Packet Loss Concealment for Audio Streaming Based on the GAPES Algorithm

Hadas Ofir, David Malah

Department of Electrical Engineering, Technion IIT, Haifa, Israel

Signal and Image Processing Lab

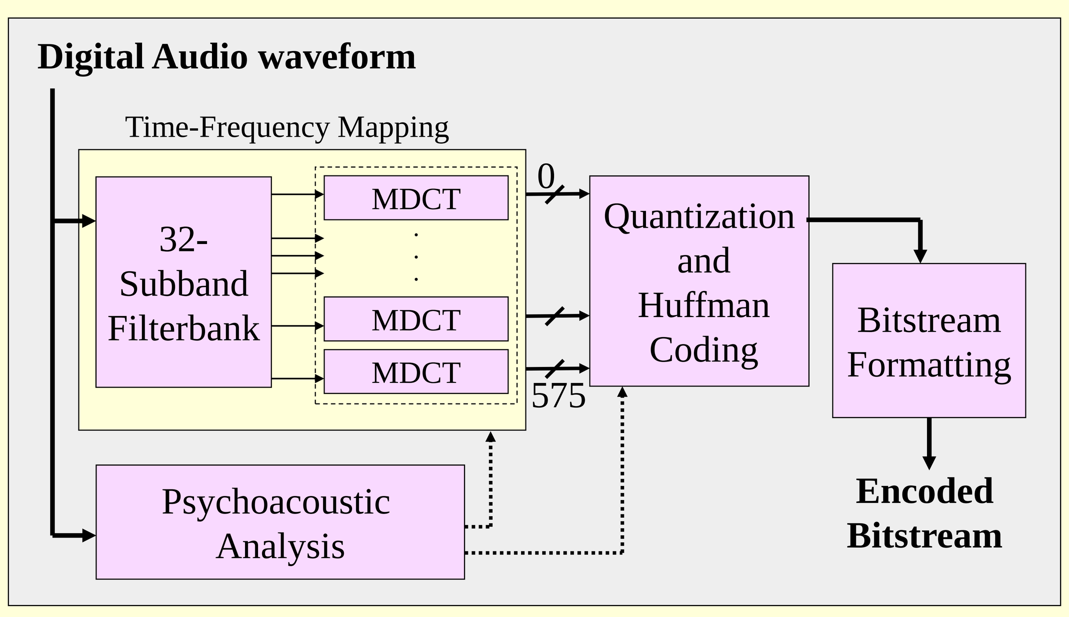
Introduction

- This work introduces a packet loss concealment algorithm for audio streaming applications that use MPEG-audio coders.
- The proposed algorithm interpolates the gap created by the lost packets, using the GAPES algorithm in the DSTFT domain, based only on the data available at the receiver's side.
- The proposed algorithm was implemented on an MPEG-1 Layer III Audio coder (a.k.a., MP3), but can easily be adapted also for MPEG-2/4 AAC.

MPEG-Audio coding

- MPEG-Audio coders use quantization in the **MDCT** domain: Specifically in the MP3 standard, every segment of 576 time-samples is mapped into 576 MDCT coefficients.
- Two segments, of 576 samples each, form a single **MP3 packet**.

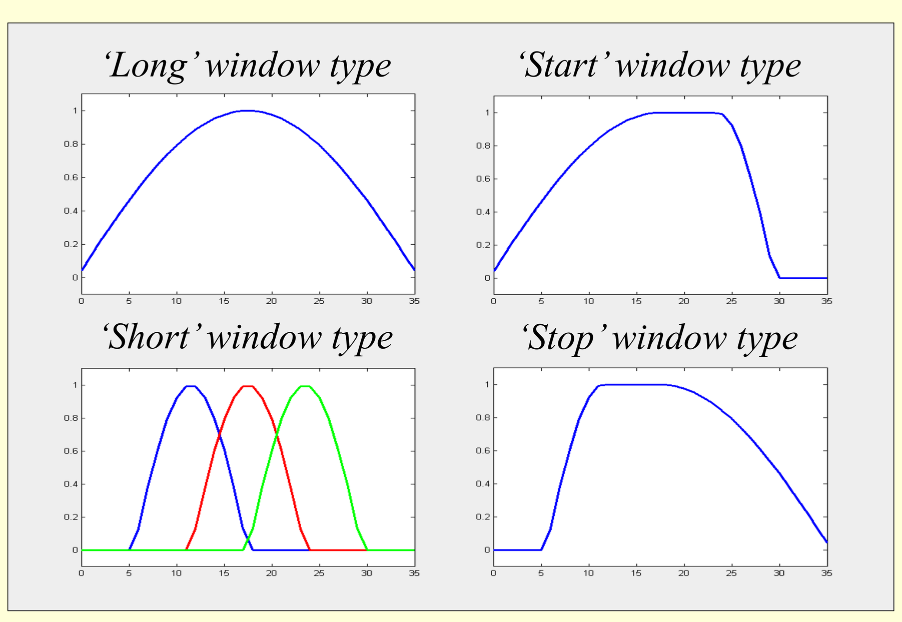
One lost MP3 packet (1152 samples) is equivalent to only 2 lost MDCT coefficients per frequency bin



A Block Diagram of the MP3 Encoder

MDCT

- The MDCT is a real-valued transform, turning 2*N* samples into *N* coefficients. The transform is lossless if certain conditions are satisfied, such as:
  - 50% overlap between consecutive windows.
  - An **overlap & add** (OLA) procedure is applied on the output of the Inv-MDCT.
- The MP3 standard defines 4 possible window functions to be used when applying the MDCT, according to the signal's local characteristics: **'Long'**, **'Short'**, **'Start'** and **'Stop'**.

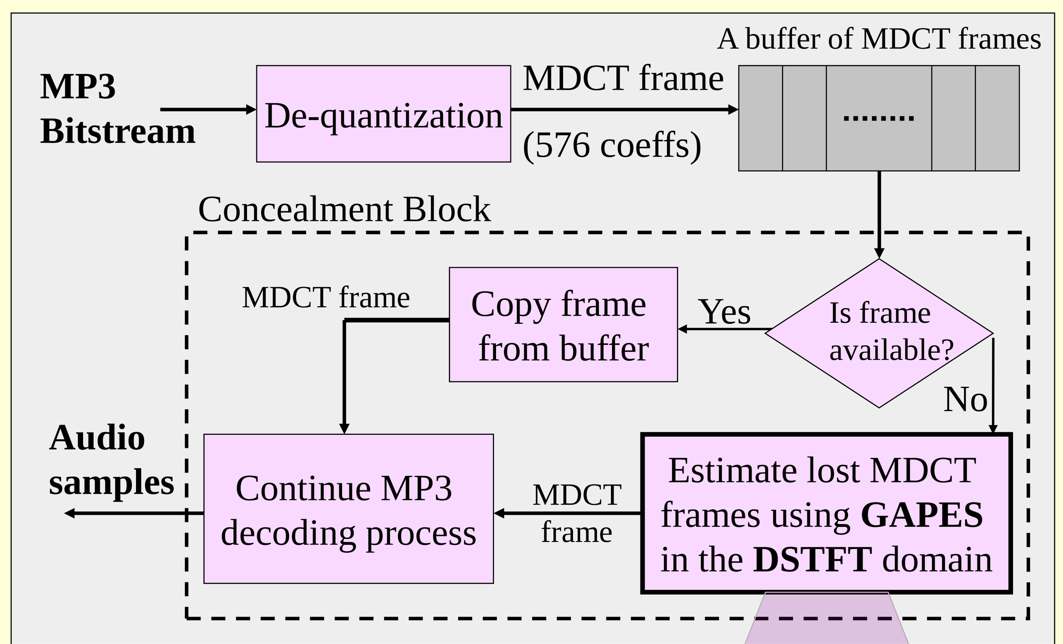


Limitations Of Previous Works

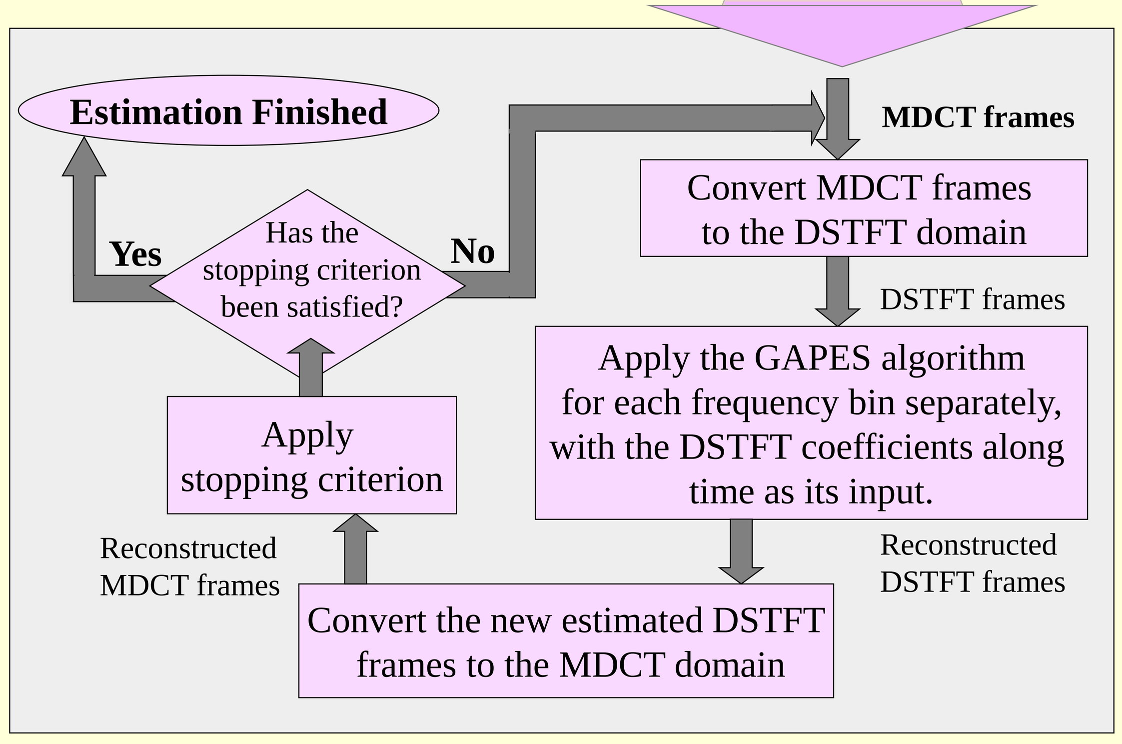
- **Packet Repetition** (MP3 standard): A naive method, data-independent.
- **Statistical Interpolation** (Quackenbush & Driessen, 2003) restores the lost data in the MDCT domain:
  - The MDCT coefficients typically show rapid sign changes along time.
  - Need special treatment in cases of short window type (due to different frequency resolutions).
  - The restoration algorithm assumes the signal is auto-regressive, and it can't support random loss patterns.

Proposed Concealment Method

- An iterative algorithm which reconstructs the MDCT coefficients of the lost packets by estimating the corresponding DSTFT coefficients.
- Working in the DSTFT domain serves two goals:
  - The phase and envelope of the coefficients along time are more consistent.
  - Using a single window type for the DSTFT assures constant frequency resolution.



The proposed decoding process, including a concealment block



A block diagram of the estimation process

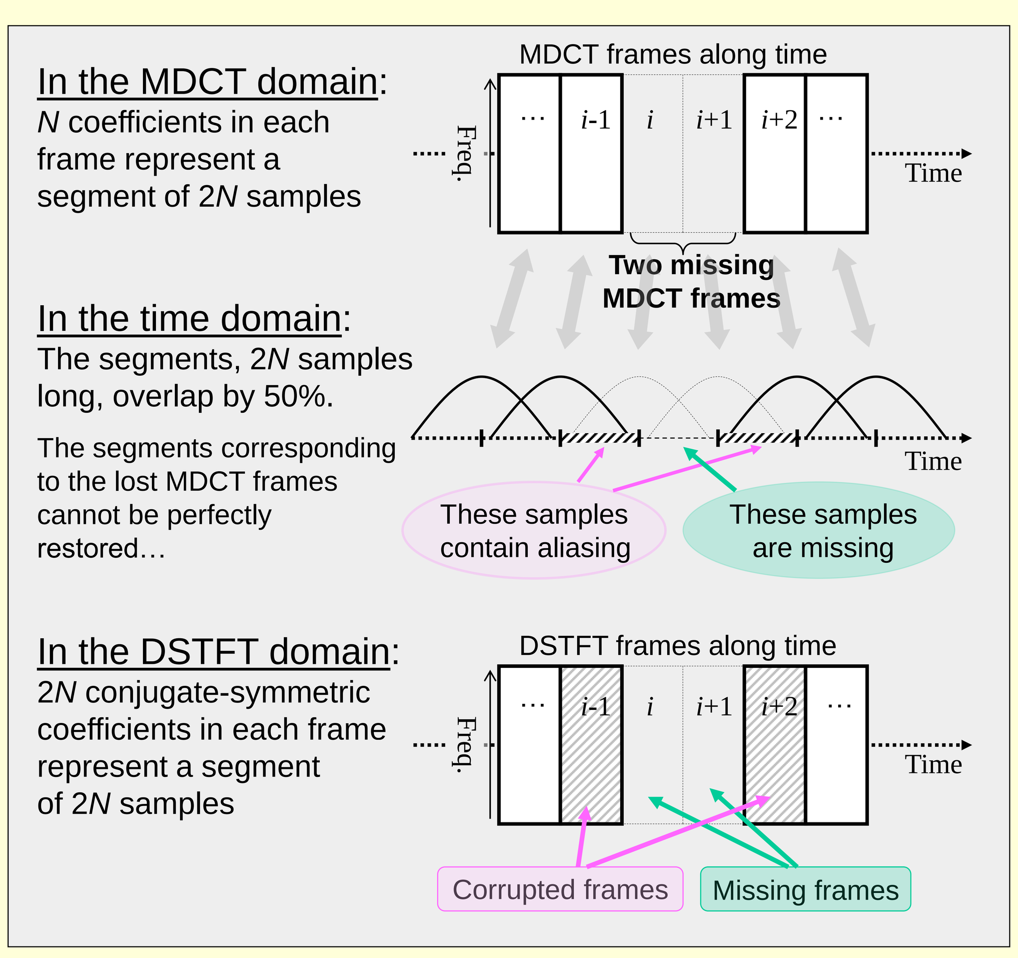
MDCT ↔ DSTFT Conversion

- There is a single expression for each conversion direction. Conversion of one frame requires its two closest neighbours, because of the overlap & add.
- For example, MDCT ➔ DSTFT Conversion:

$$X_n^{DSTFT}[m] = \sum_{k=0}^{N-1} X_n^{MDCT}[k]\cdot\left(g_d^1[m,k] + (-1)^m \cdot g_r^2[m,k]\right) + \sum_{k=0}^{N-1} X_{n-1}^{MDCT}[k]\cdot g_d^2[m,k] + \sum_{k=0}^{N-1} X_{n+1}^{MDCT}[k]\cdot\left((-1)^m \cdot g_r^1[m,k]\right), \quad 0 \le m \le N$$

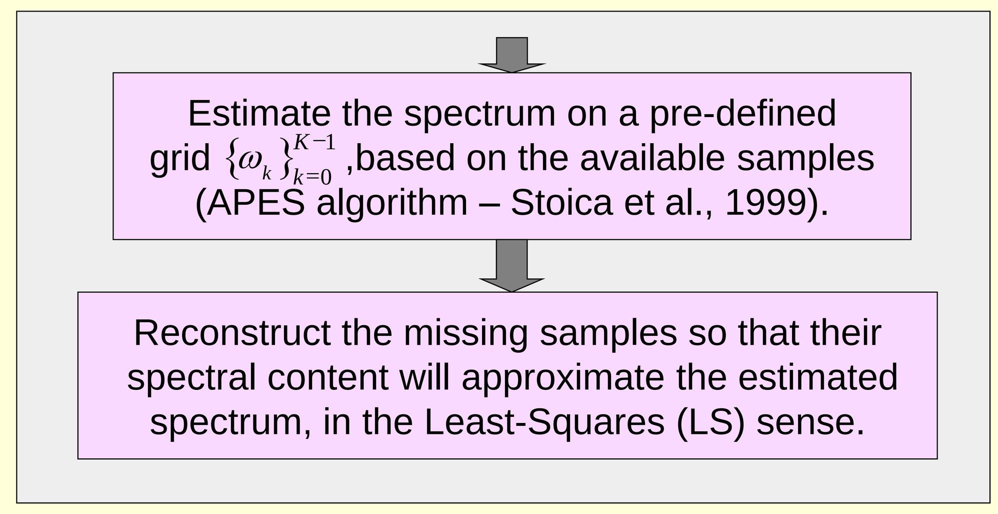
- The $g_{d/r}^1[m,k], g_{d/r}^2[m,k]$ functions are selected according to the window types of the frames X_{n-1}, X_n and X_{n+1}. There are in total 12 such possible functions which can be calculated off-line.

L lost MDCT frames correspond to L missing DSTFT frames and 2 corrupted frames



The GAPES algorithm

- The **GAPES** algorithm (Stoica & Larsson, 2000) reconstructs missing data, assuming it has the same spectral content as the available data.



The GAPES algorithm

- The **missing samples**, $\underline{x}_m$, are restored by solving the following minimization problem:

$$\min_{\underline{x}_m, \{\hat{\alpha}(\omega_k), \underline{h}_k\}} \sum_{k=0}^{K-1} \sum_{t=0}^{L-1} \left| \underline{h}_k^* \underline{x}(t) - \hat{\alpha}(\omega_k) \cdot e^{j\omega_k t} \right|^2$$

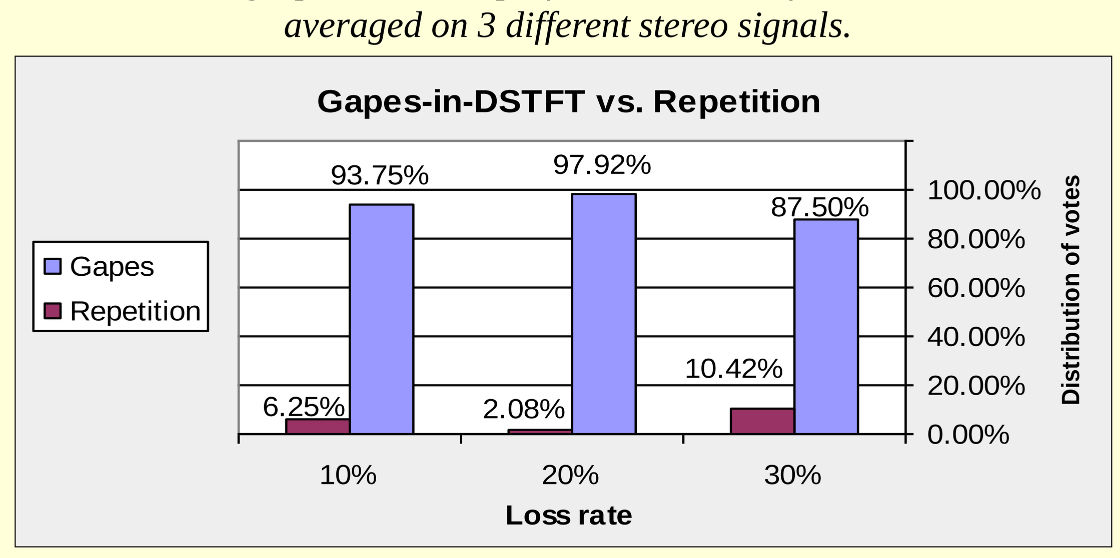
  Where for each frequency ω_k, $\hat{\alpha}(\omega_k)$ is the spectral component at this frequency, and $\underline{h}_k$ is a narrow band filter designed to pass only this frequency.
- The solution is obtained by an iterative algorithm that contains two steps:
  - Minimize with respect to $\{\hat{\alpha}(\omega_k), \underline{h}_k\}$.
  - Then, minimize with respect to $\underline{x}_m$.

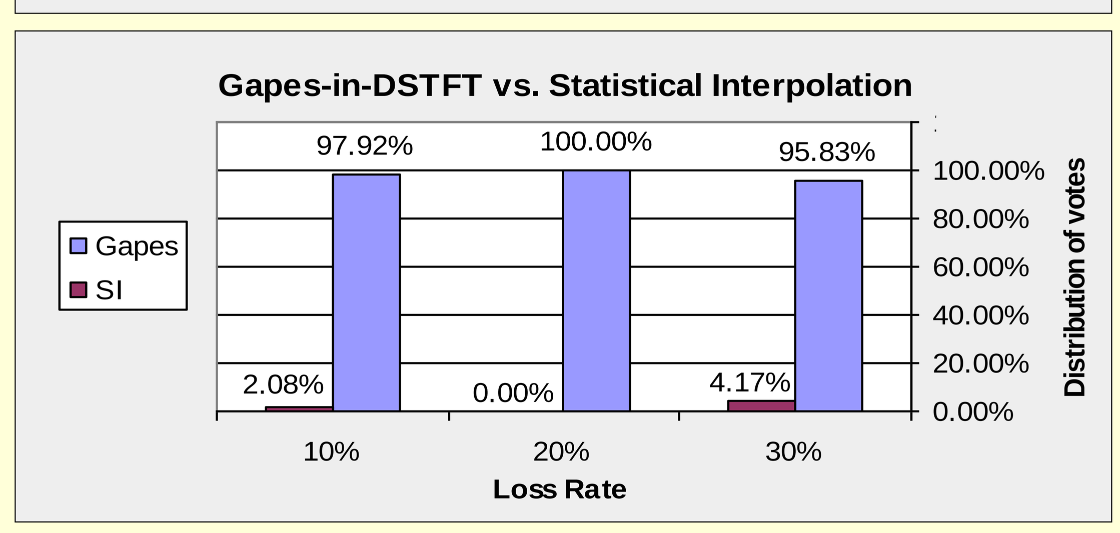
  In our algorithm a single GAPES iteration is activated in each iteration of the algorithm.
- In the first iteration missing data is not considered for the spectral estimation. In the next iterations the previously estimated data is used as part of the available data for the re-estimation.

Results

Comparative subjective test results of :
GAPES-in-DSTFT vs. Previous works

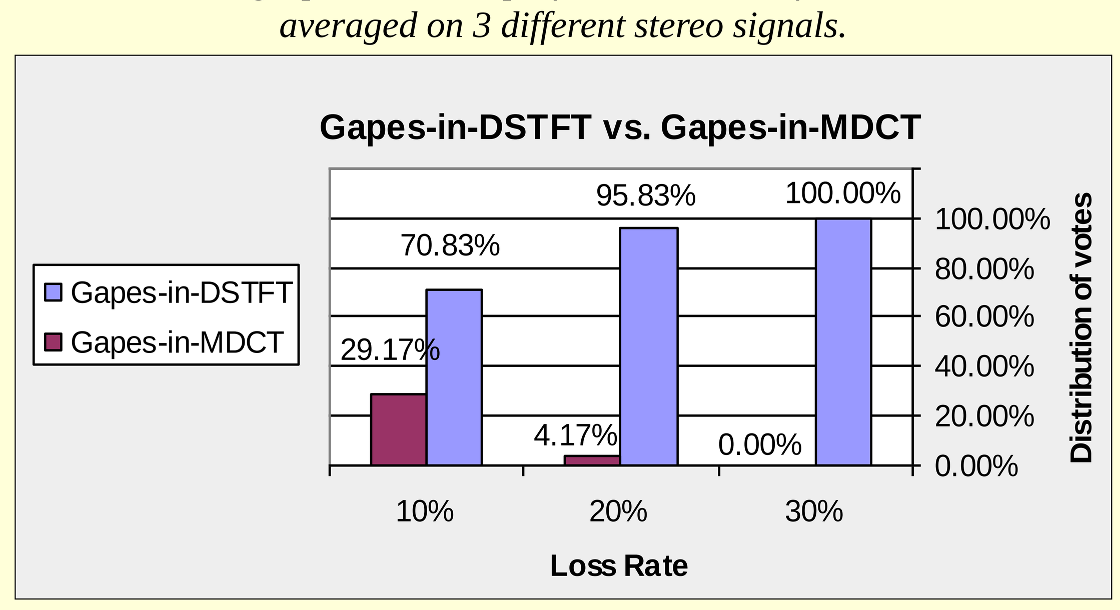
The graph shows the preference results of 16 voters, averaged on 3 different stereo signals.





Comparative subjective test results of :
GAPES-in-DSTFT vs. GAPES-in-MDCT

The graph shows the preference results of 8 voters, averaged on 3 different stereo signals.



☑ Stereo files used are 15-17 seconds long, sampled at 44.1 kHz coded by the LAME MP3 encoder at 128 kbps per channel. Packet loss patterns were random.