

Geometrically Invariant Digital Watermarking Using Robust
Feature Detectors

by

Xiao-Chen Yuan

Doctor of Philosophy in Software Engineering



Faculty of Science and Technology

University of Macau



Geometrically Invariant Digital Watermarking Using Robust
Feature Detectors

by

Xiao-Chen Yuan

SUPERVISOR: Prof. Chi-Man Pun

Department of Computer and Information Science

Doctor of Philosophy in Software Engineering

2013

Faculty of Science and Technology

University of Macau



Author's right 2013 by
YUAN Xiao-Chen





Acknowledgements

I would like to take this opportunity to express my gratitude towards everyone who contributed towards the successful completion of this thesis, especially my supervisor, Prof. Chi-Man Pun, for his constant encouragement, and the belief that he showed in my abilities. Besides his instructive advice and useful suggestion on my thesis, I am also deeply grateful for his help in the completion of this thesis. He has walked me through all the stages of the writing of this thesis. Without his consistent and illuminating instruction, this thesis could not have reached its present form.

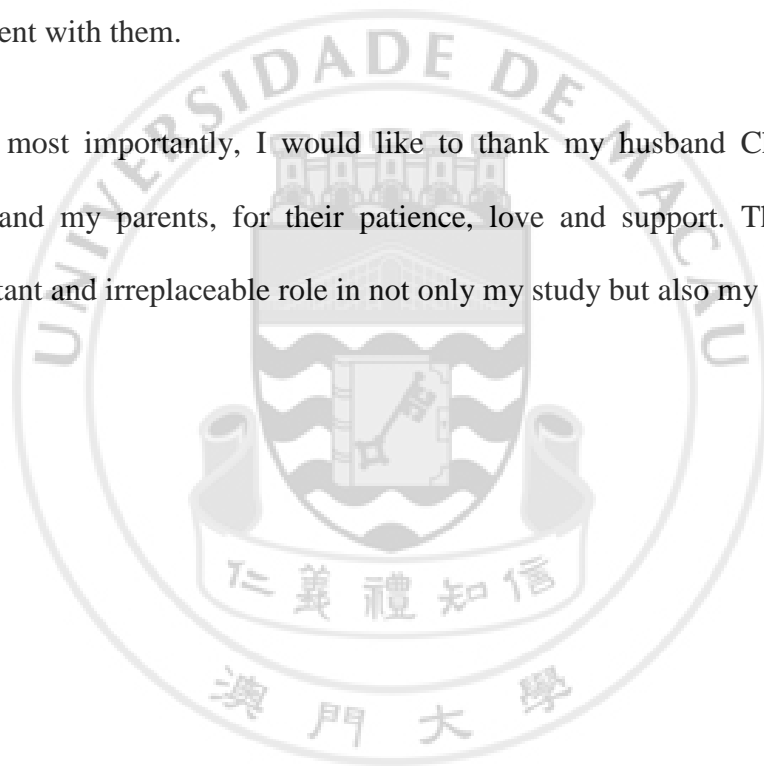
During my Ph.D. studies in University of Macau, I was provided with the best campus life, with first-class hardware and software infrastructures and a friendly environment for studying and researching. I want to thank the University of Macau for providing me the opportunity with abundant resources to conduct my research, such as a rich database of resources.

In particular, I would like to express my great thanks to Prof. C. L. Philip Chen, Prof. Yuan-Yan Tang, Prof. En-Hua Wu, Prof. Zhi-Guo Gong, Prof. Jing-Zhi Guo, Prof. Yi-Cong Zhou, and Prof. Long Chen, who gave me a lot of suggestions during my studies and works.

I would also like to thank William Sio, the lab technician who helped me a lot in my daily work, and Prof. Xiao-Lin Tian, professor in Macau University of Science and

Technology, who led me into the world of image processing. And I also owe my sincere gratitude to my friends and my colleagues: Hong-Min Zhu, Cong Lin, and Ning-Yu An, who gave me their help and time in listening to me and helping me work out my problems during the difficult course of the thesis. I have benefited a lot from the time spent with them.

Finally and the most importantly, I would like to thank my husband Chin-Ming Jimmy, Huang and my parents, for their patience, love and support. They have played an important and irreplaceable role in not only my study but also my life.



Abstract

Geometrically invariant digital watermarking schemes based on robust feature detectors are proposed in this thesis. First, three types of feature detectors are proposed for digital image watermarking: the Edge Based Feature Detector, the SIFT Based Feature Detector, and the Adaptive Harris Based Detector. The Edge Based Feature Detector is proposed based on edge detection and it can extract a unique feature in the specific region. The SIFT Based Feature Detector is proposed by improving SIFT algorithm to produce more robust feature points for digital image watermarking, and it can extract number of feature points. The Adaptive Harris Based Detector is proposed by revamping and enhancing the Harris corner detector and it can also extract a number of reliable feature points.

The three detectors are proven to be highly robust against both geometric attacks and also common signal processing. After locating the features for watermarking, two watermarking methods for different types of watermark are proposed: the histogram distribution based watermarking method, for a sequence of watermark data bits. And the Zernike transform based watermarking method for embedding data sequence of specific distribution and detecting its existence during the watermark extraction process.

Besides digital images, the feature extraction based watermarking scheme can also be applied in digital audio clips as well. The Robust Audio Feature Detector is proposed to extract features from digital audio clips. Then, the Stationary Wavelet Transform is applied to the extracted regions, and thus the regions are decomposed into approximation and detail coefficients. Afterwards, the watermark is embedded /

extracted into / from the approximation coefficients with the spread spectrum communication techniques.

Experiments are conducted to evaluate the performance of the proposed watermarking schemes. The proposed algorithms are proven to be robust against most of the attacks, including common signal /audio processing and geometric distortions. Furthermore, they outperform the existing representative works when under common signal / audio processing and geometric distortions.



Declaration

I declare that the thesis here submitted is original except for the source materials explicitly acknowledged and that this thesis as a whole or any part of this thesis has not been previously submitted for the same degree or for a different degree.

I also acknowledge that I have read and understood the Rules on Handling Student Academic Dishonesty and the Regulations of the Student Discipline of the University of Macau.





Table of Contents

Acknowledgements	i
Abstract	iii
Declaration	v
Table of Contents	vii
List of Figures	x
List of Tables	xvi
List of Abbreviations	xvii
Chapter 1 Introduction	1
1.1 General Background	1
1.2 Specific Background.....	4
1.3 Research Goals and Objectives	8
1.4 Research Methodology and Design.....	10
1.5 Potential Contributions.....	13
1.6 Organization of the Thesis.....	14
1.7 Statement of Originality.....	16
Chapter 2 Related Work	20
2.1 Invariant-Domain-Based Watermarking Schemes	21
2.2 Histogram-Based Watermarking Schemes	23
2.3 Template-Based Watermarking Schemes.....	25
2.4 Features-Based Watermarking Schemes	27
2.5 Decomposition-Based Watermarking Schemes	30
Chapter 3 Robust Feature Detectors for Digital Image Watermarking	33

3.1	Edge Based Feature Detector.....	35
3.2	SIFT Based Feature Detector.....	37
3.2.1	Scale Invariant Feature Transform Algorithm	38
3.2.2	SIFT Based Feature Detector Algorithm	42
3.3	Adaptive Harris Based Detector	46
3.3.1	Harris Corner Detector	47
3.3.2	Adaptive Harris Based Detector Algorithm	49
Chapter 4 Geometrically Invariant Watermarking Methods		54
4.1	Histogram Distribution Based Watermarking	55
4.1.1	Watermark Embedding Procedure	56
4.1.2	Watermark Extraction Procedure	60
4.2	Zernike Transform Based Watermarking	63
4.2.1	Zernike Moments and Invariance Properties	64
4.2.2	Watermark Embedding Procedure	67
4.2.3	Watermark Extraction Procedure	70
Chapter 5 De-Synchronization Resilient Audio Watermarking		75
5.1	Robust Audio Feature Detector	77
5.2	Stationary Wavelet Transform Based Audio Watermarking	86
5.2.1	Stationary Wavelet Transform	87
5.2.2	Watermark Embedding Procedure	90
5.2.3	Watermark Extraction Procedure	91
Chapter 6 Experimental Results for Digital Image Watermarking		94
6.1	Edge Based Feature Detector and Zernike Transform Based Watermarking Results.....	95

6.1.1	Watermarking Performance under Different Distortions.....	97
6.1.2	Performance Comparison.....	103
6.2	SIFT Based Feature Detector and Zernike Transform Based Watermarking Results.....	105
6.2.1	Watermarking Performance under Different Distortions.....	109
6.2.2	Performance Comparison.....	115
6.3	Adaptive Harris Based Detector and Histogram Distribution Based Watermarking Results.....	119
6.3.1	Performance under Different Distortions.....	123
6.3.2	Performance Comparison.....	137
Chapter 7	Experimental Results for Digital Audio Watermarking	140
7.1	Performance under Different Distortions	143
7.2	Performance Comparison	149
Chapter 8	Conclusions	162
8.1	Summarization.....	162
8.2	Limitations of Current Study.....	164
8.3	Perspectives for Future Work.....	165
	References	167
	Curriculum Vitae	178

List of Figures

Figure 1.1: Diagram of a watermarking system	2
Figure 1.2: Framework of the Research	10
Figure 2.1: Zheng’s invariant domain-based watermarking algorithm.....	22
Figure 2.2: Lin’s histogram-based watermarking algorithm.....	24
Figure 2.3: Pereira and Pun’s template-based watermarking method.....	26
Figure 2.4: Tang and Hang’s feature-based watermarking algorithm	28
Figure 2.5: Feature extraction by Mexican Hat Wavelet scale interaction	29
Figure 2.6: Xin’s decomposition-based watermarking algorithm.....	31
Figure 3.1: Circular Patch Extracted by EBFD under Various Attacks.....	37
Figure 3.2: SIFT Feature Points’ Generation	40
Figure 3.3: Feature Points Descriptor. (a) computation of the gradient magnitude and orientation at each image sample point in a region around the feature point location (b) the 4x4 descriptors computed from a 16x16 sample array	41
Figure 3.4: Flow Chart of SBFD	44
Figure 3.5: Feature Extraction by SBFD When under Various Attacks. (a1), (b1), (c1), and (d1): original ‘Baboon’, ‘Bridge’, ‘Lena’, and ‘Pepper’. (a2), (b2), (c2), and (d2): 45° rotation with cropping. (a3), (b3), (c3) and (d3): 20% vertical shearing. (a4), (b4), (c4), and (d4): 20% horizontal shearing. (a5), (b5), (c5) and (d5): 10% affine transformation. (a6), (b6), (c6), and (d6): scaling with the scale factor as 0.5. (a7), (b7), (c7) and (d7): 4x4 median filtering. (a8), (b8), (c8) and (d8): JPEG compression with the quality factor as 50.	46
Figure 3.6: Feature Points Extraction Comparisons (a1)-(c1) Feature points extracted with the traditional Harris Detector, from ‘Baboon’, ‘Boat’, and ‘Lena’, respectively. (a2)-(c2) Feature points extracted with the proposed AHBD, from ‘Baboon’, ‘Boat’, and ‘Lena’, respectively.	50
Figure 3.7: Feature Points Extracted by AHBD when under Various Distortions (a)-(f) respectively shows the feature points detection under the circumstance of: (a) Original image, (b) Flipping, (c) JPEG compression, quality factor = 30, (d) Scaling, scale factor = 0.5, (e)	

3x3 Gaussian low-pass filtering, standard deviation = 1.5, and (f) ‘Salt & Pepper’ noise addition, variance = 0.02.....	52
Figure 4.1: Flow Chart of Histogram Distribution Based Watermark Embedding Procedure.....	56
Figure 4.2: Embedding Region Demonstration. (a) Extracted feature points in host image, (b) Embedding Region.....	58
Figure 4.3: Histogram Modification. X-axis means the pixel intensity level, and Y-axis means the number of pixels for each intensity level. IL_l is the minimum intensity-level value, IL_h is the maximum intensity-level value of pixels in the corresponding region, IL_m is intensity-level value of the pixel in the middle position after pixels in the corresponding region are sorted. $IL_{l'}$ and $IL_{h'}$ are the two margins calculated to cause a pixel to move unambiguously. MP1 and MP2 are the two values calculated to define the pixels to be moved.	60
Figure 4.4: Flow Chart of Histogram Distribution Based Watermark Extraction	62
Figure 4.5: Flow Chart of Zernike Based Watermark Embedding Procedure.....	69
Figure 4.6: Flow Chart of Zernike Based Watermark Extraction Procedure.....	72
Figure 5.1: Segments extraction of RAFD when under various attacks (a) The original audio clip ‘My Heart Will Go On.wav’, $N = 16$; the symbols ‘o’ represent the feature points detected from the original audio clip with the RAFD. (b) The distorted audio clip attacked by 16 kHz resampling, $NCDP = 16$; the symbols ‘o’ represent the corresponding feature points which are correctly detected when the audio clip is distorted by the resampling. (c) The distorted audio clip attacked by 110% resample TSM, $NCDP = 16$; the symbols ‘o’ represent the corresponding feature points which are correctly detected when the audio clip is distorted by the resample TSM. (d) The distorted audio clip attacked by 90% pitch invariant TSM, $NCDP = 16$; the symbols ‘o’ represent the corresponding feature points which are correctly detected when the audio clip is distorted by the pitch invariant TSM.	83
Figure 5.2: Segments extraction of Canny detector when under various attacks (a) The original audio clip ‘My Heart Will Go On.wav’, $N = 16$; the symbols ‘o’ represent the feature points detected from the original	

audio clip with the Canny. (b) The distorted audio clip attacked by 16 kHz resampling, $NCDP = 16$; the symbols ‘○’ represent the corresponding feature points which are correctly detected when the audio clip is distorted by the resampling. (c) The distorted audio clip attacked by 110% resample TSM, $NCDP = 14$; the symbols ‘○’ represent the correctly detected feature points, the symbols ‘⊗’ represent the wrongly detected feature points, and the symbols ‘⊕’ represent the locations where the feature points should be detected from, when the audio clip is distorted by the resample TSM. (d) The distorted audio clip attacked by 90% pitch invariant TSM, $NCDP = 15$; the symbols ‘○’ represent the correctly detected feature points, the symbols ‘⊗’ represent the wrongly detected feature points, and the symbols ‘⊕’ represent the locations where the feature points should be detected from, when the audio clip is distorted by the pitch invariant TSM. 84

Figure 5.3: Segments extraction of Marr-Hildreth detector when under various attacks (a) The original audio clip ‘My Heart Will Go On.wav’, $N = 16$; the symbols ‘○’ represent the feature points detected from the original audio clip with the Marr-Hildreth. (b) The distorted audio clip attacked by 16 kHz resampling, $NCDP = 16$; the symbols ‘○’ represent the corresponding feature points which are correctly detected when the audio clip is distorted by the resampling. (c) The distorted audio clip attacked by 110% resample TSM, $NCDP = 15$; the symbols ‘○’ represent the correctly detected feature points, the symbols ‘⊗’ represent the wrongly detected feature points, and the symbols ‘⊕’ represent the locations where the feature points should be detected from, when the audio clip is distorted by the resample TSM. (d) The distorted audio clip attacked by 90% pitch invariant TSM, $NCDP = 14$; the symbols ‘○’ represent the correctly detected feature points, the symbols ‘⊗’ represent the wrongly detected feature points, and the symbols ‘⊕’ represent the locations where the feature points should be detected from, when the audio clip is distorted by the pitch invariant TSM. 85

Figure 5.4: One-dimensional SWT decomposition. (a) Decomposition steps. (b) Filters up-sampling.	89
Figure 5.5: Flow Chart of Watermark Embedding	90
Figure 5.6: Flow Chart of Watermark Extraction	92
Figure 6.1A: Experimental Results When Under Geometric Attacks – Rotation and Scaling. (a1), (b1), (c1), (d1), and (e1) Rotation, for ‘Elaine’, ‘Lena’, ‘Jet’, ‘Pepper’, and ‘Tank’, respectively. (a2), (b2), (c2), (d2), and (e2) Scaling, for ‘Elaine’, ‘Lena’, ‘Jet’, ‘Pepper’, and ‘Tank’, respectively.	98
Figure 6.1B: Experimental Results When Under Geometric Attacks – Shearing and Cropping. (a3), (b3), (c3), (d3), and (e3) Affine transformation of vertical shearing, for ‘Elaine’, ‘Lena’, ‘Jet’, ‘Pepper’, and ‘Tank’, respectively. (a4), (b4), (c4), (d4), and (e4) Cropping, for ‘Elaine’, ‘Lena’, ‘Jet’, ‘Pepper’, and ‘Tank’, respectively.	99
Figure 6.2: Experimental Results When Under Common Signal Processing. (a1), (b1), (c1), (d1), and (e1) JPEG compression, for ‘Elaine’, ‘Lena’, ‘Jet’, ‘Pepper’, and ‘Tank’, respectively. (a2), (b2), (c2), (d2), and (e2) Median filtering, for ‘Elaine’, ‘Lena’, ‘Jet’, ‘Pepper’, and ‘Tank’, respectively. (a3), (b3), (c3), (d3), and (e3) Gaussian low-pass filtering, for ‘Elaine’, ‘Lena’, ‘Jet’, ‘Pepper’, and ‘Tank’, respectively.	102
Figure 6.3: RFPD Extracted Features and Watermarked Images. (a1) ‘Baboon’, (a2) watermarked ‘Baboon’; PSNR=39.10dB, AVG_PSNR=32.43dB. (b1) ‘Bridge’, (b2) watermarked ‘Bridge’; PSNR=39.55dB, AVG_PSNR=32.78dB. (c1) ‘Lena’, (c2) watermarked ‘Lena’; PSNR=39.96dB, AVG_PSNR=32.61dB. (d1) ‘Pepper’, (d2) watermarked ‘Pepper’; PSNR=38.93dB, AVG_PSNR=32.24dB. (e1) ‘Blurry Scene’, (e2) watermarked ‘Blurry Scene’; PSNR = 39.889dB, AVG_PSNR=32.45dB. (f1) ‘Blurry Jet’, (f2) watermarked ‘Blurry Jet’; PSNR = 38.85dB, AVG_PSNR=32.55dB.	106
Figure 6.4: Feature Extraction by RFPD When under Various Attacks. (a1), (b1), (c1), and (d1) Original ‘Baboon’, ‘Bridge’, ‘Lena’, and ‘Pepper’; (a2), (b2), (c2), and (d2) 45° rotation with cropping; (a3), (b3), (c3), and (d3) 20% vertical shearing; (a4), (b4), (c4), and (d4) 20% horizontal shearing; (a5), (b5), (c5), and (d5) 10% affine transformation; (a6),	

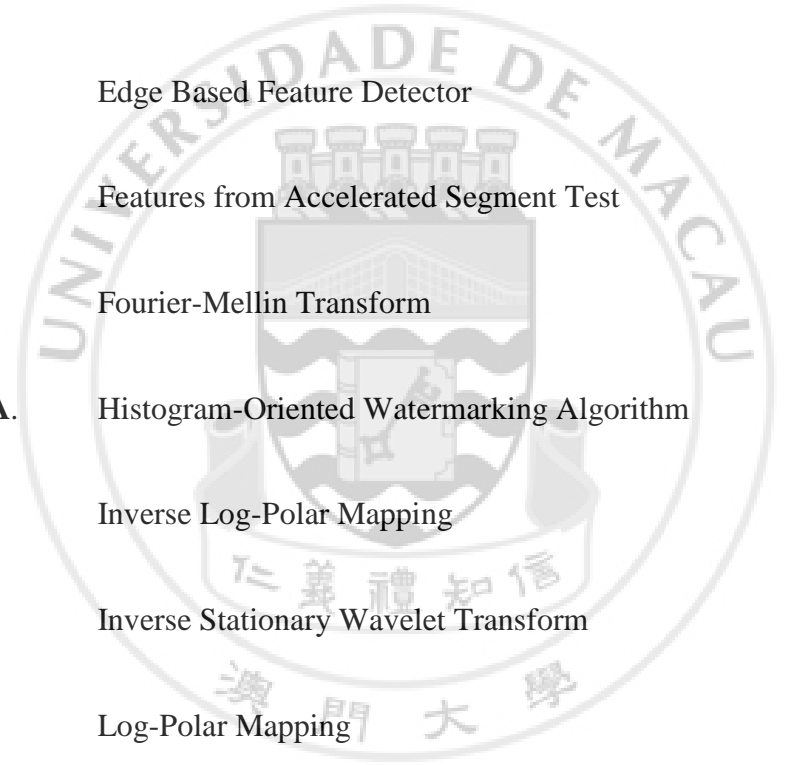
(b6), (c6), and (d6) scaling with the scale factor as 0.5; (a7), (b7), (c7), and (d7) 4×4 median filtering; (a8), (b8), (c8), and (d8) JPEG compression with the quality factor as 50.....	108
Figure 6.5: Experimental Results against Geometric Attacks (a) rotation with cropping (b) scaling (c) affine transformation of vertical shearing (d) affine transformation of horizontal shearing.....	110
Figure 6.6: Experimental Results against Common Signal Processing. (a) JPEG compression (b) median filtering (c) Gaussian low-pass filtering.....	111
Figure 6.7: Mixed Attacks Demonstration (a) 30° rotation with cropping, detection ratio = 7/10, 7/9 for ‘Pepper’ and ‘Lena’, respectively (b) 15% affine transformation, detection ratio = 9/10, 7/9 for ‘Pepper’ and ‘Lena’, respectively.....	113
Figure 6.8: Bit-error Rate against Strength of Various Attacks (a) JPEG Compression (b) rotation (c) scaling (d) affine transformation of shearing.....	114
Figure 6.9: Relationships between Capacity and Transparency.....	120
Figure 6.10: Relationships between Capacity and Robustness.....	121
Figure 6.11: Test Images and Watermarked Images. (a1)-(e1) Original ‘Baboon’, ‘Boat’, ‘Lena’, ‘Pepper’, and ‘Tank’ (a2)-(e2) Extracted feature points from the corresponding test image (a3)-(e3) Watermarked ‘Baboon’, ‘Boat’, ‘Lena’, ‘Pepper’, and ‘Tank’.....	122
Figure 6.12: Results Demonstration under Watermarked and Un-watermarked Image.....	125
Figure 6.13: Various Attacks and Corresponding Extracted Correct Bits. (a) the original image ‘Lena’ (b) image rotation, rotation angle = 45° , correct bits = 13 (c) image scaling, scale factor = 0.3, correct bits = 13 (d) JPEG compression, quality factor = 10, correct bits = 13 (e) median filtering, neighborhood = 12×12 , correct bits = 13 (f) 3×3 low-pass Gaussian filtering, standard deviation = 1.5, correct bits = 14 (g) ‘Salt & Pepper’ noise pollution, density = 0.5, correct bits = 15 (h) ‘Gaussian’ noise pollution, mean = 0, variance = 0.05, correct bits = 12.....	126

Figure 6.14: Correctly Extracted Bits When under Various Attacks (a) Rotation, (b) Scaling, (c) Cropping, (d) JPEG compression, (e) ‘Salt & Pepper’ noise pollution, and (f) Gaussian low-pass filtering.	128
Figure 7.1: Watermark detection results when under common audio signal processing (a) 16 kHz resampling, $RCEs = 16/16$; (b) 8kHz low-pass filtering, $RCEs = 16/16$; (c) 40% echo with 100ms delay, $RCEs = 11/16$	141
Figure 7.2: Watermark detection results when under synchronization geometric distortions (a) 120% resample TSM, $RCEs = 15/16$; (b) 80% pitch invariant TSM, $RCEs = 14/16$; (c) 90% pitch shifting, $RCEs = 2/16$	142
Figure 7.3: Number of segments where the watermarks have been correctly detected when the audio clips are distorted by Resample TSM, with the Similarity Rate varies from: (a) 50% to 100%, (b) 105% to 150%.	150
Figure 7.4: Number of segments where the watermarks have been correctly detected when the audio clips are distorted by Pitch Invariant TSM, with the Length Rate varies from: (a) 50% to 100%, (b) 105% to 150%.	151
Figure 7.5: Number of segments where the watermarks have been correctly detected when the audio clips are distorted by Tempo Invariant Pitch Shifting, with the Scale Factor varies from: (a) 50% to 100%, (b) 105% to 150%.	152
Figure 7.6: Comparison of the proposed scheme and the existing scheme on ‘Piano.wav’ when under TSM with the TSM Ratio varies from -15% to +15%	154
Figure 7.7: Comparison of the proposed scheme and the existing scheme on ‘Piano.wav’ when under Common Audio Signal Processing	159
Figure 7.8: Comparison of the proposed scheme and the existing scheme on ‘Piano.wav’ when under Stirmark for Audio	160

List of Tables

Table 6.1: Watermarking Extraction Results in Different Bit-Planes.....	96
Table 6.2: Experimental Results Comparison	104
Table 6.3: Watermark Detection Results under Common Signal Processing.....	116
Table 6.4: Watermark Detection Results under Geometric Distortion	117
Table 6.5: Experimental Results Comparisons	118
Table 6.6: Correct Extraction Rate for Rotation Attack.....	130
Table 6.7: Correct Extraction Rate for Scaling Attack	131
Table 6.8: Correct Extraction Rate for JPEG Compression Attack	132
Table 6.9: Correct Extraction Rate for Median Filtering Attack	133
Table 6.10: Correct Extraction Rate for Gaussian Low-Pass Filtering Attack	134
Table 6.11: Correct Extraction Rate for Noise Pollution Attack.....	135
Table 6.12: Correct Extraction Rate for Cropping Attack	136
Table 6.13: Experimental Results Comparisons	138
Table 7.1 : Quality of the Watermarked Audio Clips	144
Table 7.2: Ratio of Correctly Detected Patches (RCDP) under Resample TSM.....	145
Table 7.3: Ratio of Correctly Detected Patches (RCDP) under Pitch Invariant TSM.....	146
Table 7.4: Ratio of Correctly Detected Patches (RCDP) under Tempo Invariant Pitch Shifting.....	147
Table 7.5: Comparison of Ratio of Correctly Detected Regions (RCDR) under TSM.....	153
Table 7.6: Comparison of Ratio of Correctly Detected Regions (RCDR) under Signal Processing	156
Table 7.7: Comparison of Ratio of Correctly Detected Regions (RCDR) under Stirmark for Audio	157
Table 7.8: Comparison with Existing De-Synchronization Resilient Schemes	161

List of Abbreviations



AHBD.	Adaptive Harris Based Detector
DCT.	Discrete Cosine Transform
DFT.	Discrete Fourier Transform
DoG.	Difference-of-Gaussians
DWT.	Discrete Wavelet Transform
EBFD.	Edge Based Feature Detector
FAST.	Features from Accelerated Segment Test
FMT.	Fourier-Mellin Transform
HOWA.	Histogram-Oriented Watermarking Algorithm
ILPM.	Inverse Log-Polar Mapping
ISWT.	Inverse Stationary Wavelet Transform
LPM.	Log-Polar Mapping
LoG.	Laplacian of Gaussian
MSE.	Mean Square Estimation Error
OFPD.	Original Feature Points Dataset
PSNR.	Peal Signal-to-Noise Ratio
QIM.	Quantization Index Modulation

RAFD.	Robust Audio Feature Detector
RST.	Rotation, Scaling, Translation
SBFD.	SIFT Based Feature Detector
SDG.	Subject Difference Grade
SDMI.	Secure Digital Music Initiative
SNR.	Signal-to-Noise Ratio
SIFT.	Scale Invariant Feature Transform
SSIM.	Structural Similarity
SURF.	Speed Up Robust Features
SUSAN.	Smallest Univalue Segment Assimilating Nucleus
SWT.	Stationary Wavelet Transform
TFPD.	Trained Feature Points Dataset
TSM.	Time Scale Modification

