

Contents

Abstract	v
Acknowledgement	vii
Nomenclature	xii
1 Introduction	1
1.1 Motivation	2
1.2 Goals of the Thesis	4
1.3 Scope of the Thesis	5
1.4 Thesis' Outline	6
2 Preliminaries	9
2.1 Time Series Signals	9
2.2 Distance and Similarity Measure	11
2.3 Time Series Motif	12
2.4 Wavelet Transformations	14
2.4.1 Dual Tree Complex Wavelet Transform (DTCWT)	14
2.4.1.1 Limitations and Deficiencies	18
3 General Principles of Time Series Motif Discovery	21
3.1 Time Series Pre-Processing	21
3.2 Time Series Representation	22
3.2.1 Invariant Transformations	23
Translation-Invariant Transformations	24
Scaling-Invariant Transformations	25
3.3 Time Series Distance and Similarity Measures	26
3.4 Summary	27
4 State of the Art in Time Series Motif Discovery	29
4.1 Motif Discovery Algorithms	29
4.1.1 Time Complexity	37
4.1.2 Detecting Ill-Known Motifs	39
4.2 Research Gaps	41
4.3 Summary	43
5 Distortion-Invariant Motif Discovery	45
5.1 KITE Architecture	46
5.2 Signal Pre-Processing for Motif Discovery	48
5.2.1 Motif Length Definition	48

5.3	Invariant Time Series Representation	50
5.3.1	Analytic Complex Quad Tree Wavelet Packet Transform - (ACQTWP)	52
5.3.1.1	Inverse Analytic Complex Quad Tree Wavelet Packet Transform (IACQTWP)	55
5.3.1.2	Properties and Characteristics	56
5.3.1.3	Selection of the Best Basis	61
5.4	Feature Extraction from Variable Scales	72
5.5	Threshold Determination for Similarity Detection	75
5.6	Significant Motif Discovery	79
5.6.1	Excluding Misleading Motifs	80
5.6.2	Representative Motifs	82
5.7	Time Complexity Analysis	83
5.8	Summary	86
6	Evaluation	89
6.1	Validation Principles	89
6.1.1	Feature Selection	89
6.1.2	Quality Measures	90
6.2	Design of the Experiments	91
6.2.1	Test Cases	92
6.2.1.1	Synthetic Data	92
6.2.1.2	Real-World Data	93
6.3	Detection of Equal-Length Motifs	94
6.3.1	Equal-Length Motif Discovery on Synthesis Data	94
6.3.2	Equal-Length Motif Discovery on Real-World Data	102
6.3.3	Equal-Length Motif Discovery Summary	108
6.4	Detection of Variable-Length Motifs	110
6.4.1	Variable-Length Motif Discovery on Synthesis Data	110
6.4.2	Variable-Length Motif Discovery on Real-World Data	116
6.4.3	Variable-Length Motif Discovery Summary	119
6.5	KITE Robustness Toward Noise	120
6.6	Scalability Experiments	123
6.7	Case Studies	124
6.7.1	Anomaly Detection via Time Series Motif Discovery	124
6.8	Summary	125
7	Conclusion and Outlook	127
7.1	Conclusion and Contributions	128
7.2	Perspectives and Future Directions	130
8	Appendix A	133
8.1	Function and Signal Space	133
8.2	Transformations and Representation	133

8.3	Wavelet Transform	134
8.3.1	Scaling Function	134
8.3.2	Multiresolution Analysis	135
8.3.3	Discrete Wavelet Transform	136
8.3.4	Filter Bank Structure	137
8.3.5	Synthesis Filter Bank	139
9	Appendix B	141
9.1	Nobel Identities	141
9.2	Design of the q-shift Filters	142
9.3	Proof Lemma 1	145
9.4	Proof Lemma 2	147
10	Appendix C	147
10.1	Proof of Parseval’s Theorem	147
11	Appendix D	149
11.1	Equal-length Motif Discovery on Lightning Data Set	149
11.2	Variable-length Motif Discovery on Lightning Data Set	152
11.3	KITE’s Time Complexity and Distance Measures	153
11.4	KITE’s Performance Under Noise	156
11.5	Applications of KITE in Higher Dimension	159
11.5.1	Two Dimensional Analytic Complex Quad Tree Wavelet Packet	159
11.5.1.1	Motif Discovery on Image Data Set	162
	Bibliography	167
	List of Tables	197
	List of Figures	203