Video Error Concealment: A Brief Presentation

Rui Fernandes[1]

[1]Instituto Politécnico de Bragança
Escola Superior de Tecnologia e Gestão
Departamento de Electrotecnia

rvpf@ipb.pt, mpt09015@fe.up.pt

Abstract. *Typical error control techniques are not very well suited for video transmission. On the other hand, video transmission over error prone channels has increased greatly, e.g., over IP and wireless networks. These two facts combined together provided the necessary motivation for the development of a new set of techniques (error concealment) capable of dealing with transmission errors in video systems. These techniques can be categorized according with the approach they take to solve the problem. This categorization is presented, describing the assumptions in which they are based and giving a few examples in each category. Finally, the advantages and disadvantages of each category are presented.*
key words: *error concealment*

1. Introduction

Every communication system has to deal with the problems that may arise during transmission, such as adulteration (bit insertion, deletion or inversion) or loss of the transmitted signal. Traditionally, this problem is treated by applying error control techniques (FEC - Forward Error Correction; ARQ - Automatic Repeat reQuest) on the communication system, however, those techniques are not very well suited for video transmission. For example, live video transmission needs to have very strict transmission delays (cannot afford many retransmissions - ARQ). On the other hand, video transmission is growing even more popular via mobile phones and over the internet, which use noisy channels for the transmission. To solve this problem, a set of techniques were developed, whose purpose was to minimize the influence of the transmission errors at the decoder, taking in consideration the characteristics of the video signal. These techniques are called Error Concealment Techniques and can be divided according to the element, of the transmission system, that has the major part in its implementation [Wang and Zhu 1998]:

- Forward Error Concealment – performed by the encoder;
- Postprocessing Error Concealment – performed by the decoder;
- Interactive Error Concealment – performed jointly by the encoder and decoder.

In fact, Forward Error Concealment and Interactive Error Concealment techniques can be viewed has an extension, that takes in consideration the specificities of the video transmission, of the traditional FEC and ARQ error control techniques, respectively [Wang et al. 2002]. In this sense, Postprocessing Error Concealment techniques are the ones that brought a new way of analyzing this kind of problem. Furthermore, the change of the coding structure, from a pixel based for a object based oriented, implied the birth of Error Concealment Techniques that were also object based oriented (the so called 2nd generation error concealment techniques [Chen and Chen 2002]).

The rest of this paper is organized as follows. Section 2 describes Forward Error Concealment. Section 3 addresses Postprocessing Error Concealment. Section 4 describes Interactive Error Concealment. Section 5 addresses 2nd Generation Error Concealment. Section 6 draws some concluding remarks.

2. Forward Error Concealment

These techniques can be implemented using different approaches, nevertheless, all of them introduce some level of redundancy at the codification stage, with the intention to simplify the error recovery process at the decoder (ideally, to eliminate the need of this error recovery process by the decoder). What can vary from one technique to another is where and how that redundancy is introduced. This can be done at the source coder or at the channel coder. Some examples of this kind of techniques are Layered Coding with Transport Prioritizaton [Khansari and Vetterli 1995], Multiple Description Coding [Wolf et al. 1980], Joint Source and Channel Coding, Robust Entropy Coding. The next subsections explain how some of these techniques work.

2.1. Layered Coding with Transport Prioritization

This technique divides the video signal in layers which will be transported with different levels of priority (quality of service - QoS). It assumes two types of layers: base and enhancement layers. The base layer is transported with the maximum priority and can produce, by itself, a video signal of acceptable quality. The enhancement layers are transported with lower levels of priority and contain information that allow the improvement of the video sequence quality obtained from the base layer. The priority levels used are in accordance with the characteristics and possibilities of the system. The layers can be created by performing spatial, temporal and SNR scalability and data partitioning. Like this, enhancement layers correctly received allow the decoder to perform spatial, temporal and amplitude (quantization) refinement and a better frequency domain partition, respectively. With this architecture, the decoder generates the video according to what it has received correctly (base or base plus enhancement layers). More enhancement layers correctly received give rise to a video sequence with higher quality. An example of this technique is a transmission over an ATM system (a bit in the cells header defines the priority) where the cells with lower priority are discarded first when traffic congestion occurs.

2.2. Multiple Description Coding

This technique assumes that a channel may be disconnected or under the influence of transmission errors. However, considering different channels, the existence of errors in one specific channel is independent of the existence of errors in all other channels, thus, the probability that errors occur simultaneously in all channels is very low. Having this in mind, the coder structure is altered to produce several bit streams (*descriptions*). Each description will be transmitted thru a different channel. Since the use of a large number of several multiple channels has the downside of increasing greatly the quantity of information to be sent, it was taken into account that using few channels could present a situation in which only one description would be correctly received. This fact implies that each description has to have the necessary information to, by itself, represent a video signal with acceptable quality at the decoder. A consequence of this is that the descriptions will have redundant information between them, causing a decrease in the coding efficiency and therefore this technique is only suitable when the used channel has high losses.

3. Postprocessing Error Concealment

These techniques are applied at the decoder. This means that the base of their work is the transmitted bit stream (corrupted, or not, by the transmission errors) that arrives at the decoder. The idea is, when errors occur, to make the best possible with the correctly received information. The way to work in this situation is to realize that a real video signal varies very smoothly both in time and space, which means that spatial and temporal information correctly received, in the neighborhood of the affected area of an image, can be used to dissimulate the effects of the transmission errors. These techniques also use the characteristics of the Human Visual System, which tolerates a higher degree of distortion at high frequencies. Some examples are Motion-Compensated Temporal Prediction [Kieu and Ngan 1994], Maximally Smooth Recovery [Wang et al. 1993], Projection onto Convex Sets, Spatial and Frequency-Domain Interpolation, Recovery of Motion Vectors and Coding Modes.

3.1. Motion-Compensated Temporal Prediction

This technique only uses the temporal redundancy of the video sequence and in case an error occurs, it substitutes the affected area by information of the previous frame. The simplest implementation uses the same spatial area, of the previous frame, to replenish the affected area. However, this method only achieves good results when there is slow movement over the video sequence, otherwise, it can give unpleasant visual results. Better results can be achieved if the motion compensated block is used instead of the block in the same spatial area. This compensated block is the one identified by the motion vectors of the affected block. Under these assumptions, this technique needs that the encoder send extra information such as the motions vectors - MV. If this is corrupted during transmission, the technique can see its performance decrease since this is critical information for the technique to work properly. This means the coder has to *protect* this information as much as it can and that, if this fails, the decoder has to estimate this data, to be able to apply the technique (can use the MV of spatial neighbors blocks to do it). The simplicity behind this technique makes it widely used - it is incorporated in various video standards, e.g., the MPEG standard family.

3.2. Maximally Smooth Recovery

This technique uses the temporal and spatial redundancies of the video sequence and minimizes the spatial and temporal energy difference between the pixels of the affected block and the temporal and spatial neighbor blocks. The idea behind it is to have the smoothest video signal possible (temporal and spatially) and the function to minimize is a weighted sum of temporal and spatial difference measures. For that, it performs three kind of interpolations: frequency (within the block), temporal and spatial using the received coefficients of the block, the prediction block from the previous frame and the adjacent blocks of the actual frame. If all coefficients of the block were lost, the problem reduces itself to the calculation of the temporal and spatial interpolations. If the weight of the spatial difference measure is zero, then the technique only performs the motion-compensated temporal prediction. If the weight of the temporal difference measure is zero, it only performs a spatial interpolation with the blocks received in the actual frame, to recreate the affected block.

4. Interactive Error Concealment

These techniques use interaction between coder and decoder to enhance the error concealment performance of the system. This interaction requires a feedback channel for the decoder to provide information to the coder about what is happening at its side. The coder can then use this information to adapt the source coding or the transport mechanisms (e.g., it can decide to retransmit the erroneous data received by the decoder). Because of this interaction, updated information exists at both sides of the communication system, producing better results than the forward and postprocessing error concealment techniques. However, it is necessary to have a feedback channel to perform these kind of techniques. Some examples are Selective Encoding for Error Concealment [Wada 1989], Adaptative Transport for Error Concealment, Retransmission Without Waiting [Zhu 1996], Prioritized Multicopy Retransmission With Application to Internet Video Stream.

4.1. Selective Encoding for Error Concealment

This technique establishes that the decoder should give feedback to the encoder about the spatial location of the affected area. Having this information, the encoder can treat these areas differently to try to eliminate or, at least, reduce the error propagation effects (due to prediction coding) over the video stream at the decoder. The simplest implementation would be sending the next frame without prediction coding (intramode), whenever an error is detected.

From the decoder point of view, whenever it discovers an error, it sends the location information of the affected area to the encoder, performs postprocessing error concealment techniques in that area and continues with normal decoding. From the encoder point of view, after it has received the feedback from the decoder it chooses from two methodologies. In the first one, the present affected picture area is determined from the frame the decoder indicated up until the actual frame the encoder is working on. This area is not used for prediction coding, at present time. In the second one, the encoder performs the same postprocessing techniques done by the decoder and reexecutes a local decoding from the concealed block up until the present frame to be encoded. The new decoded frame is used for the present prediction coding.

4.2. Retransmission Without Waiting

The traditional retransmission system implies that the decoder has to wait for the retransmitted data to proceed. This introduces a delay that can be critical and that, if it surpasses a certain threshold, may create freezes in the decoded video sequence. The decoder can take two options after having received the retransmission: to decode faster until the decoded frame and its presentation time correspond once again (only a few frames will be shown out of their right presentation time and will result in a quick visual fast forward at the decoded video signal) or to continue to decode at normal speed (a fixed delay is introduced at the decoding process corresponding to the time that took to perform the retransmission). Either way, a delay appears at the decoding level. This particular technique allows the elimination of any delay associated to the retransmission.

In this technique, after a retransmission request the decoder doesn't wait and applies postprocessing techniques to the damaged area and continues the normal decoding (and displaying) while, in parallel, performs and records a trace of the error affected area (it spreads throughout the image due to the prediction coding). Once the retransmission arrives, the decoder uses this information and the trace made before to correct the present affected area. From that point, it can continue as if no errors have ever occurred. With this scheme, this technique achieves lossless recovery except during the time the information got erroneous and the retransmitted data arrives to the decoder.

5. Second Generation Error Concealment Techniques

The coding algorithm were initially established in a pixel oriented scheme. Later, it was introduced a new scheme of codification (an object based orientated one) where the superposition of various objects is what generates an image - Figure 1.

Figure 1. Object based generation of an image.

The error concealment techniques had to adapt themselves to this new coding scheme and had to be oriented to an object based configuration also. The idea is to use statistical models (2-D or 3-D) to shape the objects that appear in the image and use them to replenish the lost data. Examples of this kind of techniques are [Chen and Chen 2002] and [Turaga and Chen 2002].

5.1. Model-Based Error Concealment for Wireless Video

This technique is based in two steps: first, projection onto the model to obtain the reconstruction for the object; second, replace the lost data using the reconstruction created in the previous step. Like this, a model has to be established and for that, models that involve linear combinations of a set of basis vectors to reconstruct the object are used. Hence, the set of reconstructions obtained, when using these models, is convex. Furthermore, the set of objects or regions of interest is also convex. Like this, both operations are projections onto convex sets, and the process can be iterated to give better results. This technique can be used with any object-based video standard such as the MPEG-4 standard.

6. Concluding Remarks

The different approaches to solve the error transmission problem, in video communications, were presented giving some examples in every category (forward, postprocessing, interactive and 2nd generation error concealment).

Forward error concealment techniques trade off coding efficiency for error resilience and usually need more complex encoders. If the used channel has a very low error rate, the overhead these techniques introduce is almost useless and it may pose no advantage in using them.

Postprocessing error concealment techniques do not involve extra redundancy and can be applied at any video system. They make use of the video signal characteristics and temporal/spatial interpolations are easy to implement with visible improvements. To achieve better results, they can be used together with the forward error concealment techniques. The price to pay is the decreased coding efficiency introduced by these ones.

Interactive error concealment techniques require feedback information (that may be not available, e.g., video broadcast systems) to establish the cooperation between coder and decoder. Since the information is updated at both ends (coder, decoder) these are the ones that, up until now, produced the best results. The downside is that, depending on the system, they may not be applicable (no feedback channel) and need more complex encoders and decoders.

Second generation error concealment techniques establish statistical models for the various objects that compose the image and use them, whenever an error occurs, to replenish the affected area. Since the model is specific to an object, they can capture more accurately the variations of that object along the video sequence and, therefore, are the ones that produce the best results among all techniques. Another advantage is that the model can be updated online so that all relevant information is being used all the time. The disadvantage is the complexity degree that is introduced to create the various statistical models for the objects.

References

Chen, T. P.-c. and Chen, T. (2002). Second-generation error concealment for video transport over error prone channels. *Image Processing. 2002. Proceedings. 2002 International Conference on*, 1:I–25–I–28.

Khansari, M. and Vetterli, M. (1995). Layered transmission of signals over power-constrained wireless channels. *Image Processing, 1995. Proceedings., International Conference on*, 3:380–383.

Kieu, L. H. and Ngan, K. N. (1994). Cell-loss concealment techniques for layered video codecs in an atm network. *Image Processing, IEEE Transactions on*, 3(5):666–677.

Turaga, D. S. and Chen, T. (2002). Model-based error concealment for wireless video. *Circuits and Systems for Video Technology, IEEE Transactions on*, 12(6):483–495.

Wada, M. (1989). Selective recovery of video packet loss using error concealment. *Selected Areas in Communications, IEEE Journal on*, 7(5):807–814.

Wang, Y., Ostermann, J., and Zhang, Y.-G. (2002). *Video Processing and Communications*. Prentice Hall.

Wang, Y. and Zhu, Q.-F. (1998). Error control and concealment for video communication: A review. *Proceedings of the IEEE*, 86(5):974–997.

Wang, Y., Zhu, Q.-F., and Shaw, L. (1993). Maximally smooth image recovery in transform coding. *Communications, IEEE Transactions on*, 41(10):1544–1551.

Wolf, J. K., Wyner, A., and Ziv, J. (1980). Source coding for multiple descriptions. *Bell Syst. Tech. J.*, 59:1417–1426.

Zhu, Q.-F. (1996). Device and method of signal loss recovery for real-time and/or interactive communications. *U.S. Patent 5 550 847*.