

Sprachsynthese im Zeitbereich unter Verwendung von Phon- und Diphonclustern

von
Dipl.-Ing. Carsten Jürgens

**Wissenschaft und Technik Verlag
Berlin**

INHALTSVERZEICHNIS

Zusammenfassung	5
1 Einleitung	7
1.1 Sprachausgabesysteme	7
1.2 Anwendungsgebiete der Sprachsynthese.....	8
1.3 Zielsetzung und Gliederung der Arbeit	10
2 Grundbegriffe der Phonetik und Phonologie	13
2.1 Vorgang der menschlichen Artikulation.....	14
2.2 Möglichkeiten der Lauteinteilung	16
2.3 Phonetische und graphemische Einheiten	19
2.4 Charakteristika von Sprachsignalen	21
3 Bekannte Verfahren der Sprachsynthese	23
3.1 Linguistisch-phonetische Transkription	23
3.1.1 Klassifikationsmöglichkeiten	24
3.1.2 Clusterbasierte Transkription	25
3.1.3 Transkriptionsfehler	29
3.2 Phonetisch-akustische Umsetzung	31
3.2.1 Klassifikationsmöglichkeiten	31
3.2.1.1 Artikulatorische Modelle	32
3.2.1.2 Verkettung von Lautelementen	32
3.2.2 PSOLA-Verfahren.....	34
3.2.2.1 TD-PSOLA	35
3.2.2.2 FD-PSOLA	38
3.2.2.3 LP-PSOLA.....	40
3.2.2.4 Vor- und Nachteile der verschiedenen PSOLA-Verfahren	43
3.3 Ansätze für eigene Untersuchungen.....	44

4 Untersuchungen zur Optimierung der Lautübergänge	45
4.1 Inhärente Prosodie.....	46
4.1.1 Glättung von Grundfrequenzverläufen bei abgespeicherten Lautelementen	46
4.1.2 Glättung der Abstände der Grundperiodenmarken.....	49
4.2 Amplitudendiskontinuitäten	50
4.2.1 Energiediskontinuitäten.....	51
4.2.2 Excitement-Muting.....	54
4.3 Frequenzdiskontinuitäten	55
4.3.1 Impulse im Frequenzbereich	55
4.3.2 Sprünge in der Einhüllenden des Spektrums.....	58
4.3.3 Einfügen einer interpolierten Übergangszone	63
4.3.4 Weitere Verfahren zur spektralen Glättung.....	65
4.4 Phasendiskontinuitäten.....	66
4.5 Konkatenation und Grundperiodenmarken	67
4.5.1 Mittelwertmodul.....	68
4.5.2 Automatische Anpassung der PSOLA-Analysefensterlänge	68
4.6 Zusammenfassung	69
 5 Prosodiesteuerung	 71
5.1 Intonationsmodelle.....	71
5.2 Intonationsmodell von Fujisaki.....	72
5.3 Regelgesteuerte Prosodie	77
5.4 Vergleich mit natürlichen Intonationskurven	82
5.5 Zusammenfassung	91
 6 Lautelemente-Bibliothek	 93
6.1 Lautelemente für die Sprachsynthese	93
6.2 Segmentierung von Lautelementen	95
6.2.1 Manuelle Segmentierung.....	95
6.2.2 Automatische Segmentierung.....	97
6.2.2.1 Methode der Parameteränderung	98
6.2.2.2 Methode der Parameterkorrelation.....	104
6.2.2.3 Implementiertes automatisches Verfahren	107
6.2.2.4 Segmentierungsergebnisse	111

6.2.2.5 Weitere Ansätze	116
6.3 Verfahren zur Bestimmung der Grundfrequenz	118
6.3.1 Grundfrequenzbestimmung mit Hilfe der Kurzzeitanalyse.....	120
6.3.2 Grundfrequenzbestimmung im Zeitbereich	123
6.3.3 Grundfrequenzbestimmung bei TUBSY	126
6.3.4 Ergebnisse der Grundfrequenzanalyse	128
6.4 Verfahren zur Sprachgütemessung synthetischer Sprache	132
6.4.1 Testverfahren zur Beurteilung der Verständlichkeit.....	132
6.4.2 Testverfahren zur Beurteilung der Prosodie.....	135
6.5 Vergleich verschiedener Lautelemente für die Sprachsynthese.....	136
6.5.1 Erstellung des Testmaterials.....	136
6.5.2 Durchführung der Sprachgütemessung	137
6.5.3 Ergebnisse und Auswertung des Hörtests.....	138
6.6 Lautinventar für TUBSY	142
6.7 Zusammenfassung	148
7 Sprecherauswahl	149
7.1 Bestimmung von apparativen Stimmerkmale	150
7.1.1 Sprachmaterial.....	151
7.1.2 Ergebnisse der physikalischen Messungen.....	151
7.1.3 Zusammenfassung	156
7.2 Bestimmung von auditiven Stimmerkmale.....	157
7.2.1 Auswahl der Beurteilungskriterien.....	157
7.2.2 Sprachmaterial und Versuchspersonen.....	158
7.2.3 Versuchsaufbau	158
7.2.4 Ergebnisse	159
7.2.5 Zusammenfassung	164
7.3 Vergleich von apparativen und auditiven Stimmerkmale.....	165
7.3.1 Stimmlage.....	165
7.3.2 Stimmhärte	167
7.3.3 Klarheit.....	170
7.3.4 Gehauchtheit	172
7.3.5 Auswertung	174
7.4 Auswahl von repräsentativen Sprechern	177

7.5 Durchführung und Auswertung des Hörtests	180
8 Ausblick	185
9 Literaturverzeichnis	189
Anhang	195
A Graphem-Grundclusterliste	195
B Phoneminventar der deutschen Sprache.....	198
C Zusammensetzung der Lautelemente-Bibliothek.....	199
D Realisierung des Sprachsynthese-Systems	217