HYBRID TRANSFORM,
SPATIAL DECORRELATION AND
UNIFIED CODING SYSTEM FOR
IMAGE AND VIDEO COMPRESSION

LEE KA CHUN KENNETH

DOCTOR OF PHILOSOPHY
CITY UNIVERSITY OF HONG KONG
SEPTEMBER 2004
Special Consideration

The content in this thesis is currently under patent application and also to be used for patent application.

The thesis will be withheld from the City University Library for a period of at least 2 years.

Prof. Y. K. Chan would be very much appreciated if the content of this thesis is not disclosed to any third party.
A Hybrid Transform and Unified Coding Scheme for Image and Video Compression

Abstract

In this thesis, a novel compression model for still images and video was developed. Instead of coding image indiscriminately as a whole, as in conventional compression scheme such as JPEG and JPEG2000, the new algorithm analyzes image features, classifies into different types and processes each portion accordingly.

When there is spatial coherence, spatial decorrelation is performed prior to applying transform with either Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT), or their combination (hybrid transform).

When spatial coherence does not exist, either DCT, DWT or Hybrid Transform is chosen as the transform technique for best available compression efficiency.

In short, the model combines available spatial domain coherence, DCT and short-tap-length DWT in a non-trivial manner to achieve high compression efficiency.

Operationally, the model exploits the fact that typical images are always consisting of various image features. Specifically, there are edge-regions, smooth-regions, patterns and irregular regions, which behave rather differently in terms of bit-rate and Signal-to-Noise-Ratio (SNR) when compressed. The model selectively applies pure DCT, pure DWT or the hybrid transform to various
regions with respect to their corresponding image-feature characteristics. As a result, the psycho-visual perception of the decompressed image is significantly better than pure DCT schemes but yet the computation overhead for inverse transform is less than typical 9-7 tap wavelet and comparable to classical DCT schemes.

The model takes advantage of spatial domain correlation and could manage to further reduce bit-rate by feeding back image feature information and processed image data to the decode loop. In conjunction with the reduced bit-rate, the visual quality is also better.

Owing to the hybrid transform scheme and correlation-reduction using processed image data and image-feature information, the transformed coefficients now become much smaller in magnitude and better defined. A new entropy coding scheme for transformed coefficients was developed to take advantage of the structure of coefficients. The scheme fits equally well to Motion-Estimated frames, which is also characterized by the small magnitude coefficients.

When comparing with the start-of-the-art JPEG2000 implementations, the compression scheme developed in this thesis is competitive in terms of bit-rate / SNR for “still” and “video” frames. In most cases, the proposed scheme out-performs JPEG2000 for various still images especially the well-known “Barbara” image and “Bike” image, to name a few. For video frames, the scheme always out-performs JPEG2000, especially for fast motion video sequence. The coding scheme, as it can be applied to both “still images” and “video frames”, is referred to as an unified coding scheme.
Contents

Chapter 1 Introduction .............................................................. 1

1.0 Overview of this Introduction.................................................. 1
1.1 Motivation....................................................................................... 2
1.2 Contributions of this Thesis......................................................... 4
  1.2.1 Fast DCT transform with Intel MMX SIMD Instruction set ......... 5
  1.2.2 Hybrid Transform combining advantages of DCT and DWT.......... 6
  1.2.3 Preprocessing for decorrelation............................................. 6
    1.2.3.1 Directional Prediction ...................................................... 7
    1.2.3.2 Intra-Image Prediction ................................................. 7
  1.2.4 Unified Entropy Coding Scheme ........................................... 8
  1.2.5 Feature Oriented Coder with Unified Scheme ....................... 8
1.3 Outline of this thesis ............................................................... 9

Chapter 2 Background and Review on Image Compression..... 10

2.0 Overview of this chapter........................................................... 10
2.1 Still Image Compression ............................................................ 11
  2.1.1 JPEG - Classical DCT based coder ....................................... 11
  2.1.2 JPEG2000 - Start-of-the-art Wavelet based coder.................. 13
  2.1.3 Concept of Hybrid, Image-Feature-Oriented coding ............... 15
  2.1.4 Non-mainstream image compression schemes ...................... 17
2.2 Video compression........................................................................ 18
  2.2.1 MPEG...................................................................................... 18
  2.2.2 Advanced Video Coding / H.264 ......................................... 20

Chapter 3 Implementation of fast software-based DCT routine
with SIMD instruction set............................................................ 21

3.0 Overview of this chapter........................................................... 21
3.1 Introduction .................................................................................. 22
3.2 Operation of Discrete Cosine Transform...................................... 24
3.3 Efficient DCT routine with SIMD instructions............................ 26
  3.3.1 First Pass of DCT ................................................................. 26
  3.3.2 Second Pass of DCT.............................................................. 29
List of Figures

Figure 2.1: "F-16" Image compressed by JPEG (left) and JPEG2000 (right) at 0.25bpp.......................................................... 15

Figure 2.2: Original "Woman" Image on left hand side and JPEG2000 compression on right hand side at 0.25bpp ......................... 16

Figure 3.1: Accumulation of table entries in first pass DCT. In each 16-bit word, 4 bits are allocated as fraction .............................................. 29

Figure 3.2: Accumulation of table entries in second pass DCT. Vectors of lower-order bit-group contribute more bits to sub-integer accuracy. ........................................................................................................ 33

Figure 3.3: Interleaving of Odd (64bit MMX) and Even (32bit Register) DCT coefficients. .......................................................... 35

Figure 3.4: Random 8x8 data block for testing (left) and its transformed result (right)............................................................................. 39

Figure 3.5: Output error achieved by the new DCT routine (left) and that from IJG (right)................................................................. 39

Figure 4.1: Feathers (left) and Shoulder (right) of Lena Image, compressed by LLEC at 0.25bpp ......................................................... 44

Figure 4.2: Primitive patterns of 8x8 DCT......................................................................................................................... 45

Figure 4.3: Applying 4-tap Daubechies kernel to 8x8 block .................... 46

Figure 4.4: Flow of the forward hybrid transform process ...................... 49

Figure 4.5: 8x8 block at (72,16) of the original Lena image ..................... 61

Figure 4.6: Portion of Lena Image compressed with typical DCT scheme (left) and the hybrid DCT/WT scheme (right) ......................... 61

Figure 4.7: Block distribution of Lena Image, showing decompressed image of DCT/WT scheme. DCT blocks are blackened ...................... 63

Figure 4.8: Lena Image compressed with DCT (left) and DCT/WT (right) ... .............................................................................................. 64
Figure 5.1: Commonly used directions for directional decorrelation ........ 67
Figure 5.2: The two strategies for "reference pixel" selection, along the 45 degree axis ................................................................. 68
Figure 5.3: Part of the "Barbara" image using typical prediction scheme. Angles are intentionally forced to 63 degree ......................... 70
Figure 5.4: Part of the "Barbara" image using weighted average of neighbors Angles are intentionally forced to 63 degree ......................... 72
Figure 5.5: Eight more directions for directional decorrelation ........ 73
Figure 5.6: Predicted image of Lena, using 8 directions (left) and 16 directions (right) respectively .................................................. 74
Figure 5.7: Predicted image of Barbara, using 8 directions (left) and 16 directions (right) respectively .................................................. 75
Figure 5.8: Typical scheme: The same intensity propagates along direction ................................................................. 76
Figure 5.9: Operation steps for intensity interpolation ....................... 77
Figure 5.10: Direction prediction on almost-straight line .................. 80
Figure 5.11: Prediction of the line in Figure 5.10, using 45 degree (left) and the prediction error (right) .................................................. 80
Figure 5.12: Prediction of the line in Figure 5.10, with predicted block shifted by half pixel ................................................................. 81
Figure 5.13: Original of Lena image (left), residue with origin at (0,0) (middle) and residue with origin shifted by half pixel (right) .............. 82
Figure 5.14: Images having multi-directional characteristics: Lena (left), Barbara (middle) and Boat (right) ........................................... 83
Figure 5.15: Feather of Lena's hat: Uni-directionally predicted (left), Original (middle) and Multi-directionally predicted (right) .............. 84
Figure 5.16: Residue of uni-directional prediction (left) and multi-directional prediction (right) ................................................................. 85
Figure 5.17: Tablecloth of Barbara, by uni-directional prediction (left) and multi-directional prediction (right) ........................................... 86
Figure 5.18: Predicted image of Lena with existing scheme .......................... 87

Figure 5.19: Predicted image of Lena with all the techniques presented in this thesis................................................................. 88

Figure 6.1: LZSS compressor, which discovers repeated structure in history ................................................................. 92

Figure 6.2: Self-similarity in various test images ...................... 93

Figure 6.3: Barbara image predicted with intra-image prediction ....... 97

Figure 6.4: Residue of the predicted Barbara image in Figure 6.3 ........ 98

Figure 6.5: Extended search for Motion Estimation, considering blocks at rotated axis (figure shows the counter-clockwise directions only) ........................................................................ 100

Figure 6.6: Frame 0 and Frame 1 of Football sequence (352x240), showing massive motion.................................................. 101

Figure 6.7: Frames of "Lady" sequence (320x200), showing slight motion ................................................................. 101

Figure 6.8: Frames of "Parade" sequence (320x200), showing moderate motion (both translation and rotation).......................... 102

Figure 6.9: Frames of "Basketball" sequence (352x288) (captured from TV sport show), showing massive motion ...................... 102

Figure 7.1: Barbara Image (left) and the corresponding residue(right) .... 106

Figure 7.2: Comparison of the magnitudes of transformed coefficients (Image vs. Residue) ............................................................. 107

Figure 7.3: Comparison of the number of non-zero ACs inside each transformed 8x8 block.......................................................... 107

Figure 7.4: Comparison of coding efficiency (LLEC vs. new scheme) on typical 8x8 block (left) and rare 8x8 block (right) ........... 112

Figure 7.5: New coding scheme achieves lower bit size for blocks with regular order. ................................................................. 113

Figure 8.1: Flow of the FOCUS encode process ........................................... 119
Figure 8.2: Wallpaper from BANDAI used in testing. The image is used to evaluate the behavior of different compression schemes on computer graphics ................................................................. 129

Figure 8.3: Captured screen on Windows 98. The image is used to test the behavior of compression schemes on computer generated GUI with text .................................................................................. 129

Figure 8.4: Scanned image of one of our to-be-submitted papers, printed with laser printer. The image is used to test the behavior of different compression schemes on hardcopy documents ......................... 130

Figure 8.5: Frame 0 of Matrix sequence (720x480) ............................................ 133

Figure 8.6: Residue by ordinary motion estimation scheme (left), and the residue using intra-directional prediction (right) ....................... 136
List of Tables

Table 3.1: Runtime comparison of different DCT routines at 20,000,000 iterations.................................................................37

Table 3.2: File size and PSNR of various images compressed by JPEG with different DCT routines ........................................40

Table 4.1: File size and PSNR comparison between DCT and Hybrid DCT/WT on various 512x512 test images..............................62

Table 5.1: LLEC file size and PSNR of direction decorrelated residues, using "direct pixel duplication" and "weighted average" method respectively........................................................................72

Table 5.2: LLEC file size and PSNR of direction decorrelated residues, using 8 directions and 16 directions respectively ..............75

Table 5.3: LLEC file size and PSNR of direction decorrelated residues, using intensity interpolation..............................................79

Table 5.4: LLEC file size and PSNR of direction decorrelated residues, with origin shifted by half pixel........................................83

Table 5.5: LLEC file size and PSNR of decorrelated residues, with uni-directional and multi-directional technique respectively.......86

Table 5.6: LLEC file size and PSNR of decorrelated residues, using all the enhancement techniques presented in thesis....................88

Table 6.1: PSNR and LLEC file size of various test images, using intra-image decorrelation method..............................................98

Table 6.2: PSNR and LLEC file size of various video frames, using rotated motion estimation.....................................................102

Table 7.1: Variable Length Code employed in LLEC scheme ...................... 108

Table 7.2: Block Type classification in the new coding scheme ............ 110

Table 7.3: Compressed file size of decorrelated image by LLEC and the new coding scheme ................................................... 114

Table 7.4: Compressed file size of motion estimated residue by LLEC and the new coding scheme ............................................. 114
Table 8.1: PSNR achieved by state-of-the-art still image compression schemes at a target bitrate of 0.25bpp on photorealistic images .......... 123

Table 8.2: PSNR achieved by state-of-the-art still image compression schemes at a target bitrate of 0.50bpp on photorealistic images .......... 124

Table 8.3: PSNR achieved by state-of-the-art still image compression schemes at a target bitrate of 1.00bpp on photorealistic images .......... 125

Table 8.4: PSNR achieved by FOCUS and JPEG2000 at a target bitrate of 0.5bpp ................................................................. 131

Table 8.5: PSNR achieved by FOCUS and JPEG2000 at a bitrate of 0.75bpp ................................................................................. 132

Table 8.6: Compression bitrate achieved by FOCUS on video frames....... 134

Table 8.7: Compression bitrate achieved by FOCUS on Foreman frame 0 and frame 15................................................................. 135

Table 8.8: Compression bitrate achieved by FOCUS on Football frame 0 and frame 15................................................................. 135

Table 8.9: Memory Consumption of various compression schemes (Decoder side) ................................................................. 138

Table 8.10: Memory consumption of Cosman and Zeger’s scheme disclosed in [35] ................................................................................. 139