Circuit-Switched Mobile Video

3G-324M is an umbrella standard for support of real-time multimedia services over a wireless circuit-switched network. This standard comprises several sub-protocols that handle multiplexing and demultiplexing of speech, video, user, and control data (H.223) and in-band call control (H.245).

H.223 Multiplexing and Demultiplexing

In order to provide different degrees of error-resilient transport, 3G-324M defines levels of H.223 transport:

- **Level 0**, or baseline H.223, provides support for synchronization and bit stuffing. Level 0 allows 16 different multiplexing patterns to assemble media, control, and data packets. Multiplexing patterns are negotiated between the endpoints. The error resilience capabilities of Level 0 are limited. Bit errors can break the high-level data link controller (HDLC) protocol, can interfere with bit stuffing, and can be the cause of flag emulations in the payload.

- **Level 1**, defined in H.223 Annex A, provides a synchronization mechanism that considerably improves performance over error-prone channels. HDLC is replaced by a more robust framing and framing flag of larger length.

- **Level 2**, defined by H.223 Annex B, is a further enhancement of Level 1, including a more robust MUX-PDU framing.

- **Level 3**, defined in H.223 Annex C, defines the most robust delivery scheme. It includes modified multiplex and adaptation layers. It includes forward error correction (FEC) and retransmission (ARQ) schemes.

In H.223, every level comprises an adaptation, multiplexing, and demultiplexing layer. AL1, AL2, and AL3 are the three defined adaptation layers.

- **AL1** is designed for data transfer and is typically used to transport user data and H.245 control messages. It relies on upper layers for error control and handling.

- **AL2** provides an 8-bit cyclic redundancy check (CRC) and optional sequence numbering to allow loss detection. AL2 has the capability to handle AL service data units (SDUs) of variable length and is the preferred adaptation layer for the transport of audio data.

- **AL3** is designed primarily for video and includes a 16-bit CRC and optional sequence numbering. It has the capability to handle AL SDUs of variable length and provides an optional retransmission procedure.

H.245 Call Control

H.245 employs a simple retransmission protocol (SRP) that can be optionally numbered (NSRP) and provides a control channel segmentation and reassembly layer (CCSRL) to ensure reliability in error-prone environments. The use of SRP vs. NSRP and CCSRL depends on the negotiated level. H.245 relies on Abstract Syntax Notation 1 (ASN.1) for the definition of its messages. Furthermore, the messages are binary encoded according to the packed encoding rule (PER).

When parties start an H.245 conversation, it is necessary to decide which endpoint has the right to provide resolution in case of conflict, or to be the master. Endpoints may have different capabilities in terms of H.223 multiplexing/demultiplexing, supported video and speech codecs, data sharing, and other optional features. H.245 provides a capability exchange functionality to allow the negotiation of capabilities and to identify a set of features common to both endpoints.

The media and data flows are organized in logical channels that require support for their control. H.245 provides logical channel signaling to allow logical channel open/close and parameter exchange operations.

In H.245, the choice of codecs and their parameters are decided at the transmitter side on the basis of the capabilities that the receiver has advertised. If the receiver has a preference within its capability, it can signal it to the transmitter via a mode request.

Finally, H.245 provides a variety of call control commands and indications to allow flow control, user input indications, video codec control, jitter indication, and skew.

Speech Codecs

Under 3G-324M, the adaptive multi-rate (AMR) codec is the mandatory speech codec. G.723.1 is an optional legacy codec included in the 3rd Generation Partnership Project (3GPP) recommendation for compatibility with standards such as H.323.

AMR can operate at different rates between 12.2 and 4.75 kbps. It also supports comfort noise generation (CNG) and a discontinuous transmission (DTX) mode. It can dynamically adjust its rate and error control, providing the best speech quality for the current channel conditions. The AMR codec also supports unequal error detection and protection (UED/UEP). This scheme partitions the bit stream into classes on the basis of their perceptual relevance. An AMR frame is discarded if errors are detected in the most perceptually relevant data, otherwise it is decoded and error concealment is applied.

Video Codecs

The 3G-324M standard specifies H.263 baseline level 10 as mandatory and MPEG-4 simple profile level 0 as recommended video codecs. H.263 is a legacy codec that is used by existing H.323 systems and has been kept for compatibility. MPEG-4 is more flexible than H.263 baseline and offers advanced error detection and correction schemes.

Both codecs typically use Quarter Common Intermediate Format (QCIF) input picture format. MPEG-4 provides a series of toolkits to boost error resilience. They include data partitioning, reversible variable length codes (RVLC), resynchronization markers, and header extension codes (HEC).

Data partitioning separates Discrete Cosine Transform (DCT) coefficients and motion vector information by markers to prevent an error in one set from affecting the decoding of the other one. As an example, if errors are detected only in the DCT coefficients of a given macro-block, it will be possible to reconstruct the macro-block
using the correct motion information while concealing the errors in the DCT coefficients. This results in a higher visual quality of the decoded picture than replacement of that macro-block with the corresponding one in the temporally previous frame.

- RVLCs allow for the decoding of a given data block starting from its beginning (forward) or from its end (backward). This capability increases the chances of partial correct decoding of a set of corrupted data.

- Resynchronization markers are codes inserted in the bit steam that allow the decoder to resynchronize the decoding process.

- Header extension codes (HEC) allow a more effective resynchronization of the decoding process, extending the resynchronization marker to also include timing information.

**User Input Indication**

H.245 User Input Indication (UII) plays a key role in all the services that require user interaction. For video messaging, typical uses of UII include selection of user preferences, message recording and retrieval, and typical mailbox management functions. The H.245 signaling protocol is reliable, which ensures that the messages (i.e., DTMF tones) are guaranteed to be delivered. H.245 UII provides two level of representation of user indications — alphanumeric and the addition of duration information to the alphanumeric string to represent, for example, how long a given key has been pressed.

**Media Adaptation Overview**

Multimedia mobile communications enabled by technologies such as 3G offer the promise of multimedia information access anytime, anywhere, on any networked multimedia-enabled terminal. The problem, however, is how to provide content and services in an acceptable format to a wide variety of terminals. The capabilities of these terminals differ in computing power, display capabilities, network access, and supported network bandwidth.

Media adaptation provides the framework to dynamically alter the content, e.g., picture size, encoding format, and organization of the multimedia content, while ensuring that the content semantics are kept as coherent and faithful as possible to the original. Adaptation is performed on the basis of terminal capabilities and user preferences. The media adaptation framework utilizes several components:

- A multimedia information model for adaptation, which includes a hierarchy of content descriptors for the different modalities in which the multimedia content can be presented and delivered.

- An adaptation strategy decision component that analyzes the context information and computes and selects the proper adaptation strategy.

- Media processing techniques for manipulating, translating, transcoding, and rearranging the multimedia content.

Standards such as MPEG-7 and more recently MPEG-21 cover the definition of multimedia information models in support of media adaptation. Telecom operators typically define the viable adaptation strategies on the basis of the available media processing and delivery resources.

**Processing Techniques for Media Adaptation**

There are at least two classes of techniques to perform media adaptation — intra-media adaptation and cross-media adaptation.

- Intra-media adaptation relies on the characteristics of the particular encoding scheme used for a given media to perform the adaptation. For example, it is possible to operate on the video compression attributes, such as the video frame rate and the picture formats or the specific intra-frame and inter-frame quality, to meet a specific data size and format target. Similarly, content can be transrated to meet specific terminal bandwidth constraints. Transcoding can also be employed to adapt to the different terminal coding standards. For applications and services based on 3G-324M, H.263 and MPEG-4 transcoding is required because both video formats are part of the standard. This adaptation model, however, has an inherent lower bound which is dictated by the lowest perceptually acceptable quality of the specific media.

- Cross-media adaptation can be used to overcome this limitation. In the context of cross-media adaptation it is possible to substitute a given media type with a “semantically equivalent” one, meaning that the substitution of media minimally affects the user-perceived content of the presentation.

As an example, a given TV-format (720 x 480 pixels) video clip can be transformed into a sequence of “key” still pictures, sampled at scene changes or when considerable motion is detected in the scene, resized to QCIF format (176 x 144 pixels), synchronized with a lower rate coded version of the audio, and packaged as an MMS message that can be delivered to a 2.5G phone.

Because the mobile environment has the most stringent set of presentation capability and network bandwidth constraints, it is expected that both intra- and cross-media adaptation will play an important role in the delivery of content for mobile-video-enabled services.