CCIT Report #526

April 2005

Packet Loss Concealment for Audio Streaming Based on the GAPES Algorithm

Hadas Ofir[1], David Malah[2]

[1] Department of Electrical Engineering Technion IIT,Haifa 32000, Israel
hadaso@tx.technion.ac.il

[2] Department of Electrical Engineering Technion IIT,Haifa 32000, Israel
malah@ee.technion.ac.il

ABSTRACT

In this work we present a novel approach for audio packet loss concealment, designed for MPEG-Audio streaming, based only on the data available at the receiver. The proposed method is based on the GAPES (Gapped-data Amplitude and Phase Estimation) algorithm for replacing the missing data, using interpolation in the spectral domain. The MPEG standard uses the Modified Discrete Cosine Transform (MDCT) for compression. However, better interpolation results are obtained by converting the data to the Discrete Short-Time Fourier-Transform (DSTFT) domain. This conversion is done directly using an efficient procedure developed in this work. This technique was tested subjectively and was found to provide better performance than previously reported works, even with a packet loss rate of 30%.

1. INTRODUCTION

With the growing popularity of the internet and the advancement in modem technology, there is increasing interest in using the internet for multimedia broadcasting, such as audio and video. This kind of broadcasting is usually referred to as *streaming media*, and its low cost and convenience make it appealing to both media distributors and consumers.
Audio streaming operates by first compressing a digital audio file and then breaking it into small packets, which are consecutively sent over the internet. When the packets reach their destination, they are decompressed and reassembled into a form that can be played by the user's system. To maintain the illusion of seamless playing, the packets are "buffered". That is, a number of them are downloaded to the user's machine before playback. As those buffered packets are played, more packets are being downloaded and queued up for playback. This way, the client experiences only a small delay of a few seconds, waiting for the buffer to build up, instead of waiting several minutes, or even hours, for the complete files to be downloaded.

However, since internet delivery doesn’t assure quality of service, data packets are often delayed or discarded during network congestions. When the stream of arriving packets becomes too slow, the client’s audio player has nothing to play, thus an annoying gap is created in the streamed media.

Previous researches on this subject [1-3] show that most of the receivers in such internet connections experience a mean loss rate of about 10% or lower, but it is still possible that at certain times the loss rate will go to the extremes, such as up to 20% or 30% loss rate, or no loss at all. At small loss rates (10% or less) the packet losses are random and become more and more correlated as the loss rates go higher [2]. Hence, when the network load is low to moderate, the packet losses are usually isolated audio packets [1]. Each loss, unless concealed in some way, produces an annoying disturbance. The common approach for dealing with such cases is to interpolate the gap, approximating the original waveform, so that a human listener will not notice the disturbance. Packet loss concealment algorithms are usually divided into two categories: *receiver-based* methods where only the data available at the receiver is used for the concealment and *sender-based* methods, where the sender changes the encoded bitstream, adding some redundancy or additional side information that the receiver can later use for the concealment process. In this work we focus on a receiver-based solution.