

ELEKTRONISCHE MESS- UND
DIAGNOSETECHNIK



Vortrag

“Implementation eines CELP Vocoders
auf einem dsPIC microcontroller”

Vortragender: Daniel Gühne

Gliederung

- Zielstellung
- Grundkonzept CELP Encoder/Decoder
- FS-1016 Standard
- Demo-Vocoder CELP 3.3C
- Konzeption CELP Vocoder auf dsPIC
- Qualitätsbewertung via PESQ-Standard
- Integration CELP Vocoder in Spisa-Infrastruktur
- Zusammenfassung der Ergebnisse
- Ausblick

Zielstellung

- Erarbeitung eines lauffähigen CELP Vocoders für den dsPIC30F4013
- Höhere Sprachqualität als bei bisherigen LPC10 (bisher 3 Implementationen vorhanden)
- Reduzierung des Arbeitsspeicherbedarfs auf <2kByte des dsPIC
- Speicherung der Sprach-/CELP-Daten auf SD-Karte
- Integration des Vocoders in vorhandene Spisa-Infrastruktur (Steuerung und Daten über PC)
- Nutzung des FS-1016 Standards und dessen Demo-Implementation als Basis und Vergleich

Gliederung

- Zielstellung
- Grundkonzept CELP Encoder/Decoder
- FS-1016 Standard
- Demo-Vocoder CELP 3.3C
- Konzeption CELP Vocoder auf dsPIC
- Qualitätsbewertung via PESQ-Standard
- Integration CELP Vocoder in Spisa-Infrastruktur
- Zusammenfassung der Ergebnisse
- Ausblick

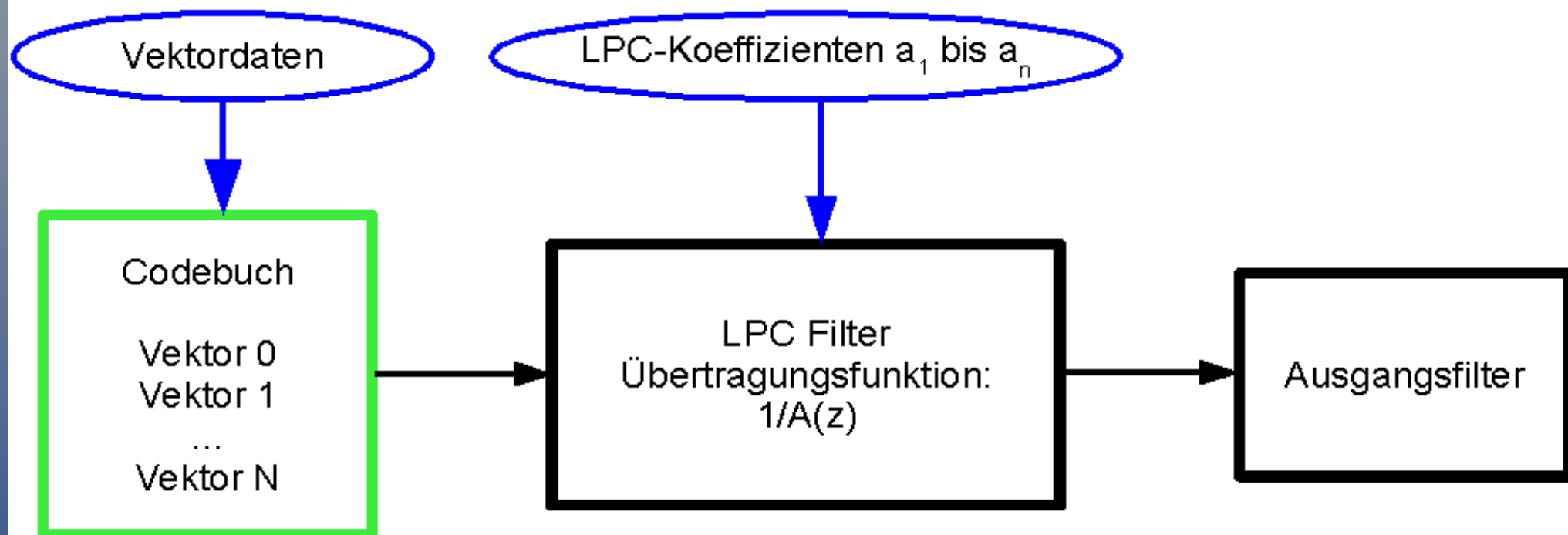
Grundkonzept CELP

Grundkonzept CELP Encoder/Decoder:

- Struktur Decoder
- Struktur Encoder
- LPC Analyse
- Perceptual Weighting
- Selektierung Excitation Vektor

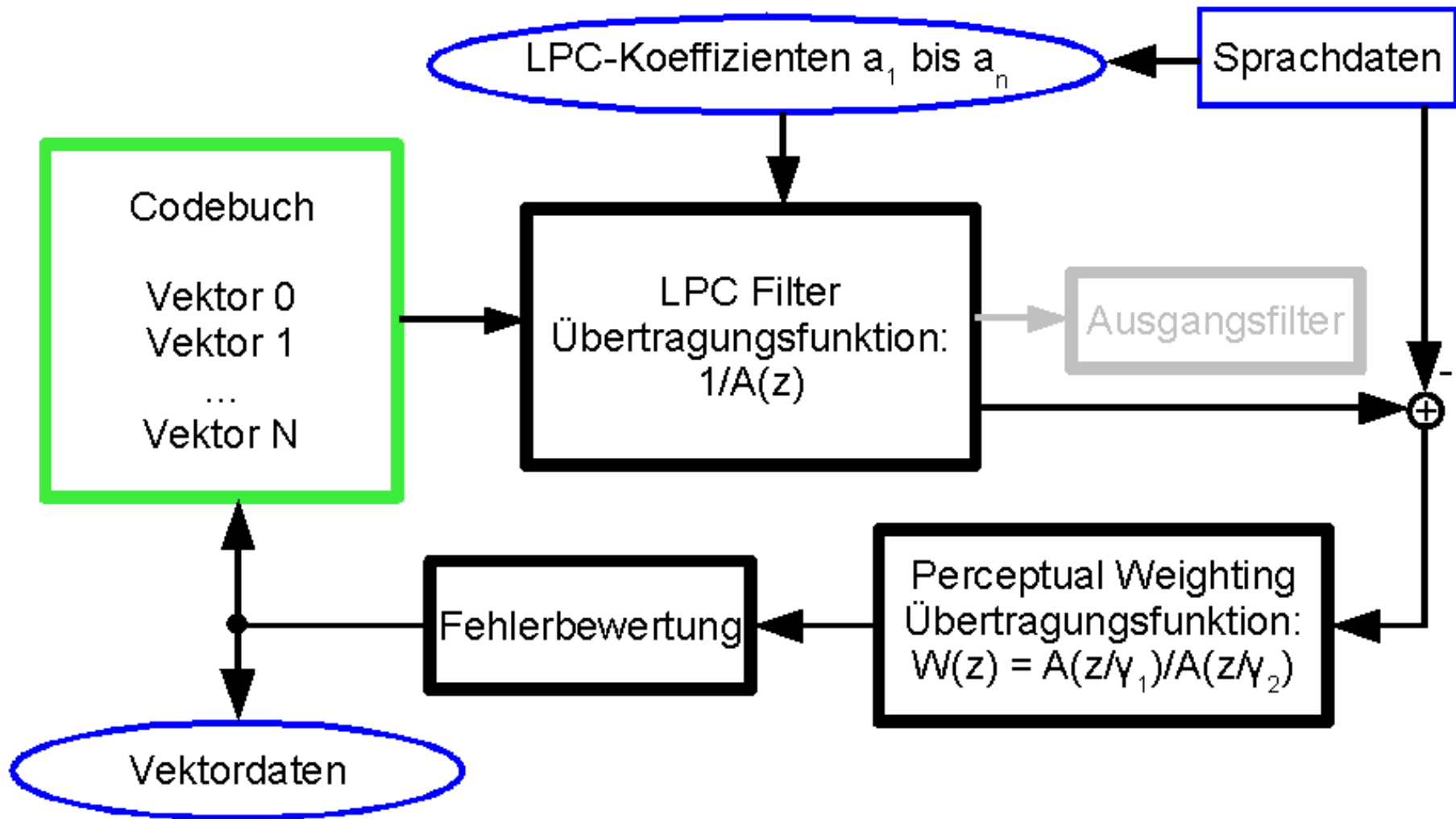
CELP : Code-Excited Linear Prediction

Struktur Decoder



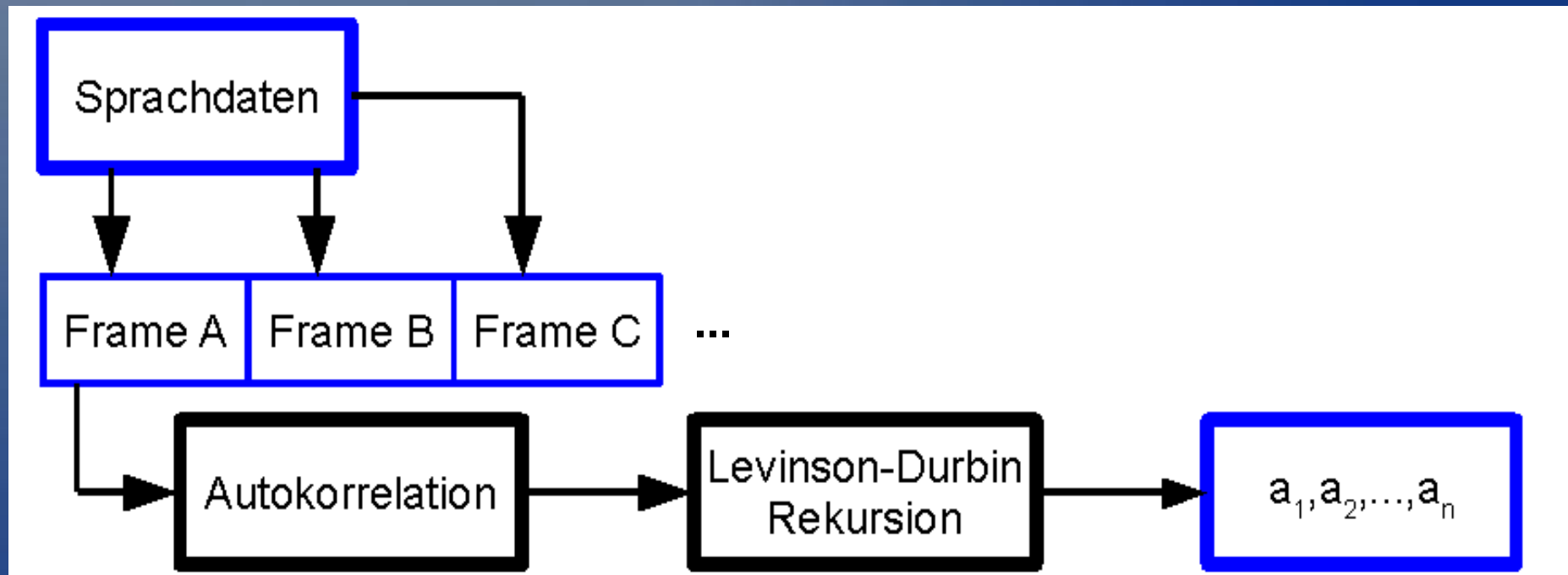
$$A(z) = 1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_n z^{-n}$$

Struktur Encoder



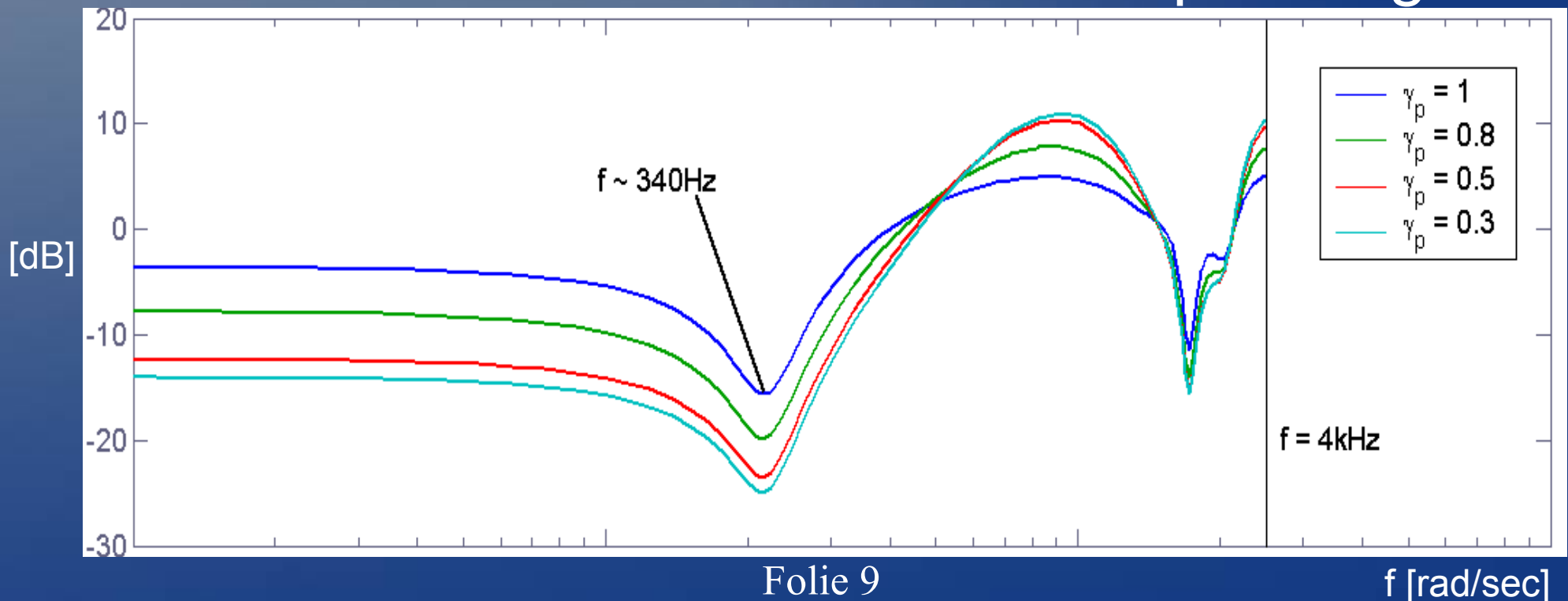
LPC Analyse

- Analyse erfolgt Frame für Frame
- Autokorrelation mit Verschiebung $0, 1, 2, \dots, n$
- Levinson-Durbin Rekursion zur Lösung des Filter-
-Problems



Perceptual Weighting

- Filterkoeffizienten a_1 bis a_n werden expandiert
$$a_{ex,i} = a_i \cdot \gamma^i$$
- Filter reduziert im Idealfall Eingangsdaten auf die wahrnehmbaren Bestandteile des Sprachsignals



Selektierung Excitation Vektor

- Brute-Force: Ermittlung der Fehlerfunktion für jeden Vektor des Codebuchs
- Fehlerfunktion muss „minimal“ sein
- typische Bewertungen für Minimum:
 - Energieminimum (Autokorrelation)
 - Amplitudenabweichung von Gewichtsfunktion
 - ...

Gliederung

- Zielstellung
- Grundkonzept CELP Encoder/Decoder
- FS-1016 Standard
- Demo-Vocoder CELP 3.3C
- Konzeption CELP Vocoder auf dsPIC
- Qualitätsbewertung via PESQ-Standard
- Integration CELP Vocoder in Spisa-Infrastruktur
- Zusammenfassung der Ergebnisse
- Ausblick

FS-1016 Standard

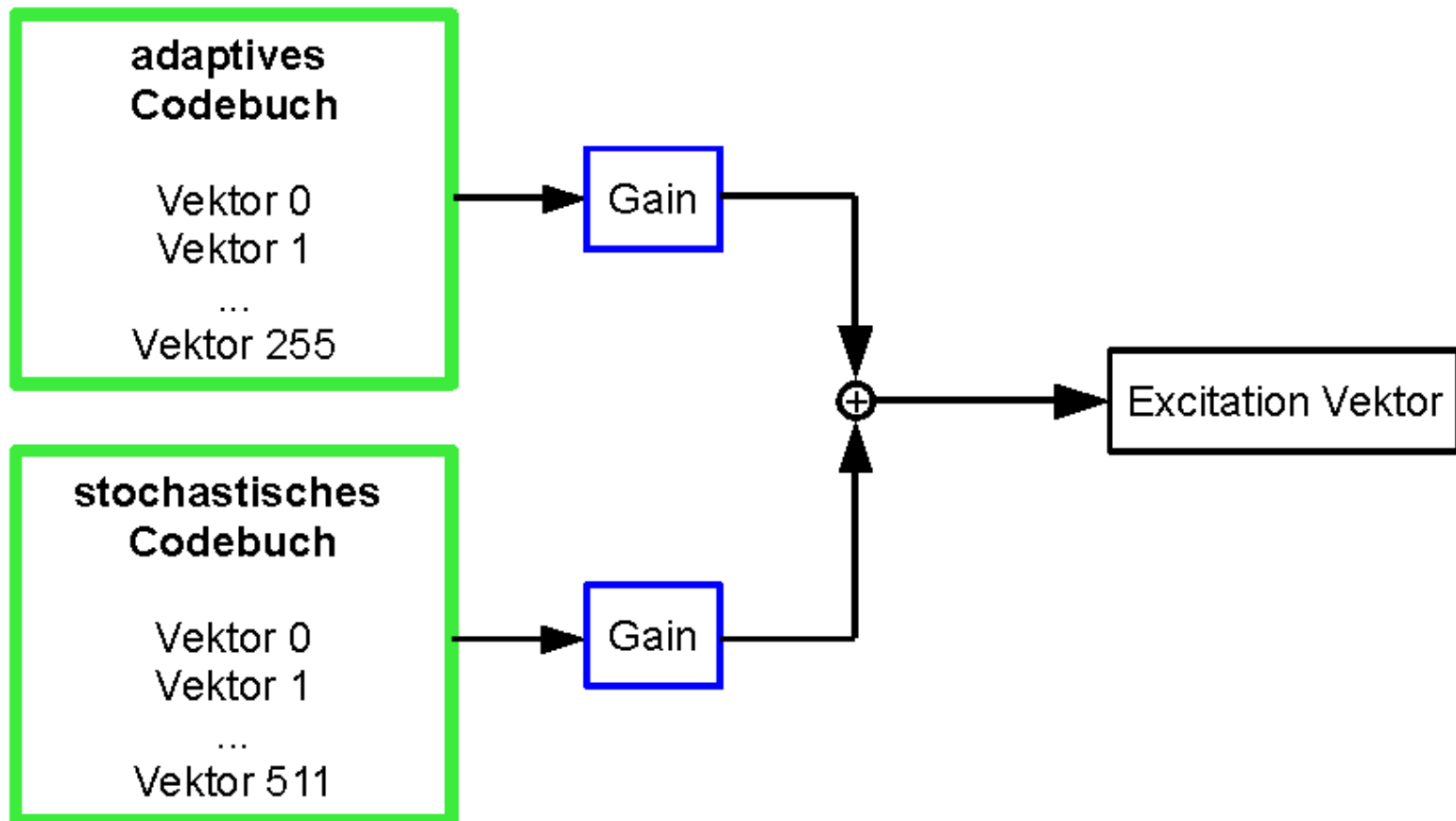
FS-1016 Standard:

- Kenndaten
- Codebuch Konzept
- Adaptive Codes
- Stochastische Codes
- Line Spectral Frequencies (LSF/LSP)
- Interleaving Line Spectral Parameter

FS-1016 Kenndaten

- Datenformat der Sprachdaten:
 - vorzeichenbehaftet 16bit Werte
 - Sampling-Frequenz 8kHz
- 30ms Frame-Länge (240 Samples)
- Analyse in 4 Sub-Frames (7.5ms/60 Samples)
- 2 Codebücher: adaptives und stochastisches
- 144Bit/Frame --> 4800 Bit/s Datenrate
- LPC Filter 10.Ordnung

Codebuch Konzept



Adaptive Codes (I)

- 256 adaptive Codes
- 128 ganzzahlige und 128 fraktale Codes
- Codebuch enthält 147 Elemente
- ganzzahlige Codes werden aus den 147 Elementen direkt gebildet
- fraktale Codes werden durch Interpolation aus dem benachbarten, ganzzahligen Code gebildet
- Codebuch wird mit dem zuletzt genutzten Vektor aktualisiert

Adaptive Codes (II)

Adaptives Codebuch

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
30	31	32	33	34	35	36	37	38	39	40	41	42	43	44
45	46	47	48	49	50	51	52	53	54	55	56	57	58	59
60	61	62	63	64	65	66	67	68	69	70	71	72	73	74
75	76	77	78	79	80	81	82	83	84	85	86	87	88	89
90	91	92	93	94	95	96	97	98	99	100	101	102	103	104
105	106	107	108	109	110	111	112	113	114	115	116	117	118	119
120	121	122	123	124	125	126	127	128	129	130	131	132	133	134
135	136	137	138	139	140	141	142	143	144	145	146			

Adaptive Codes (III)

- Codes 0..39 werden durch Wiederholung der ersten(neuesten) 20...59 Elemente des Codebuchs gebildet

Vektor 0

0	1	2	...	19	20	21	22	23	...	52	53	54	55	56	57	58	59
19	18	17	...	0	19	18	17	16	...	7	6	5	4	3	2	1	0

Vektor 40

0	1	2	...	19	20	21	22	23	...	52	53	54	55	56	57	58	59
59	58	57	...	40	39	38	37	36	...	7	6	5	4	3	2	1	0

- Codes 40..127 werden ohne Wiederholung ab Element 60 des Codebuchs gebildet

Stochastische Codes

- 3 Werte je Element zulässig: -1, 0, 1
- Vektoren sind analog zu adaptiven Codes verschränkt
- Codebuch enthält 1082 Elemente
- Codebuch wird nicht zur Laufzeit modifiziert

Vektor 0

0	1	2	...	19	20	21	22	23	...	52	53	54	55	56	57	58	59
0	1	2	...	19	20	21	22	23	...	52	53	54	55	56	57	58	59

Vektor 1

0	1	2	...	19	20	21	22	23	...	52	53	54	55	56	57	58	59
2	3	4	...	21	22	23	24	25	...	54	55	56	57	58	59	60	61

Line Spectral Frequencies

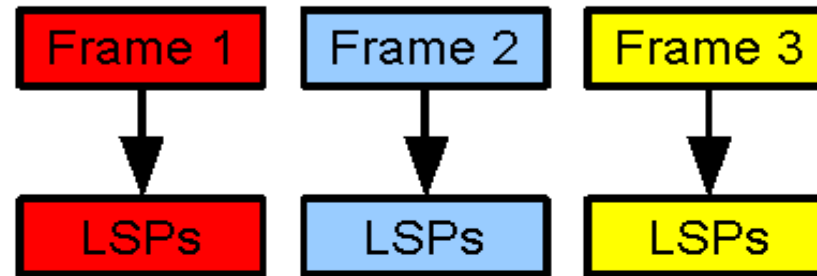
- Konvertierung der LPC-Koeffizienten a_1 bis a_{10} in LSP 1...10 (Line Spectrum Parameters/Pairs)
- LSPs bilden immer einen stabilen Filter!!!
- LSPs sind definiert als die Wurzeln folgender Polynomiale:

$$F_1 = A(z) - z^{-11} \cdot A(z^{-1})$$

$$F_2 = A(z) + z^{-11} \cdot A(z^{-1})$$

- keine direkte mathematische Lösung
- Lösung über numerische Annäherung

Interleaving LSP



Subframe 3	Subframe 4	Subframe 1	Subframe 2	Subframe 3	Subframe 4	Subframe 1	Subframe 2
1/8	3/8	5/8	7/8	7/8	5/8	3/8	1/8
7/8	5/8	3/8	1/8	1/8	3/8	5/8	7/8

$$LSP = f \cdot LSP_{alt} + g \cdot LSP_{neu}$$

$$f = \frac{7}{8}, \frac{5}{8}, \frac{3}{8}, \frac{1}{8}$$

$$g = \frac{1}{8}, \frac{3}{8}, \frac{5}{8}, \frac{7}{8}$$

Gliederung

- Zielstellung
- Grundkonzept CELP Encoder/Decoder
- FS-1016 Standard
- Demo-Vocoder CELP 3.3C
- Konzeption CELP Vocoder auf dsPIC
- Qualitätsbewertung via PESQ-Standard
- Integration CELP Vocoder in Spisa-Infrastruktur
- Zusammenfassung der Ergebnisse
- Ausblick

Demo-Vocoder CELP 3.3C

Demo-Vocoder CELP 3.3C:

- Varianten
- Konzepte zur Rechenzeitminimierung
- Korrekturkonzepte adaptive Vektorsuche
- Korrekturkonzepte stochastische Vektorsuche

Varianten CELP 3.3C

3 Varianten sind über symbolische Konstanten compilierbar:

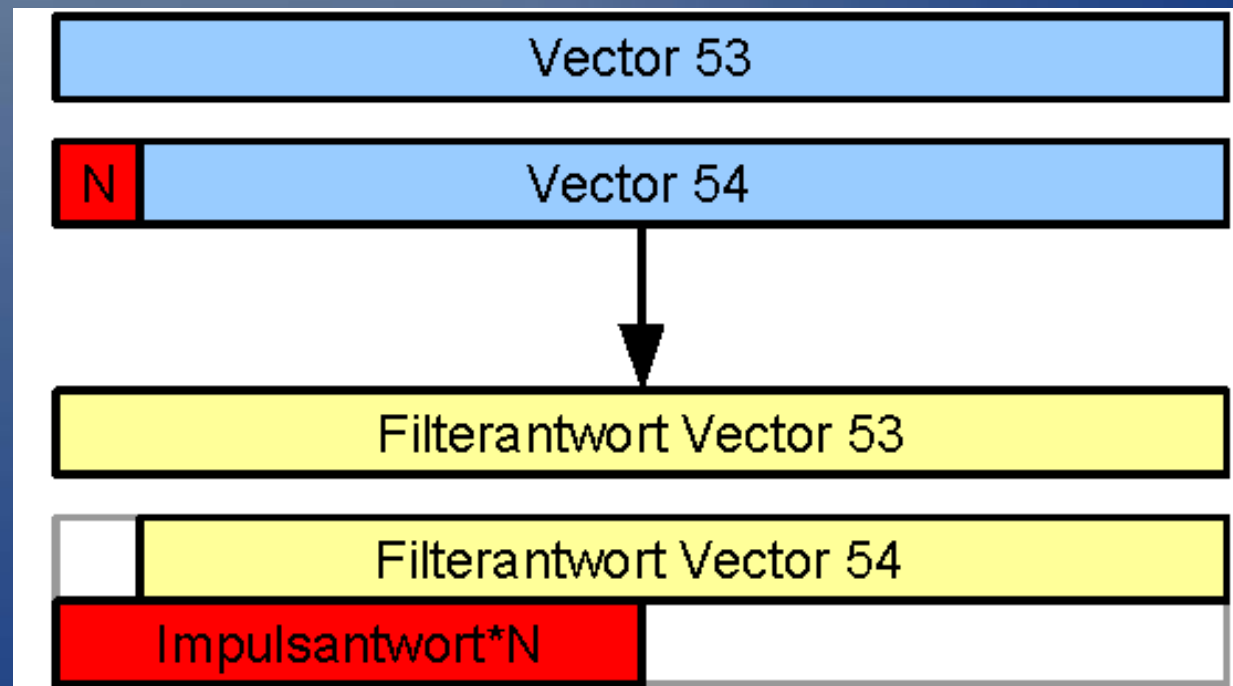
- alle 256 adaptiven Vektoren werden durchsucht
- nur die ganzzahligen Vektoren werden durchsucht
- ganzzahlige adaptive Vektoren und Nachbarschaftssuche

Nachbarschaftssuche:

- Programmierfehler vorhanden, benutzt fraktale Vektoren obwohl über symbolische Konstante deaktiviert

Konzepte Rechenzeitminimierung

- reduzierter Vektorraum für 2. und 4. Subframe bei der adaptiven Vektorsuche
- Ermittlung der Filterantwort anhand der gekürzten Impulsantwort des Filters



Modifikationen adaptive Suche

- im 2. und 4. Subframe Suchraum $-31/+32$ vom letzten adaptiven Vector ausgehend
- Suche nach kürzestem Delay (Raum ± 5 um berechnetem Delay/2, /3, /4)

Modifikationen stochastische Suche

- berechnetes Gain wird über Korrelation Zielvektor adaptive Suche und gefundene adaptive Lösung nicht-linear skaliert (empirisch ermittelt)
- Impulsantwort wird bei adaptiver Lösung mit Delays < 40 (Wiederholungselemente) mit Impulstrain-Signal neu ermittelt

Gliederung

- Zielstellung
- Grundkonzept CELP Encoder/Decoder
- FS-1016 Standard
- Demo-Vocoder CELP 3.3C
- **Konzeption CELP Vocoder auf dsPIC**
- Qualitätsbewertung via PESQ-Standard
- Integration CELP Vocoder in Spisa-Infrastruktur
- Zusammenfassung der Ergebnisse
- Ausblick

Konzeption CELP auf dsPIC

Konzeption CELP auf dsPIC:

- Speicherbedarf
- Konzepte zur Speicherbedarfsminimierung
- Qualitätsbeeinflussung des Encoders

Speicherbedarf

- adaptives Codebuch 588 Byte
- stochastisches Codebuch mind. 128 Byte, real 512
- Eingangsdaten 480 Byte
- Zwischenspeicher Eingangsdaten 240 Byte
- CELP-Daten (10 LSPs, 8 Vektoren, 8 Gains)
mind. 18 Byte, real 30 Byte
- alte LSPs mind. 5 Byte, real 10 Byte
- Impulsantwort 120 Byte
- Ausgangsdaten Impulsantwort 120 Byte
- Zwischenspeicher für alte Ein- und Ausgangswerte der Filter

--> 2 kByte RAM sind definitiv zu wenig

Speicherbedarf Minimierung

- A) stochastisches Codebuch und alle Festwerte ins Flash verlagern
- B) Ein- und Ausgangswerte der Filter auf 16bit-Werte begrenzen
- C) adaptives Codebuch nimmt durch B) 16bit Werte auf
- D) keine Suche nach kürzestem Delay

Qualitätseinstellung Encoder

Ziel der Qualitätseinstellungen: Rechenzeit sparen

- Reduzierung des Vektorraums in 6 Stufen
- stochastischer Suchraum: 512, 256, 128, 32, 16, 2
- adaptiver Suchraum: 128, 128, 128, 128, 64, 32

Rechenzeit je Frame reduziert sich von 5.28s auf 1.38s

Gliederung

- Zielstellung
- Grundkonzept CELP Encoder/Decoder
- FS-1016 Standard
- Demo-Vocoder CELP 3.3C
- Konzeption CELP Vocoder auf dsPIC
- Qualitätsbewertung via PESQ-Standard
- Integration CELP Vocoder in Spisa-Infrastruktur
- Zusammenfassung der Ergebnisse
- Ausblick

Qualitätsbewertung via PESQ

Qualitätsbewertung via PESQ-Standard:

- Grundlagen
- PESQ Werte der einzelnen Implementationen

PESQ Grundlagen

- ITU-T P.862 Standard (02/2001)
- objektives, automatisches Verfahren zur Bewertung der Sprachqualität
- PESQ Skala umfasst die Werte MOS von -0.5 bis 4.5
- PESQ 1.7 entspricht etwa MOS 1.0
- PESQ 4.5 entspricht MOS 5.0

Perceptual Evaluation of Speech Quality (PESQ)
Mean Opinion Score (MOS)

PESQ Werte

Datei	CELP 3.3C voller Suchraum	CELP 3.3C ganzzahliger Suchraum	TUBCELP PC-basiert	TUBCELP dsPIC- basiert
original.wav	2.831	2.795	2.840	2.742
track49.wav	3.237	3.188	3.155	
track50.wav	3.368	3.381	3.323	
track51.wav	3.244	3.215	3.195	
track52.wav	3.469	3.467	3.406	
track53.wav	3.326	3.271	3.208	
track54.wav	3.557	3.582	3.540	

Gliederung

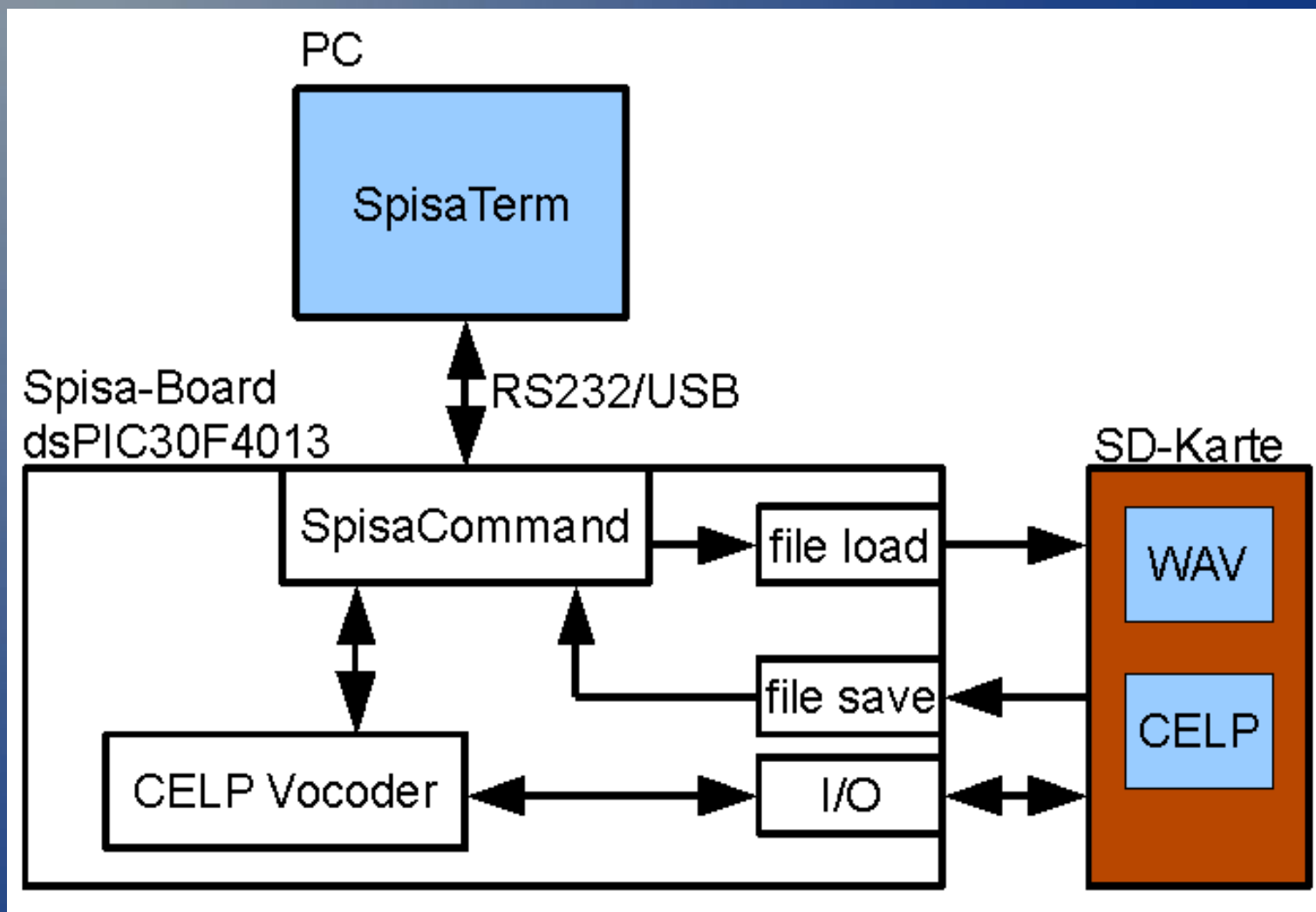
- Zielstellung
- Grundkonzept CELP Encoder/Decoder
- FS-1016 Standard
- Demo-Vocoder CELP 3.3C
- Konzeption CELP Vocoder auf dsPIC
- Qualitätsbewertung via PESQ-Standard
- **Integration CELP Vocoder in Spisa-Infrastruktur**
- Zusammenfassung der Ergebnisse
- Ausblick

Spisa Integration

Integration CELP Vocoder in Spisa Infrastruktur:

- Spisa
- Nutzung SD-Karte
- Benutzerinterface

Spisa



Nutzung SD-Karte

- SD-Karte wird ohne Dateisystemaufsatz genutzt
- Audiodaten im 1., CELP im 2. Teil der Karte gespeichert

SD-Karte 64MB

Sektor 0 Dateilänge	WAV-Datei, 8kHz Mono, max. 31,74MiB
--> max. 2079s oder 34,66min.	
Sektor 65000 Dateilänge	TUBCELP-Datei, 4533Bit/s, int-Vektoren

Benutzer Interface

- Nutzung der vorhanden „file load/save“ Funktionen zum Laden/Sichern der Daten
- zusätzliche Funktionen „celp encode“ und „celp decode“ zur Kodierung/Dekodierung nach/von CELP
- Funktion „celp quality=x“ zum Setzen der Encoder-Qualität ($x = [1,2,3,4,5,6]$)

Gliederung

- Zielstellung
- Grundkonzept CELP Encoder/Decoder
- FS-1016 Standard
- Demo-Vocoder CELP 3.3C
- Konzeption CELP Vocoder auf dsPIC
- Qualitätsbewertung via PESQ-Standard
- Integration CELP Vocoder in Spisa-Infrastruktur
- Zusammenfassung der Ergebnisse
- Ausblick

Zusammenfassung

- PC-basierte Implementation erreicht identische Qualität wie CELP 3.3C
- CELP auf dsPIC funktioniert und liefert gute Sprachqualität
- Datenspeicherung auf SD-Karte funktioniert
- Integration in Spisa-Umgebung funktioniert
- Qualitätseinstellung arbeitet auf Basis von Rechenzeit-Bedarf

Ausblick

- Erhöhung der Qualität über zusätzliche Zwischenspeicher in den Filtern möglich
- Verringerung des Speicherbedarfs (mind. 240 Byte für halben Frame) durch 2-stufiges Kodierungsverfahren
- Beschleunigung des Vocoders durch Nutzung des DSP-Teils des Microcontrollers

Vielen Dank für die Aufmerksamkeit!