多媒体会议中的快速实时自适应混音方案研究

樊 星, 顾伟康, 叶秀清

樊 星, 顾伟康, 叶秀清

(浙江大学 信息与电子工程学系,浙江 杭州 310027)

作者简介: 樊星(1976－),男,重庆人,博士生,主要研究领域为图像处理,视频处理,音频处理,多媒体通信;顾伟康(1939－),男,教授,博士生导师,主要研究领域为计算机视觉,图像处理,模式识别;叶秀清(1936－),女,教授,主要研究领域为计算机视觉,图像处理,视频处理.

联系人: 樊 星 Phn: +86-571-87951529, Fax +86-571-87951529, E-mail: starfan@mail.hz.zj.cn, http://www.zju.edu.cn

Abstract

In multimedia conferencing, multi-point controlling unit (MCU) provides the capabilities to process audio, video and data stream for multi-point conference. The capability of audio processing is basic and requires more for real-time criteria. This paper categorizes and analyzes the schemes, and a new multi-point speech audio mixing scheme using align-to-self weighted algorithm is provided to meet the demand of the practical need of multi-point speech processing. By applying the adaptive mixing algorithms, these high-performance processing schemes do not use the saturation operation which is widely used in multimedia processing. Therefore, no new noise will be added to the output, and they have low complexity and good hearing perceptibility. In the mean time, the schemes are designed for parallel processing, so they can be easily implemented with hardware, such as DSPs, and widely applied in multimedia conferencing systems.

Fan X, Gu WK, Ye XQ. Fast real-time adaptive audio mixing schemes in multimedia conferencing. *Journal of Software*, 2005,16(1):108-115.
http://www.jos.org.cn/1000-9825/16/108.htm

摘要

多媒体会议中多点控制单元(multi-point controlling unit,简称MCU)在多点会议中提供音频、视频和数据等的集中处理能力,其中音频处理能力是最基本的,也是实时性要求最高的要素.针对多点多媒体会议的实际应用需求,归类并分析了多种自适应多点语音混合处理方案,提出了采用自对齐加权的高性能混音方案.该方案不使用在实时多媒体处理中广泛运用的饱和运算,所以不引入新的噪声,因而具有较低的算法复杂度,其混合处理结果具有良好的听觉主观舒适感.同时,这套方案具有较好的并行处理特性,使用DSP等硬件较易实现,可以广泛应用在多媒体会议系统的实现中.

References:

[1] Yang ST, Yu SS, Zhou JL. A multipoint real-time speech mixing and scheduling algorithm based on packet networks. Journal of Software, 2001,12(9):1413-1419 (in Chinese with English abstract).

[2] Daigle JN, Langford ID. Model for analysis of packet voice communications systems. IEEE Journal on Selected Areas in Communications, 1986,4(6):847-855.

[3] Venkat RP, Harrick MV, Srinivas R. Communication architectures and algorithms for media mixing in multimedia conferences. IEEE/ACM Trans. on Networking, 1993,1(1):20-30.

[4] Agustín JG, Hussein AW. Audio mixing for interactive multimedia communications. In: Wang P, ed. Proc of the JCIS'98. NC: Research Triangle, 1998. 217-220.

[5] ITU-T. Packet-Based multimedia communication system. ITU-T Rec H.323 v4, 2000.

[6] Schulzrinne H, Caner S, Frederick R, Jacobson V. RTP: A transport protocol for real-time applications. IETF RFC 1889, IETF, 1996.

[7] Tu WP, Hu RM, Ai HJ, Xie X. Audio MP in video conference. Geomatics and Information Science of Wuhan University, 2002, 27(1):98-101 (in Chinese with English abstract).

附中文参考文献：
[1] 杨树堂,余胜生,周敬利.基于分组网络的多点实时语音混合及调度算法.软件学报,2001,12(9):1413-1419.

[7] 涂卫平,胡瑞敏,艾浩军,谢兄.视频会议中音频多点处理的研究.武汉大学学报(信息科学版),2002,27(1):98-101.