Unequal Error Protection Based on Expanding Window Fountain for Object-Based 3D Audio

YANG Cheng¹,²,⁴, HU Ruimin¹,², SONG Yucheng¹, SU Liuyue¹, WANG Xiaochen¹,³, CHEN Wei¹,²
¹. National Engineering Research Center for Multimedia Software / School of Computer Science, Wuhan University, Wuhan 430072, Hubei, China;
². Hubei Key Laboratory of Multimedia and Network Communication Engineering (Wuhan University), Wuhan 430072, Hubei, China;
³. Research Institute of Wuhan University in Shenzhen, Shenzhen 518000, Guangdong, China;
⁴. School of Physics and Electronic Science, Guizhou Normal University, Guiyang 550001, Guizhou, China
© Wuhan University and Springer-Verlag Berlin Heidelberg 2017

Abstract: This paper proposes an unequal error protection (UEP) coding method to improve the transmission performance of three-dimensional (3D) audio based on expanding window fountain (EWF). Different from other transmissions with equal error protection (EEP) when transmitting the 3D audio objects. An approach of extracting the important audio object is presented, and more protection is given to more important audio object and comparatively less protection is given to the normal audio objects. Objective and subjective experiments have shown that the proposed UEP method achieves better performance than equal error protection method, while the bits error rates (BER) of the important audio object can decrease from $10^{-3}$ to $10^{-4}$, and the subjective quality of UEP is better than that of EEP by 14%.

Key words: object-based 3D audio; unequal error protection; equal error protection

CLC number: TP 37

0 Introduction

Nowadays, more and more 3D video and audio materials are provided by the Internet and the mobile network. With the increase of the objects/channels of 3D audio, the bit rate of 3D audio is increasing linearly, which requires more bandwidth for the transmission in real time. As bandwidth is quite limited in mobile network, it is necessary to examine the issue of transmitting 3D audio materials on mobile network and achieve good reconstruction quality.

The Internet and the mobile network are typically two kinds of binary erasure channel (BEC), which can be used to model various data networks, and where packets are either received reliably or lost completely due to the unknown characters of channels and their time-varying erasure probabilities. It is usual to employ a feedback mechanism from the receiver to the sender to manage the retransmission of erased packets. However, such retransmission protocols might be burdened with an excessive number of feedback messages. One possible solution is to ensure that erasure-correcting codes require no feedback (or almost no feedback). In fact, there exists an erasure-correcting code called Fountain code featured with that.

Fountain codes are a class of erasure-correcting codes that require almost no feedback. They are rateless as the encoder can generate as many encoded symbols as needed [1]. Those advantages can be used when the channel conditions are unknown or time-varying, which meets the features of BEC. Fountain codes have low encoding and decoding complexity, but require a few more encoded symbols at the receiver for successful decoding. Luby Transform (LT) codes [2] and Raptor codes [3] are examples of Fountain codes.
Because of the attractive features of fountain codes, recently introduced unequal error protection (UEP) fountain code designs have become more popular for multimedia communications. In fact, a portion of data may need more protection than the rest of data in several applications. For example, in a MPEG stream, I-frames need more protection than P-frames. UEP has been successfully used [4] for the protection of scalable image.

Since the design of UEP-LT codes applied in video transmission has achieved a good performance [4], we are considering that the UEP-LT codes can also be applied in the transmission of 3D audio in order to obtain a good reconstruction performance and give users a uniform feeling of the received audio and video materials. Object-based 3D audio coding (D-DirAC [3], MPEG-H [6]) bit-streams are composed of several objects with different priorities, and the object with higher priority is more important than the objects with lower priority. Therefore, how to protect the object with higher priority and ensure a good reconstruction performance under the bandwidth-limited network has become an important issue.

For the reasons we mentioned above, we propose a method of UEP-LT codes and a scheme to give more protection to the object with higher priority based on object-based 3D audio. The method is called expanding window fountain (EWF) codes.

The rest of the paper is organized as follows. In Section 1, the background of 3D audio bit-streams structure and LT codes are introduced. In Section 2, the UEP design of LT codes for object-based 3D audio is introduced in detail. In Section 3, the objective simulation experiment and subjective listening experiment are implemented, and in the last section, conclusions are summarized.

1 Background

1.1 Bit-Streams Structure Features of Existing Object-Based 3D Audio

The bit-streams of object-based 3D audio coding such as D-DirAC and MPEG-H include several audio objects and the priorities of different objects are different. Our method can be utilized in these 3D audio coding.

As shown in Fig.1, the bit-streams of object-based 3D audio (MPEG-H) are composed of two parts, and they are audio objects and spatial parameters (object metadata). Different audio objects have different priorities (the priority values change from 0 to 7). The audio object with higher priority is more important than that with lower priority. Based on this feature of object-based 3D audio, we can protect the object with higher priority and ensure a good reconstruction performance under the bandwidth-limited network by using UEP LT coding.

![Fig. 1 Top level block diagram of MPEG-H 3D audio decoder](image)

1.2 Principles of LT Codes

In this section, the principles of LT codes are briefly introduced. LT codes are the first class of practical fountain codes proposed by Luby [2]. Their features are the good performance of random liner fountain codes and the low complexity of encoding and decoding.

Every encoded symbol of LT codes is generated by performing bitwise XOR (exclusive OR) operations on it and all the input symbols connecting with it. More specifically, given $n$ information symbols $i_1, i_2, \ldots, i_n$, a sequence of $m$ encoded symbols $e_1, e_2, \ldots, e_m$ is generated. Figure 2 is the encoding graph of LT codes.

![Fig. 2 Encoding graph of LT codes with $n$ is 6 and $m$ is 8](image)

This is a bipartite graph in which one side is the $n$ input symbols and the other side is the generated $m$ symbols. When generating every encoded symbol, it needs to produce a distribution of degree $d$ randomly. Each encoded symbol is connected with $d$ information symbols. The process of encoding is shown as the following four steps:

1. Choose a value of degree $d$ randomly in terms of the designed degree distribution;
2. Choose $d$ information symbols from $n$ original information symbols with equal probability.