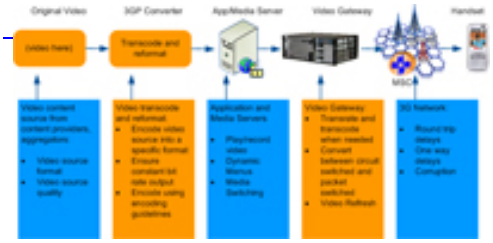
Creating Quality Mobile Video Telephony Applications

By Marwan Jabri, PhD, Dilithium Networks



[1]

3G video-telephony is commonly used today in interactive video value-added services. Such services can take advantage of the 3G circuit-switched bearer because of three main reasons: the guaranteed quality of service (QoS), the real-time interactive response, and the ability to access such services by dialing short codes (4-digit numbers). 3G circuit-switched (VT) and packet-switched services are becoming complementary. Video applications that are less sensitive to QoS use the 3G packet-switched connection, where as real-time applications use the circuit-switched bearer.

Most operators and service providers around the world have either launched or are about to launch VT-based value-added services (VAS – peer to/from machine services). For example, video portals, video on-demand (video snacks), video blogging, participation TV/radio, video surveillance and video ringback tones.

Characteristics of a 3G to 324M Video

The deployment of quality VT VAS requires good understanding of the characteristics of the underlying protocol (3G-324M) and the transmission bearer. Packet-based video services have greater but less guaranteed bandwidth. 3G VT has less but more guaranteed bandwidth, and the underlying 3G-324M bearer has the following properties: &#149For WCDMA and TD-SCDMA, the bearer is 64kbps and carries the 3G-324M multiplexed stream, which includes audio/voice, video, data, command and control, and any framing and filling flags. &#149Voice is typically GSM-AMR (also known as Narrowband AMR) and is carried at 12.2kbps (about 19%). &#149Command/Control: typically only at beginning of session (session setup). Not counted in steady state operation as it is minimal. &#149Framing overheads: Typically at about 13bytes/160bytes frames, so about 8.2% &#149Finally, but most importantly, video. Counting the above, this leaves about 46 kb/s (72%) for video.

The Command and Control (ITU-T H.245 protocol) is used to set up the session (codecs, bitrates, etc.). The session-setup time in baseline 3G-324M can stretch to

about 5-to-8 seconds, which may significantly impair the customer experience. As a result, in 2006 the International Telecommunication Union/ITU Telecommunication Sector (ITU-T) introduced H.324 Annex K which defines Media Oriented Negotiation Acceleration (MONA), a technique used to reduce session-setup time to around 1 second, similar to audio calls. MONA is increasingly being adopted by handset manufacturers and many operators mandate its support in their services.

For video, although H.263 is the mandatory codec, most handsets today support MPEG4-part 2, and an increasing number of handsets support H.264. The main advantage of H.264 over MPEG4-part2 and H.263 is the vastly improved video quality, and that's essential for VT VAS.

From a VT perspective, the video characteristics of H.264 are: &#149Video Frame Rates: Typically 8-to-15 frames/sec &#149Key Frames (I-Frames) and I-Macro-Blocks: Most handsets would only send I-Frames on request; otherwise the handset's codec would ensure that a macroblock is coded as in intra mode, at least once every 130 frames. &#149For video VAS services, a video gateway is required to ensure video bitrates are adequate. Video gateways need to support at least real-time transrating/transcoding; or lip-synch problems, video jitter and significant frame losses occur.

Because video occupies over 72% of the bearer, transmission errors manifest themselves to the users mostly in the form of video corruption. Typical forms of video corruption are: corrupt macroblocks (16x16 pixels), a series of corrupt macroblocks, multiple rows of corrupt macroblocks, or green frames with shadows. If the transmitting handset or video gateway implements video fast-update techniques (VFU), video corruption will last only a small fraction of a second (invisible); otherwise the corruption may last up to 10 seconds.

3G VT Service Video Quality

The quality of the user experience in VT VAS hinges on system design methodology. We will use the example of delivering a video stream to a 3G VT handset as shown in the figure below. The video stream could be a greeting, a video showing a selection menu, an actual video to be watched by the user (e.g. sports, news, and comedy) or a video ringback tone. The delivery of a video stream is the key ingredient in any VAS 3G VT application.

Figure 1 shows the various steps involved in delivering a video stream optimized for 3G VT characteristics.

The key aspects of the quality of service depend on the quality of the content (from the content provider), its encoding, the media server delivery, the video gateway real-time video adaptation, and the network.

Key Issues Affecting Video Quality in VT VAS Applications

The first issue with which to be concerned is the quality of the video source (from content provider) and its adequacy to be encoded to be "VT streamable" content.

Typically, the input video content is encoded/converted to container/stream format so it can be streamed by the media server of the VAS infrastructure.

The quality of the input video source needs to be as good as possible, as when it is converted to a 3G VT stream, the quality will be reduced to that of QCIF-sized 46kbps of H.263, MPEG4 or H.264. Low bitrate content cannot be used as a source as the quality will significantly degrade after converting it to VT streamable content.

The bitrate of encoded video should hover on average around 46kbps. Exceeding 46kbps for excessive time could have severe repercussions on customer experience. For example, if the converted video bit rate overshoots to 60kbps for a duration of 5 seconds, then the excess bits (14kbps/sec) over that period would be 70kbits, and this will introduce an audio/video (lip-sync problem) of about 1.5secs (and accumulating over time). Therefore it is highly recommended that the video gateway (downstream in the processing chain) implements real-time constant bitrate adaptation with short latency to ensure that lip-synch issues are minimized. Further, as video content can be in H.264 (for best quality delivery), the video gateway needs to transcode to H.263 or MPEG4-part2 for handsets that do not support H.264.

3G handsets are equipped with the facility to request peer terminals to transmit I-frames when they encounter corrupt bitstreams. This facility is called Video Fast Update in the 3G-324M standard (similar facilities exist in SIP and other packet based protocols), and is supported by all 3G handsets on the market today.

Similarly, video gateways needs to service VFU requests from handsets and generate their own when encountering video corruption. If the video gateway does not support VFU processing or generation, video corruption will appear on the user screen and may last for up to 10 seconds.

Conclusion

Video content quality and constant bitrate encoding are essential for quality 3G VT VAS deployments. It is also essential that video gateways implement short latency transcoding/transrating to minimize lip-synch issues, and maximize the use of advanced codecs such as H.264. Finally Video Fast Update is essential in video gateways so they can provide dynamic recovery from mobile corruption while maintaining the best quality video.

Marwan Jabri, PhD is co-founder & CTO of Dilithium Networks, www.dilithiumnetworks.com, 707-792-3900.

Source URL (retrieved on *08/20/2014 - 1:45pm*):
http://www.wirelessdesignmag.com/articles/2010/03/creating-quality-mobile-video-telephony-applications?qt-digital_editions=0

Links:

[1] http://www.wirelessdesignmag.com/sites/wirelessdesignmag.com/files/legacyimages/1002/Dilithium_lrg.jpg