

Frank Schäfer

Das hermitesche Eigenwertproblem -
Implementierungsaspekte für Festkomma-SIMD-DSPs

Beiträge aus der Informationstechnik

Frank Schäfer

**Das hermitesche Eigenwertproblem -
Implementierungsaspekte für
Festkomma-SIMD-DSPs**

 VOGT

Dresden 2008

Bibliografische Information der Deutschen Bibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.ddb.de> abrufbar.

Zugl.: Dresden, Techn. Univ., Diss., 2007

Die vorliegende Arbeit stimmt mit dem Original der Dissertation „Das hermitesche Eigenwertproblem - Implementierungsaspekte für Festkomma-SIMD-DSPs“ von Frank Schäfer überein.

Besuchen Sie uns im Internet:
www.vogtverlag.de

© Jörg Vogt Verlag 2008
Alle Rechte vorbehalten. All rights reserved.

Gesetzt vom Autor
Printed in Germany

ISBN 978-3-938860-18-2

Jörg Vogt Verlag
Niederwaldstr. 36 · 01277 Dresden
Telefon: +49-(0)351-31403921
Telefax: +49-(0)351-31403918
Email: info@vogtverlag.de

Technische Universität Dresden

Das hermitesche Eigenwertproblem -
Implementierungsaspekte für Festkomma-SIMD-DSPs

Frank Schäfer

der Fakultät Elektrotechnik und Informationstechnik der Technischen
Universität Dresden
zur Erlangung des akademischen Grades eines

Doktoringenieurs

(Dr.-Ing.)

genehmigte Dissertation

Vorsitzender: Prof. Dr.-Ing. Dr. rer. nat. K. Reinschke

Gutachter: Prof. Dr.-Ing. Gerhard Fettweis

Tag der Einreichung: 06.11.2006

Prof. Dr.-Ing. habil. Renate Merker

Tag der Verteidigung: 18.07.2007

Prof. Dr.-Ing. Jürgen Götze

Danksagung

Eigenmotivation ist nicht alles, aber ohne Eigenmotivation ist alles nichts.

Die vorliegende Arbeit, die größtenteils während meiner Tätigkeit als wissenschaftlicher Mitarbeiter in der DSP-Entwicklungsgruppe des Lehrstuhls für mobile Nachrichtensysteme der Technischen Universität Dresden entstand, würde es so nicht geben, wenn ich nicht von vielen Seiten Hilfe und Unterstützung erfahren hätte:

An erster Stelle möchte ich dem Lehrstuhlleiter Herrn Prof. Dr.-Ing. Gerhard Fettweis danken, der mit seiner Art und Weise den Lehrstuhl zu führen, für eine angenehme, professionelle und kreative Atmosphäre sorgt. Die vielen fachlichen Diskussionen mit ihm über den Inhalt der Arbeit haben stets inspirierend und motivierend gewirkt.

Ebenso möchte ich mich bei allen Mitarbeitern des Lehrstuhls für die nette Zusammenarbeit und das ausgezeichnete Klima bedanken, insbesondere seien hier die Mitforscher aus der DSP-Gruppe Torsten Limberg und Dr. Thomas Richter erwähnt. Hervorzuheben ist auch die Arbeit über Projektgruppen hinweg, so unter anderem mit Dr. Katja Schwieger, Dr. Matthias Stege und Clemens Michalke. Nicht vergessen möchte ich unsere Sekretärinnen Kathrin Fromke und Sylvia Steppat, sowie die Systemadministratoren Raffael Kozerski und Rüdiger Hartmann, die mit viel Engagement den wissenschaftlichen Mitarbeitern zur Seite stehen.

Ganz besonders möchte ich an dieser Stelle meinen Eltern danken, die mich auf meinem gesamten Lebensweg unterstützt und durch die Förderung meiner Interessen entscheidend zu meinem beruflichen Werdegang beigetragen haben.

Nicht zuletzt möchte ich mich für die Geduld und das Verständnis meiner Lebensgefährtin und unserer beiden Kinder während der Fertigstellung der Arbeit bedanken.

Freital, im Oktober 2006

Frank Schäfer

Inhaltsverzeichnis

1	Einleitung	1
2	CATS DSP-Plattform	5
2.1	Verbindungsnetzwerk	7
2.2	DSP-Datenpfad	8
3	Mathematische Grundlagen und Algorithmen-Evaluierung	11
3.1	Definitionen und Sätze	12
3.2	Karhunen-Loève-Transformation	13
3.3	Algorithmen zur Lösung des hermiteschen Eigenwertproblems	16
3.3.1	Einfach- und Multivektoriterationen	16
3.3.2	Transformationsbasierte Methoden	20
3.3.3	Weitere Verfahren und Zusammenfassung	30
4	Parallelisierung der Algorithmen	35
4.1	Grundlagen der Parallelverarbeitung	36
4.1.1	Ebenen der Parallelisierung	36
4.1.2	Kenngrößen der Parallelisierung	38
4.2	Matrix-Matrix-Produkt	40
4.2.1	Eindimensionale Parallelisierung	41
4.2.2	Zweidimensionale Parallelisierung	41
4.2.3	Dreidimensionale Parallelisierung	42
4.2.4	Zusammenfassung	42
4.3	Einseitiger JACOBI-Algorithmus	43
4.3.1	Speicherbelegung des SIMD-DSP	44
4.3.2	Ordnungsschemata	46
4.3.3	Kleine und große Matrizen, $n \neq P$	55
4.4	Zweiseitiger JACOBI-Algorithmus	56
4.5	QR-Algorithmus	57
4.6	Zusammenfassung	59

5	Parameterberechnung	61
5.1	Zahlendarstellung in DSPs	63
5.1.1	Darstellung von Festkommazahlen	63
5.1.2	Darstellung von Fließkommazahlen	64
5.2	Division	65
5.2.1	Bitweise Verfahren	66
5.2.2	Funktional-iterative Verfahren	69
5.2.3	Vergleich	71
5.3	Inverse Quadratwurzel	72
5.4	Phasenberechnung	74
5.5	Direkte Parameterberechnung	75
5.6	Newtonverfahren zur Parameterberechnung	75
5.7	Division mit geringer Bitbreite	79
5.8	CORDIC-basierte Parameterberechnung	80
5.9	Zusammenfassung	87
6	Tracking der Eigenpaare	89
6.1	Motivation	89
6.2	Einleitung	90
6.3	Tracking von Eigenräumen	91
6.3.1	Übersicht über etablierte Tracking-Algorithmen	91
6.3.2	Das Prinzip des JACOBI-Tracking	92
6.3.3	Quantifizierung der Systemänderung	92
6.3.4	Ergebnisse	94
6.4	Komplexitätsvergleich	95
6.5	Anwendung des JACOBI-Trackings auf MIMO-Systeme	95
6.6	Zusammenfassung	96
7	Implementierung auf Plattform-DSP	99
7.1	Simulationsumgebung	100
7.2	Skalierung	102
7.3	Runden	103
7.4	Implementierungsergebnisse	104
7.4.1	Direkte Parameterberechnung	105
7.4.2	Lookup-basierte Parameterberechnung	106
7.5	Auswertung der Simulationsergebnisse	110
	Zusammenfassung	113

A	Wichtige Beweise	115
A.1	Orthogonalität von Spalten- und Zeilenvektoren	115
A.2	Erhalt der Frobenius-Norm	116
A.3	Eigenschaften von Shuffle-Matrizen zur Erzeugung von Vektorelementpaaren	117
B	Herleitung wichtiger Gleichungen	119
B.1	Ableitung der Parameter für komplexwertige JACOBI-Rotationen	119
B.1.1	Ableitung der Parameter für zweiseitige, komplexwertige JACOBI- Rotationen	119
B.1.2	Ableitung der Parameter für einseitige, komplexwertige JACOBI- Rotationen	121
B.1.3	Grenzwertberechnung der Fehlerfunktion	123
C	Randbedingungen der Simulationsergebnisse	125
C.1	Simulationsparameter für JACOBI-Algorithmus	125
C.2	Simulationsparameter für JACOBI-Algorithmus mit approximierten τ . . .	126

Abbildungsverzeichnis

1.1	Raum-Zeit-Empfänger mit Dimensionreduktion	2
2.1	DSP-Plattform Konzept [22]	7
2.2	Datenpfad eines Streifens [22]	10
3.1	Dekorrelation durch Anwendung der KLT	16
3.2	Konvergenzverlauf des QR-Algorithmus	22
3.3	Konvergenzverlauf des zweiseitigen, zeilenzyklischen JACOBI-Algorithmus .	29
3.4	Konvergenzverlauf des einseitigen, seriellen JACOBI-Algorithmus	30
4.1	Parallelisierung einer komplexwertigen Multiplikation auf Wortebene	37
4.2	Verlust an Parallelität	39
4.3	Matrix-Matrix-Multiplikation auf einem linearen systolischen Array	41
4.4	Matrix-Matrix-Multiplikation auf einem quadratischen systolischen Array .	42
4.5	Matrix-Matrix-Multiplikation auf einem Hypercube	43
4.6	Beeinflussungsschema bei der einseitigen seriellen JACOBI-Rotation	45
4.7	Beispiele für Beeinflussungsschemata bei der einseitigen parallelen JACOBI- Rotation	45
4.8	Varianten der Speicherbelegung	47
4.9	Mögliche Anordnungen der Datenpfade (PEs) bei einem SIMD-DSP	51
4.10	Round-Robin Schema für verschieden Datenpfadanordnungen	52
4.11	Up-Down Schema für zirkulare Datenpfadanordnung	53
4.12	Odd-Even Schema für verschiedene Datenpfadanordnungen	54
4.13	Konvergenzverlauf des einseitigen, parallelen JACOBI-Algorithmus für un- terschiedliche Ordnungsschemata	55
4.14	Beeinflussungsschema bei einer zweiseitigen Verbundrotation	57
4.15	Abbildung des zweiseitigen JACOBI-Algorithmus auf ein systolisches Array	58
4.16	Parallelisierungsschemata für den QR-Algorithmus	59
5.1	$\sin(\phi)$ und $\cos(\phi)$ als Funktion von τ	63
5.2	Zahlendarstellung einer N-Bit Zahl im Q-Format	64
5.3	Zahlendarstellung einer 32-Bit Zahl im Fließkommaformat nach IEEE-754	65

5.4	Veranschaulichung der SRT-Division mit normiertem Divisor	68
5.5	Relativer Fehler bei der Reziprokwertbildung mit dem Newton-Raphson-Verfahren in Abhängigkeit des Argumentes D	70
5.6	Relativer Fehler bei der Berechnung der reziproken Quadratwurzel mit dem Newton-Raphson-Verfahren in Abhängigkeit des Argumentes D	73
5.7	3D-Darstellung des Newtonverfahrens	76
5.8	Relativer Fehler der Verbunditeration zur Berechnung der Kosinus- und Sinuswerte mit dem Newton-Verfahren in Abhängigkeit des Argumentes τ .	78
5.9	Relativer Fehler der Verbunditeration zur Berechnung der Kosinus- und Sinuswerte mit dem vereinfachten Newton-Verfahren in Abhängigkeit des Argumentes τ	79
5.10	Vergleich zwischen exaktem und approximiertem Funktionsverlauf bei einer Wortbreite von $N = 16$ Bit für verschiedene Blockgrößen N_B	81
5.11	Konvergenzverlauf des einseitigen JACOBI-Algorithmus mit approximierten τ mit Wortbreiten von $N = 24, 32, 40$ Bit für verschiedene Blockgrößen N_B	82
5.12	Veranschaulichung der zwei CORDIC-Algorithmen	85
5.13	Relativer Fehler der Kosinuswerte bei der Berechnung mit CORDIC	85
6.1	Vergleich des Konvergenzverlaufs zwischen dem konventionellen JACOBI-Algorithmus und dem JACOBI-Tracking	93
6.2	Nebendiagonalnorm δ_T als statistische Abhängigkeit der Unterraumdistanz zwischen $S(k)$ und $S(k + 1)$	94
7.1	Simulationsumgebung [30]	101
7.2	Logikschaltung zur Ermittlung der größten signifikanten Bitstelle [30] . . .	103
7.3	Logikschaltung eines Barrel-Shifters mit Runden [30]	104
7.4	Konvergenzverlauf des einseitigen JACOBI-Algorithmus mit direkter Parameterberechnung und Rundung für eine Matrixgröße von $n = 16$ und verschiedene Wortbreiten	108
7.5	Konvergenzverlauf des einseitigen JACOBI-Algorithmus mit lookup-basierter Parameterberechnung für eine Matrixgröße von $n = 16$ und eine Wortbreite von $w = 40$	108
7.6	Anzahl und Verteilung der Taktzyklen für den einseitigen JACOBI-Algorithmus in Anhängigkeit der Matrixgröße	111
A.1	Beispiel für Paarbildung innerhalb eines Sweeps, $n = 6$	117

Tabellenverzeichnis

3.1	Vergleich der EVD-Algorithmen	33
4.1	Flächenverteilung für den M3-DSP [63]	40
4.2	Vergleich der Parallelisierungen einer Matrix-Matrix-Multiplikation	44
4.3	Reduktion der zu betrachtenden Permutationen	50
4.4	Vergleich der EVD-Algorithmen	60
5.1	Vergleich der Verfahren zur Berechnung der Rotationsparameter	87
6.1	Gegenüberstellung des numerischen Aufwands für Vormultiplikation und einen Sweep	96
7.1	Übersicht der implementierten Verfahren und deren Parameter	106
7.2	Implementierungsergebnisse mit direkter Parameterberechnung	107
7.3	Flächen- und Leistungsbedarf des Datenpfades mit direkter Parameterbe- rechnung	109
7.4	Implementierungsergebnisse mit lookup-basierter Parameterberechnung	109
7.5	Flächen- und Leistungsbedarf des Datenpfades ($w = 40\text{bit}$) mit lookup- basierter Parameterberechnung	109

Algorithmenverzeichnis

3.1	Power-Methode	17
3.2	Inverse Power-Methode	18
3.3	Rayleigh-Quotient-Iteration	19
3.4	Orthogonale Iteration	20
3.5	QR-Algorithmus	21
3.6	QR-Zerlegung	23
3.7	Zweiseitiger, zeilenzyklischer JACOBI-Algorithmus	28
3.8	Einseitiger, serieller JACOBI-Algorithmus	31
5.1	Restoring-Algorithmus	67
5.2	Non-Restoring-Algorithmus	67
5.3	SRT-Algorithmus	69
5.4	Goldschmidt-Algorithmus	71
5.5	CORDIC-Algorithmus	86

Liste häufig verwendeter Symbole und Abkürzungen

Symbole

A	beliebige Matrix
<i>a</i>	Matrizelement der Matrix A
B	beliebige Matrix
<i>b</i>	Matrizelement der Matrix B
C	Korrelationsmatrix
\mathbb{C}	Menge der komplexen Zahlen
<i>c</i>	Matrizelement der Korrelationsmatrix C
I	Einheitsmatrix
<i>i</i>	Zeilenindex einer Matrix
<i>j</i>	Spaltenindex einer Matrix
<i>m</i>	Anzahl der Zeilen einer Matrix
<i>n</i>	Anzahl der Spalten einer Matrix
<i>N</i>	Anzahl der Sweeps
\mathbb{R}	Menge der reellen Zahlen
S	Singulärwertmatrix
U	Eigenvektormatrix
u	Eigenvektor
X	Matrix, mit Eingangsdaten der KLT
<i>x</i>	Matrizelement der Matrix X
Y	Matrix, mit Daten nach der KLT
<i>y</i>	Matrizelement der Matrix Y
λ	Eigenwert
σ^2	Varianz
μ	Schätzwert für einen Eigenwert

Abkürzungen

ALU	engl. <i>A</i> rithmetic <i>L</i> ogic <i>U</i> nit
ASIC	engl. <i>A</i> pplication <i>S</i> pecific <i>I</i> ntegrated <i>C</i> ircuit
CATS	engl. <i>C</i> oncept for <i>A</i> pplication <i>T</i> ailored <i>S</i> ignalprocessors
CG	<i>C</i> ode- <i>G</i> ruppe
CORDIC	engl. <i>C</i> oordinate <i>R</i> otation <i>D</i> igital <i>C</i> omputer
DCT	engl. <i>D</i> iscrete <i>C</i> osine <i>T</i> ransform
DFT	engl. <i>D</i> iscete <i>F</i> ourier <i>T</i> ransform
DSP	engl. <i>D</i> igital <i>S</i> ignal <i>P</i> rocessor
EVD	engl. <i>E</i> igenvalue <i>D</i> ecomposition
FFT	engl. <i>F</i> ast <i>F</i> ourier <i>T</i> ransformation
FIR	engl. <i>F</i> inite <i>I</i> mpulse <i>R</i> esponse
FPGA	engl. <i>F</i> ield <i>P</i> rogrammable <i>G</i> ate <i>A</i> rray
HDL	engl. <i>H</i> ardware <i>D</i> escription <i>L</i> anguage
ICU	engl. <i>I</i> nterconnection <i>U</i> nit
IFFT	engl. <i>I</i> nverse <i>F</i> ast <i>F</i> ourier <i>T</i> ransformation
ISA	engl. <i>I</i> nstruction <i>S</i> et <i>A</i> rchitecture
KLT	<i>K</i> arhunen- <i>L</i> oève- <i>T</i> ransformation
LAN	engl. <i>L</i> ocal <i>A</i> rea <i>N</i> etwork
M3	engl. <i>M</i> obile <i>M</i> ultimedia <i>M</i> odem
MAC	engl. <i>M</i> ultiply <i>A</i> ccumulate
MIMO	engl. <i>M</i> ultiple <i>I</i> nput <i>M</i> ultiple <i>O</i> utput
PE	<i>P</i> rozessorelement
QRD	engl. <i>Q</i> R- <i>D</i> ecomposition
SIMD	engl. <i>S</i> ingle <i>I</i> nstruction <i>M</i> ultiple <i>D</i> ata
SISD	engl. <i>S</i> ingle <i>I</i> nstruction <i>S</i> ingle <i>D</i> ata
SoC	engl. <i>S</i> ystem <i>o</i> n <i>C</i> hip
SVD	engl. <i>S</i> ingular <i>V</i> alue <i>D</i> ecomposition
UMTS	engl. <i>U</i> niversal <i>M</i> obile <i>T</i> elecommunications <i>S</i> ystem
VHDL	engl. <i>V</i> ery <i>H</i> igh <i>S</i> peed <i>I</i> ntegrated <i>C</i> ircuit <i>H</i> ardware <i>D</i> escription <i>L</i> anguage
VHPI	engl. <i>V</i> HDL <i>P</i> rogramming <i>I</i> nterface
VLIW	engl. <i>V</i> ery <i>L</i> ong <i>I</i> nstruction <i>W</i> ord
TVLIW	engl. <i>T</i> agged <i>V</i> ery <i>L</i> ong <i>I</i> nstruction <i>W</i> ord

Kapitel 1

Einleitung

Den Anlass für die Beschäftigung mit dem hermiteschen Eigenwertproblem gab die Dissertation „Dimensionsreduktion des Empfangsraums von Systemen mit mehreren Empfangsantennen und deren Anwendung in der Raum-Zeit-Verarbeitung“ von Jens Jelitto [25], die während seiner Tätigkeit als wissenschaftlicher Mitarbeiter am Vodafone Stiftungslehrstuhl mobile Nachrichtensysteme an der TU-Dresden entstand. In der Arbeit wird ein neues Verfahren der dimensionsreduzierten Raum-Zeit-Verarbeitung für Mobilfunksysteme vorgestellt, welches die Karhunen-Loeve-Transformation (KLT) nutzt, um die Bandbreiteneffizienz zu steigern. In Abbildung 1.1 ist ein solcher Raum-Zeit-Empfänger mit Dimensionreduktion vereinfacht dargestellt. Die Verarbeitung der empfangenen n Signalströme (Vektoren), die zur Empfangsmatrix $\mathbf{X} \in \mathbb{C}^{n \times m}$ zusammengefasst werden können, findet blockweise (m Blocklänge) statt. In einem ersten Schritt werden die Eingangssignale durch eine orthogonale Transformation, was nichts anderes bedeutet als die Multiplikation mit der orthogonalen Matrix \mathbf{U}^H ($\mathbf{U} \in \mathbb{C}^{n \times n}$), entkorreliert, womit sich unter anderem Signal- von Rauschraum trennen lassen. Dieser Verarbeitungsschritt der signalangepassten Entkorrelation inklusive der Berechnung der Matrix \mathbf{U} wird als KLT bezeichnet. Eine Möglichkeit, die Matrix \mathbf{U} zu ermitteln, ist die Berechnung der Eigenvektormatrix der hermiteschen Autokorrelationsmatrix $\mathbf{C}_X \in \mathbb{C}^{n \times n}$. Im darauffolgenden Schritt werden einige (n_S) der Signalvektoren von $\mathbf{Y} \in \mathbb{C}^{n \times m}$ zur Weiterverarbeitung mit dem Ziel der Dimensionsreduktion bei möglichst geringem Informationsverlust ausgewählt. Der numerische Aufwand der nachfolgenden Empfängeralgorithmen steigt in manchen Fällen quadratisch mit der Anzahl der Eingangssignale, was deren Reduktion besonders interessant macht.

Simulativ und anhand von Demonstratoren konnte die Leistungsfähigkeit dieses Prinzips nachgewiesen werden. Die Vorteile des Verfahrens sind unbestritten, allerdings stellt eine Echtzeit-Implementierung z.B. bei Wireless LAN¹ oder UMTS² aufgrund der hohen Abstraten erhebliche Anforderungen an Rechenleistung und Energieeffizienz solcher Systeme.

Besonders die Eigenwertzerlegung (EVD³) der hermiteschen Matrix \mathbf{C}_X besitzt durch

¹LAN: engl. *Local Area Network*

²UMTS: engl. *Universal Mobile Telecommunications System*

³EVD: engl. *Eigenvalue Decomposition*

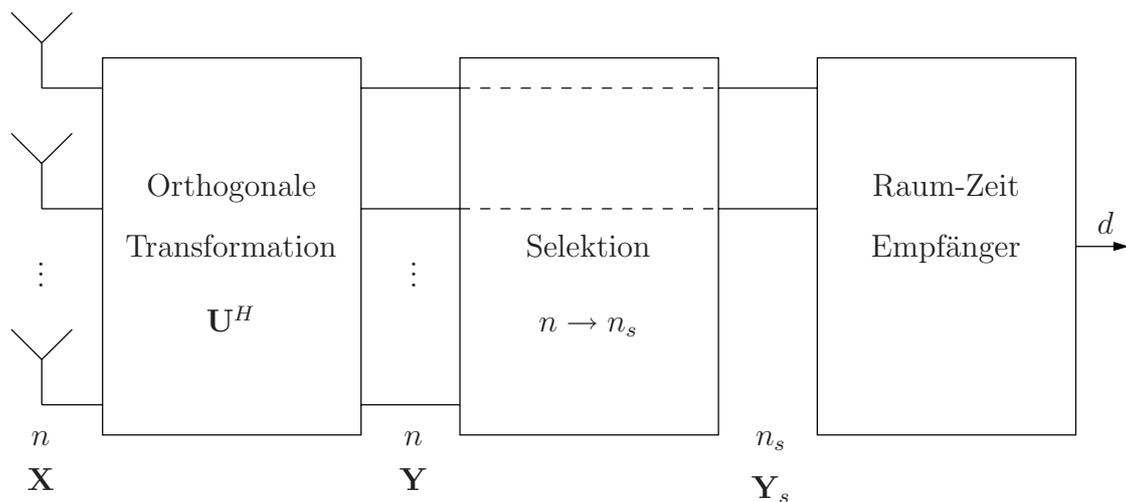


Abbildung 1.1: Raum-Zeit-Empfänger mit Dimensionreduktion

eine geschickte, hardwareangepasste Implementierung Optimierungspotenzial bezüglich des Rechenaufwandes.

Für eine Echtzeit-Implementierung bieten sich prinzipiell ASICs¹, FPGAs² und DSPs³ an. Da ein weiterer bedeutender Forschungsschwerpunkt am Lehrstuhl mobile Nachrichtensysteme die Weiterentwicklung von SIMD⁴-DSP-Architekturen ist, führte dies in Zusammenhang mit der dargestellten Anwendung zu dem Thema „Das hermitesche Eigenwertproblem - Implementierungsaspekte für Festkomma-SIMD-DSPs“.

Die KLT und die Lösung des hermiteschen Eigenwertproblems finden natürlich auch in anderen Disziplinen Anwendung, so z.B.

- in der Bild- und Videoverarbeitung zur Datenkompression und Mustererkennung,
- auf dem Gebiet der künstlichen Intelligenz zur automatischen Klassifizierung,
- bei der Analyse akustischer Signale zur Maschinenüberwachung und
- zur Berechnung von Eigenfrequenzen mechanischer System sowie
- bei Stabilitätsbetrachtungen in der technischen Mechanik.

¹ASIC: engl. *Application Specific Integrated Circuit*

²FPGA: engl. *Field Programmable Gate Array*

³DSP: engl. *Digital Signal Processor*

⁴SIMD: engl. *Single Instruction Multiple Data*

Die Arbeit ist wie folgt aufgebaut: Im nächsten Kapitel wird in Kurzform die CATS¹ DSP-Plattform eingeführt. Besonders hervorgehoben werden der Datenpfad des DSP sowie das Kommunikationsnetzwerk zwischen den parallelen Datenpfaden.

In Kapitel 3 werden die mathematischen Grundlagen der KLT und der Eigenwertzerlegung behandelt. Verschiedene Algorithmen werden dargestellt und auf Tauglichkeit für die Implementierung auf einem Plattform-DSP untersucht.

Das Kapitel 4 widmet sich der Parallelisierung der Algorithmen für eine linear-parallele Struktur, die durch den DSP vorgegeben ist.

Die Berechnung der Parameter für die Matrixoperationen unter dem Gesichtspunkt der Festkomma-Arithmetik wird in Kapitel 5 diskutiert.

Im Kapitel 6 wird eine eigene Idee zum Tracking von Eigenpaaren vorgestellt, die auf dem zweiseitigen Jacobi-Algorithmus basiert.

Das Kapitel 7 beschreibt dann die Ergebnisse der Umsetzung des Algorithmus auf einem DSP-Modell unter Nutzung der vorher theoretisch herausgearbeiteten Erkenntnisse.

¹CATS: engl. *Concept for Application Tailored Signalprocessors*