Adaptive Loss Concealment for Internet Telephony Applications

Henning Sanneck

GMD FOKUS, Berlin

sanneck@fokus.gmd.de

Supported by the USMInt (DFN) and Multicube (ACTS) projects

Overview

- Motivation
- Receiver-Based Concealment
- **Adaptive Packetization / Concealment** (Sender/Receiver operation)
- Properties (packet sizes / header overhead, delay)
- Subjective Test
- Conclusions

Motivation: Loss of Speech Packets

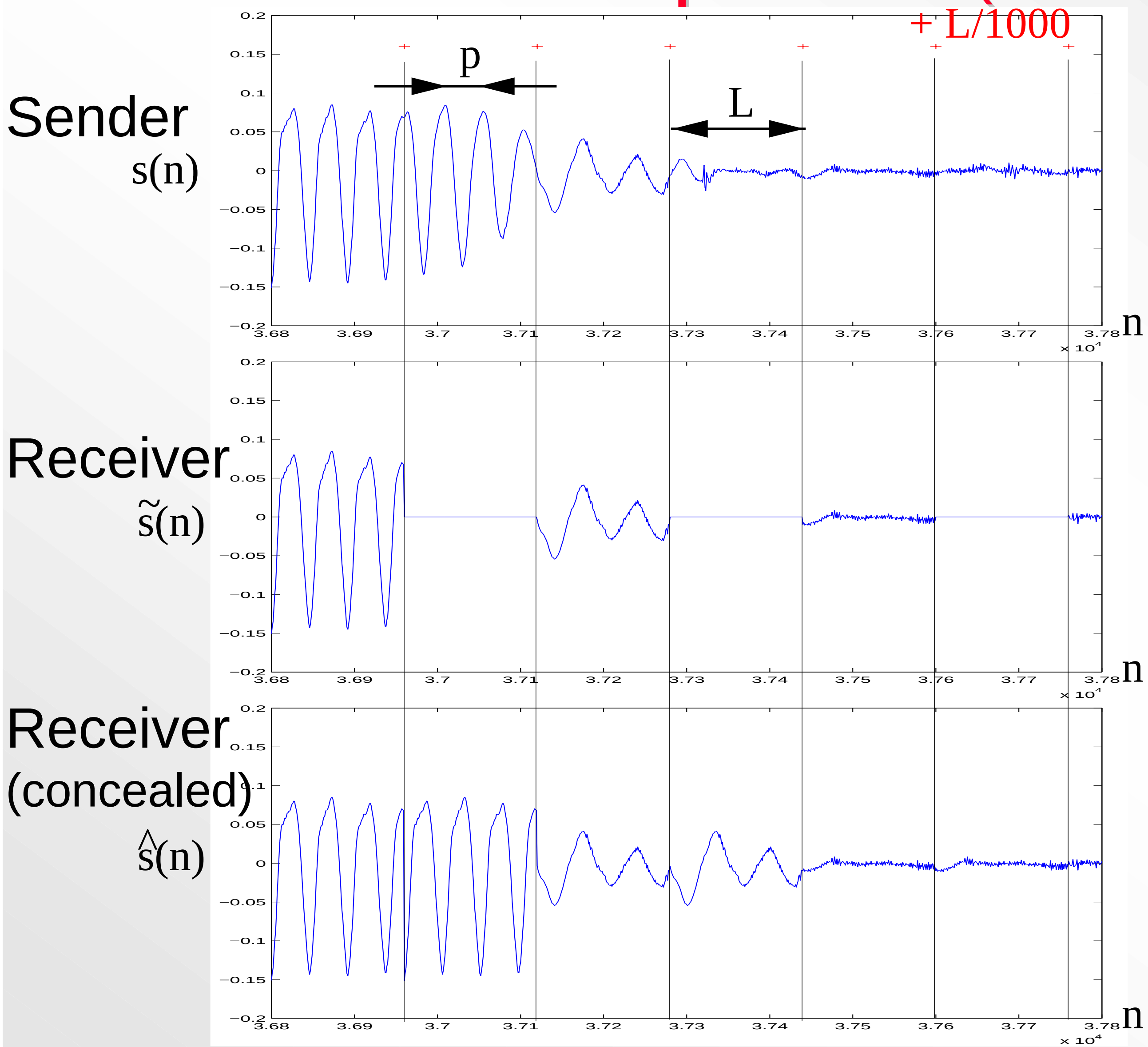
- Congestion in the Internet / Mbone
  ⇨ Packet Loss
  ⇨ speech signal dropouts
  ⇨ need to enhance speech quality
- Solutions:

  bandwidth adaptation, resource reservation,

  differential services, redundancy/FEC,

  interleaving, ***receiver-based*** **concealment**

Packet Repetition (Receiver-Based)



n: sample number

p(n): pitch period

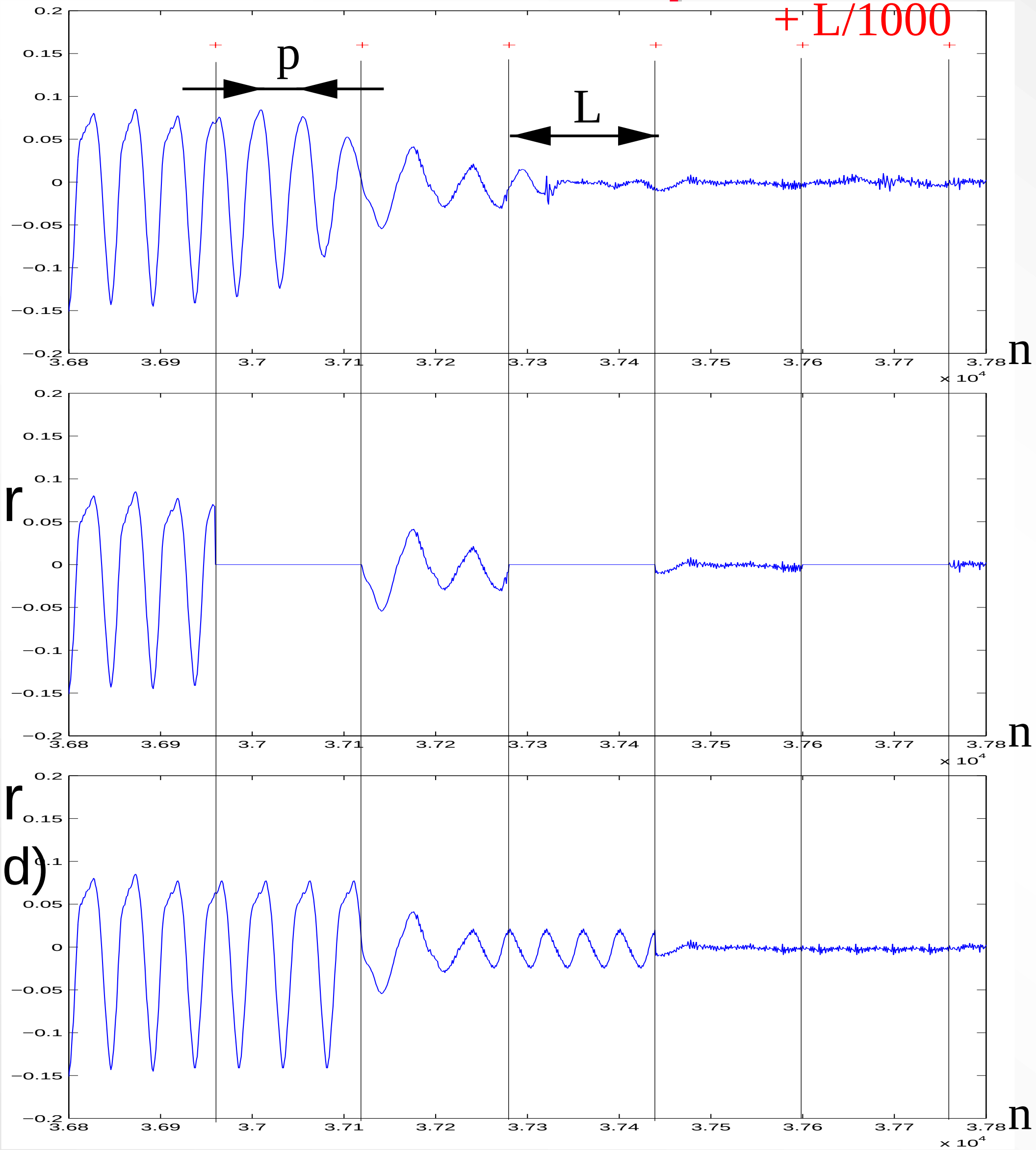
L: packet size

Pitch Waveform Replication (Receiver-Based)

Sender
s(n)

Receiver
$\tilde{s}(n)$

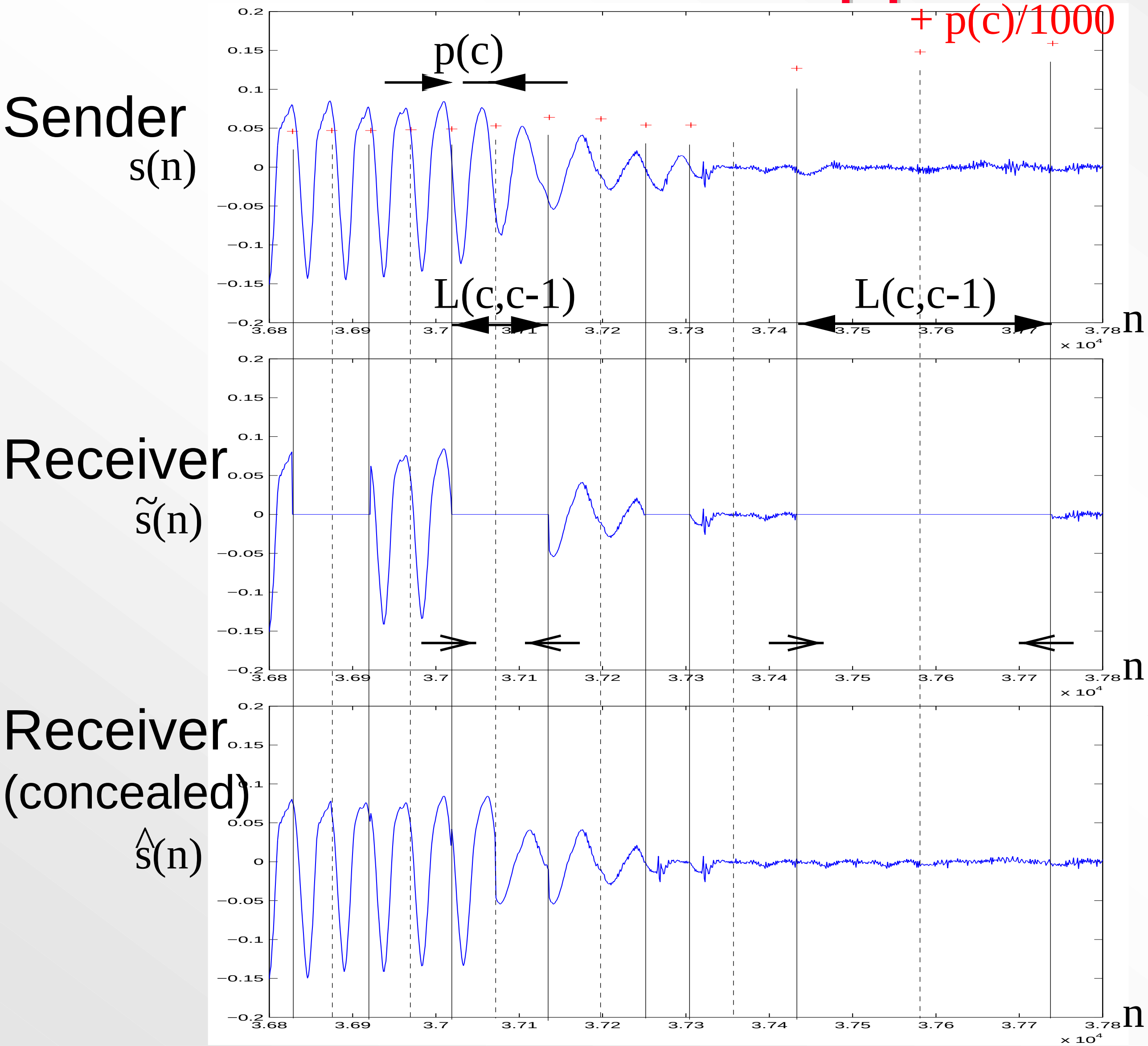
Receiver
(concealed)
$\hat{s}(n)$



p(n): pitch period

L: packet size

New approach



c: „chunk“ number

p(c): pitch period

L(c,c-1):

packet size

Adaptive Packetization / Concealment

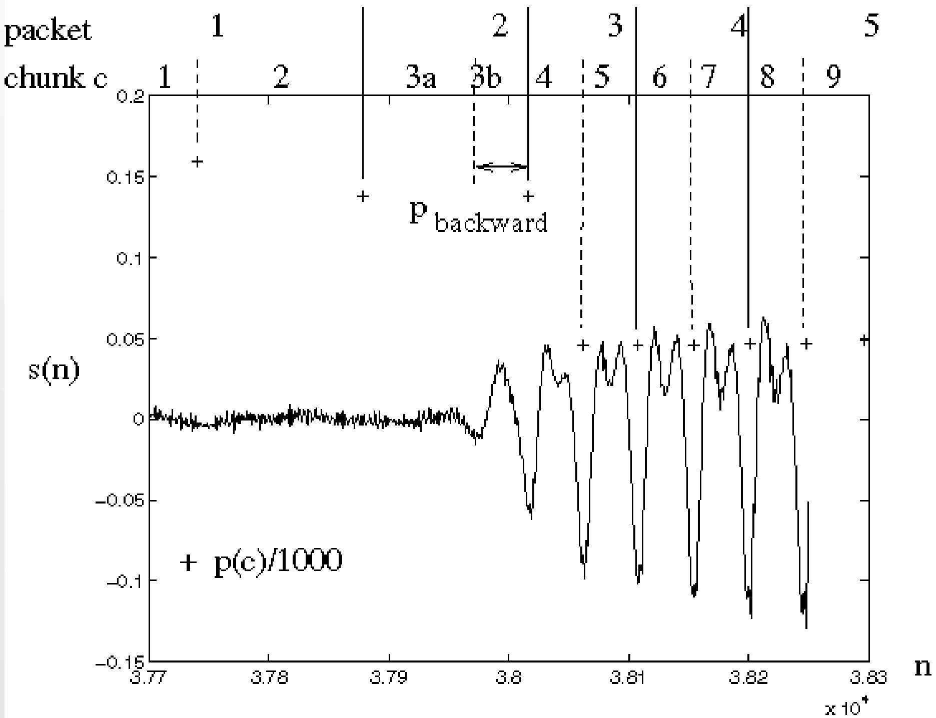
Sender-supported concealment:

choose packetization interval adaptively

- packet size (size of lost segment) relates to „importance“ of packet content
- pre-processing of the *undistorted* signal
- enables *simple* concealment operation at the receiver (high probability that adjacent packets contents resemble each other)

Sender: Adaptive Packetization

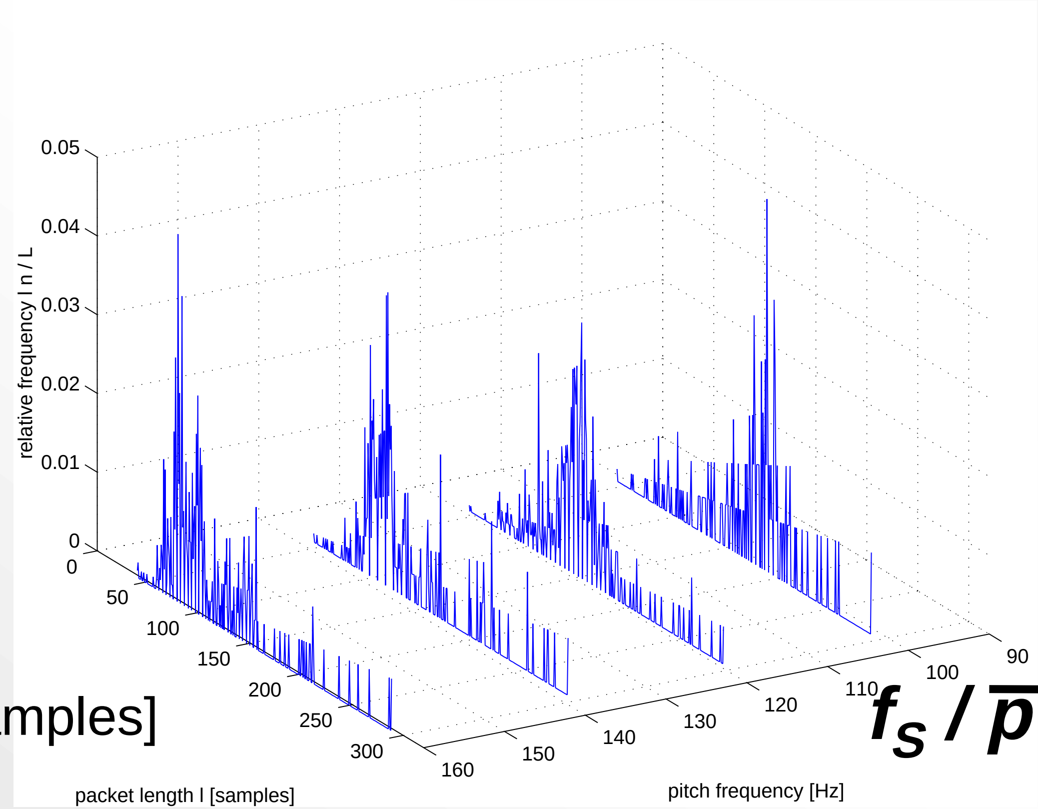
- Auto-correlation of signal ⟹ partitioning („chunks“)
- speech content transition check: voiced/unvoiced
- packetization: 2 chunks/packet (header overhead)



(110 ms speech)

Packet Size Frequency Distribution

- Packet size is now dependent on speaker's pitch (range: $\boldsymbol{p_{min}}$= 30, $\boldsymbol{2p_{max}}$= 320 samples; *4...40ms*)



$\boldsymbol{l}$ [samples]

$\boldsymbol{f_S / \overline{p}_V}$ [Hz]

(Relative frequency $\boldsymbol{n}$ weighted with size $\boldsymbol{l}$)

GMD FOKUS

Relative Packet Header Overhead

Typical value for IP Telephony: $O = 20\%$

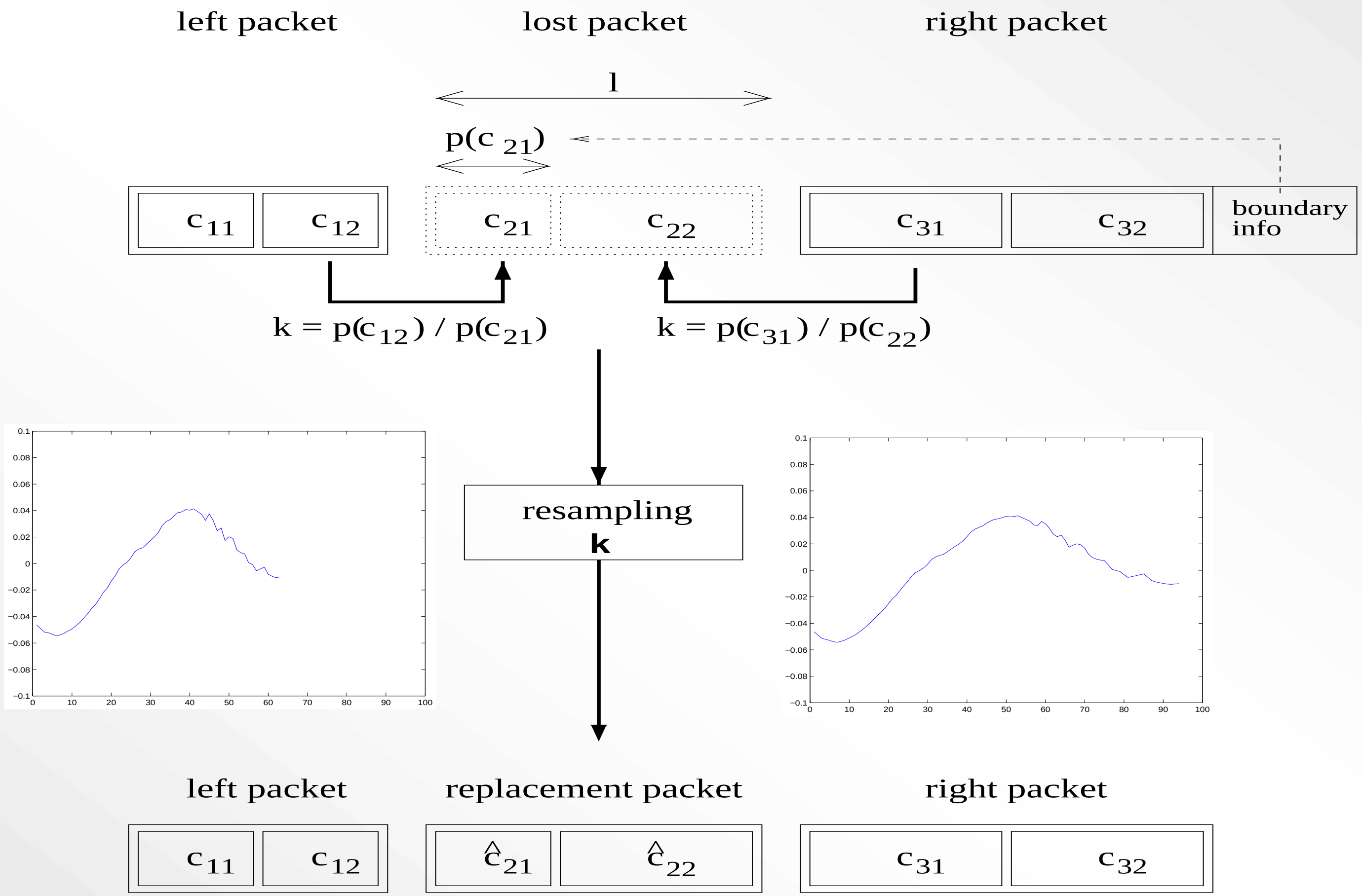
- fixed packetization interval: 160 samples [20ms],
- RTP/UDP/IP per packet overhead: $o = 40$ octets

AP/C:

Speaker	**Estimated overhead** $o/(o+2p_v)$	**Measured overhead** O **[%]**
Male low	20.16	20.14
Male high	22.97	22.83
Female low	25.72	24.84
Female high	28.62	27.98

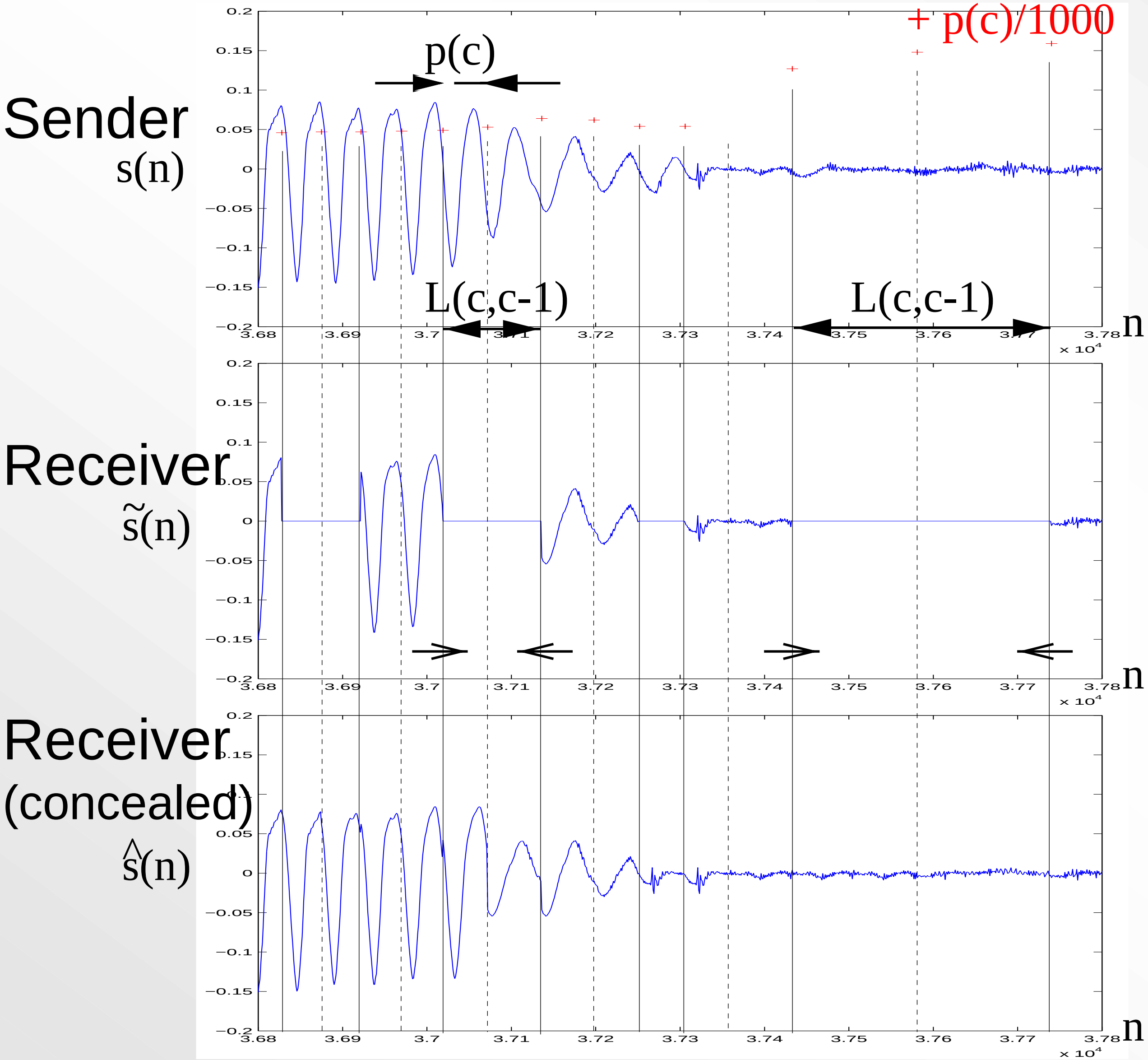
(mean pitch period: p_v)

Receiver: Concealment



resampling: no specific distortions introduced (like e.g. with PWR)

Receiver: Concealment (contd.)



c: „chunk“ number

p(c): pitch period

L(c,c-1):
packet size

Discussion

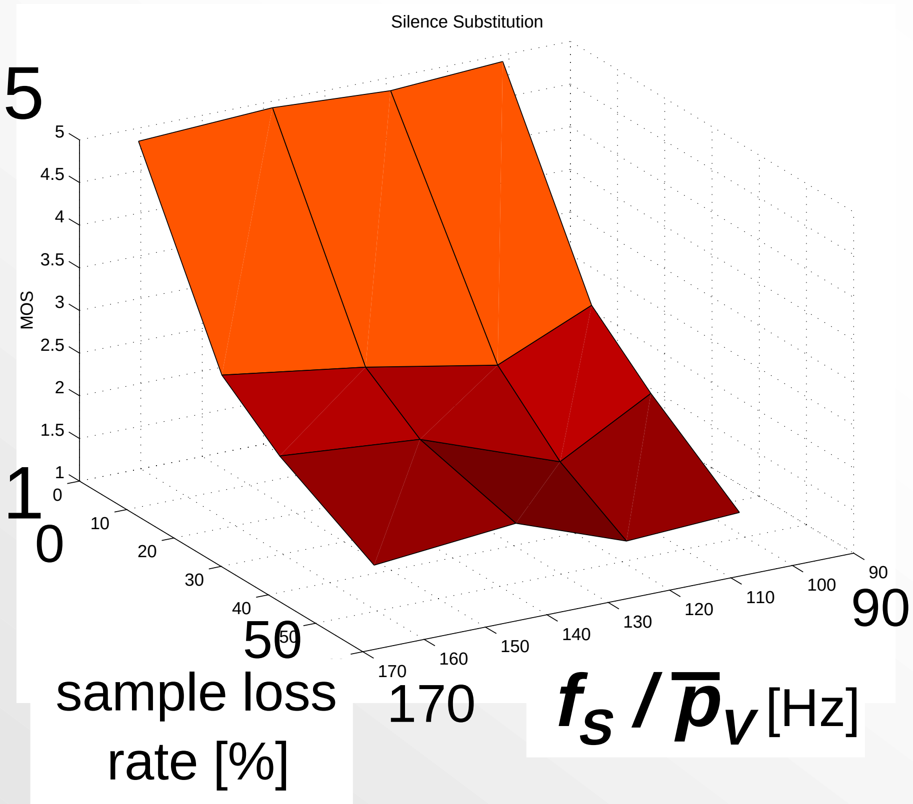
- characteristic“ information: 2 octets (own and following intra-packet boundary)
- additional **delay** (buffering):
  sender: [p_{max}, $2p_{max}$-p_{min}]= *20...36ms*
  receiver: [p_{min}, $2p_{max}$]= *4...40ms* (on loss only)
- computational complexity/processing delay: low
- backwards compatible with existing tools

Subjective Test: Test Procedure

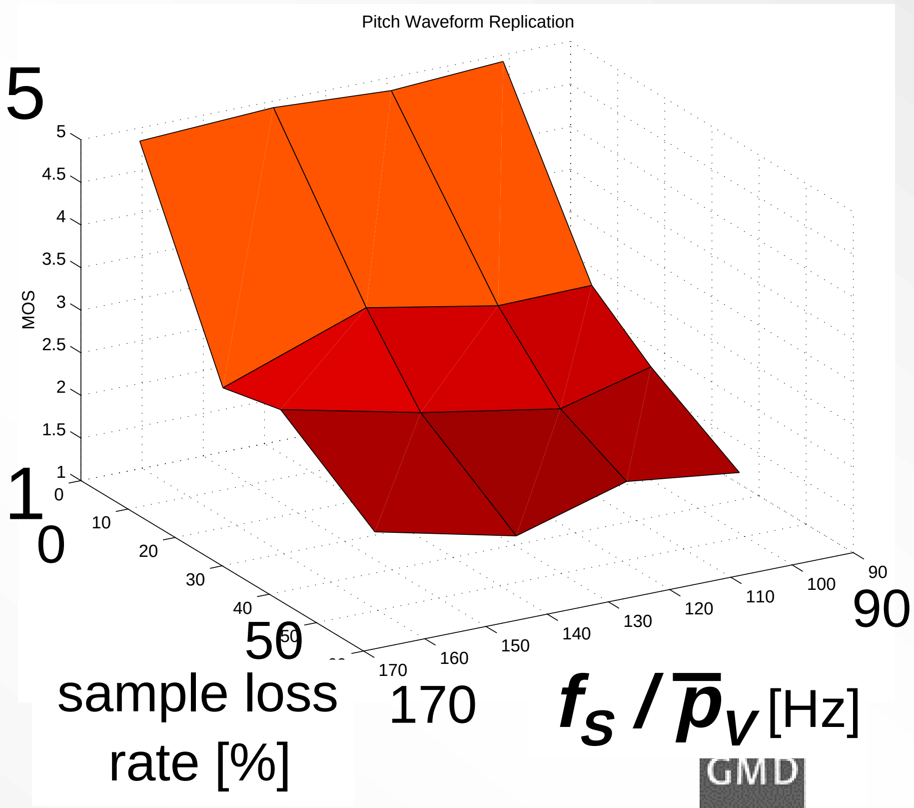
- Four signals (different speakers), PCM 16 bit linear, 8 kHz
- comparison with „silence substitution“ and PWR
- random, yet *isolated* packet losses
- 40 test conditions: 4 speakers x (3 algorithms x 3 loss rates + original)
- thirteen non-expert listeners judged on MOS scale
- Anchoring: Original=5, „Worst Case“=1 (50% loss)
- test conditions in rapid, random sequence

Subjective Test: Results

MOS: Silence Substitution



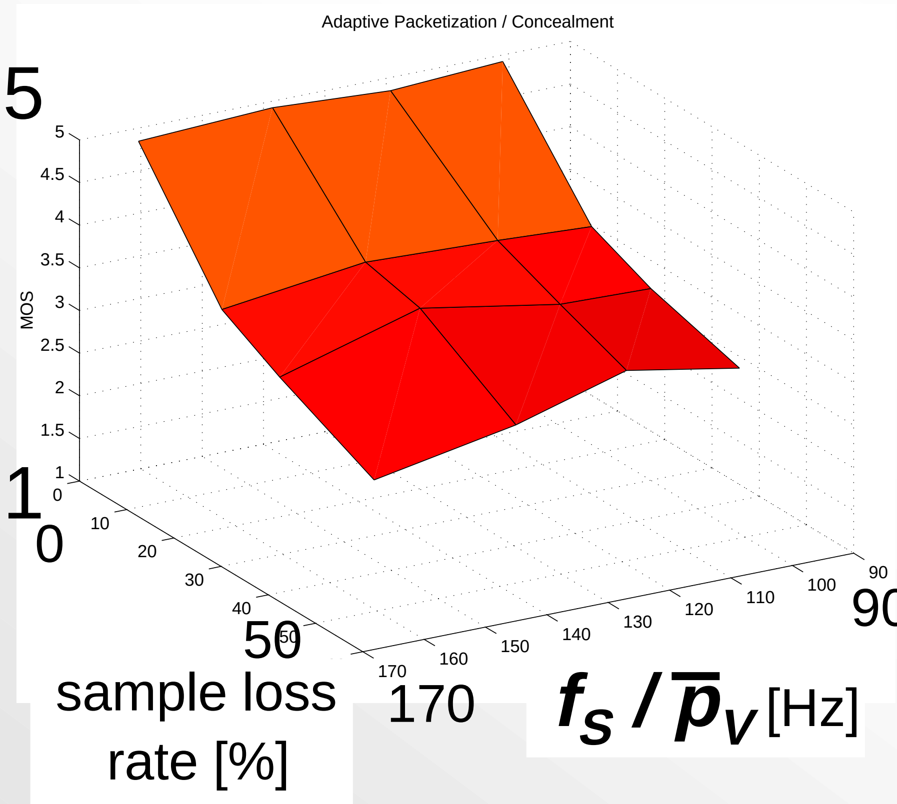
MOS: Pitch Waveform Replication



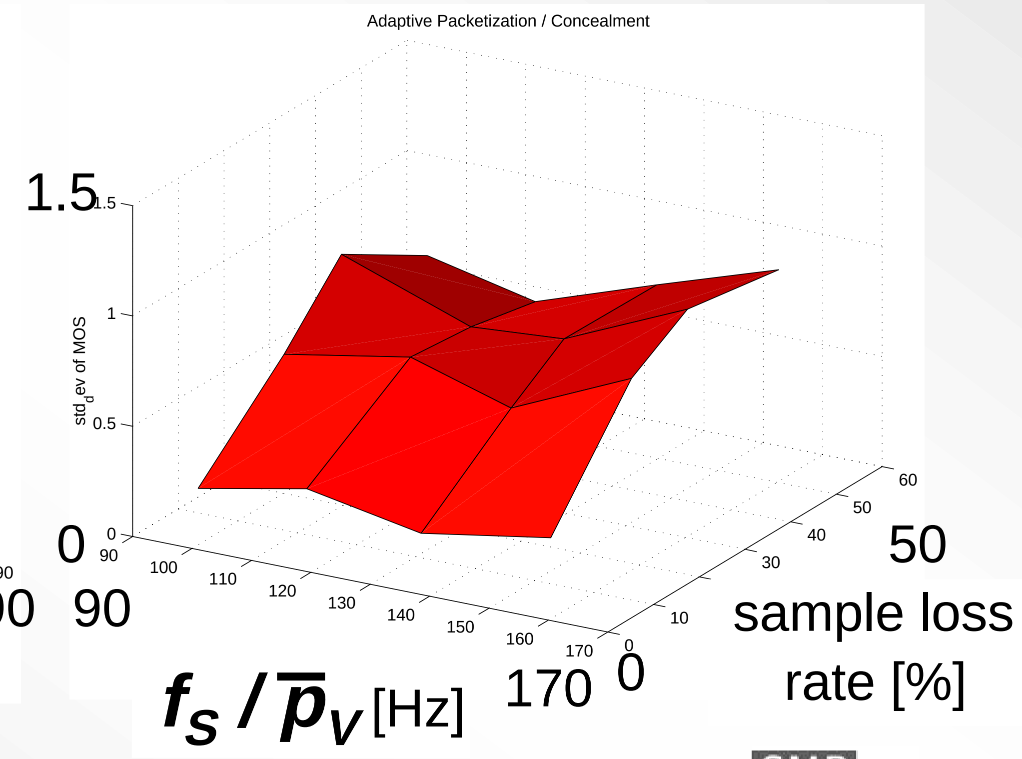
GMD FOKUS

Subjective Test: Results (contd.)

MOS: Adaptive Packetization/ Concealment



Standard deviation of MOS (AP/C)



Conclusions

- Sender preprocessing *(Adaptive Packetization)*
  ⇨ pre-defined parts of the signal are dropped
  ⇨ less perceptible distortion, simple concealment
- low overhead
  (data, delay, processing, deployment)
- Future/ongoing work:
  frame-based codec support / integration
  complement end2end mechanism with queue management at routers (loss burstiness !)
- **`http://www.fokus.gmd.de/research/cc/glone/products/voice/apc`**