the HR image that is required to be estimated. Motivated by the potential of this particular set of algorithms, we turned our focus into the wavelet domain methods to achieve image enlargement. We introduced the basic methods utilising the wavelet transformation and then described the more advanced wavelet domain methods proposed in the literature. While doing so, we also recited the potential problems associated with these methods. Moreover, we mentioned the lack of any formal classification of wavelet domain methods and infrequent citing of other similar algorithms. We aimed to fill this gap with an extensive review of the methods available and by categorizing them according to the essence of their approach.

In the following chapters, we proposed three new methods to achieve the enlargement. The first proposed method involves estimation of unavailable high-frequency subband coefficients. The algorithm operates in a quad-tree decomposition framework and exploits wavelet coefficient correlation in a local neighbourhood sense. The proposed method employs linear least-squares regression to estimate the unknown detail coefficients. The estimator is generated by utilising further levels of wavelet decomposition and using the coefficient correlation in these available subbands, eliminating the need to train using ground-truth data.

The next method that we proposed utilises cycle-spinning. We introduced the cycle-spinning as used in the wavelet domain denoising and demonstrated the promising results that are obtained when cycle-spinning is adapted to the problem in hand. Then, motivated with the fact that edges get wider with this process, we proposed that these results could further be improved upon by rectification of the edges in reconstructed images. We utilised a parametric edge model to classify edges and proposed an edge rectification algorithm using linear least squares regression for each of these clusters.
In the final method, we propose a directional variant of cycle-spinning methodology. The observation that the dominant ringing artefacts occurs normal to an edge motivates an approach where a directionally weighted cycle-spinning scheme is used. In this method, we use the information from wavelet subbands to measure directional activity and influence the cycle-spinning in this particular direction. By applying the cycle spinning to eliminate the artefacts and preventing processing along the edges, over-smoothing is prevented, as well as reducing the computational load.

As well as proposing these methods, we also brought the basic wavelet domain zero-padding method into light and showed that it is capable of over performing conventional techniques.

For all these proposed methods we showed that they deliver robust results and achieve consistent improvements for a variety of images by illustrating the simulation results. We stated that the algorithms work better particularly on images having abundance of edge information and fall short in the areas lacking salient image features.

In the previous chapter, we presented a comparative assessment of the proposed methods as well as a number of other methods presented in the literature and/or used in implementation of commercial imaging software. As well as putting the proposed algorithms into context by facilitating comparison with other established methods, this chapter also aims to present an evaluation of other relevant work as the evaluation criteria -and in many cases results in terms of PSNR for the same algorithms- in the literature often vary between different papers and a direct comparison is not always possible.

While the objective performances of the proposed methods in terms of PSNR are similar; visibly, the reconstructed images are significantly different. Cycle-spinning based methods generate smoother images while the coefficient estimation based method generates more detailed-looking images. When choosing a particular algorithm, the input image characteristics, intended purpose of use for the reconstructed high-resolution image and personal preference should be considered. For example, when trying to recognise a person, the coefficient estimation based method would potentially be more suitable. On the other hand, for printing images, cycle-spinning based methods would be preferable as they are free from artefacts due to introduction of estimated high-frequency domain information.

In cycle-spinning based methods, the directionally adaptive method would be preferable in many cases to the method utilising cycle-spinning followed by edge rectification due to the following reasons: (i) the directional method is computationally less expensive as it doesn’t require all the processing (ii) it doesn’t require edge rectification, hence avoiding the necessity of a stage for learning the edge rectification parameters.
MATERIAL REDACTED AT REQUEST OF UNIVERSITY
Reconstructions using the proposed methods are illustrated in Figure 66 and Figure 67. We attempted resolution enhancement on the image in Figure 1, which was printed in a newspaper. Although the original low-resolution image wasn’t available, we generated a low-resolution version by first filtering and down-sampling of the image in hand by a factor of 4x and then reconstructed the images by using nearest neighbour interpolation and each of the methods proposed in this thesis.

The limits of the enlargement is illustrated in Figure 68 where 8x and 16x enlarged extracts from Lena image is shown. Although the proposed algorithms generate better results than bilinear interpolation at both for 8x and 16x enlargement factors, the image quality starts to deteriorate below acceptable quality levels. Even though 8x enlargement could still be useful in some cases, the usefulness of 16x enlargement factor is questionable as the image details start to disappear.

8.2 Future Work

In the context of this thesis, we proposed three independent algorithms for image resolution enhancement. Although these methods are complete and as we demonstrated generates considerably good results, there are some potential enhancements to develop these further. We discuss these potential enhancements and speculate about probable directions for the resolution enhancement techniques in the remainder of this section.

The coefficient estimation method introduced in chapter 4 utilises a linear estimation method to combine neighbouring coefficients and generate an estimated higher frequency subband coefficient. A potential direction of research would be investigation of utilisation of non-linear methods for this purpose. For example, a set if estimator weights could be used for step edges and a different set for the line edges, which is more suitable to the edge characteristics. Step and line edges are illustrated in Figure 69.

![Figure 69: Illustration of a step edge (left) and a line edge (right)](image)

For the method introduced in chapter 5, a more sophisticated edge reconstruction method could be considered for the edge rectification part. A possibility would be to utilise an object based algorithm where the objects in the image are recognised and enlarged individually with utilisation of the futures specific to the objects rather than considering a general image model. A recognition
based super-resolution technique (also called hallucination) is proposed in (Baker and Kanade, 2002), although not robust enough to use reliably has found some attention recently. The method has been applied to face and text images and utilised super-resolution methodology. Similar learning based approaches could be developed for single image resolution enhancement if the characteristics of the image to be enlarged are known in advance. The system could be trained using particulars of these images and prior information could be utilised.

The directional cycle-spinning method in chapter 6 does a weighted combination of cycle-spinning in different directions, these weights are determined by the amount of directional activity in a particular direction. Different ways of combining these intermediary images could be an interesting subject to investigate. The combination method could be adaptive regarding other considerations, such as image prior information. Maximum-a-posteriori (MAP) could be utilised to maximise the likelihood using such information.

Algorithms utilising a single image are the only way forward when multiple images are not available or even when available, the object of interest does not appear in more than one image or might have been pictured from different views, prohibiting use of super-resolution methods. Also as already mentioned in chapter 2, super-resolution methods have some drawbacks such as difficulties caused by objects going through non-rigid motion and registration problems. However, if these problems could be overcome, in the availability of multiple images, super-resolution techniques could provide advantages over methods using single images.
Appendix A - Least Squares Estimation

Least squares estimation defined as (Mathworld, 2005):

*A mathematical procedure for finding the best fitting curve to a given set of points by minimising the sum of the squares of the offsets ("the residuals") of the points from the curve.*

Let $y$ is the solution, i.e. the desired value, in this case it is the real coefficient value. $x_1, x_2, \ldots, x_k$ are the observation values, which are the initial estimates at the neighbourhood. Then the multiple linear-regression could be formulated as below:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k$$

The objective is to estimate the $\beta$ values which minimises the sum of squared error (SSE):

$$E = \sum_{i=1}^{n} (y - \beta_0 - \beta_1 x_1 - \beta_2 x_2 - \ldots - \beta_k x_k)^2$$

Using matrix algebra formulation:

$$Y = X\beta$$

Where $Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$, $X = \begin{bmatrix} x_{1,1} & x_{1,2} & \ldots & x_{1,k} \\ x_{2,1} & x_{2,2} & \ldots & x_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n,1} & x_{n,2} & \ldots & x_{n,k} \end{bmatrix}$ and $\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}$

Then the SSE could be written as

$$E = (Y - X\beta)^\top(Y - X\beta)$$

To minimise the error, we can take the derivative of the error and equate to zero which leads to the solution:

$$\beta^* = (X^\top X)^{-1}X^\top Y$$
Appendix B - Subjective Evaluation Criteria

For the measurement of distortion, we employed the commonly used Peak Signal-to-Noise Ratio (PSNR) metric which gives a measure of similarity to the original signal. A higher PSNR value indicates that the reconstructed image is more similar to the original. For an input image $X$ and an output image $Y$ each having $mxn$ pixels with a luminance range of 0 (black) to 255 (white) PSNR(dB) is defined as:

$$PSNR = 20 \log_{10} \frac{255}{RMSE}$$

where Root Mean Square Error (RMSE) is defined as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{m} \sum_{j=1}^{n} (X(i,j) - Y(i,j))^2}{mn}}$$
MATERIAL REDACTED AT REQUEST OF UNIVERSITY
Bibliography


Bibliography


Bibliography


