Packet Loss Concealment for Audio Streaming based on GAPES and MAPES Algorithms

Hadas Ofir and David Malah

15-17 Nov. 2006 IEEE Conference – Eilat, Israel
Presentation Outline

- Introduction to Packet Loss
- Previous Works
- MPEG-audio compression
- Proposed concealment algorithm
  - Concealment domain
  - GAPES and MAPES interpolation algorithms
- Subjective tests results
- Conclusion
Internet Audio Streaming

- A Real-Time application.
- Connectionless protocol:
  Each packet may use a different route.

Audio signal frames are compressed into data packets

The packets are consecutively sent over the internet

The receiving packets are reassembled, decompressed and played
Packet Loss

- Internet broadcasting doesn’t assure quality of service (QoS).
  - Data packets are often delayed or discarded during network congestions.

- Each loss, unless concealed, produces an annoying disturbance.
  - Example: Original 10% loss 30% loss
    - 30% Repaired
Packet Loss Concealment

- Fill-in the gap with an approximation to the original signal.

- **Goal**: Generate a good enough replacement.
  - Good enough = won’t be noticed by a human listener.

- **Problem**
  - A typical audio packet takes around 1000 samples.
  - Even a single lost packet creates a very wide gap that is difficult to interpolate.

- So, what is the best concealment method?...
Previous works

- Receiver-based techniques:
  - Waveform substitution (off-line)
  - Packet repetition
  - Statistical Interpolation
  - GAPES in the DSTFT domain
MPEG-Audio Coder

- Perceptual audio coders.
- MP3: Every two frames form an **MP3 packet**.

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

- \(2N\) time samples \(\rightarrow N\) MDCT real-valued coefficients
- **Lossless** transform if windows overlap
- Reconstruction using overlap & add

- **MP3** defines 4 window functions for the MDCT.
  - **Long**: better frequency resolution for stationary segments.
  - **Short**: better time resolution for transients.
Concealment in MDCT domain

Q consecutive missing packets are equal to...

- **Time domain**: \((2Q+1) \cdot 576\) missing samples.
- **MDCT domain**: 2Q missing coeff. per frequency bin.

However...

- MDCT coefficients along time show rapid sign changes.
  
  **Solution**: Use a domain with less signal fluctuations.

- Different windows \(\rightarrow\) different frequency resolutions.
  
  **Solution**: Use a single window type when converting the data to another domain.

**Our choice**: The DSTFT domain
Example: Solution in MDCT domain

- **Statistical Interpolation (SI)**  
  
  [Quackenbush and Driessen, 115th AES conv., Oct. 2003]

- Each frequency bin along time is reconstructed separately.

- **Benefits:** Applied directly in the compressed domain.

- **Limitations:** Limited loss patterns, high complexity, assumes parametric model.
MDCT ↔ DSTFT Conversion

Issues:
- FFT cannot be used in conversion.
- 4 MDCT windows → 12 different conversion expressions.

Solution
A single expression for each conversion direction.

For example, MDCT → DSTFT Conversion:

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

\( g_d^1 [m,k], g_d^2 [m,k] \) (calculated off-line) are selected according to window types of frames \( X_{n-1}, X_n, X_{n+1} \).

Efficient conversion: operations number reduced by factor 3.
# APES-based Interpolation Algorithms

- **APES**: Amplitude and Phase Estimation (Stoica & Li, 1999).
  - An algorithm for spectral estimation.

- **GAPES**: Gapped-data APES (Stoica & Larsson, 2000).
  - Uses an adaptive filter-bank approach.

- **MAPES**: Missing-data APES (Stoica & Wang, 2005).
  - Uses an ML-estimator approach.
  - MAPES has lower complexity and can handle more loss patterns.

## Comparison to SI

<table>
<thead>
<tr>
<th>Benefits</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Can handle many loss patterns.</td>
<td>Higher complexity</td>
</tr>
<tr>
<td>Doesn’t assume parametric modeling.</td>
<td></td>
</tr>
<tr>
<td>Can be applied on complex signals.</td>
<td></td>
</tr>
</tbody>
</table>
The APES Algorithm

Let \( \{x_n\} \) be a data-sequence of length \( P \).

**Problem**

Estimate spectral component at frequency \( \omega_0 : \alpha(\omega_0) \).

**Solution**

- \( x_n \) is modeled as: \( x_n = \alpha(\omega_0) \cdot e^{j\omega_0 n} + e_n(\omega \neq \omega_0) \)
- Design a narrow-band filter \( h(\omega_0) \), of length \( M \):
  - The filter should pass the frequency \( \omega_0 \) without distortion.
  - The filter should attenuate all the other frequencies.
- By filtering \( \{x_n\} \) with the filter \( h(\omega_0) \) we get:
  \[
  h(\omega_0) \ast x_n \approx \alpha(\omega_0) \cdot e^{j\omega_0 n}
  \]
- Use DFT on the filtered data to estimate \( \alpha(\omega_0) \).
Description as a minimization problem:

\[ \min \quad \left\| \omega \right\| \quad \text{subject to} \quad h(\omega_0) \cdot a(\omega_0) = 1 \]

Where:

- \( h(\omega_0) \in \mathbb{C}^{M \times 1} \) is a data-dependent narrow-band filter, centered at \( \omega_0 \).
- \( a(\omega) \triangleq [1, e^{j\omega}, \ldots, e^{j(M-1)\omega}]^T \in \mathbb{C}^{M \times 1} \) is a vector of exponents.
The GAPES Algorithm

- APES is expanded to a frequency grid: \( \{ \omega_k \} \), \( 0 \leq k \leq K \).

- The **missing samples**, \( \{ x_m \} \), are restored by solving the following minimization problem:

\[
\min_{\{ x_m \}, \{ \alpha_k, h_k \}_{k=0}^{K-1}} \sum_{k=0}^{K-1} \sum_{l=0}^{P-M} \left| h_k^H \cdot x_l - \alpha_k \cdot e^{j\omega_k l} \right|^2, \text{ subject to } h_k^H \cdot a_k = 1
\]

- \( \alpha_k \triangleq \alpha(\omega_k) \) is the spectral component at frequency \( \omega_k \).
- \( h_k \triangleq h(\omega_k) \) is a data-dependent narrowband filter, centered at \( \omega_k \).
- \( a_k \triangleq a(\omega_k) \)

- Iterative algorithm. A single iteration:

**Estimate the spectrum** using the APES algorithm. (lost samples are initiated to zero)

**Reconstruct missing samples** to match the estimated spectrum, in the Least-Squares sense.
The APES Algorithm – Different Approach

- APES has also an ML-estimator interpretation.

- Assuming $\{e_l(\omega_0)\}$ are statistically-independent, zero-mean complex Gaussian random vectors, with unknown covariance matrix: $Q(\omega_0)$.
  - APES only approximates an ML-estimator since the vectors contain overlapping data!

- Under these assumptions, the ML-estimator:
The MAPES Algorithm

- The **missing samples**, \( \{x_m\} \), are restored by solving the following maximization problem:

\[
\max_{\{x_m\}, \{\alpha_k, Q_k\}_{k=0}^{K-1}} \sum_{k=0}^{K-1} \log(Pr\{\{x_l\}_{l=0}^{P-M} \mid \alpha_k, Q_k\})
\]

Where \( Q_k \) is the covariance-matrix of the \( \{e_l(\omega_k)\} \) vectors.

- Solved using an iterative algorithm that contains two steps:
  - Solve with respect to \( \{\alpha_k, Q_k\} \) by applying APES.
  - Solve with respect to the missing samples.

- Lower complexity due to simple calculation process.
Proposed Concealment Algorithm

A block diagram of the reconstruction process

1. Convert MDCT frames to the DSTFT domain
2. Apply the GAPES/MAPES algorithm for each frequency bin separately, with the DSTFT coefficients along time as its input.
3. Convert the new estimated DSTFT frames to the MDCT domain.
4. Apply stopping criterion.
5. If stopping criterion satisfied, Estimation Finished.
6. If stopping criterion not satisfied, repeat from Step 2.
Results

- **Perceptual quality evaluation**
  - Proposed solution (GAPES/MAPES) sounds better than previous works.
  - GAPES has small advantage over MAPES.

- **Complexity**
  - Proposed solution (GAPES/MAPES) is more complex than previous works.
  - MAPES has lower complexity than GAPES.

  **Example:**
  - For 4 missing frames in a buffer of 16, GAPES requires twice the number of multiplications needed by MAPES.
Audio Examples

Example 1:

Piano Original
No concealment
(30% random loss)

SI
Repetition
Proposed Algorithm
(GAPES)
Proposed Algorithm
(MAPES)

Example 2:

Flute Original
No concealment
(20% random loss)

MAPES
GAPES
Repetition
SI
A new algorithm for packet loss concealment.
- For audio streaming, encoded by MPEG-audio coders.
- By applying GAPES or MAPES in the DSTFT domain.

A direct conversion scheme was introduced: MDCT $\leftrightarrow$ DSTFT.
- Enables efficient conversion between domains.

Proposed algorithm performs better than packet repetition and statistical interpolation.

MAPES is more robust than GAPES and needs lower complexity.
Thank You

Contact Information:
Hadas Ofir
hadaso@tx.technion.ac.il