

A vision of quality

After much work on the quality of video telephony services over 3G networks, David Jack, director global support services, Dilithium Networks, reckons the road has never been straighter for this technology in waiting

Video telephony services over 3G networks are opening up a new range of possibilities for operators and service providers to allow users to communicate and interact with one another from any device, anywhere. These services, while still in their infancy, require careful design and planning to match user's high expectations of the 3G video experience.

This article identifies three key criteria that must be involved in the service redesign in order to meet the users' needs, namely 1. Call setup time, 2. Synchronisation and delay, and 3. Media quality. It also discusses the issues with packet-switched video telephony services.

Call setup time

Call setup is the time it takes before both video and voice start being transferred between two terminals. It involves the video call protocols running in the terminals and also the mobile network itself.

It is critical to the success of mobile video telephony that users are able to talk and see one another almost immediately after the call is connected and, within standards bodies, work is progressing to design and specify enhancements to the established call connection protocols to enable communications in the minimum time possible, to the level where any delay is insignificant to the users of the service.

3G Video calling is supported using circuit switched bearers in the mobile network, using the 3G-324M protocol for the negotiation of the video session.

The call setup time contributed by the



mobile network is around 4 seconds, and can be significantly longer for international calls, compared with around 2 seconds for voice-only calls. The networks are now being optimised to equalise this difference in an attempt to make video call setup as close to that seen for voice services as possible.

Synchronisation and delay

Two aspects that affect a user's perception of mobile video telephony are delay and synchronisation (lip-sync). Delay is a well-understood challenge in all forms of telephony, especially the mobile variety, and voice services on mobile networks typically aim to have an end-to-end latency of not more than 150ms. Lip-sync is a new challenge which was only introduced when video was added; it is accepted that systems must maintain a constant lip-sync of within 50ms to enable natural face-to-face conversation.

With mobile video services the underlying network delays are sometimes as much as 300ms, although work is ongoing to optimise the access network to make this as close as the standard voice services as possible.

Maintaining good synchronisation between audio and video channels is in general in direct contradiction to having low delays. It takes significantly longer to capture and encode a single frame of video for transmission than it does for the corresponding speech frames; this means that the audio data must be delayed to match the video sequence in order to maintain synchronisation between the streams, adding to the overall delay observed by the user.

The problems are most evident when using multi-media gateways or other systems, which rely on tandem transcoding, or decoding and recoding of the media stream. These types of gateways introduce lip-sync errors unless a compensatory delay is added. When used in conjunction with messaging servers, tandem transcoding problems can be further compounded

Maintaining good synchronisation between audio and video channels is in general in direct contradiction to having low delays



When errors occur, as is common in mobile networks, error-checking techniques are available to allow the terminals to quickly restore the corrupted video image

because lip-sync error is introduced on both record and playback or each message. Newer types of gateways and systems which are able to convert between media formats more efficiently, without the need to decode and recode the media stream, do not produce significant lip-synch related delays.

Media quality

Media Quality is a function of the digital compression technique (codec) used to encode the data for transmission.

Video telephony services use a 64kbps bearer channel and the terminals must be optimised to make the best use of this limited bandwidth by using advanced digital compression techniques and efficient sharing of bandwidth between audio and video. Video bit-rates are typically in the range of 45 - 55kbps, with a frame rate of 10 - 15 frames per second. Higher frame rates give better motion, but lower frame rates mean more bandwidth can be devoted to offering a sharper image, so this is an area where some balance must be made.

The digital chips used by the current generation of terminals are heavily loaded during video calls causing additional delay as well limiting the quality of video seen by the remote users. This further constrains the use of more advanced video compression standards, again reducing the quality of transmitted video.

As 3G terminals are enhanced with higher performance digital processors and are able to overcome the power limitations seen in the current generation, these quality constraints will no longer be a problem allowing higher quality compression techniques to be adopted.

This is easily demonstrated by a video

In the packet switched systems available today, there are insufficient Quality of Service controls to maintain a video call

call made between a laptop PC equipped with 3G video telephony software and UMTS data-card and a 'regular' video mobile phone. The images received by and displayed on the video mobile phone will be of a much higher quality in terms of sharpness and frame rate than those received by and shown on the laptop, revealing the level of quality that can be obtained over today's networks, and the limitations of today's mobile terminals.

When errors occur, as is common in mobile networks, error-checking techniques are available to allow the terminals to quickly restore the corrupted video image. This feature is built into all mobile terminals, but when calling with messaging servers and streaming applications there is the potential for trouble, as these systems traditionally do not support advanced error checking functions.

Advanced video gateways, designed to operate in error prone environments, are now available and these work with the messaging platform more closely to overcome these issues and maintain a higher quality service for the end user. Sadly these features are not available in all currently available video gateways.

Circuit-switched versus packet-switched

Circuit switched video telephony is currently the only viable way to deliver conversational video services, and this situation is likely to remain for a number of years. There is continued discussion on the future being with packet based video services. However, for the foreseeable future, these are unlikely to be able to offer the same quality as circuit based services. There are several reasons for this:

The available packet switched bandwidth is 384kbps downstream, and 64kbps upstream, but these figures are not guaranteed, they depend upon how many other users are accessing the packet switched services, and in heavily loaded cells significant packet loss can occur.

In the packet switched systems available today, there are insufficient Quality of Service controls to maintain a video call. When the IP overheads are applied the packet switched connection only provides around 60% of the payload capacity that is available using circuit switched based video telephony. Current services are therefore 'best efforts', providing a variable user experience.

The 3G-324M circuit switched protocols also incorporate a range of

error protection techniques specifically designed to operate in mobile environments. These provide the facility for terminals to detect and recover from error conditions as seen in the live networks very efficiently, and are not available for IP data transport (where the packets are instead re-transmitted when errors leading to excessive delays in noisy environments).

Until the packet switched services can overcome these limitations, circuit switched video services will continue to dominate through the use of optimised call connection times and improved quality video codecs. Further enhancements are underway to allow circuit switched video services to seamlessly inter-network with packet based applications such as IM and Presence services to enable a 'best of both worlds' solution.

In summary

Where once video conferencing services were restricted to dedicated facilities, or required the use of a broadband internet connections, users can now make video calls whenever and wherever they like.

Alongside the ability to make point-to-point video calls, operators are offering a range of video-enabled enhancements to network facilities, ranging from enhanced video messaging and IVR services, to portal and real time streaming services and multi-party conferencing. Customer service providers, such as banks and helpdesks, are enhancing their services with video-telephony, allowing them to offer a more personal service to their clients.

The road to widespread adoption of video telephony is now wide and straight. What was a few years ago the 'maybe' application for 3G services is now seen as a key element in the 3G portfolio offered to customers. Applications are being developed for consumer and business customers alike, and service providers are keen to take advantage of the benefits that face-to-face communications can bring in their interactions with customers.

Enhancements to Mobile UE terminals and the underlying video telephony protocols and standards, to the mobile networks and to media gateways and multi-media systems are currently underway to assist with the onset of widespread mobile video telephony, to ensure that complete interoperability exist: between the different services and to enable operators to bring services to market which will fully meet the expectations of their customers.

