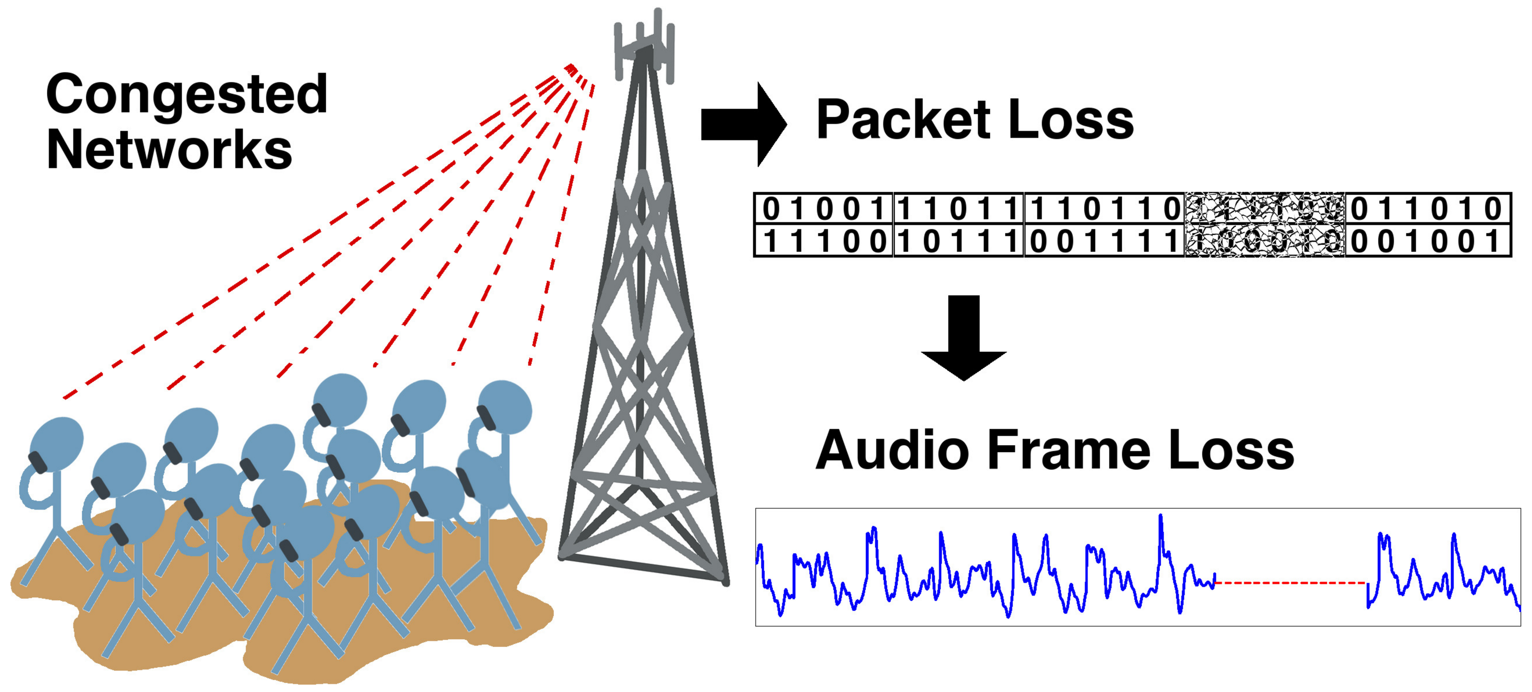
Bidirectional Cascaded Long Term Prediction for Frame Loss Concealment in Polyphonic Audio Signals

Tejaswi Nanjundaswamy and Kenneth Rose

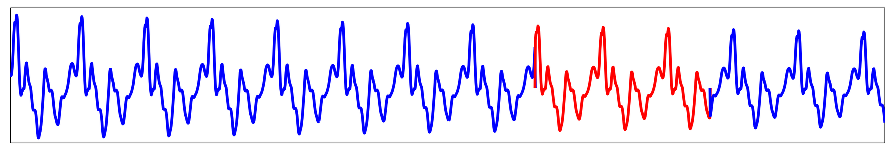
Signal Compression Lab, Department of ECE, University of California Santa Barbara

Why Frame Loss Concealment (FLC) ?



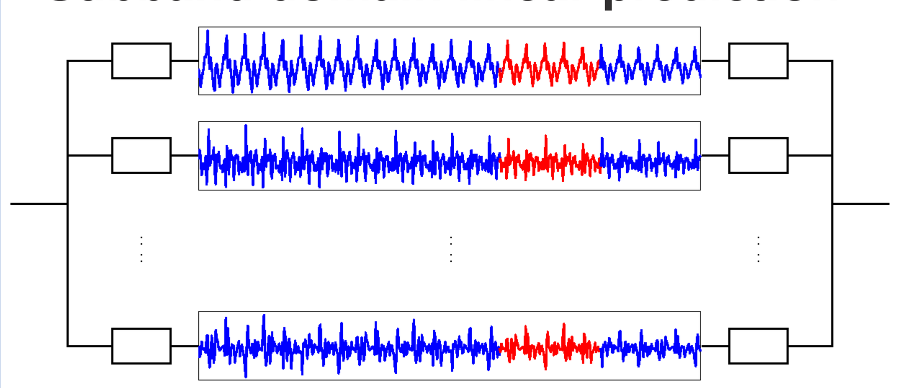
Existing FLC techniques

Waveform repetition



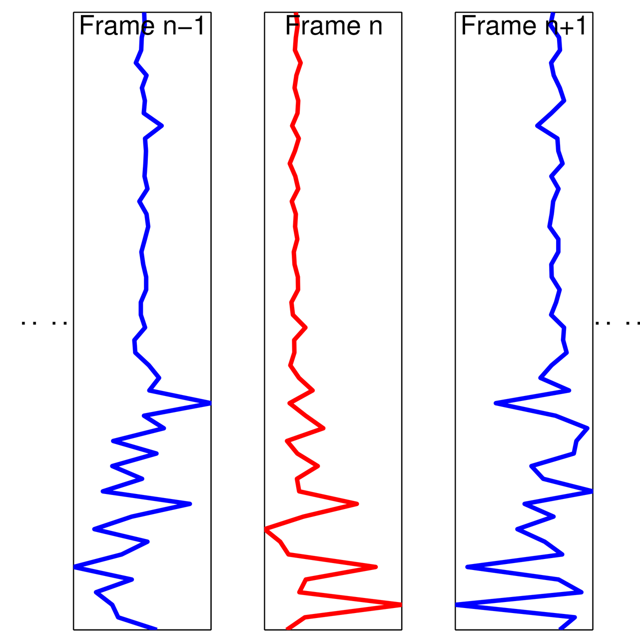
X Ineffective for most audio signals as they are polyphonic.

Subband domain linear prediction



X Ineffective for long lost blocks.

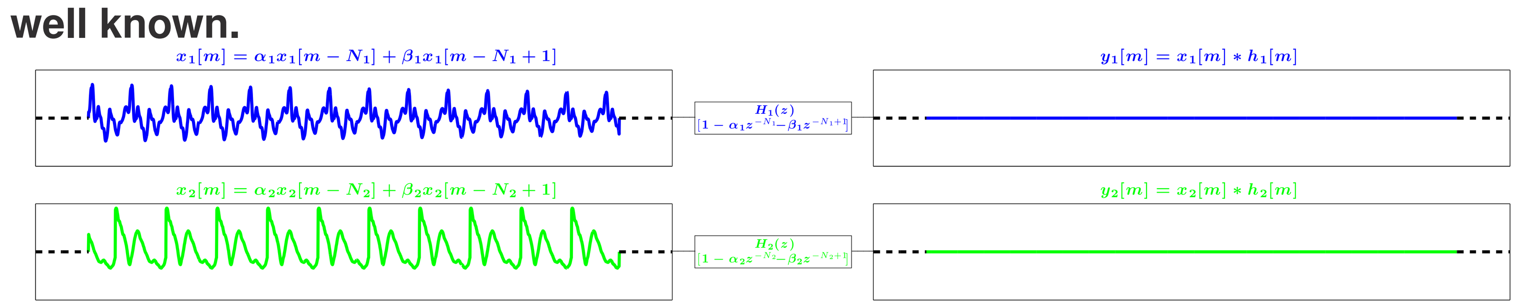
MDCT domain tonal interpolation



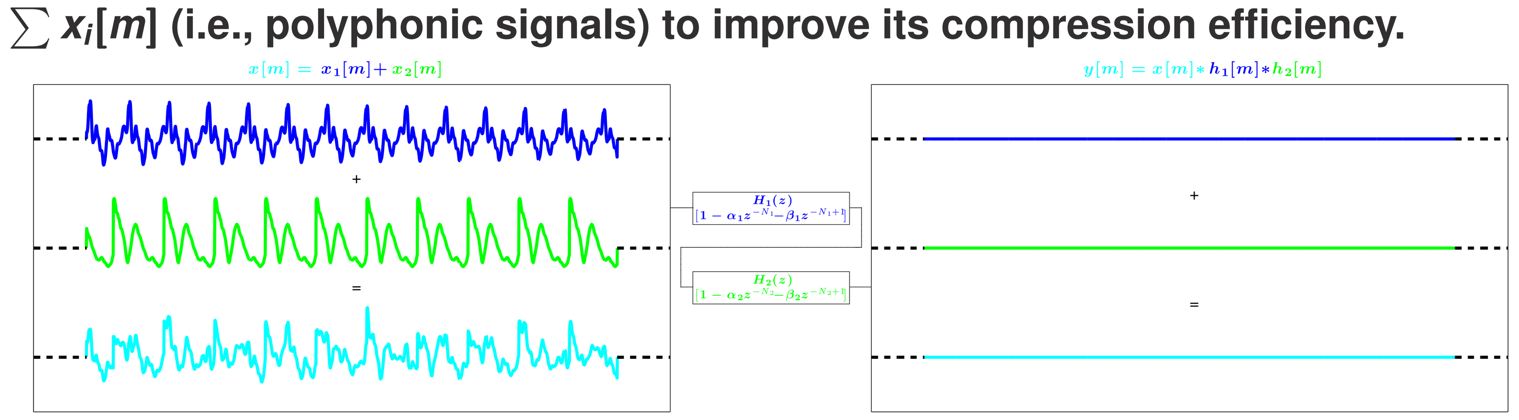
X Ineffective when MDCT fails to resolve individual tones.

Cascaded Long Term Prediction (CLTP)

- Predicting periodic signals such as $x_i[m] = \alpha_i x_i[m - N_i] + \beta_i x_i[m - N_i + 1]$ is well known.



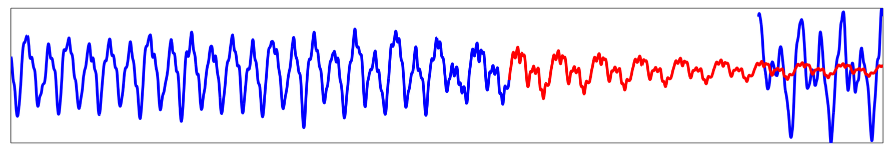
- We recently introduced CLTP for mixture of such periodic signals $x[m] = \sum x_i[m]$ (i.e., polyphonic signals) to improve its compression efficiency.



- Here we propose employing a CLTP synthesis filter, suitably adapted for FLC by utilizing the past and even the usually available future samples effectively.

Cascaded Long Term Prediction for Frame Loss Concealment

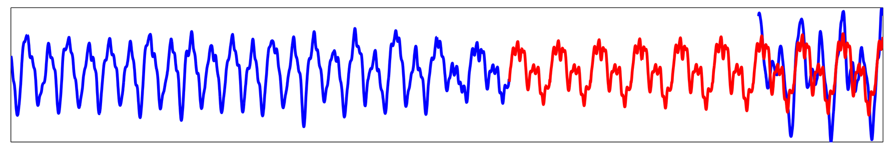
- Preliminary parameters are estimated from the past samples via a recursive technique
  - Parameters of jth filter $(1 - \alpha_j z^{-N_j} - \beta_j z^{-N_j+1})$ are estimated in the residue of filtering with all the others $\prod_{\forall i, i \neq j}(1 - \alpha_i z^{-N_i} - \beta_i z^{-N_i+1})$, via the well known technique for LTP.
  - Each filter in the cascade is estimated this way in a loop until convergence.
- Using only the past samples for the filter parameter estimate doesn't explain future samples correctly



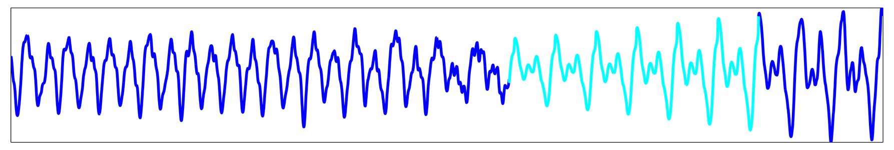
- So CLTP filter updated with multiplicative gain factors

$$H_c(z) = \prod_{i=0}^{P-1} (1 - G_i(\alpha_i z^{-N_i} + \beta_i z^{-N_i+1})).$$

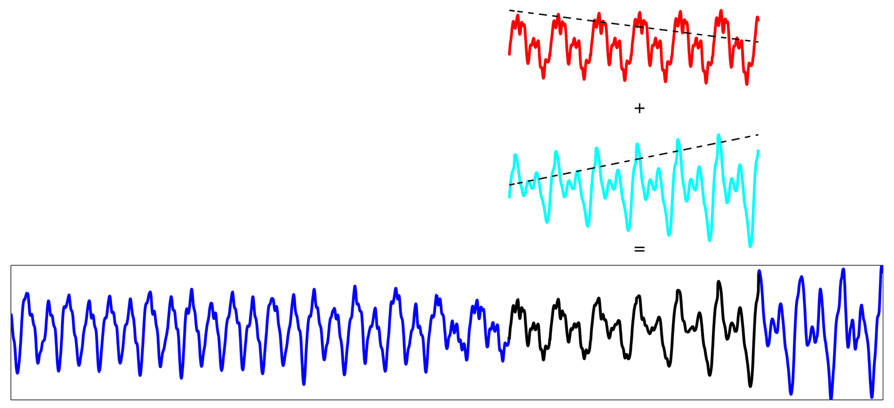
  - The gain factors are adjusted to minimize squared prediction error in the future samples.
  - As cost function has complex dependency on these factors, a generic quasi-Newton optimization called L-BFGS method is used along with backtracking line search for step sizes.



- Simply predicting from past samples doesn't ensure smooth transition into the available future samples.
- Thus lost frame samples are predicted in reverse direction from future samples with different set of CLTP gain factors.

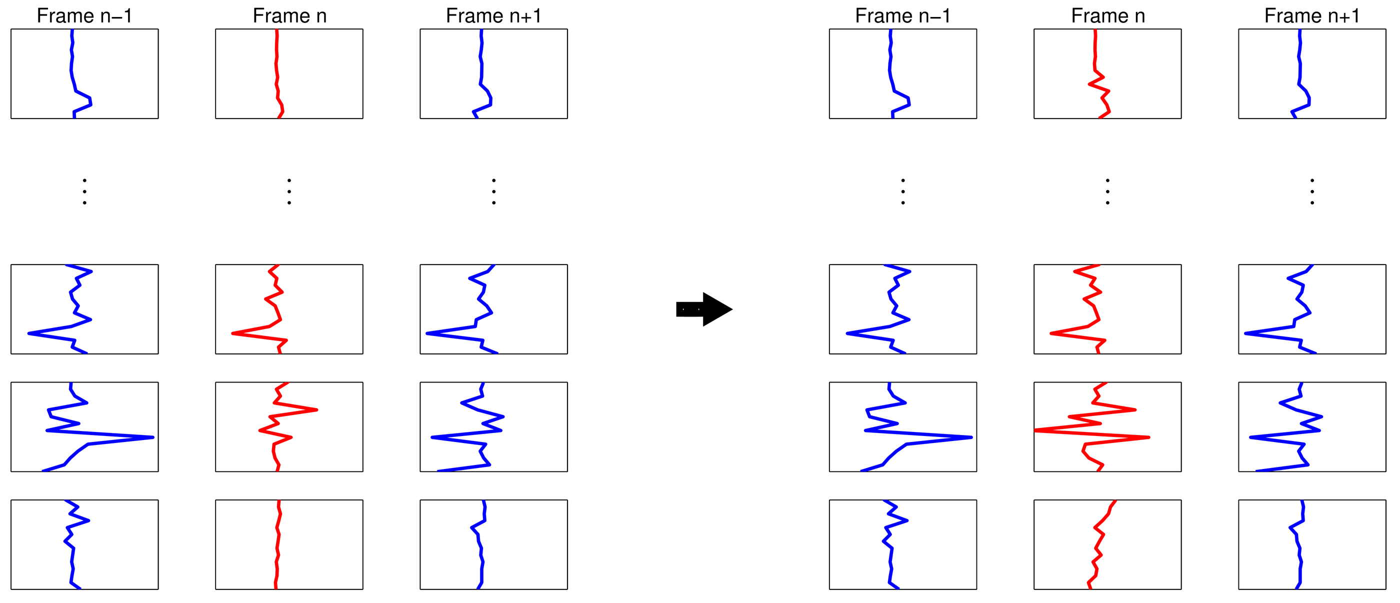


- Final reconstruction of lost frame is a weighted average of predicted samples in each direction.



- For use in MPEG AAC, the reconstructed frame is transformed to MDCT domain and energy smoothing performed in each band l, via a gain factor given as,

$$f[l] = \begin{cases} \sqrt{\dfrac{\sqrt{e_{n-1}[l]e_{n+1}[l]}}{e_n[l]}}, & \text{if } \dfrac{e_n[l]}{\sqrt{e_{n-1}[l]e_{n+1}[l]}} > T \text{ or } \dfrac{e_n[l]}{\sqrt{e_{n-1}[l]e_{n+1}[l]}} < 1/T, \\ 1, & \text{otherwise.} \end{cases}$$

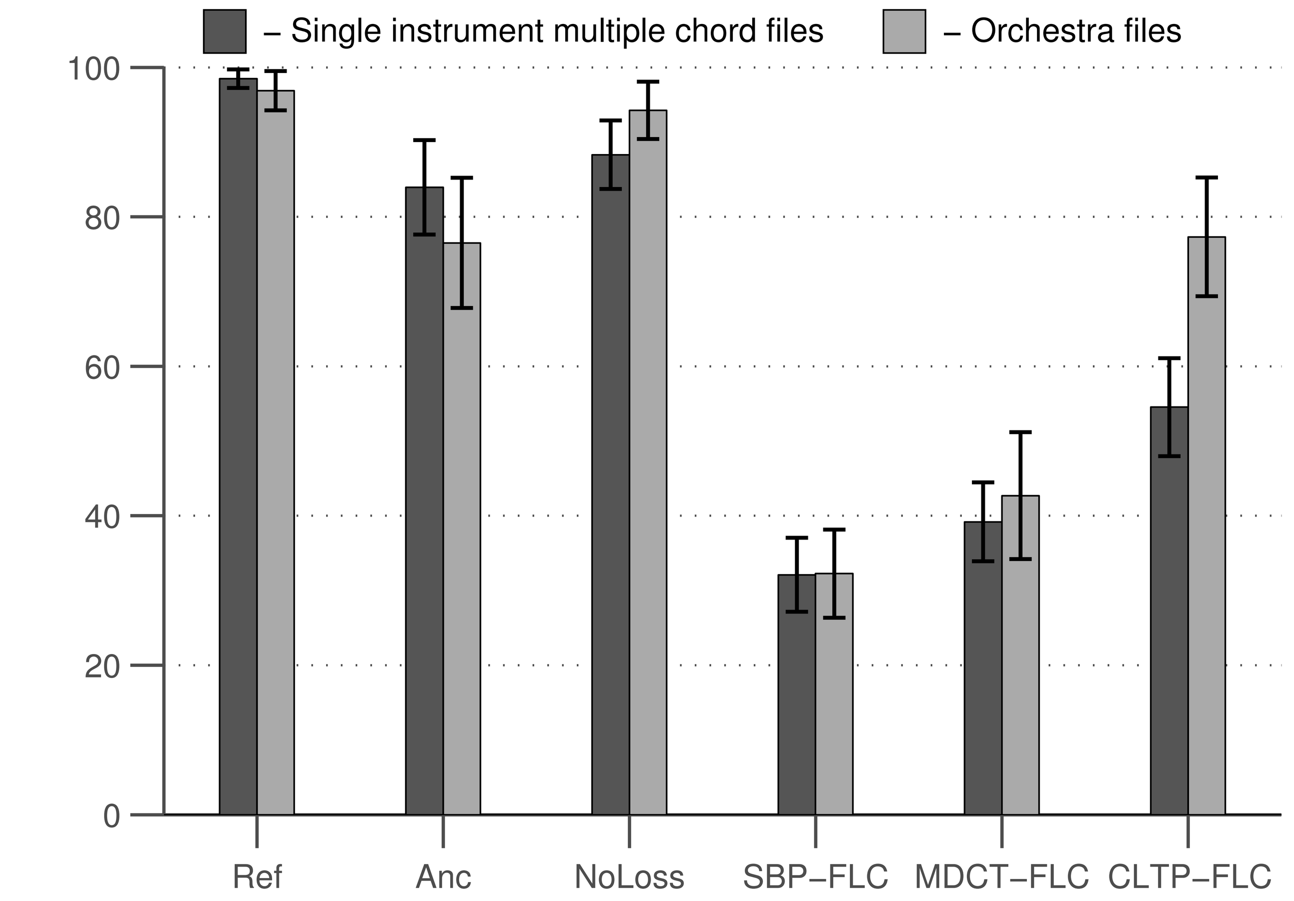


Evaluations

- MPEG reference AAC-LD encoder used to generate 64 kbps bitstreams and the following decoders compared,
  - Reference decoder with no frame loss.
  - Reference decoder with subband domain linear prediction based FLC (SBP-FLC).
  - Reference decoder with MDCT domain tonal interpolation FLC (MDCT-FLC).
  - Reference decoder with the proposed CLTP based FLC (CLTP-FLC).
- Testing data-set: 6 audio files, 4s each, mono, 44.1/48 kHz.
- Frame loss was at the rate of 10% and random.
- Objective evaluation results of Segmental SNR in dB.

Filename	SBP-FLC	MDCT-FLC	CLTP-FLC
Piano	-3.16	-0.67	5.10
Guitar	-1.95	0.19	7.15
Harp	-3.59	-1.77	3.80
Bells	-2.08	0.06	4.26
Mfv	2.27	0.34	11.53
Mozart	-2.03	1.22	8.4
Average	-1.76	-0.11	6.71 (+6.82)

- Subjective evaluation results of MUSHRA listening tests (16 listeners, plots with average and 95% confidence interval).



Conclusions

- Currently used FLC techniques sub-optimal for polyphonic audio signals.
- Bidirectional cascaded LTP proposed for significantly improved FLC, which takes into account all the available information.
- Subjective and objective evaluations substantiate these improvements.
- Future directions include developing low complexity variant and handling burst frame losses.

http://www.scl.ece.ucsb.edu/

{tejaswi, rose}@ece.ucsb.edu