Technion - Israel Institute of Technology
Department of Electrical Engineering
**Signal and Image Processing Laboratory**

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

Hadas Ofir and David Malah

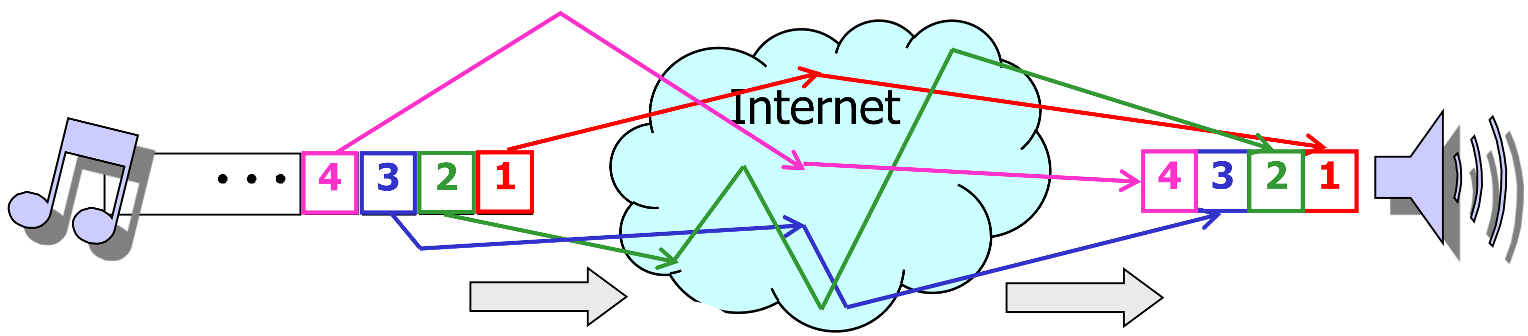
15-17 Nov. 2006 IEEE Conference – Eilat, Israel

# 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.



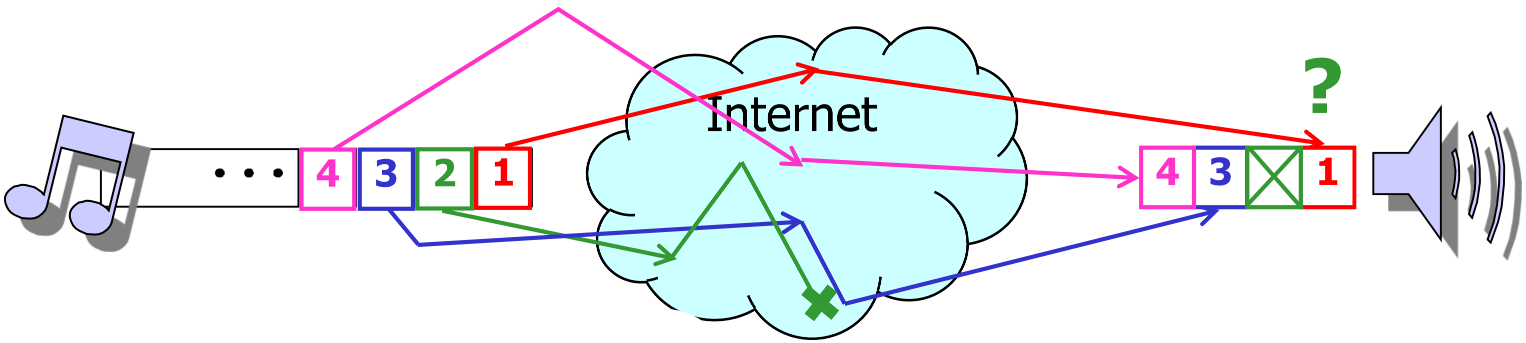

Audio signal frames are compressed into data packets

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.

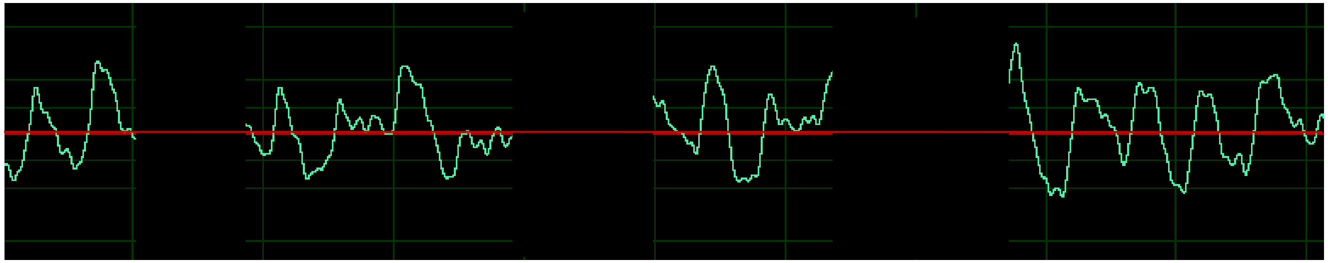



- Each loss, unless concealed, produces an annoying disturbance.
  - <u>Example</u>: Original 10% loss 30% loss

    30% Repaired

# Packet Loss Concealment

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



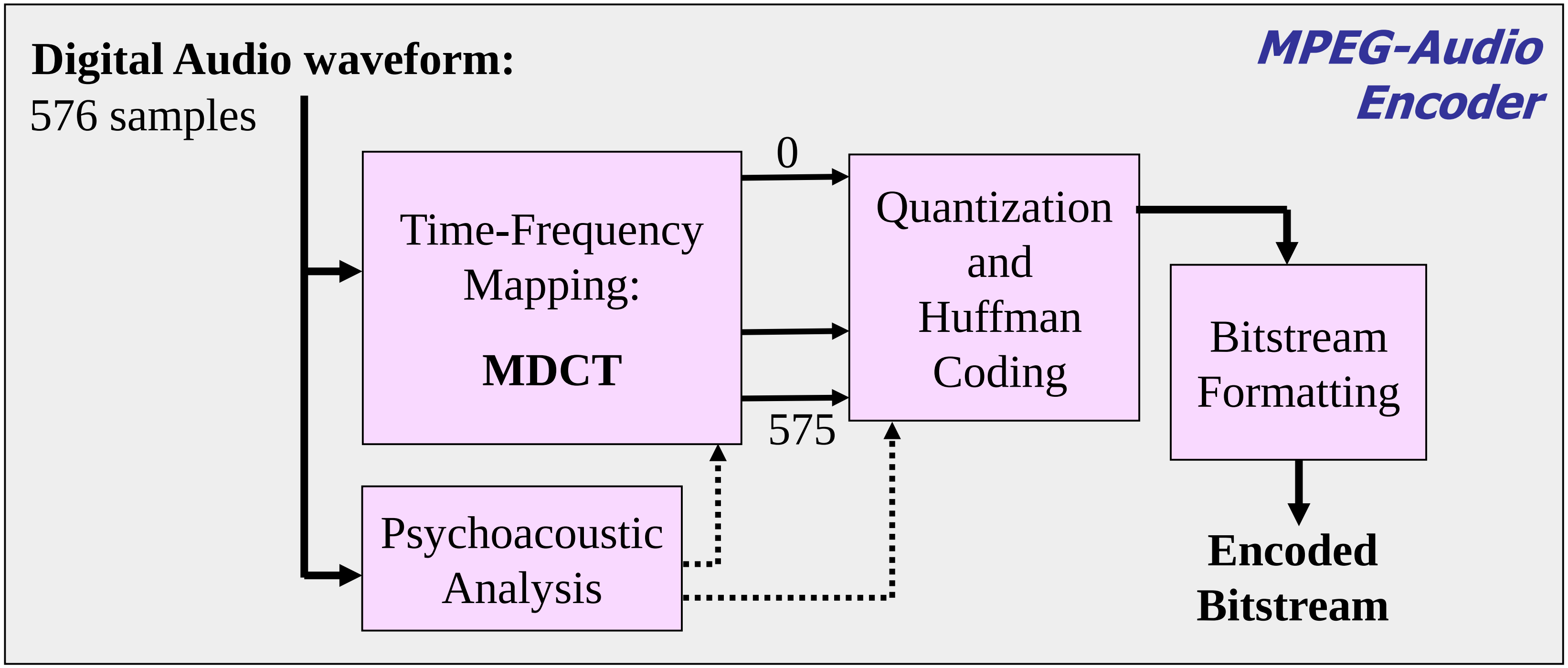
- <u>Goal</u>: Generate a good enough replacement.
  - Good enough = won't be noticed by a human listener.
- <u>Problem</u>
  - A typical audio packet takes around 1000 samples.
  - Even a single lost packet creates a very wide gap that is difficult to interpolate.
- So, what is the best concealment 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**.

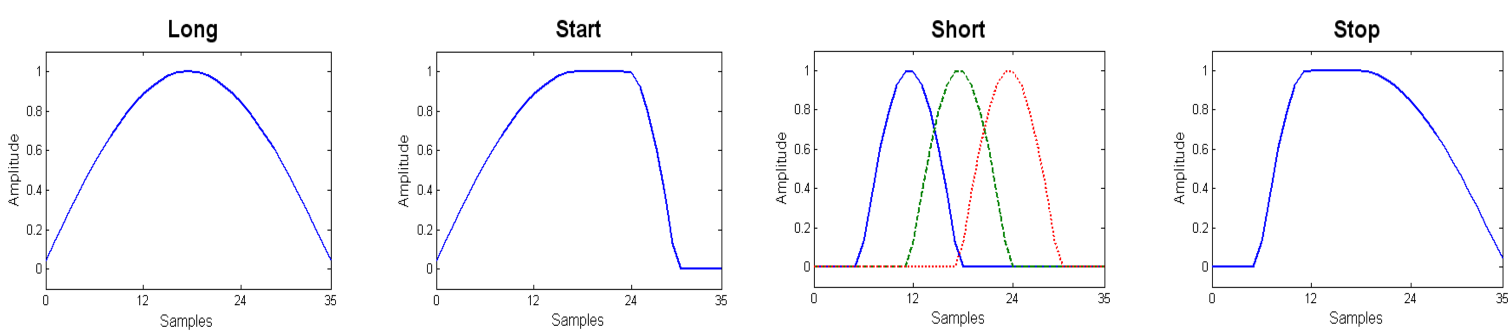



One lost MP3 packet (1152 samples) is equivalent to **only** 2 lost MDCT coefficients per frequency bin!

# MDCT

- $2N$ time samples ➔ $N$ 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)·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.

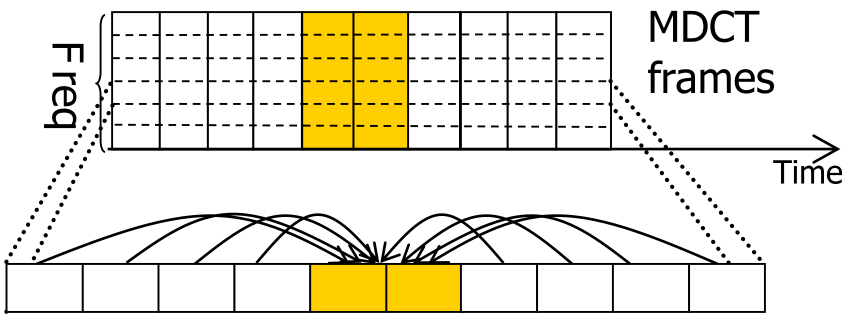
- Different windows ➔ different frequency resolutions.

Solution: Use a single window type when converting the data to another domain.

**Our choice**: The DSTFT domain

# Example: Solution in MDCT domain

- <u>Statistical Interpolation (SI)</u>
  [Quackenbush and Driessen, 115th AES conv., Oct. 2003]
  - Each frequency bin along time is reconstructed separately.




  - <u>Benefits</u>: Applied directly in the compressed domain.
  - <u>Limitations</u>: 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:

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

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

- $g_{d/r}^1[m,k], g_{d/r}^2[m,k]$ (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**

Benefits

- Can handle many loss patterns.
- Doesn't assume parametric modeling.
- Can be applied on complex signals.

Limitations

- Higher complexity

# The APES Algorithm

- Let $\{x_n\}$ be a data-sequence of length $\mathrm{P}$.

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 $\underline{h}(\omega_0)$, of length $\mathrm{M}$:
  - 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 $\underline{h}(\omega_0)$ we get:

$$h(\omega_0) * x_n \simeq \alpha(\omega_0) \cdot e^{j\omega_0 n}$$

- Use DFT on the filtered data to estimate $\alpha(\omega_0)$.

# The APES Algorithm – Cont.

- Description as a minimization problem:

$\cdot \underline{a}(\omega_0) = 1$

Where:

$\underline{h}(\omega_0) \in \mathbb{C}^{\mathrm{Mx1}}$ is a data-dependent narrow-band filter, centered at $\omega_0$.

$\underline{a}(\omega) \triangleq [1, e^{j\omega}, \ldots, e^{j\omega(\mathrm{M}-1)}]^{\mathrm{T}} \in \mathbb{C}^{\mathrm{Mx1}}$ is a vector of exponents.

# The GAPES Algorithm

- APES is expanded to a frequency grid: $\{\omega_k\}, 0 \le k \le \mathrm{K}$.
- The **missing samples**, $\{x_m\}$, are restored by solving the following minimization problem:

$$\min_{\{x_m\},\{\alpha_k,\underline{h}_k\}_{k=0}^{K-1}} \sum_{k=0}^{K-1}\sum_{l=0}^{P-M}\left|\underline{h}_k^H\cdot\underline{x}_l-\alpha_k\cdot e^{j\omega_k l}\right|^2, \text{ subject to } \underline{h}_k^H\cdot\underline{a}_k=1$$

$\alpha_k \triangleq \underline{\alpha}(\omega_k)$ is the spectral component at frequency $\omega_k$.
$\underline{h}_k \triangleq \underline{h}(\omega_k)$ is a data-dependent narrow-band filter, centered at $\omega_k$.
$\underline{a}_k \triangleq \underline{a}(\omega_k)$

- 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 $\{\underline{e}_l(\omega_0)\}$ are statistically-independent, zero-mean complex Gaussian random vectors, with unknown covariance matrix: $\mathbf{Q}(\omega_0)$.
  - APES only <u>approximates</u> an ML- estimator since the vectors contain overlapping data !
- Under these assumptions, the ML- estimator:

# The MAPES Algorithm

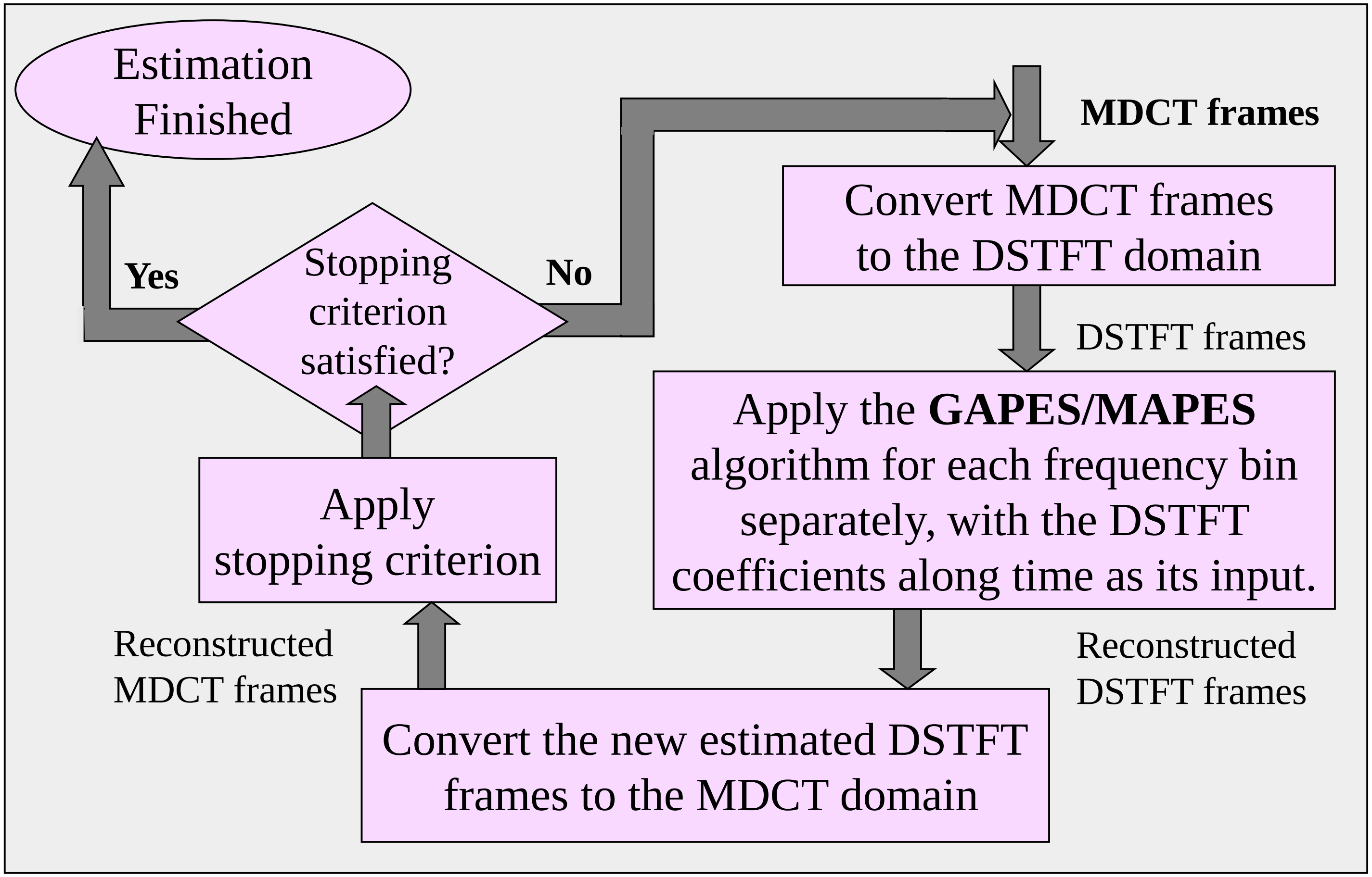
- The **missing samples**, $\{x_m\}$, are restored by solving the following maximization problem:

$$\max_{\{x_m\},\{\alpha_k,\mathbf{Q}_k\}_{k=0}^{K-1}} \sum_{k=0}^{K-1} \log\left(Pr\left\{\left\{\underline{x}_l\right\}_{l=0}^{P-M} \mid \alpha_k, \mathbf{Q}_k\right\}\right)$$

  Where $\mathbf{Q}_k$ is the covariance-matrix of the $\{\underline{e}_l(\omega_k)\}$ vectors.

- Solved using an iterative algorithm that contains two steps:
  - Solve with respect to $\{\alpha_k , \mathbf{Q}_k\}$ 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*

# 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

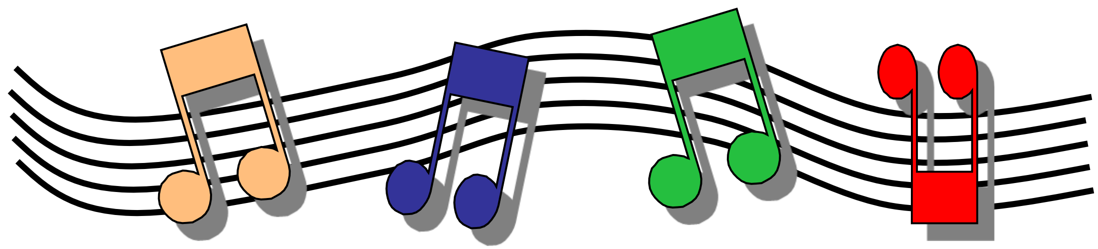
SI

Proposed Algorithm (GAPES)

No concealment (30% random loss)

Repetition

Proposed Algorithm (MAPES)



## Example 2:

Flute Original

MAPES

Repetition

No concealment (20% random loss)

GAPES

SI

# Conclusion

- 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 $\rightleftarrows$ 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