

**INTERNATIONAL ORGANIZATION FOR STANDARDIZATION
ORGANISATION INTERNATIONALE NORMALISATION
ISO/IEC JTC 1/SC 29/WG 11
CODING OF MOVING PICTURES AND AUDIO**

ISO/IEC JTC 1/SC 29/WG 11 N3233

March 2000

Source: Audio Subgroup
Title: Report on Transmission of MPEG-4 Audio over RTP
Status: Approved

1. Introduction

This document describes an interoperability test of MPEG-4 Audio RTP transmission based on one of the proposals for a RTP packet format for MPEG-4 Audio, currently in the standardization process at IETF [1]. In this RTP specification, an MPEG-4 Audio payload based on the LATM format [2] is directly mapped into the RTP payload. The interoperability test was conducted between NEC and FhG.

2. Test procedure and conditions

Figure 1 shows the block diagram for the interoperability test. At the transmitter side an audio signal is encoded and packetized into an RTP packet stream. These encoding operations were done off-line and the RTP packet stream was stored in a temporary file. The RTP packets are transmitted through the IP network, where the packets are transmitted in real-time based on the RTP timestamp. The receiver side receives the RTP packets and decodes them by means of a real-time software implementation to synthesize the reconstructed audio signal. .

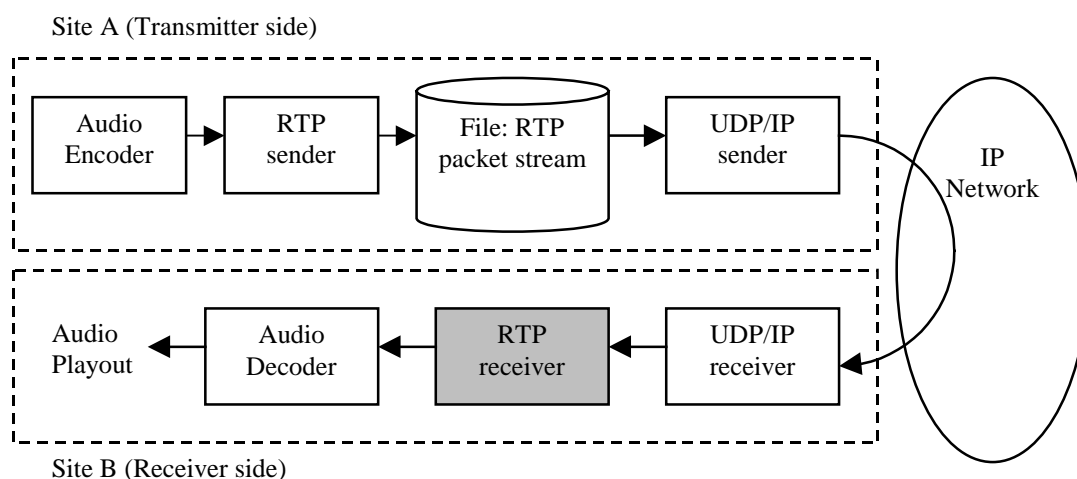


Figure 1 A block diagram for the interoperability test

Two interoperability tests were conducted using CELP and AAC objects, respectively. The software implementations of the MPEG-4 Audio encoder/decoder and the RTP transmitter/receiver were independently developed by NEC and FhG.. The test conditions are summarized in Table 1. The audio-specific and LATM configurations are pre-determined prior to the interoperability test as shown in Table 2. It is assumed that these configurations are transmitted by out-of-band means.

Table 1 Interoperability test conditions

	Audio object type	Transmitter Implementation	Receiver Implementation
Test 1	AAC	FhG (Linux)	NEC (Windows)
Test 2	CELP	NEC (Windows)	FhG (Linux)

Table 2 Audio specific and LATM configurations

Audio object	Audio specific configurations	LATM configurations
AAC (2 minutes long)	stereo, 64 kbit/s (average), 24 kHz sampling, 1024 sample/frame	Variable frame length (in bits), 1 audio-frame, 1 program, 1 layer, (MuxConfigPresent is set to 0, No other and CRC-check data)
CELP (5 minutes long)	mono, 6 kbit/s (fixed), 8 kHz sampling, 160 sample/frame, (FRC, BRS, BWS are OFF)	Fixed frame length (in bits), 1 audio-frame, 1 program, 1 layer, (MuxConfigPresent is set to 0, No other and CRC-check data)

3. Considerations of transmission overhead

In the test, the overhead due to the LATM and RTP/UDP/IP layers was also evaluated based on the above configurations. The overhead and audio payload length are summarized in Table 3.

Table 3 Examples of overhead for RTP transmission

	Audio Payload Length	Overhead		
		LATM	RTP/UDP/IP	Rate
AAC	342 bytes (average)	2 bytes	12+8+20 bytes (uncompressed) 2-4 bytes (compressed)	10.9 % 1.2-1.7 %
CELP	15 bytes (fixed)	0 bytes	12+8+20 bytes (uncompressed) 2-4 bytes (compressed)	72.7 % 11.8-21.1 %

The overhead for the LATM layer (payloadLengthInfo) depends on the audio payload. In the case of AAC, for this experiment the overhead is 2 bytes. On the other hand, there is no overhead for the LATM layer in the fixed bitrate CELP mode.

The 40-byte overhead for the RTP/UDP/IP layer is too large for a low-bitrate coder like CELP. However, if a header compression mechanism is applied, the overhead can be reduced to 2-4 bytes [3]. In this situation, the overhead rate for the CELP payload is reduced to 11.8-21.1 %.

4. Test results on LAN

Test 1 and test 2 were conducted over a Local Area Network (LAN) at NEC and FhG, respectively. The interoperability of the two different implementations was verified, as all RTP packets stored in a file were transmitted from a sender to a receiver and were unpacketized and decoded successfully. There was no packet loss and, therefore, the decoded sound was clear and undistorted by the RTP transmission.

5. Test results of Internet transmission

Test 2 was conducted between NEC and FhG through the Internet. The NEC site transmitted CELP RTP packets to the FhG site. At the FhG site, the received packets were unpacketized and decoded. The network connection was quite good and hence the packet loss rate was relatively low. The decoded sound was still clear and undistorted by the RTP transmission through the Internet. The experiment will continue.

6. Conclusion

An interoperability test of the MPEG-4 Audio RTP packet format has been conducted between two companies, NEC and FhG. MPEG-4 AAC and CELP object types have been used for this test. For these the maturity and interoperability of the MPEG-4 Audio RTP payload format based on LATM has been confirmed.

7. References

- [1] Y. Kikuchi, T. Nomura, S. Fukunaga, Y. Matsui, H. Kimata, "RTP payload format for MPEG-4 Audio/Visual streams", Internet Engineering Task Force, Internet Draft, draft-ietf-avt-rtp-mpeg4-es-00.txt, Feb. 2000.
- [2] ISO/IEC JTC1/SC29/WG11 N3058: "Text of ISO/IEC 14496-3/FDAM 1", Dec. 1999.
- [3] RFC 2508: "Compressing IP/UDP/RTP Headers for Low-Speed Serial Links", Feb. 1999.