Video QoS enhancement using perceptual quality metrics

D Hands, A Bourret and D Bayart

As the market for broadband video services matures, the ability to deliver high-value video content will become increasingly important. For the telecommunications industry to compete effectively with other video providers, it is vital that the quality of video services matches the expectations of customers. A major challenge lies in ensuring that the trade-off between price and quality is acceptable to consumers of broadband video. This paper introduces a new method for measuring in real-time the perceptual quality of video. The potential operational benefits of this method are discussed. The paper describes how a no reference video quality measurement method may be deployed as a mechanism for quality control at the point of video encoding and transmission. Further, it is proposed that a real-time video quality metric can be used to measure the quality received on end users' devices. By applying perceptual quality measurements for quality control and feedback, this mechanism can be used to ensure adequate quality is delivered to customers, make more efficient use of bandwidth and thereby reduce backhaul costs, and act as a quality assurance check on the customer's end device.

1. Introduction

High-value video services are becoming more prominent as part of broadband packages. In mainland Europe, a number of service providers (e.g. Fastweb, Free) offer high-quality video delivered over fixed-line broadband networks and telecommunications operators (e.g. France Telecom) are trialling video delivered over DSL. The success of broadband video will be primarily dependent on content and cost. However, the reproduction quality of the video will affect take-up and, critically, churn between operators and service providers. Currently, broadband is capable of providing a variety of video services, including video-on-demand (VoD) sourced from a central server, peer-to-peer downloaded video, streamed real-time video, and videoconferencing. For the highest value services, such as VoD, it is essential that consumers have a range of desirable content to purchase. Once purchased, consumers will demand that the sound and vision they experience is of good quality. For operators and providers alike, there is a need to consider the trade-off between quality and capacity, particularly at the bottle-necks, such as network routers and residential gateways.

Willis, in his paper [1], describes the range of factors affecting network quality of service (QoS). The proposition described here is concerned with QoS from an end user's perspective — in particular, the perceptual quality of broadband video. The perceptual quality of video is affected by a number of factors, including:

- video content (as the amount of detail and/or motion in a scene increases, then the video becomes more difficult to encode),
- encoding rate (typically, increasing the bit rate improves picture quality),
- coding scheme (recent advances in video-coding algorithms have resulted in improved picture quality, especially at bit rates of 2 Mbit/s or less),
- source video (a good quality source video will encode better than a poor quality source),
- network performance (packet loss can seriously degrade picture quality, and low latency is especially important for interactive video applications).

In addition to the factors listed above, video quality is influenced by customer premises equipment. For example, the performance capabilities of the video graphics card impacts on the reproduction quality of the video. The processing speed of the computer can influence perceptual video quality, although modern computer technology easily handles fast rendering of video. Finally, the properties of the display device can affect the perceptual quality of the video. A high specification display can present sharper images with high-quality colour representation.

This paper introduces a method for optimising quality and delivery of broadband video. The method is based on the use of objective perceptual quality metrics. These metrics, discussed in section 2, mirror human quality judgements of video quality. A real-time video quality metric, developed by BT, is described in section 3. Section 4 outlines how this metric is central to a perceptual quality control mechanism that can be used to deliver good-quality video to broadband customers. This new method provides a means of quality constancy, thereby providing improved customer experience compared to variable quality (constant bit-rate) services. Section 5 discusses how perceptual models may be employed in live networks for quality feedback that can be used for fault analysis and as an aid to helpdesks.

2. Perceptual quality measurement

Picture quality has been of interest to the television and video industry almost since moving pictures were shown to the public. The first reported studies of picture quality took place in the early part of the 20th Century [2] and a procession of experimental work has been published since [3—6]. During this time, the only reliable method for assessing the quality of video was to recruit a sample of subjects, present a series of test sequences to each subject, and obtain a set of subjective quality ratings for each test sequence. At the conclusion of testing, subjective quality ratings would be averaged across subjects to provide a mean opinion score (MOS) for each test condition. MOS remains the currency of image quality specialists and today subjective tests continue to play an important role in developing our understanding of the perceived quality of modern video systems and services.

Subjective tests, although valuable, are expensive in time and labour to perform. Further, in any single subjective test only a small sample of test conditions may be presented for assessment. For some twenty years, attempts have been made to define an objective method for measuring the perceptual quality of video. Objective perceptual video quality models are computational algorithms that process video files and output predicted mean opinion scores (MOSp) [7]. The goal of these models is to accurately mirror human ratings of video quality. This has proved a difficult and elusive goal. Since 1997, the work of the video quality experts group (VQEG) has been focused on testing, validating and committing for standardisation objective video quality models [8]. In 2003, the VQEG testing work identified objective models that were sufficiently accurate to merit standardisation. The resulting standard provides guidance on how to measure the quality of MPEG-2 video sequences using a full reference method [9]. Full reference methods have access to both the original (termed reference) and the processed video sequences. Both the reference and processed video are identical in terms of content, but may differ in terms of perceptual quality. Full reference methods compare the properties of the reference with those of the processed video. A MOSp is calculated by the method. BT's full reference video model is part of this new international standard.

The use of full reference models is primarily for laboratory-based performance testing of video services, applications or equipment. Full reference models tend to be non-real-time and require access to a very high quality original, ideally undegraded, video as well as the processed video. In an operational setting, full reference models are of limited use. BT has been working on a no reference (NR) video model where MOSp are obtained directly from the processed video. This NR method of video quality measurement is real-time and does not need access to the original video. The NR video model is therefore ideal for in-service applications such as quality monitoring and quality control.

3. No reference video quality model

The NR video model uses knowledge of the system being assessed or monitored to calculate MOSp. Once the compression algorithm is known, a set of detectors is activated. These detectors have been designed to specifically measure the presence and intensity of artefacts known to be produced during the bandwidth reduction and transmission process of particular coding schemes. These detectors are applied in parallel to the incoming decoded image stream under test. A generic set of detectors has been defined. The application of these detectors across different coding schemes has been developed. For different coding schemes (e.g. MPEG-2, MPEG-4, H.264) additional detectors can be called. This paper provides a general description of the NR model (see Fig 1).

- Temporal activity detection

  The temporal activity detector measures the average temporal activity present in a video sequence. This is achieved by measuring over time the number of pixels that change from one frame to the next.

  Three measures relevant to quality prediction are obtained from this detector. The first measure assesses the amount of white noise present in an image (spatially) and between adjacent frames (temporally). The presence of white noise is

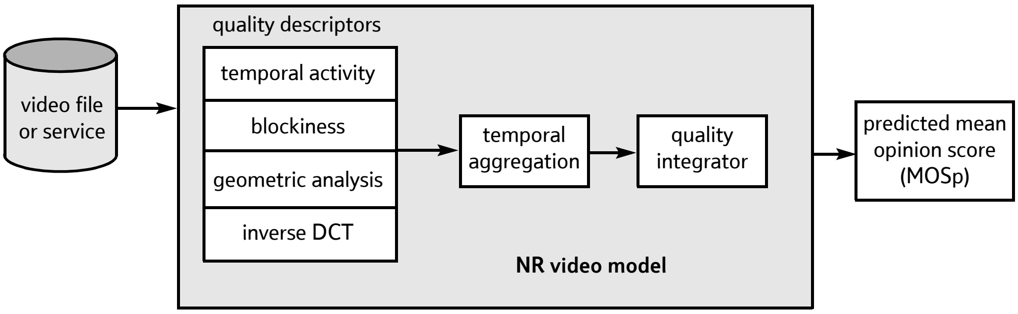

Fig 1 Schematic representation of the NR video model. This figure illustrates the general composition of the model. The model works on the decoded video information.

indicated by small variations in pixel values and is a feature of video capture using analogue cameras. As the amount of compression is increased, codecs raise the threshold necessary for a pixel change to be transmitted, removing this white noise.

A second measure tracks the temporal activity variation in a video sequence. This is specifically designed to detect the impact of the group of picture (GOP) structure used in MPEG codecs. In this family of codecs, a given frame can be either independently coded (called an I-frame), or coded with reference to other frames (B- and P-frames). In an uncompressed sequence, where motion flows naturally, the amount of change will tend to be constant, or will vary slowly from one frame to the next. When the sequence is heavily compressed, the amount of change takes a cyclical nature, the only time when a frame is entirely refreshed being when a new I-frame is transmitted.

The last process in analysing the level of temporal activity in a video sequence is to determine its frame rate, and count any missing frames. To reduce the bandwidth, the codec can decrease the number of frames sent every second. In some instances, e.g. buffer underflow, it can also decide to drop a frame. This is disturbing to the user and decreases the opinion score for the sequence.

- Blockiness

To reduce the amount of information to transmit, MPEG coders use discrete cosine transform (DCT) on $N \times N$ pixel blocks prior to thresholding and quantisation on the obtained coefficients. Using this technique, the resulting error when decoding is spread across the whole block, making it less visible. But this technique also reduces the continuity between blocks, making their boundaries more visible. This effect, called blockiness, can be measured by comparing the pixel values at the centre of blocks with the pixel values at the borders between adjacent blocks.

I-frames require more bandwidth to be transmitted, and therefore tend to be more compressed and hence more blocky. P- and B-frames contain correction information that improves the image definition on the basis of the last sent I-frame. As a result, blockiness levels vary in a cyclical way — blockiness tends to peak when I-frames are sent. Although this cycle is difficult to perceive, it can be reliably measured and constitutes a very useful indicator of the level of compression that has been applied to encode the sequence. For this reason, blockiness levels are analysed over time, in a similar way to temporal activity detection.

- Artificial geometrical features

MPEG codecs rely on a hierarchical structure starting from the pixel blocks, followed by the macroblocks, the slice, the frame and the GOP [10]. The organisation of the final bitstream, as well as some error-correction algorithms, is based on this structure. A consequence of this organisation is that any bitstream errors will have an impact on areas of the screen covering one of these elements. Most transmission errors will end up distorting blocks (8×8 pixels), macroblocks (16×16 pixels), or slices (16 pixels high, starting on a multiple of 16 pixels and propagating to the right side of the screen). The model searches the decoded image for straight lines following these characteristics.

- Inverse DCT to estimate number of coefficients

It has been shown [11] that the peak signal-to-noise ratio (PSNR) between a reference and coded sequence can be predicted by the number of quantisation steps used to simplify the representation of the DCT coefficients. This technique is used by the NR model. By recomputing the DCT coefficient for each set of $N \times N$ pixel blocks in the image, the number of quantisation steps present in the frame is estimated. This information is then used to obtain a PSNR estimate. PSNR on its own is not an especially reliable predictor of perceptual quality. However, a measure of PSNR is an important parameter when combined with other measures used to predict perceptual quality.

- Temporal aggregation

  All the measured parameters can produce values on a frame-by-frame basis. However, the properties of the human visual system, as well as judgement formation processes, do not tend to operate at such a fine level of granularity. Thus, the temporal aggregation function produces a single parameter value periodically. In the general NR model, parameter values are averaged every 12 frames.

- Integration function

  All the measured parameters are polled into a single number giving a prediction of the visual quality. This number must be as close as possible to the reproduction quality as perceived by an end user. An integration function has been defined to integrate all the outputs of the detectors to produce a single predicted quality rating. The integration function is identified using statistical regression methods, where the set of weighted parameters achieving the highest correlation between model predictions and subjective ratings is selected.

4. Model performance

The NR model was tested on the VQEG Phase I 625-line database [8]. This database contains a variety of video content including sports, film and graphics. The material was encoded using both H.263 and MPEG-2 at bit rates ranging from 768 kbit/s up to 50 Mbit/s. The database was split into a training set ($N = 80$) and a testing set ($N = 80$). The training and test datasets were separated on the basis of content. This meant that the model was tested on unknown video content. The training set was used to identify the best integration function. The performance of the NR model using the selected integration function was then evaluated against the test set.

In the training phase, parameter values were obtained for each video sequence. An iterative integration procedure was then applied to define the set of parameters that best predicted subjective quality across the training sequences. The best parameter set resulted in a correlation between subjective and objective scores of 0.74. This parameter set was used to predict the perceptual quality for the test sequences. The correlation between subjective and objective scores for the test sequences was 0.70. Across both the training and test sequences, this model produced a correlation between subjective and objective scores of 0.73. Fig 2 shows the association between subjective and objective quality scores.

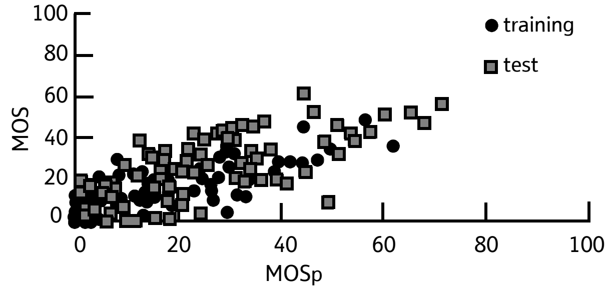

Fig 2 Correlation between subjective (MOS) and objective (MOSp) quality scores for VQEG Phase I 625-line video sequences.

It is worth stating that the accuracy of the model is expected to improve by at least 10 per cent before it is used operationally. NR models offer very powerful in-service measurement capabilities, as described below. This type of technology is already used for in-service monitoring of speech services. For example, NR speech quality measurement technology has been deployed for defining and policing SLAs [12]. NR video quality measurement models can perform a similar role for video services. The remainder of this paper presents a mechanism for controlling and monitoring video quality using the NR video-quality model. This mechanism has yet to be implemented, but is illustrative of the power offered by real-time perceptual quality metrics. Fundamental to this proposed mechanism is that measurements and operations are based on perceptual quality as opposed to more standard network measurements. If a service is attractive in price and content, then the perceptual quality becomes a key determinant of user satisfaction with the service. In addition, in a competitive market-place, perceptual quality can become a key differentiator between service providers.

5. A proposed quality control and feedback mechanism

At the beginning of this paper, two potential problems with broadband delivery of video were identified:

- providing adequate and consistent perceptual video quality to end users,

- making best use of available network capacity.

The authors are proposing a method to address both these problems. The method is based on two mechanisms.

- Control

  Firstly, a control mechanism maintains consistent perceptual video quality for transmission. Control operations are performed prior to transmission, at the video server or head-end. The commercial logic behind this proposed mechanism is that by using perceptually informed encoding decisions, end users will experience improved quality with no increase in cost to the provider or customer.

- Feedback

  Secondly, a feedback mechanism informs the provider or operator of the video quality actually experienced by the end user. Feedback measurements are taken at the point of reception (e.g. set-top box or PC in the home).

Figure 3 shows the proposed control and feedback system. The proposed mechanism uses information provided by real-time video-quality measurement tools, such as the NR video model presented above. A control unit is used to define the quality of the video service prior to transmission. Using predefined upper and lower quality threshold MOSp settings, a quality-control mechanism is provided. Encoded video is initially placed into a temporary buffer store. Measurements from the video-quality assessment tool are sent to a quality decision control unit. This unit has direct access to the video encoder and can control the bit rate at which the video stream is encoded. Whenever the quality measurements fall below the minimum threshold value, the control unit increases the encoder bit rate. If the maximum threshold value is exceeded, the control unit decreases the encoder bit rate. When the quality value falls within the threshold limits, the control unit sends a 'transmit' message to the buffer mechanism. Only when the 'transmit' message is received by the buffer mechanism is the video data transmitted.

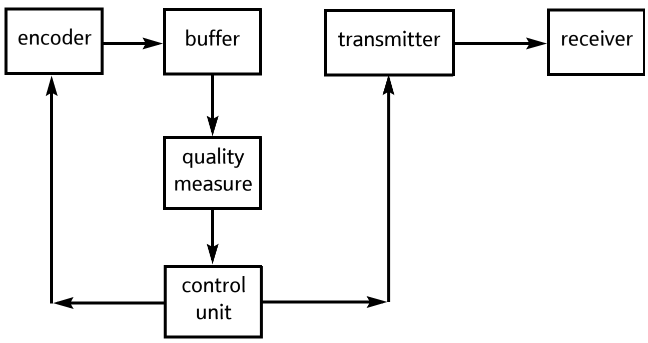

Fig 3 Schematic representation of the proposed quality control method.

If the network performance is perfect, then the quality of the video at source should be identical to the quality of video at the receiver. Where network performance is impaired, or if there are problems with customer premises equipment, then the quality at the receiver can be lower than at the point of transmission. Thus, the procedure at the receiver becomes:

- measure the quality at the receiver,
- consider whether the quality falls below some predefined acceptability threshold,
- if it does, feed back the quality value to the service provider or network operator in the form of an alarm.

How the service provider responds to alarms will be considered in section 6.

5.1 Video quality control

The control component works using the following principles. A video encoder is linked to a quality control unit. The encoder is responsible for compressing the video bit-stream prior to transmission. Typical encoders suitable for broadband applications include WM9 [13] and H.264 [14]. The quality control unit contains information regarding the upper and lower quality limits acceptable for the video service and a record of the current bit rate used by the encoder. The service provider is responsible for setting the threshold limits to accord with their performance requirements. The upper and lower thresholds will be defined in terms of MOSp. So, for example, a service provider aiming to offer a high-quality service may set the thresholds at MOSp = 3.5 (lower threshold) and MOSp = 4.5 (upper threshold). The control unit is connected to the encoder unit and can read and set the bit rate at which the video encoder operates.

Video services are delivered by firstly selecting the starting encoding bit rate, together with other encoding parameters (e.g. the image resolution, frame rate, single or dual pass encoding, and so on). The encoder then begins compressing the video data prior to transmission. Once the video data has been encoded, it is placed into a temporary data buffer. A real-time quality measurement tool, in this instance the NR perceptual video quality model, analyses the quality of the decoded video data held in the frame buffer. Quality measurements, reported for every *N* frames (e.g. based on group of pictures used by the encoder), are then passed to the quality control unit.

Note that, for meaningful perceptual quality measures to be obtained, the NR model relies on access to a sequence of frames. Typically, the GOP sequence is selected by the model as this provides useful information about how the video is encoded as well as being perceptually meaningful. The control unit uses these quality measurements to test whether the quality of the video service meets the performance criteria set by the operator.

If the quality measurement, as passed to the control unit, exceeds the upper quality threshold limit, the control unit changes the bit rate downward by some predetermined amount (e.g. 5%). For example, if the bit rate was initially 2 Mbit/s and the quality is measured as being greater than the upper threshold limit, the control unit sets the encoder to operate at 1.9 Mbit/s. If the quality measurement is less than the lower quality threshold limit, the control unit changes the bit rate upward by some predetermined amount (e.g. 5%). Using the example above, if the encoder initially encodes at 2 Mbit/s and the resulting quality measure is below the lower threshold value, the control unit sets the encoder to operate at 2.1 Mbit/s. This procedure runs iteratively until the quality falls within the threshold limits. When the measured video quality is within the upper and lower threshold values, the control unit sends a message to the buffer unit to transmit the video data. It is worth noting that the most effective means of performing the quality-control process is still under investigation. For example, the number of possible iterations will depend on the service. If video is being prepared for storage or archiving then a large number of iterations may be possible. For services where there is a short delay between encoding and transmitting, then only a few iterations will be possible. For instances where only a short delay is permissible between encoding and transmission, some shortcut will be needed to guide the control unit. One such shortcut, namely using a look-up table, is introduced in section 5.3.

5.2 Thresholding

In the description of the control mechanism above, it was noted that thresholds would be manually set. Thresholds should be set using the MOSp, as this is the basis of measurement produced by the perceptual model. In the simple scenario, single upper and lower thresholds are applied. Multiple thresholding can also be useful, for example in designing a service with differential pricing for quality. Presently, if a provider wishes to offer three levels of quality then the only effective method of achieving this is using bandwidth. So, a low-cost 'bronze' service may guarantee bit rates above 32 kbit/s. A medium priced 'silver' service may offer guaranteed bit rates above 128 kbit/s. A premium 'gold' service may guarantee bit rates always above 356 kbit/s. A more effective method for price differentiation, and indeed a method that can offer better customer satisfaction for bandwidth hungry video services, is to offer price differentiation based on perceptual quality. For example, minimum quaranteed MOSp values of 2, 3 and 4 for bronze, silver and gold services respectively. A European Commission 5th Framework project investigated the theory behind providing such a price-differentiated service [15]. In the proposed system, the quality control unit could be initiated with pairs of lower/upper thresholds, each threshold pairing linked to a particular class of service.

5.3 Limitations and applications of perceptual quality control

For streaming video, the maximum allowable delay between encoding and transmitting the video data will influence the number of possible quality adaptation iterations. There is scope for circumventing the iteration problem for streaming or real-time video applications. It is reasonable that experiments in the laboratory can be performed to examine the relationship between content, bit rate and coding scheme. A look-up table may then be compiled based on the experimental data. In the operational quality-control mechanism, video that is measured to be outside the quality bounds can be adjusted by making reference to the look-up table. For example, if the initial bit rate is 2 Mbit/s and the measured quality is below the minimum threshold value, based on knowledge that the video has a high spatial and temporal value, the look-up table may identify minimum acceptable quality obtained at 2.6 Mbit/s. Thus, the look-up table can act as a quality heuristic. Work is needed to define the look-up table. It should be noted here that the look-up table is providing a best-guess regarding the most suitable encoding rate. It is expected that the full iterative control procedure described above will provide the most accurate mechanism for controlling quality. Studies are necessary to quantify the trade-off between using the full iteration control mechanism and using the look-up table.

For encoding video for storage, the number of iterations is not time-limited and more efficient encoding can be achieved. For example, when preparing video for on-demand services it is preferable to ensure that the compression of the video content is optimised. This optimisation allows operators and providers to benefit in two ways. Firstly, video with consistent perceptual quality can be generated. Experimentation has shown that end users value stable quality over variable quality [16]. Thus, by using a perceptual quality tool to encode at constant quality should provide an immediate gain in terms of user satisfaction. Secondly, the use of the perceptual quality tool linked to a video encoder promises real benefits in bandwidth savings. It is well known that the quality of digitally encoded video is content dependent [17]. If a decision is made to encode at 5 Mbit/s across all content types, then there will be significant waste of network resource.

5.4 Quality control for multiplexing video channels

The control mechanism described above is for a single channel video service. The quality control mechanism is also applicable to optimise encoding between multiple channel video services. Using the same basic system described above, but adding differential thresholding and prioritising between channels, enables the control unit to optimise the quality of several video channels simultaneously.

Consider a video multiplexer that encodes ten channels. The service provider has finite bandwidth within which these ten channels may be transmitted. The control mechanism can optimise encoding between channels by applying the following procedure. The service provider first sets the upper and lower thresholds for each channel. The service provider may also assign a more stringent lower quality threshold to those channels that are of greatest value (e.g. those channels that incur a high subscription charge, such as film or sports channels).

The channels must then be prioritised. Equal prioritisation is acceptable, either for a few or for all channels. It is expected that in most cases some prioritisation will be present. For example, consider a video service provider who is delivering four premium-rate subscription channels, two non-premium sports channels, two free news channels and two free general content channels. The service provider may choose to protect the quality of the premium rate channels at the expense of all other channels, and to further protect the sports channels ahead of the news and general content channels.

In a multi-channel environment, the control unit must have some method for deciding which channels require quality protection. Consider the situation where the maximum available bandwidth is insufficient for even the minimum quality threshold to be met across all ten channels. If all channels are assigned equal priority, then the control unit will attempt to get each channel's quality level as close to the lower minimum threshold value as possible. Where one or more channels are assigned higher priority status, the control unit will first aim to get these channels' level of quality to (at least) meet the lower threshold value. The control unit will try to achieve a quality level for the remaining channels as close to the lower threshold value as possible.

5.5 Feedback of perceptual quality

The feedback component operates at the receiving unit where the video service terminates. Once the video data has been decoded, the data is passed into a frame buffer. Measurements produced by the NR model are passed to a feedback unit located on the receiving device. The feedback unit sends a message, using for example an IP connection, to the control unit located at the video source. The feedback unit's message may contain the following information:

- identity of receiver (e.g. IP address),
- frame number for which the quality measurement was extracted,
- quality measurement value (MOSp),
- video service from which the measurement was obtained,
- video player application,
- any statistics obtained by the video player (e.g. packet loss, latency, throughput).

The control unit, which stores a record of frame number and associated quality measurement prior to transmission, then records the following information:

- quality measure at source (MOSp(source)),
- quality measure at destination (MOSp(destination)),
- change in quality during transmission (MOSp(source) – MOSp(destination)).

Based on this data, the control unit can determine perceptual quality on an end-to-end basis. The presence and extent of any transmission errors can be assessed. Feedback of perceptual quality information is especially valuable for monitoring of service level agreements. By having access to measures of perceptual quality at source and destination, the feedback mechanism enables true insight into the effect of compression and transmission errors on user experience. Experimentation has shown that knowledge of packet loss can not be used to predict actual subjective experience of video quality. Similarly, perceptual quality cannot be predicted simply by having knowledge of how a video was compressed and the degree of compression. By calculating the perceptual quality of video services, together with information on network performance and the video compression applied, does allow for accurate and informative measurement of user experience to be obtained. This end-to-end quality monitoring capability can be used to ensure predefined quality guarantees (e.g. promised by a network operator) are met. Further, patterns of video quality variation across time will be readily available as will a capability to build profiles of end-user service usage and acceptability in relation to video quality.

6. Quality monitoring for fault analysis and helpdesks

Up to now, the focus has been on using NR video models for controlling the quality at the server and for feedback of quality delivered to the client. The NR model has other, more passive applications. In particular, applying the NR model for in-service quality monitoring provides a mechanism for fault identification and diagnosis and can be used as an aid to helpdesks. Basic network parameters can be used to identify problems with the delivery of services. Such network performance details are not suitable for obtaining a qualitative assessment of the effect of network errors. The use of perceptual models does provide a method for qualifying problems and this capability is particularly useful for classifying faults and onward reporting to helpdesks. The following discussions will centre on helpdesks, although the basic argument can be applied to other fault analysis tasks (e.g. reporting faults to engineers so that more effective repair decisions may be taken). Figure 4 provides a general overview of how perceptual quality metrics may be used to aid helpdesk operations.

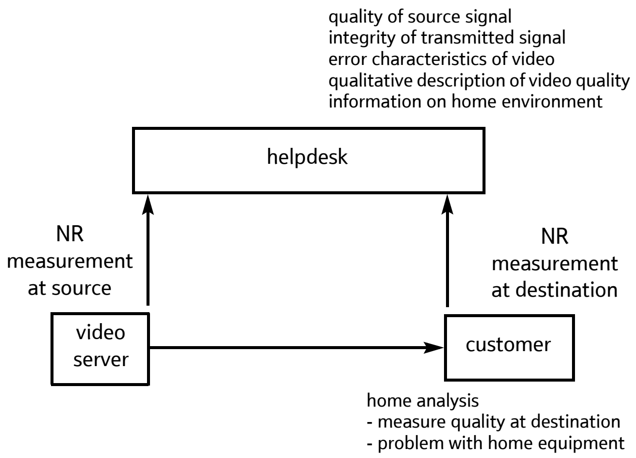

Fig 4 Example of how perceptual quality measures can be used to provide quality analysis to helpdesks. In this example, perceptual quality measures are taken at the point of transmission and point of reception. The software can then analyse the perceptual quality of the video and identify the source of the problem.

The problem for helpdesks is that operators can have difficulty interpreting service performance data when dealing with calls from customers. NR video quality metrics can be used to extract information at both source and destination regarding perceptual quality. Thus, a qualitative measure of end-to-end perceptual performance of the signal can be calculated by comparing the NR measure at source and NR measure obtained at destination. Further, the NR tool is capable of providing specific information regarding the nature of degradations (e.g. blockiness, jitter, blur).

Using the NR tool in conjunction with network performance measures, it is possible to identify whether the problem with a service is due to the network or to the equipment installed in the home. For example, if the video is good quality at source but the perceptual quality is inadequate at the destination, then it is important to identify the cause of the problem. Reference to network statistics can determine whether the cause was packet loss or latency in the network. If the network performance is intact, then examination of the home network can be performed.

The use of perceptual quality tools for helpdesk applications has four main benefits. Firstly, helpdesk operators will have easy-to-interpret data on quality of service delivered to customers. This information is expressed in mean opinion score terms rather than complex statistical data. Secondly, the perceptual tools can extract specific information describing degradations present in a service. This will assist operators in dealing with customer complaints, for example in both understanding the precise nature of any problem and in advising engineers sent to a customer's premises. Thirdly, the tool can help identify the root of problems. This can be especially important where the problem is located in the customer's home. Finally, where simple fixes can be applied (e.g. if the problem is due to some incorrect setting of the application), it is possible to either automatically resolve the problem or simply inform the customer via a screen message to change the setting. This will remove the need for the customer to contact the helpdesk at all.

7. Conclusions

In this paper an NR video quality model has been described. The performance of the model is good and as this technology matures the accuracy and reliability of the measures should improve yet further. The applications of this NR video model are widespread, for example it may be used as a quality control regulator to provide efficient encoding of video for storage or archiving, or for modifying video encoding bit rate to ensure the perceptual quality remains within some pre-set limits for live services, or as a quality monitoring agent on receivers. Here, a mechanism using an NR model for controlling the perceptual quality of both single and multiple video channels has been proposed. The control method here is based on the output of a video encoder being measured and then refined. Refinements are only made if the initial quality is outside some pre-set perceptual quality limits. At the end user's device, the NR model is used to measure the video quality delivered to the customer. The ability to measure perceptual quality at both the point of transmission and the point of reception offers service providers and operators a powerful end-to-end QoS analysis tool. The real power of this tool is that it offers perceptually based measurements of quality. As such, the end-to-end quality monitoring system can be used to improve customer satisfaction by identifying where and when quality falls below some standard of acceptability — engineers can then be informed so that problems can be resolved quickly and efficiently. The ability to control quality prior to transmission has the benefit of ensuring consistent and acceptable quality as well as making most efficient use of bandwidth.

References

1 Willis P J: ‘An introduction to quality of service’, BT Technol J, 23, No 2, pp 13—27 (April 2005).

2 Engstrom E: ‘A study of television image characteristics’, Proceedings of the Institute of Radio Engineers, 21, pp 1631—1651 (1933).

3 Allnatt J W: ‘Transmitted-picture Assessment’, Wiley, London (1983).

4 Winkler S, Dufaux F: ‘Video quality evaluation for mobile streaming applications’, Visual Communications and Image Processing 2003, SPIE-Int Soc Opt Eng Proceedings of SPIE — The International Society for Optical Engineering, 5150, No 1, pp 593—603 (2003).

5 Corriveau P, Gojmerac C, Hughes B and Stelmach L: ‘All subjective scales are not created equal: the effects of context on different scales’, Signal Processing, 77, No 1, pp 1—9 (August 1999).

6 Hands D: ‘Temporal characterisation of the forgiveness effect’, Electronics Letters, 37, pp 752—754 (2001).

7 Voelcker R, Hands D, Davis A, Bourret A and Bayart D: ‘Video quality measurement — predicting subjective quality using perceptual models’,The Journal of the Communications Network, 2, Part 1, pp 40—45 (March 2003).

8 Video Quality Experts Group — http://www.vqeg.org/

9 ITU-T Recommendation J.144: ‘Objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference’, Geneva (May 2003).

10 Ghanbari M: ‘Video coding: an introduction to standard codecs’, The Institution of Electrical Engineers, London (1999).

11 Bourret A J and Ghanbari M: ‘Video quality metric’, Patent Number: WO04054274 (2002).

12 Psytechnics — http://www.psytechnics.com/

13 Srinivasan S, Hsu P, Holcolb J, Mukerjee K, Regunathan S L, Lin B, Liang J, Lee M-C and Ribas-Corbera J: ‘Windows Media Video 9: overview and applications’, Signal Processing: Image Communication, 19, No 9, pp 851—875 (October 2004).

14 ITU-T Recommendation H.264: ‘Advanced video coding for generic audiovisual services’, Geneva (May 2003).

15 European Union Fifth Framework Project Number IST-1999-11429: ‘Market Managed Multiservice Internet’, — http://www.m3i.org/

16 Bouch A and Sasse M A: ‘The case for predictable network service’, in Nahrstedt K and Feng W (Eds): ‘Proc MMCN’2000’, San Jose, CA, pp 188—195 (January 2000).

17 Aldridge R, Davidoff J, Ghanbari M, Hands D and Pearson D: ‘Measurement of scene-dependent quality variations in digitally coded television pictures’, IEE Proceedings on Vision, Signal and Image Processing, 142, pp 149—154 (1995).

David Hands is currently a group leader in BT at Adastral Park. Since joining BT in 1997 he has worked on various projects associated with quality measurement of video and multimedia services.

He has been actively involved in both subjective and objective video quality measurement research for over 10 years.

He is co-chair of the VQEG working group on multimedia and is Associate Rapporteur for Study Group 9 of the ITU-T.

Alex Bourret graduated from the ENIC (National School of Telecommunication Engineering, Villeneuve D’Ascq, France) in 1995. After a placement with BT at Adastral Park, working on stereo vision techniques, he joined the University of Essex as a research assistant in the Department of Electronic Systems Engineering.

Since joining BT in 1997 he has been researching objective methods for the assessment of the perceived quality of digital images. He has been co-chair of the VQEG-RRNR group since 2004.

Damien Bayart has been working on issues regarding perceptual video quality assessment since joining BT at Adastral Park in the summer of 2001 after graduating from the National School of Telecommunication Engineering (ENIC), France.

He is currently working in the Perceptual Engineering Group, part of the Broadband Applications Research Centre. He worked on the development of the intrusive perceptual quality model. He is now working on the integration of the non-intrusive model.