AMISP: A Complete Content-Based MPEG-2 Error-Resilient Scheme

Pascal Frossard and Olivier Verscheure

Presented by
Chen-hsiu Huang

January 17, 2002
Abstract

- This paper addresses a new error-resilient scheme for broadcast quality MPEG-2 video streams to be transmitted over lossy packet networks.
- A new scene complexity adaptive mechanism, namely Adaptive MPEG-2 Information Structuring (AMIS) is introduced.
- End-to-end video quality = encoding quality + degradation due to data loss
MPEG-2 Sensitive to Data Loss

• An MPEG-2 video stream is highly hierarchically structured. Like any other compressed data, compressed video is highly sensitive to data loss.

• Temporal Propagation: when loss occurs in a reference picture (I or P picture), the lost macroblocks will affect the predicted macroblocks in subsequent frame(s).

• Spatial Propagation: in case of transmission error, an MPEG-2 decoder skips all video information up to the next slice header.
Data loss propagation in MPEG-2 video streams

Error-resilience Schemes

- Error-concealment techniques try to estimate missing video data using information available at the receiver.
- Resynchronization/error-localization: limiting spatial/temporal error propagation
- Unequal error-protection scheme try to efficiently recovery the missing video information.
Error Concealment

- The error concealment algorithms include, for example, spatial interpolation, temporal interpolation and early resynchronization.

- The MPEG-2 standard proposes an elementary error concealment algorithm based on motion compensation. This simple technique is certainly not optimal, but it offers a satisfying decoding quality.

- Some MPEG-2 encoder chips automatically produce concealment motion vectors for all macroblocks.
Motion Compensation

- The motion compensation concealment estimates the motion vectors of the lost macroblock by using the motion vectors of neighboring macroblocks in the affected picture.
- The motion compensation concealment estimates the motion vectors of the lost macroblock by using the motion vectors of neighboring macroblocks in the affected picture.
- However, an obvious problem with lost macroblocks whose neighbors are intra-coded. To get around this problem, the encoding can also include motion vectors for intra macroblocks.
The robustness of compressed MPEG-2 video may be dramatically increased by judiciously inserting resynchronization points in the bit stream. These can be obtained by extra slice headers to limit spatial propagation and intra-coded macroblocks to stop temporal propagation.

However, the addition of extra slice headers and/or intra-coded macroblocks is not costless.
Resynchronization (2/2)

Defect:

1. The larger the number of slices, the bigger the overhead. The larger the number of intra-coded macroblocks, the higher the overhead.

2. It resets the differential coding of the DC values and motion vectors.

3. Under the same video traffic constraints, extra resynchronization points therefore reduce the amount of bits available to code pure video information, and thus decrease the quality of the reconstructed video.
Problem Formulation

The problem addressed in this paper consists in finding the optimal tradeoff between video information and error protection.

- A random insertion of extra resynchronization points in the bit stream or regular forward-error-correction (FEC) packets is not optimal – the efficiency strongly depends on the content of the corresponding video area.

- There is no need to insert resynchronization points where the impact of data loss would not affect the video quality (under a given error-concealment technique).

- Moreover, the protection level has also to be adapted to the network performances or the expected loss probabilities.
Problem Considered

Given the expected PLR or dynamically measured through RTCP (Real-Time Control Protocol):

- MPEG-2 Structuring
  Determine the most appropriate MPEG-2 structure in terms of resynchronization points location.

- MPEG-2 Protection
  Derive a content-based FEC scheme to protect areas where structuring is not sufficient.
A macroblock may be damaged in any of the three following cases:

- It belongs to an RTP packet that has been lost during transmission
- It belongs to a slice that has been affected by a packet loss (spatial propagation)
- It is temporally dependent on a damaged macroblock of a previous reference frame (temporal propagation).
Transmission Error

- Without any other information about the packet loss process, every RTP packet has the same average probability to be lost: 
  \[ \theta = PLR \]

- Under the assumption that a macroblock is lost as soon as part of it is missing, the probability \( \lambda_n(i, j) \) for the macroblock to be lost is given by

  \[ \lambda_n(i, j) = \theta N_n(i, j), \]

  \[ \forall (i, j) | 1 \leq i \leq B_{row}, 1 \leq i \leq B_{column} \]

where \( N_n(i, j) \) is the number of RTP packets containing the macroblock \( B_n(i, j) \)
Spatial Propagation

- In MPEG-2 decoder, when a macroblock is lost within a slice, all subsequent macroblocks of the same slice are considered as being damaged, even if they do not belong to the lost RTP packet.

- Thus, for a given frame $n$, the probability $P_n(i, j)$ for a macroblock not to be correctly decoded (transmission error spatial propagation) is given by

$$P_n(i, j) = \lambda_n(i, j) + \theta M_n(i, j) = \theta [N_n(i, j) + M_n(i, j)],$$

$$\forall (i, j) | 1 \leq i \leq B_{row}, 1 \leq i \leq B_{column}$$

where $M_n(i, j)$ represents the number of RTP packets within the same slice before the first packet related to $B_n(i, j)$.
• Illustration of $P_n(i, j)$. The numbers in each macroblock represent $M_n + N_n$ values.
**Temporal Propagation (1/4)**

- The granularity of the loss probability matrix must be refined to the pixel level. The loss probability matrix due to spatial propagation for every pixel of frame is called $\mathcal{P}_n$.

$$\mathcal{P}_n = I_{16} \otimes P_n$$

The resulting matrix $\mathcal{P}_n$ has therefore the same size as the video frame.

- The B frames are not considered for additional intra-coded macroblock insertion. Indeed, B frames do not propagate degradation, since they are never referenced. Therefore, the impact of data loss in B frames is barely visible.
Temporarl Propagation (2/4)

• Thus $\varepsilon_{n-k}^n$ obviously depends on $P_m$, with $m = n - k, n - k + 1, \ldots, n$. Moreover, $\varepsilon_{n-k}^n$ needs to be computed in a recursive manner.

• The probability for all the pixels in frame $n$ to be damaged by losses occurring in frame $(n - 1)$. The mapping operation can be denoted by $M_n(P_{n-1})(i,j)$.

• Therefore

$$\varepsilon_{n-1}^n(i,j) = M_n(P_{n-1})(i,j)(1 - P_n(i,j))$$

$\forall(i,j)|1 \leq i \leq W, 1 \leq j \leq H$

where $\varepsilon_{n-1}^n(i,j)$ represents the probability for pixel at $(i,j)$ in frame $n$ to be damaged in reference of frame $n - 1$. 
Temporal Propagation (3/4)

- The probability matrix for pixels not to be spatially lost could be written as

\[ \bar{P}_n = I_{W,H} - P_n \]

- Taking into account losses in any of the \( k \) previous reference frames, with \( k \leq n \). The generic loss probability matrix due to temporal propagation, \( \varepsilon_{n-k}^{n-k} \), can be obtained via recursion.

\[ \varepsilon_{n-k}^{n-k+1}(i,j) = M_{n-k+1}(\varepsilon_{n-k}^{n-k})(i,j)\bar{P}_{n-k+1}(i,j) \]

\[ \forall (i,j)|1 \leq i \leq W, 1 \leq i \leq H \]

with

\[ \varepsilon_{n-k}^{n-k}(i,j) = P_{n-k}(i,j) \]
Temporal Propagation (4/4)

• The process can then be generalized

\[ \varepsilon_{n-k+m}^{n-k}(i, j) = M_{n-k+m}(\varepsilon_{n-k+m-1}^{n-k})(i, j) P_{n-k+m}(i, j) \]

\[ m = 1, 2, 3, ..., k \]

• Finally, \( \varepsilon_{k}^{n-k} \) is obtained when \( m = k \).

• \( \varepsilon_{k}^{n-k} \) represents the generic loss probability matrix for frame \( n \)
due to temporal propagation of data loss in frame \( (n - 1) \).
Adaptive MPEG-2 Information Structuring

- An extra resynchronization point is inserted in the bitstream whenever hypothetical data loss, following a uniform loss process, would lead to video degradation above a desired threshold, after error concealment.

- The mean luminance difference (MLD) has first been chosen as distortion measure.

- The distortion is computed between the current macroblock after encoding (generally given by the encoding scheme) and the same macroblock impaired by loss. This one is obtained by simulating losses and concealment in the encoder.
The MLD for $B(i, j)$ is defined as follows:

$$\delta(i, j) = \left| \frac{1}{256} \sum_{p=1}^{256} B^P(i, j) - \frac{1}{256} \sum_{p=1}^{256} \tilde{B}^P(i, j) \right|$$

where

- $(i, j)$ macroblock position in the frame
- $p$ pixel position in the corresponding macroblock
- $B^P(i, j)$ pixels of the correctly (error-free) decoded macroblock
- $\tilde{B}^P(i, j)$ pixels of the corresponding damaged macroblock at position $(i, j)$
AMISP: A Complete Content-Based MPEG-2 Error-Resilient Scheme

AMIS by Parts

AMIS is divided in two distinct parts:

1. Spatial part: deals with slice headers insertion
2. Temporal part: in charge of deciding when a macroblock should be intra-coded.
The Spatial part of AMIS aims at limiting the spatial error propagation, or at least its visible degradation.

It introduces an extra slice header as soon as the distortion due to hypothetical loss reaches a given threshold, $\Delta_s$.

Clearly a new slice is inserted as soon as

$$\sum_{B_n(i,j) \in S} \delta^s_n(i,j) P_n(i,j) \geq \Delta_s$$

where $B_n(i,j)$ is the current macroblock belonging to slice $S$ and $\delta^s_n(i,j)$ corresponds to the expected MLD in case $B_n(i,j)$ was damaged.
AMIS-Temporal (1/2)

- AMIS-Temporal analyzes every single macroblock and decides whether or not to intra-code it. Again, this decision depends on the macroblock distortion due to temporal propagation of data loss.

- The distortion due to temporal error propagation is weighted by the corresponding loss probability matrix and compared to a threshold $\Delta_t$.

- This weighted distortion is obtained by summing the effects of uniformly distributed packet losses in every single previous reference frame, up to the last intra-coded picture.
Finally, the condition for a macroblock to be intra-coded in frame is given by

\[
\sum_{k=1}^{I} \left( \frac{1}{256} \sum_{p \in B_n(i,j)} \varepsilon_{n-k}^{i,j} \delta_{n,k}^{i,j} \right) \geq \Delta_t
\]

The expected MLD between the current MB correctly decoded and its substitute in case of loss the reference frame \( k \) is given by \( \delta_{n,k}^{t} \).
AMISP: FEC-Based Protection

- Data loss may still include unacceptable degradation in the reconstruction video.
- FEC means that redundancy is added to the data so that the receiver can recover from losses or errors without any further intervention from the sender.
- FEC schemes build $n_{FEC}$ packets blocks where $k_{FEC}$ video packets are protected by means of $n_{FEC} - k_{FEC}$ redundancy packets.
AMISP: A Complete Content-Based MPEG-2 Error-Resilient Scheme

**AMISP: Adaptive Protection Algorithm**

- During the encoding process, a packet is marked to be protected whenever its hypothetical loss would introduce an unacceptable degradation.

- The loss probability weighted distortion is compared to a third threshold $\Delta_{FEC}$

$$
\sum_{B_n(i,j) \in P} \delta_s(i,j) \theta \geq \Delta_{FEC}
$$

- When it decides to protect a packet, it triggers the underlying network adaptation layer (NAL). The NAL starts counting $k_{FEC}$ video packets and then insert $n_{EFC} - k_{EFC}$ EFC packets in the MPEG-2 bit stream.
Conclusion

• A new adaptive error-resilient scheme for TV-resolution MPEG-2 video streams interactive delivery, namely AMISP. It includes a media-dependent FEC algorithm relying on an MPEG-2 syntactic structuring technique.

• A judicious combination of protection redundancy, MPEG syntactic data, and pure video information were shown to greatly improve the final quality under a given bit budget.

• AMISP does not significantly increase the MPEG-2 encoding complexity. However, the protection part of AMISP requires the underlying layer (NAL) to provide FEC capabilities.