International Journal of Innovative Computing, Information and Control
Volume **7**, Number **12**, December **2011**
 ISSN 1349-4198
pp. 6773–6783

AN ADAPTIVE PACKET LOSS RECOVERY METHOD BASED ON REAL-TIME SPEECH QUALITY ASSESSMENT AND REDUNDANT SPEECH TRANSMISSION

Jin Ah Kang and Hong Kook Kim

School of Information and Communications
Gwangju Institute of Science and Technology (GIST)
1 Oryong-dong, Buk-gu, Gwangju 500-712, Korea
{ jinari; hongkook }@gist.ac.kr

Received August 2010; revised February 2011

Abstract. *In this paper, an adaptive packet loss recovery (APLR) method that improves the speech quality of a real-time speech streaming (RSS) system over IP networks is proposed. The proposed APLR method estimates the packet loss rate (PLR) of network via a real-time speech quality assessment (RSQA) at the receiver side of the RSS system, and then requests the opposite RSS system to transmit redundant speech frame data (RSD). Thus, it assists the speech decoder employed in the RSS system to reconstruct lost speech signals when the estimated PLR is high. In particular, according to the estimated PLR, the proposed APLR method then controls the bitrate of speech coding for the RSS system. In other words, a speech packet combines the bitstreams of the current speech frame data (CSD) and the RSD for a high PLR. Otherwise, for a low PLR, a speech packet consists of the CSD bitstreams alone. The effectiveness of the proposed APLR method is finally demonstrated by using an adaptive multirate-narrowband (AMR-NB) speech codec and ITU-T Recommendation P.563 as the scalable speech codec and RSQA, respectively. It is shown from experiments that an RSS system employing the proposed APLR method significantly improves the speech quality under packet loss conditions.*
Keywords: Packet loss recovery, Real-time speech quality assessment, Redundant speech transmission, Adaptive multi-rate speech coding

1. **Introduction.** Due to the rapid development of Internet protocol (IP) networks over the past few decades, audio and video streaming services are increasingly available via the Internet. Moreover, as these services are extended to wireless networks, the quality of service (QoS) of audio and video streaming is becoming ever more critical. To this end, there have been a number of previous studies reported that have attempted to guarantee the quality of speech that could be impaired due to environmental factors such as network conditions [1, 2] or noise [3]. Specifically, real-time speech streaming (RSS) systems require a minimum level of speech communication quality. For example, the guarantee of speech quality is important for e-learning applications, where such speech quality is largely related to packet losses and end-to-end packet delays [4]. When speech streaming is performed via user datagram protocol/IP (UDP/IP) networks, however, packets may be lost or arrive too late for playback due to inevitable delays. In this case, a typical RSS system can only tolerate a few packet losses for real-time services, though these packet losses frequently occur in wireless networks due to bandwidth fluctuations [5].

Several packet loss recovery methods, implemented via the Internet and wireless networks, have been proposed for RSS systems. For instance, the techniques proposed in [6, 7] were sender-based packet loss recovery methods using forward error correction (FEC). In regards to wireless networks, the techniques proposed in [8, 9] were based on unequal error