Intelligent Bandwidth Allocation of IPTV Streams with Bitstream Complexity Measures

Sandro Moiron
Instituto de Telecomunicações, Lisboa, Portugal
smoiro@essex.ac.uk

Rouzbeh Razavi
Bell Laboratories, Alcatel-Lucent, Dublin, Republic of Ireland
rouzbeh.razavi@alcatel-lucent.com

*Martin Fleury
University of Essex, Colchester CO4 3SQ, United Kingdom
fleum@essex.ac.uk

Mohammed Ghanbari
University of Essex, Colchester CO4 3SQ, United Kingdom
ghan@essex.ac.uk

*Contact author: Martin Fleury
E-mail: fleum@essex.ac.uk

Abstract

IPTV video services are increasingly being considered for delivery to mobile devices over broadband wireless access networks. The IPTV streams or channels are multiplexed together for transport across an IP core network prior to distribution across the access network. According to the type of access network, prior bandwidth constraints exist that restrict the multiplex data-rate. This paper presents a bandwidth allocation scheme based on content complexity to equalize the overall video quality of the IPTV sub-streams, in effect a form of statistical multiplexing. Bandwidth adaptation is achieved through a bank of bit-rate transcoders. Complexity metrics serve to estimate the appropriate bandwidth share for each stream, prior to distribution over a wireless or ADSL access network. These metrics are derived after entropy decoding of the input compressed bit-streams, without the delay resulting from a full decode. Fuzzy-logic control serves to adjust the balance between spatial and temporal coding complexity. The paper examines constant and varying bandwidth scenarios. Experimental results show a significant overall gain in video quality in comparison to a fixed bandwidth allocation.

Intelligent Bandwidth Allocation of IPTV Streams with Bitstream Complexity Measures

INTRODUCTION

Internet Protocol TV (IPTV) media streams are delivered across converged telecommunications networks to a home network end device such as a set-top box or PC, or increasingly to mobile devices (Singh et al., 2008). The use of managed networks and IP framing differs from (Maisonneuve et al., 2009) Internet or Web TV, which is streamed over best-effort IP networks. Before distribution to individual devices, multiple videos streams will share a multimedia sub-channel. Other types of data will form other sub-channels or pipes. TV channels may be multiplexed onto an MPEG-2 transport stream (TS), and it is also possible (Wagner et al., 2009) to employ the Real-time Transport Protocol (RTP) to encapsulate MPEG-2 packets (usually between 6 and 7 MPEG-2 TS packets, each representing a TV channel, per RTP packet). Broadcasters have generally employed a Constant Bit-Rate (CBR) multiplex of streams (Böröczy et al., 1999) previously stored at a high quality. This paper is essentially about how to achieve the bandwidth allocation of the TV channels or video sub-streams within the multimedia sub-channel.

When the video sub-streams within the multimedia sub-channel leave a managed core-network, they will be accessed over various network types such as Asymmetric Digital Subscriber Line (ADSL) (Zheng & Lin, 2000), broadband wireless including IEEE 802.16 (WiMAX) (She et al., 2007), or a passive optical network (Hermsmeyer et al., 2007). For example, in Lee et al. (2009) the architecture of an IPTV scheme for mobile WiMAX is illustrated. The IPTV content passes over the core IP network before reaching an access control router to a set of WiMAX base stations. At that point, in the solution of Lee et al. (2009) a controller for the WiMAX Multicast and Broadcast Service (MBS) intervenes at the access control router. Popular content is extracted from the IPTV multimedia sub-stream is extracted for distribution via MBS as near-Video-on-Demand (NVoD), while other content is sent as true VoD by unicast to mobile devices. Where IPTV differs from terrestrial and normal satellite distribution is that the MPEG2-TS or RTP multiplex content mix can be made flexible through session control feedback mechanisms.

As the IPTV bandwidth may be constrained by a particular type of access network technology, a practical solution, (Kasai et al., 2005) is to employ a transcoder bank to resize the video streams' bit-rates to fit within the constraints of the target access network, especially when the access network is a mobile one. Transcoding can dynamically and selectively modify the bit-rate of each stream within the multimedia sub-channel in order to fit the available bandwidth of the access network channel, a form of statistical multiplexing. In this way, IPTV can be extended to mobile devices.

However, allocating bandwidth to video streams simply on the basis of efficient usage and fair distribution of bandwidth, for example in Jain et al. (1996), is not necessarily wise, because the delivered video quality of some video streams will be more affected by a reduction in bandwidth than by others. Both unwarranted degradation of quality and unnecessarily high video quality may arise. This is also the reason why allocating bandwidth based on the past statistics of data-rates may be ill-advised, as it fails to account for the impact of such allocations on the delivered video quality. A further weakness of statistical allocation for variable bit-rate video is that such approaches are not appropriate if the input data-rates have high variances. Smoothing of the date-rates may be applied to remove the impact of intra-coded I-frame data but this can affect latencies. Thus, video streams within a multimedia sub-channel cannot have bandwidth allocated in the same way as other data but should take account of the video stream content in a dynamic manner.

The goal of our paper's statistical multiplexing scheme is to dynamically adjust the bandwidth share between several concurrent streams based upon their content complexity in order to equalize their delivered video quality. Ideally, the quality of all video streams will then fall within an acceptable range, being neither too high nor too low in quality. Broadcast quality video normally falls within the range 30–38 dB. As an illustration, consider the rate-distortion (R-D) curves in Figure 1 for three reference video sequences. At an initial target input rate of 1 Mbps, the objective quality of the *Mobile* video sequence is on the boundary of that range while the quality of both the *Highway* and *Bridge-closed* sequences exceeds that range. Hence, there is a need to equalize their objective quality. The proposed scheme computes spatial and temporal complexity measurements by extracting the transform coefficients and motion vectors after entropy decoding has taken place but before a full decode has occurred. This procedure allows the use of frequency-domain transcoders (Nam et al., 2006), which reduce the latency and computational complexity of the joint transcoding system.

In our statistical multiplexing scheme, R-D analysis is turned on at H.264/AVC encoders so that all rate decisions are optimized according to their effect on video quality. The problem of lookahead is resolved by directly transcoding each video stream Group of Pictures (GOP) according to a joint estimation across the concurrent streams without the need for complex forward inspection of complexity (two-pass encoding) or potentially erroneous predictions. A fuzzy logic controller (FLC) merges the spatial and temporal complexity metrics. Notice that temporal complexity has been included. Depending on the original target bit-rate (Seeling & Reisslein, 2005) the motion vectors may have limited impact on the overall bit-rate but a comprehensive scheme cannot neglect the effect of temporal complexity. The result of the implemented scheme is that within a desired quality range there is a significant gain in overall video quality for the video streams within the multimedia channel.

This paper considers allocation of bandwidth within the multimedia sub-channel to each of the video streams. It also assesses the robustness of that allocation after the individual video streams are distributed over a particular access network type. The paper's main contributions are a method that makes predictions of future bandwidth requirements based on coding complexity rather than statistical analysis of video stream data-rates and the design of the FLC that results in an overall allocation gain in video quality. The complexity indexes employed can be retrieved from the coded bitstream without a full decode or cross-layer interaction with the source, and as such are compatible with existing systems. A preliminary conference version of this paper (Moiron et al., 2010) has been expanded to: improve the presentation of statistical multiplexing. The FLC has also been described in further detail. Further tests have verified the operation of the proposed scheme across a range of scenarios.

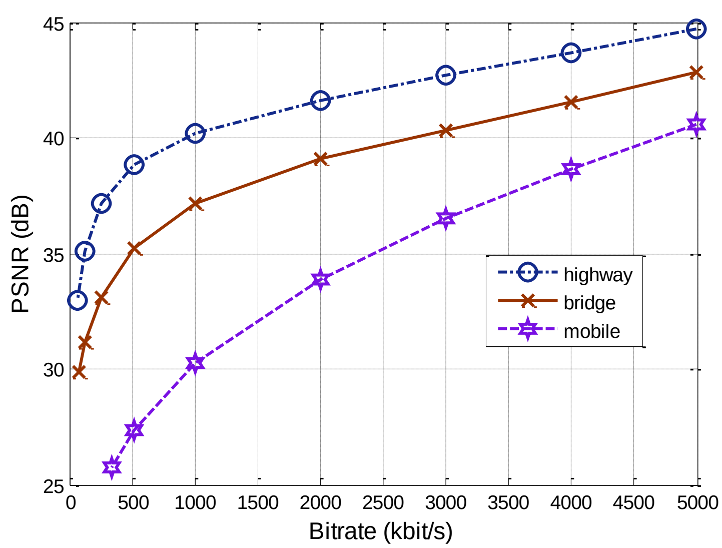

Figure 1. R-D curves for three example video sequences.

For the purposes of this paper, we consider the available bandwidth for cable and ADSL access networks as essentially fixed. Even if the multimedia sub-channel available bandwidth is considered as essentially fixed there is still a need to allocate that bandwidth within the sub-channel. In the case of point-to-multipoint (PMP) broadband wireless access, available bandwidth is assumed to vary over time owing to the arrival of other types of traffic (peer-to-peer file download, web pages, voice, …) sharing the wireless channel, apart from the video streams delivered over the network core by the multimedia sub-channel. Allocation of available bandwidth is made for the multimedia sub-channel passing across the network core prior to distribution over the access network, and as such is oblivious to the physical channel characteristics of the access network. Nevertheless, our results consider the impact of subsequent ADSL and wireless errors on the delivered video quality, because the allocation should be sensitive to their impact. Of course, application-layer forward error correction is also possible but its contribution is beyond the scope of this paper.

The compressed video streams forming the IPTV sub-channel will not necessarily have the same bandwidth requirements, as their coding complexity will vary over time with changes in their spatial and temporal complexity. In the long term, this variation is determined by the video genre, such as sport, cartoon, soap, and so on (Agboma & Liotta, 2006) but there are also changes over a shorter time period, caused by such factors as the type of frame and whether there is a shot change or a scene cut. Consequently, multiple video streams sharing the sub-channel can each be adaptively allocated a proportion of the available bandwidth according to their coding complexity. To reduce decision latency and to create a more direct way of judging the coding complexity, we use metrics derived from the encoded bitstream. Entropy decoding is required but this is a small overhead compared to a full decode. For example in H.264/Advanced Video Codec (AVC) bit-stream parsing and Context Adaptive Variable Length decode on average take up only 13% of the computational complexity of a full decode (Horowitz et al., 2003).

In the paper, temporal complexity is indicated by a count of per-frame non-zero motion vectors summed across a Group of Pictures (GOP), whereas spatial complexity is found (Rosdiana & Ghanbari, 2000) by averaging a Scene Complexity Index (SCI) (ITU-T, 1992) across the GOP. It is also possible to make decisions at scene change boundaries or through a GOP-sized sliding window, at a cost in complexity but with a gain in reaction time. Because a large proportion of the bit-stream’s length is contributed to by quantized transform coefficients, the weighting given to the SI metric is increased through the decision rules for the rate controller. From the input video streams (coded in Constant Bit-Rate (CBR) mode) the average per-frame quantization parameter (QP) is modulated by the target bit rate to find the SCI. Consequently, multiple video streams sharing the sub-channel can each be adaptively allocated a proportion of the available bandwidth according to their instantaneous spatial and temporal complexity. The target output rate of the video streams was changed accordingly at the encoder after each decision point (every GOP).

The translation of joint spatial and temporal complexity into future bandwidth needs is an uncertain or imprecise process, suggesting that fuzzy logic control is appropriate for this task, rather than a mathematical formulation. An FLC is suited to hardware implementation, implying that real-time operation is possible. Within video coding fuzzy ogic control has already found an application (Grant et al., 1997) in maintaining a constant video rate by varying the encoder quantization parameter according to the output buffer state, which is a complex control problem without an analytical solution. In fact, fuzzy logic control for video applications has gained acceptance (Jammeh et al., 2008; Rezaei et al., 2008) for congestion control and rate control.

For distribution to a broadband wireless access network, the available bandwidth for the IPTV sub-channel is assumed to vary according to a four-state Markov chain. A similar four-state model was adopted in the research reported in Ma & El Zarki (1999) for bandwidth reservation of fixed wireless channels in a video-on-demand service. However, the number of retransmissions of each video frame acts as an adaptive bandwidth reservation indicator, with no account taken of video content. In our work, a linear predictive filter (LPF) serves to predict the likely available bandwidth based on monitoring the history of prior availability of bandwidth to the IPTV sub-channel. The research in both Ma & El Zharki (1999) and Adas (1998) also employed linear prediction for similar purposes.

In general, too rapid (per-frame) available bandwidth share allocations are avoided in order not to cause an unsettling subjective effect on the viewer, when the matched video streams are adjusted to the allocation. Adjusting the bandwidth share based on an average (over a GOP) of past Temporal Indices (TIs) and SIs also overcomes possible signaling latency in adjustment of the video streams' allocation within the IPTV sub-channel. In the scheme modeled, rate control of CBR streams is achieved directly within the encoder but, as previously mentioned, it is also possible to derive the spatial and temporal metrics after entropy decode. In which case, the method can be applied to pre-encoded video through H.264 bit-rate transcoding (Nim et al., 2006).

Much of the previous work on allocation of bandwidth to video streams was conducted for Asynchronous Transfer Mode (ATM) networks and for Digital Video Broadcast-Terrestial (DVB-T) television. The following Section briefly discusses this background. The proposed bandwidth allocation method is then given in detail, while the following Section evaluates the methods across the selected three access network technologies. The final Section of the paper draws some conclusions and suggest future developments.

BACKGROUND

Video stream bandwidth allocation has some affinity to statistical multiplexing except that, rather than output compressed Variable Bit-Rate (VBR) video to a common buffer, in bandwidth allocation the CBR bit-rate is adapted according to the available bandwidth. An early study of videophone statistical multiplexing (Haskell, 1972) showed that multiplexing encoder outputs to a shared buffer resulted in significant multiplexing gain. Numerous studies have modeled video output as a finite-state Markov chain to such an extent that Heeke (1993) considered altering the encoder to behave in this way to aid the design of ATM network admission control. A fluid-flow model (Antick et al., 1982) is an alternative form of VBR video output model and in Bashworth & Williamson (1998) video streams are modeled as self-similar flows at the GOP level. However, these are statistical models, which do not directly account for the encoder's behavior.

Closer to the work of this paper, in Böröczy et al. (1997) bandwidth allocations between CBR streams at GOP boundaries occurred, with a refinement to include scene-change boundaries. In Böröczy et al. (1997), joint control through complexity statistics (spatial only) was also applied to a set of R-D controlled video encoders (for MPEG-2 before rate-distortion was built in to encoders). However, the rate decisions made were deterministic, based on the complexity statistics, whereas in this current work an FLC makes decisions. The scheme (Böröczy et al., 1997) also suffered from the problem of scene changes occurring within a GOP inspection window, as the complexity may change significantly within a GOP. Some allowance for this problem was made by a sliding window GOP prediction method. As an alternative to R-D analysis it is also possible to predict complexity prior to encoding as in Guha &

Reininger (1994), though this method apart from its complexity cannot be adapted to bit-rate transcoding of pre-encoded video. In our system, R-D analysis is turned on in the H.264/AVC encoder so that all rate decisions are optimized according to their effect on video quality, apart from the relative adjustment in quality between the video streams within the multimedia sub-channel.

Statistical multiplexing techniques vary according to their complexity. In Wang & Vincent (1996), a relatively simple form of statistical multiplexing occurred, in which the same QP was applied to all video frames within a multiplexed group to achieve an overall target bit rate. A binary chop search across the range of available QPs was conducted. This procedure in the tests presented (Wang & Vincent, 2996) appeared to achieve its objective, even though no direct account was taken of content complexity. The alternative is to partially decode future frames, as occured in He & Wu (2008). Unfortunately, in He & Wu (2008), only the temporal complexity measure was found by partial decode, while the spatial complexity was predicted from a previous frame. Overall, it is reported in Vukadinovic & Huschke (2008) that little prior research has been conducted on statistical multiplexing of H.264/AVC streams, even though this has now become the preferred codec for emerging national applications of HDTV and within wireless systems such as 3GPP's Multimedia Broadcast Multicast Service (Afzal et al., 2006).

METHODOLOGY

A top-level diagram of the proposed statistical multiplexing system is presented in Figure 2. Video stream input is either from live video streams, such as from a sporting event, or from pre-encoded video, possibly adapted for TV display. The complexity measures (TI and SI) are output by the encoders for input to the FLC. The FLC also receives notice of the available bandwidth for the multimedia sub-channel from the managed IP network control layer. Final access links may be cable, ADSL, or broadband wireless. Another candidate technology is a passive optical network (PON), in which, as we also assume for cable, following (Muntean, 2006), the physical channel is virtually error free. ADSL is subject to burst errors following the Repetitive Electrical Impulse Noise (REIN) model (Luby et al., 2008) and broadband wireless is characteristically affected by error bursts arising from slow fading, along with superimposed RF noise and interference from fast fading. At the Network Distribution Point in Figure 2, the multimedia sub-channel emerges from the core network and is split into its constituent video streams in the case of cable and ADSL, while broadband delivery is PMP (broadcast). An important point to note is that the FLC is unaware of the ‘last-mile’ technology at the time of IPTV bandwidth allocation.

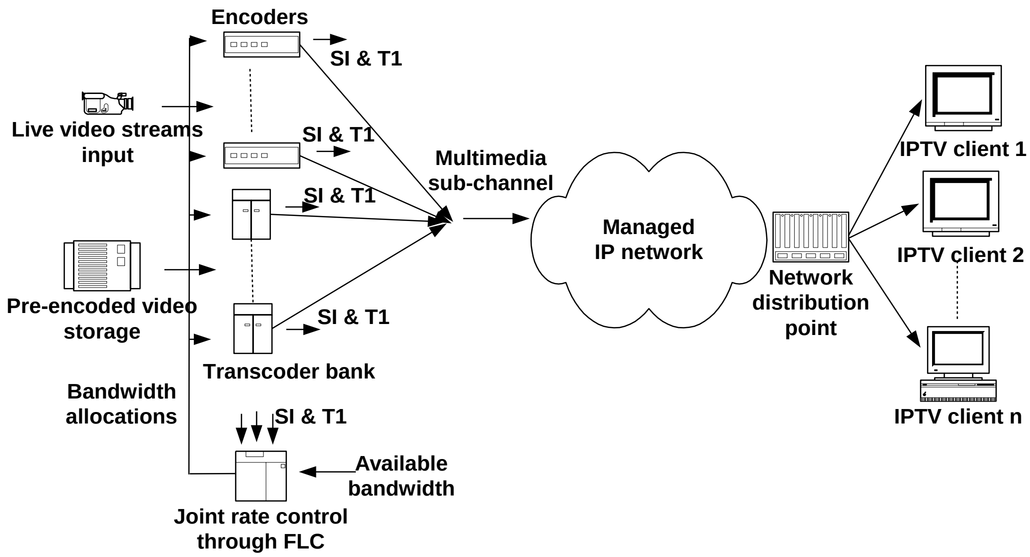

Figure 2. IPTV streams over a broadband network with FLC bandwidth allocation.

Once the FLC has determined the proportions allocated to each video stream sharing the IPTV sub-channel, the video bit-stream rates are jointly adjusted, either by bit-rate transcoders or, if live video, by direct adjustment of the QP. Because the current complexity metrics are based on live video, for pre-encoded video the video complexity data are either derived from the bit-stream or possibly pre-calculated and sent as side information.

Input video characteristics

For tests, 900 frames from three well-known sequences were selected (Table 1) with approximate coding complexity characterized as difficult, medium, and easy. The JM v. 14.2 codec was used with Common Intermediate Format (CIF)-30 Hz @ 1 Mbps, 4:2:0 sampling and a GOP size of 15. An IPPP… GOP structure was set with Instantaneous Decoder Refresh (IDR) frames configured. R-D control was set for the CBR output with initial QP at 28 and 4 × 4 DCTs only.

From an R-D plot of the test videos, Figure 1, as mentioned in the Introduction, it is apparent that comparing the quality rankings, there is a significant difference between the videos. Therefore, the goal of this work is to adjust the quality between the streams relative to their coding complexity rates over time. To do so, a desired quality range was aimed at and in this work that range was 30–38 dB. At an initial target input rate of 1 Mbps, the quality of *Mobile* is on the boundary of that range, while the quality of both *Highway* and *Bridge* exceeds the range.

Figure 3 is a plot of the per GOP number of non-zero motion vectors over time. For ease of representation, an average of the first 20 GOPs is plotted, though naturally the entire 60 GOP sequences were examined in our investigations. The TI measure for *Mobile* fluctuates over time within this short excerpt, though recall that the T1 emphasis is reduced in the FLC compared to the SI. Figure 4 is a matching plot for the SCI over time. In general, the SCI is found as:

$$SCI = \frac{1}{1+p+b}\left[IQ_I + \sum_{j=1}^{p} P_j Q_{Pj} + \sum_{j=1}^{b} B_j Q_{Bj} \right] \qquad (1)$$

where p, b are the number of P-, B-pictures per GOP respectively and QI,, QP and QB are the respective average (over the picture's macroblocks) quantization step sizes for the I-,P-, and B- pictures. I, P and B are the corresponding per picture type target bit-rates, though in practice there is just one target rate that can be set in the JM H.264/AVC configuration. Again, in Figure 4 it is the *Mobile* sequence that shows the most fluctuation. For comparison Figure 5 shows another candidate SI: the non-zero coefficient numbers for the three sequences. The advantage of including QPs into (1) is that the changing dynamic range of the coefficients is taken into account. Comparing Figures 4 and 5, it is apparent that including QPs trims the relative offset between *Bridge* and *Mobile*, though of course this is just one example comparison. For completeness, Figure 6 illustrates the average per-GOP QP reported by the encoder software. The reader should compare between Figures 4, 5 and 6 to judge the effect of including QPs.

Figure 7 shows the input Peak Signal-to-Noise Ratios (PSNRs) for the three sequences, showing fluctuations in behavior. Notice that, as PSNR is relative to each individual sequence (Huynh-Thu & Ghanbari, 2008), no direct comparison of objective video quality between the sequences should be inferred, though it is clear that there is an inverse ranking according to coding complexity.

Table 1. Input test sequences.

Sequence Name	*Original Length (frames)*	*Selection procedure*	*Category*
Mobile	300	Repeat 3 times	difficult
Bridge-close	2000	First 900 frames	medium
Highway	2000	Range [600-1500]	easy

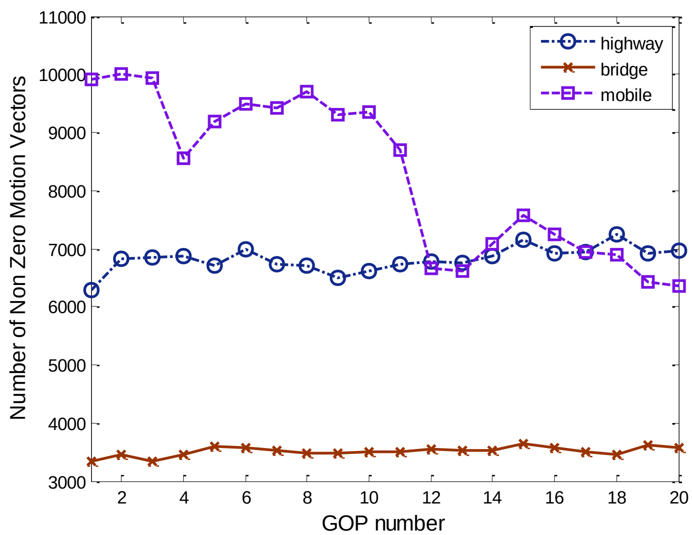

Figure 3. Per GOP count of non-zero motion vectors.

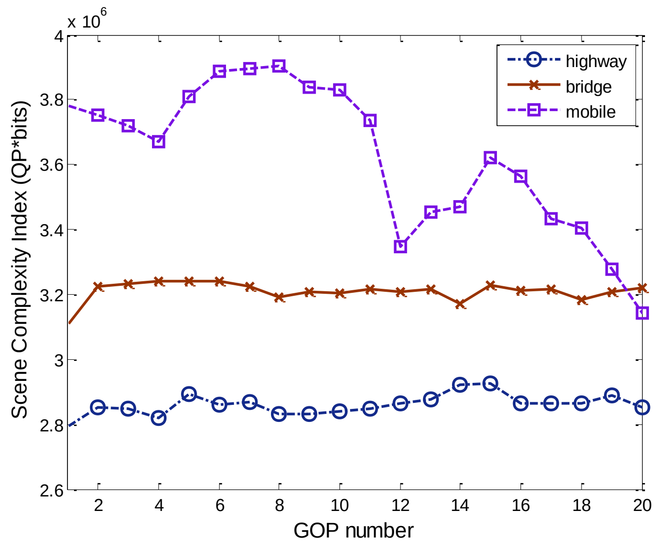

Figure 4. GOP no. versus average Scene Complexity Index.

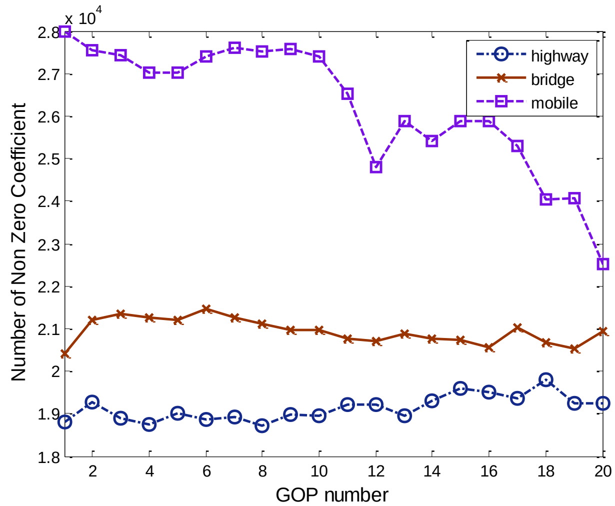

Figure 5. Per GOP no. of non-zero coefficients.

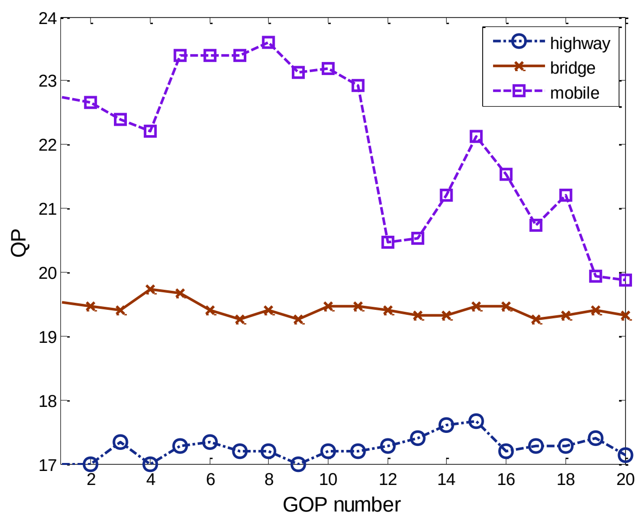

Figure 6. Average per-GOP QP reported by H.264 encoder**.**

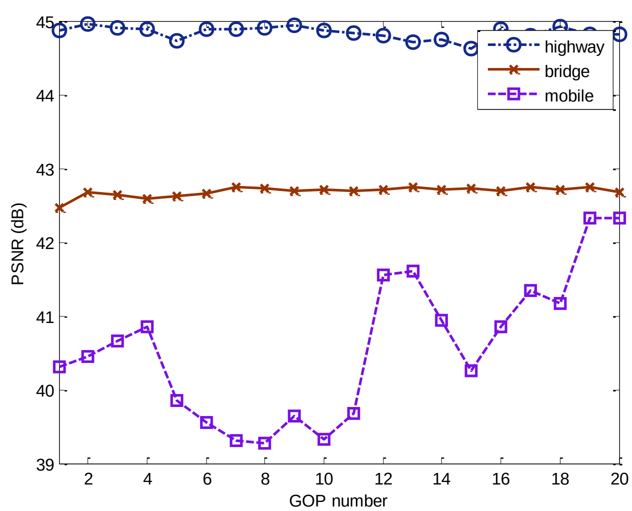

Figure 7. Changing input PSNRs showing the fluctuation of PSNR for individual sequences.

Fuzzy logic control

Before detailing the FLC for bandwidth allocation, this Section first introduces the general form of an FLC. A fuzzy set is expressed as a set of rules which take the form of linguistic expressions. These rules express human experience of tuning an application and, in the methodology, are captured in a knowledge database. An inference engine is the intelligence of an FLC, with the capability of emulating the human decision making process, based on fuzzy logic, by means of the knowledge database and embedded rules for making those decisions. Lastly, the de-fuzzification process converts inferred fuzzy control decisions from the inference engine to a crisp or precise value that acts as a control signal.

In a fuzzy subset, each subset member is an ordered pair, with the first element of the pair being a member of a set *S* and the second element being the possibility, in the interval [0, 1], that the member is in the fuzzy subset. This should be compared with a Boolean subset in which every member of a set *S* is a member of the subset with probability taken from the set {0, 1}, in which a probability of 1 represents certain membership and 0 represents non-membership. As a simple example, in a fuzzy subset of (say) ‘tall’, the possibility that a person with a given height taken from the set *S* of heights may be called tall is modeled by a membership function, which is the mapping between a data value and possible membership of the subset. Notice that a member of one fuzzy subset can be a member of another fuzzy subset with the same or a different possibility. Membership functions may be combined according to a set of ‘if . . . then’ rules to make inferences such as if x is tall and y is old then z is happy, in which tall, old and happy are membership functions of the matching fuzzy subsets and x, y, z are linguistic variables (names for known data values). A crisp control value (Ctrl) is found through the process of de-fuzzification, with (2) representing de-fuzzification by the standard center of gravity method.

$$Ctrl = \frac{\sum_{i=1}^{M} S_i K_i}{\sum_{i=1}^{M} K_i} \qquad (2)$$

where M is the number of rules, S_i is the value of the output for rule *i*, K_i is the inferred weight of the *i*th output membership function.

More specifically, S_i is the value at the middle of the range of data values that are possible members of the *i*th fuzzy subset. K_i is the area under the *i*th output membership function, clipped by the minimum possibility of membership of the input membership function of the ith rule. As more and more rules may apply, (2) is in the form of a weighted average of the values arising from the different rules that are applicable, with outputs of zero for those remaining fuzzy subsets for which the data value is not a member.

The inputs to the FLC were the two complexity measures, TI and SI, which were used to determine the bandwidth allocation of any one video stream. The TI and SI inputs to the fuzzy models were first normalized by dividing by the largest value of TI and SI respectively for the set of samples from each of the current frames of all video streams sharing the sub-channel.

The fuzzy models for inputs TI and SI, being a number of overlapping membership functions are shown in Figure 8. These are typically triangular at a cost in smoothness to allow rapid calculation of output, possibly with a hardware unit (Batarone et al., 2000). The inputs were combined according to the common Mamdani inference method (Mamdani & Assilia, 2000) to produce the output values from triangular output membership functions similar to those of the input models, according to the rule set given in Table 2. For example, if TI is ‘high’ and SI is ‘medium’ then output is ‘medium’. The membership value of the output in the ‘high’ output subset is determined by inference method. As can be observed from Table 2, greater weighting is given to the SI input than to the TI input.

The FLC’s behavior itself was examined through Matlab Fuzzy Toolbox v. 2.2.4. The behavior can be predicted from its output surface, Figure 9, formed by knowledge of its rule table and the method of de-fuzzification. Matlab’s toolbox allows a set of output data points to be calculated to a given resolution, allowing interpolation of the surface. By means of a look-up-table derived from the surface, a simple hardware implementation becomes possible.

The crisp outputs formed by repeated application of the fuzzy logic model to the SI and TI inputs of each video stream results in a control value for each video stream. These control values are converted to fractions of the available bandwidth by division of each by the total of the output values. An average is subsequently taken over an epoch of a GOP. The average forms the control signal to a video encoder (or transcoder) to adjust the bandwidth share for a particular video stream over the next GOP. The FLC’s output is a normalized proportion of the available bandwidth.

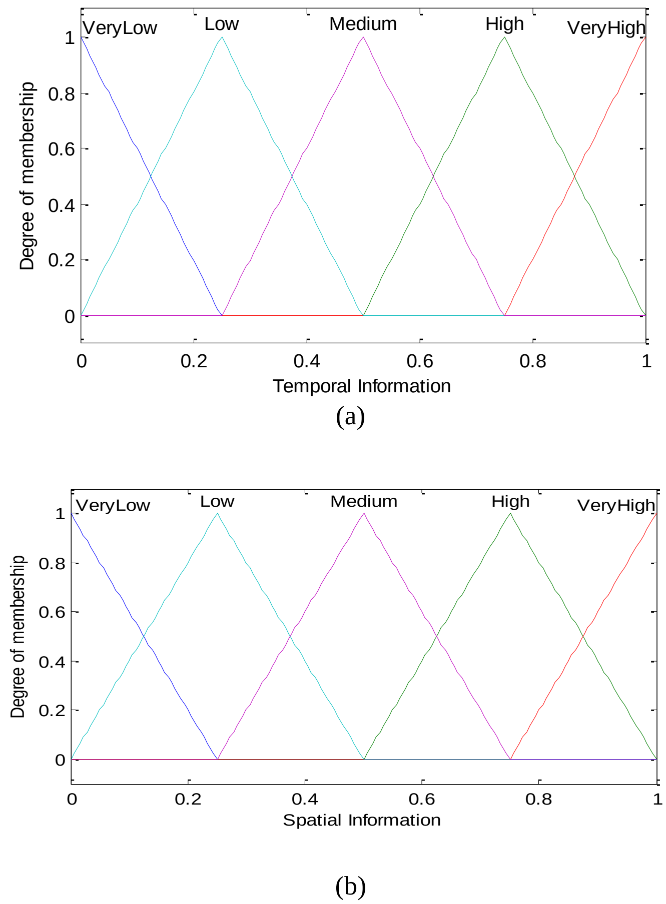

Figure 8. Fuzzy membership functions for a) TI and b) SI input.

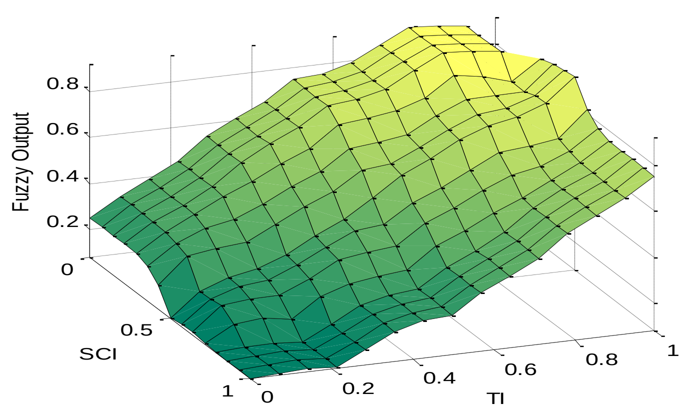

Figure 9. Output surface giving the available bandwidth proportion for any one video stream.

Table 2. FLC If..then rules used to identify output fuzzy subsets from inputs.

		TI				
		Very Low	*Low*	*Medium*	*High*	*Very High*
SI	*Very low*	Very low	Very low	Very low	Very low	Low
	Low	Very low	Low	Low	Low	Medium
	Medium	Low	Medium	Medium	Medium	Medium
	High	Medium	High	High	High	Very High
	Very high	High	Very high	Very high	Very High	Very High

Predicting available bandwidth

Available bandwidth for the IPTV sub-channel is predicted by a *P*-order LPF (Adas, 1998), with an order-eight filter adopted by us. We start by concisely summarizing the algorithm. The *P*-order LPF prediction filter is represented by

$$X(m+1) = \sum_{k=1}^{P} w_k . X(m-k+1) \tag{3}$$

where $X(m+1)$ is the predicted available bandwidth of the IPTV sub-channel, estimated from P previous monitored values of available bandwidth over sample instances m, while the w_k are the P adaptive filter weights indexed by k. The weights are estimated through:

$$\mathbf{w}(m+1) = \mathbf{w}(m) + \frac{e(m).\mathbf{X}(m)}{\left\|\mathbf{X}(m)\right\|^2}, \tag{4}$$

where $\mathbf{w}$ is the length P column vector of weights and $\mathbf{X}$ is a length P column vector of available bandwidth measurements over time, as in (5).

$$\mathbf{X}(m) = \left[X(m), X(m-1), \ldots, X(m-P+1)\right]^T \tag{5}$$

where T represents vector transpose. The variable $e(m)$ is the error between the monitored and the previously predicted available bandwidth value.

The algorithm to form the prediction filter can be described as follows. In equation (3), the coefficients, w_k, vary over time. A new coefficient at sampling instance $m + 1$ is formed by adjusting the previous value for that coefficient at time m. The adjustment in turn depends on P previous measurements over time of available bandwidth. Thus, the prediction error at time *m*, that is *e(m),* is fed back in equation (4) to improve the previous value of each coefficient w_k so that the prediction of the available bandwidth at time $m + 1$ is more accurate. It can be shown (Adas, 1998) that if the monitored values' behavior in a statistical sense is stationary over time then equation (4) will result in the w_k converging to form an optimal prediction. Moreover, for the form of equation (4) selected, for example the normalization term in the denominator of the second term on the right-hand-side of equation (4), convergence does not result in fluctuations in the predicted values once convergence has taken place.

Available bandwidth model

A four-state Markov chain modeled the available bandwidth in the sub-channel over time. Each state was directly reachable from every other state. The mean available bandwidth in state 1, 2, 3, and 4 was set to 3, 2.5, 2, 1.5 Mbps respectively. Each state on average was maintained for 2 s, which was set to be equivalent to 2000 monitoring points. If T_s=2000 then the probability of being in any one state is:

$$P_s = 1 - \frac{1}{T_s} = 0.9995 \tag{6}$$

and given that the probability of going to any other of three states is equi-probable and equal to (1-0.9995)/3 the state transition matrix is:

$$\begin{bmatrix} 0.9995 & 1.66 \times 10^{-4} & 1.66 \times 10^{-4} & 1.66 \times 10^{-4} \\ 1.66 \times 10^{-4} & 0.9995 & 1.66 \times 10^{-4} & 1.66 \times 10^{-4} \\ 1.66 \times 10^{-4} & 1.66 \times 10^{-4} & 0.9995 & 1.66 \times 10^{-4} \\ 1.66 \times 10^{-4} & 1.66 \times 10^{-4} & 1.66 \times 10^{-4} & 0.9995 \end{bmatrix} \tag{7}$$

In the model of available bandwidth, it is supposed that perturbations occur to the mean available bandwidth in any one state. For example, if the mean available bandwidth in the sub-channel was 4 Mbps then this could be perturbed in a positive or negative-going direction by a small amount, for example no more than 0.15 Mbps in either direction. To generate the amplitude of the perturbation, samples were taken from a symmetrical Uniform distribution.

Physical channel models

In Luby et al. (2008), the REIN model for ADSL was applied to IPTV. This is a simple model of fixed-length error bursts, the duration of which was set to 8 ms. The bursts were randomly placed to achieve a loss rate in the range 5×10^{-2} to 5×10^{-3}. In Luby et al. (2008), a larger error range was simulated but at the lower ends of that range the error rate was effectively close to being error free.

Broadband wireless channel errors are usually ‘bursty’ and dependent in time, rather than independent and identically distributed. A two-state, discrete-time Markov model served to model available bandwidth, with the Bit Error Rate (BER) in the good and bad states set as 10^{-4} and 10^{-3} respectively. Each state modelled an Additive White Gaussian Noise (AWGN) source. With time in the good state and bad state set respectively as 2 s and 0.25 s and with 1000 monitoring per second, the probability of remaining in a good state (from (6)) is 0.9995 and in a bad state is 0.9980.

RESULTS

In the tests, packetization was on the basis of one H.264/AVC Network Adaptation Layer unit (NALU) per packet, with each row forming a slice to be encapsulated in a NALU. Error-free, fixed bandwidth (3 Mbps) allocation, as appropriate to cable access is firstly considered.

Fixed bandwidth allocation

In Figure 10 shows the time-wise allocation of bandwidths after application of the FLC, based on the SI and TI metrics. The allocation follows approximately the coding complexity of the test clips in the sense that a more complex sequence receives a larger proportion of bandwidth. Figure 11 is a histogram of the per-frame frequency that the video sequences fell within the desired quality range (30–38 dB), compared to the same allocation if no adjustment to the initial CBR rates was made.

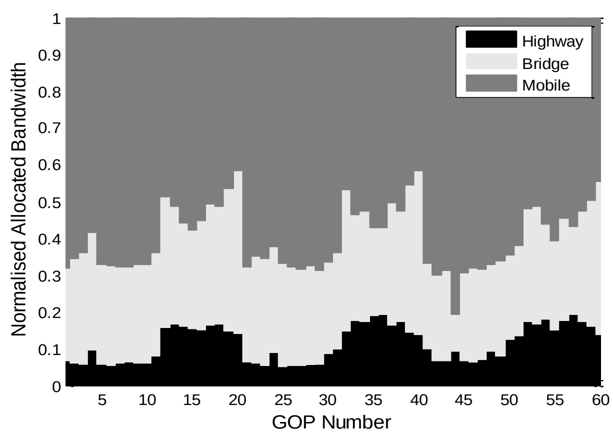

Figure 10. Bandwidth allocation over time for fixed bandwidth scenario.

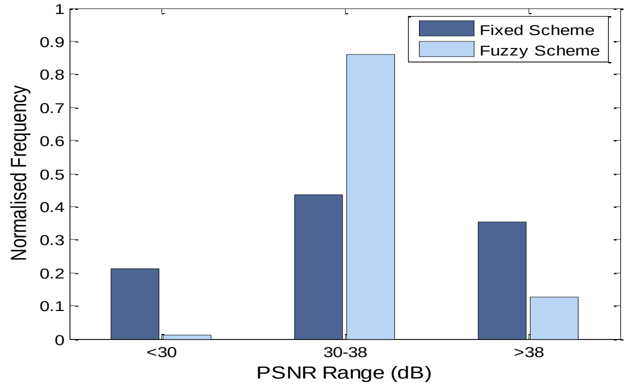

Figure 11. Normalised per–frame frequency that video quality for the three test clips is maintained within a desired quality range for FLC and equal CBR allocation (fixed) for fixed available bandwidth.

The allocation over time for each in turn of the three clips is illustrated in Figures 12–14. It is apparent that for *Mobile* much of the time the FLC allocation results in a higher video quality than a CBR scheme would do, whereas for *Highway* and *Bridge-closed*, the video quality (which is already high) is somewhat reduced. Table 3, summarizes the average video qualities resulting from the FLC and equal CBR allocations.

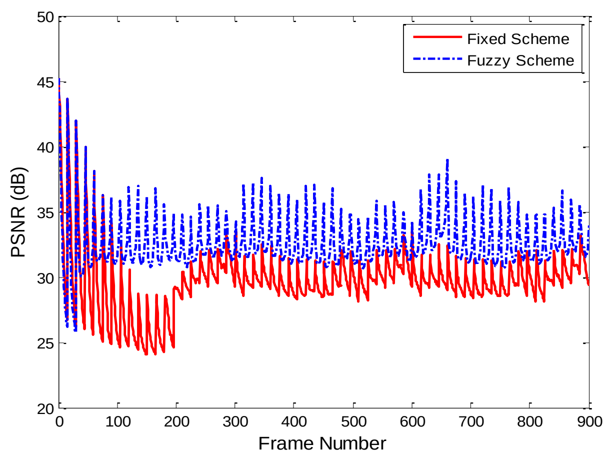

Figure 12. Comparative video quality over time achieved for *Mobile* between FLC and equal CBR (fixed) schemes.

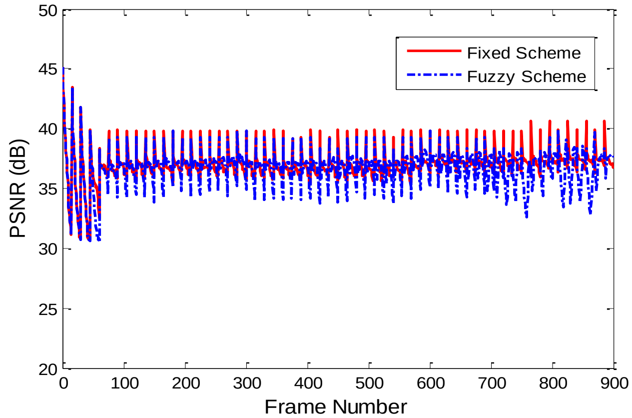

Figure 13. Comparative video quality over time achieved for *Bridge* between FLC and equal CBR (fixed) schemes.

Table 3. Summary results of comparative bandwidth allocation between FLC and equal CBR schemes with fixed available bandwidth.

	FLC		CBR	
	Bitrate (kbit/s)	PSNR (dB)	Bitrate (kbit/s)	PSNR (dB)
Highway	338.51	36.03	1000	40.60
Bridge	857.10	36.62	1000	36.96
Mobile	1802.90	32.53	1000	29.61

The effect of physical channel degradation on video according to the REIN model for ADSL was ascertained for a fixed available bandwidth (of 3 Mbps), after subsequent distribution of the video streams to individual users. Figure 15 reports the relative video quality between equal CBR share and FLC allocation. In Figure 15, each data point is the result of forty runs in order to achieve convergence. In particular, the video quality of *Mobile* is improved by the FLC allocation. At higher BERs, *Mobile*'s quality becomes poor in both schemes, though the CBR allocation can result in unwatchable TV. Perhaps, a weakness of the FLC allocation is that *Highway* also drops out of the desired quality range, though at high BERs.

Variable available bandwidth allocation

The available bandwidth was also varied according to the four-state available bandwidth model. Table 4 shows summary results and should be compared with Table 3. It will be apparent that, for the FLC allocation, all video sequences' quality is within the desired range, whereas again for equal allocation (with varying available bandwidth) the video quality is either excessively high or low, so that *Mobile*'s quality drops outside the desired range. In general, as a result of the changing available bandwidth, delivered video quality is reduced in Table 4 compared to Table 3's results. As a visual comparison of the allocation, Figure 16 should be compared with Figure 11, when it will be seen that the FLC scheme maintains its advantage when there is a variable available bandwidth.

The physical channel model for 'bursty' wireless errors was also applied to the variable available bandwidth model. Therefore, this scenario assesses how robust the FLC allocation is after subsequent introduction of error bursts during PMP distribution. From Figure 17, it is apparent that there is little gain for *Bridge* and *Highway* but *Mobile* shows some gain, though in this set of experiments the resulting video quality was much reduced. Because the bit-rate of *Bridge* and *Highway* is reduced the packet size is smaller and, consequently they are less prone to AWGN errors. This was not the case for the fixed length error bursts arising from the REIN model, which explains the relative improvement in PSNR over ADSL access.

Table 4. Summary results of comparative bandwidth allocation between FLC and equal CBR schemes with variable available bandwidth.

	FLC		CBR	
	Average Bitrate (kbit/s)	PSNR (dB)	Average Bitrate (kbit/s)	PSNR (dB)
Highway	405.50	35.57	996.81	39.73
Bridge	872.39	35.53	997.11	36.38
Mobile	1640.47	31.44	996.57	28.96

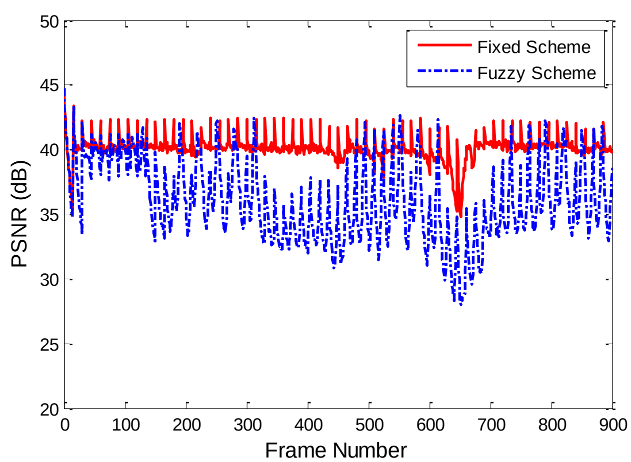

Figure 14. Comparative video quality over time achieved for *Highway* between FLC and equal CBR (fixed) schemes.

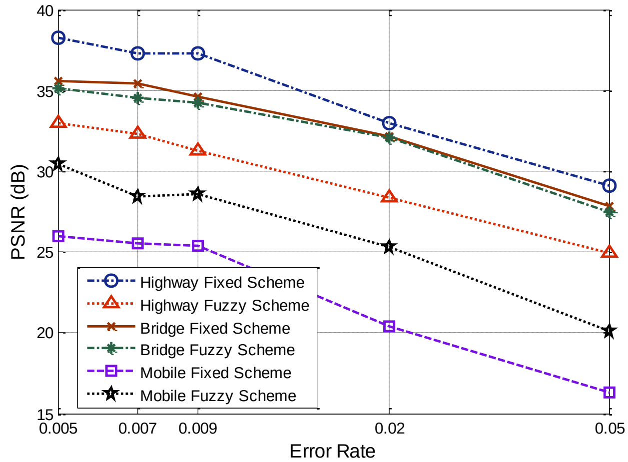

Figure 15. Video quality according to FLC and equal CBR (fixed) schemes for the ADSL REIN model.

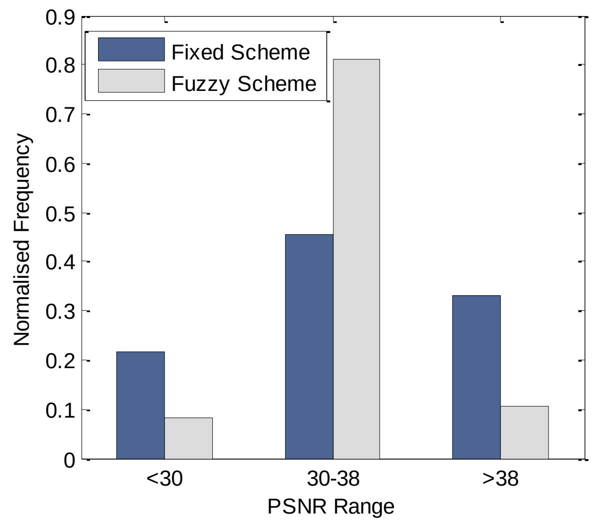

Figure 16. Normalised per–frame frequency that video quality for the three test clips is maintained within a desired quality range for FLC and equal CBR allocation (fixed) for variable available bandwidth.

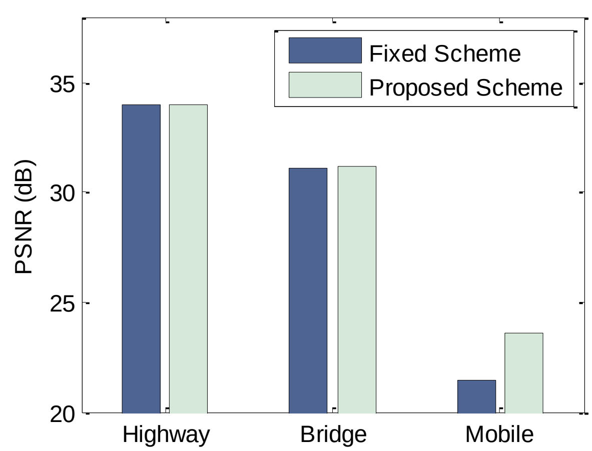

Figure 17. Comparative video quality for the three test clips after FLC (proposed) and equal CBR allocation (fixed) bandwidth allocation, with variable available bandwidth and 'bursty' channel error model.

CONCLUSION

Bandwidth allocation control aims to equalize the quality of a set of video streams sharing a common multimedia sub-channel. The quality should also, as far as possible, fall within an acceptable range. The danger of statistical control of data rates is that it does not take account of the varying coding complexity of video streams. In this paper, dynamic adjustments were jointly made to the target video data rates in response to prior input of spatial and temporal compression metrics. These are either output by the encoder itself or can be extracted from the encoded bit-stream. Fuzzy logic control was shown to consistently outperform equal allocation of bandwidth. The impact of a varying available bandwidth and differing physical channel conditions was investigated and again the proposed method was superior. A practical system would introduce application-layer error control but, due to the variety of physical channel types, it is probably preferable to not take account of the access network channel conditions in sub-channel bandwidth allocation. Instead, if need be, cross-layer adjustment to conditions can take place at the access network distribution point. There is also clearly a need for real-time, video-quality metering at various monitoring points in IPTV networks to present a more accurate record of the likely viewer experience than PSNR can report. Further investigation should consider VBR input, as this will reduce coding latency arising from two-pass CBR encoding and reduce storage requirements by not maintaining the bitrate if the coding complexity does not merit it. However, true H.264/AVC VBR video may result in too frequent bit-rate oscillations, especially with the increased coding efficiency of H.264/AVC codec standard codecs. Therefore, video smoothing may well be required before input into the multiplexor. We have employed a transcoder bank to alter the video rate. However, the scalable video coding extension to H.264/AVC allows possible rate-change points to be embedded into a pre-encoded video bitstream. Thus, the scheme could be applied to other forms of adaptive streaming. The fuzzy-logic system demonstrated in this paper applies type-1 fuzzy logic. In type-1 fuzzy logic, the models are static and do not account for inaccuracy in the original design. There are essentially two ways that this weakness can be addressed: the first is to employ some form of adaptive fuzzy models, possibly through neuro-fuzzy modeling; and the second is by employing type-2 fuzzy logic, in which the models themselves can be fuzzy. Therefore, improvements in the fuzzy system can be sought for in the future.

REFERENCES

Adas, A. M. (1998). Using adaptive linear prediction to support real-time VBR video under RCBR network service model. IEEE/ACM Transactions on Networking, 6(5), 635:644.

Agboma, F., & Liotta, A. (2006). User centric assessment of mobile contents delivery. In Mobile Computing and Multimedia, (pp. 121-130).

Anick, D., Mitra, D., & Sondhi, M.M. (1982). Stochastic theory of a data-handling system with multiple sources. Bell System Technical Journal, 61(8), 1871:1894.

Afzal, J., Stockhammer, T., Gasiba, T., & Xu, W. (2006). Video streaming over MBMS: A system design approach. Journal of Multimedia, 1(5), 25–35.

Bashworth, B.N., & Williamson, C.L. (1998). Simulation study of self-similar video streams: Simulation study and performance results. In International Symposiium on Modeling, Analysis and Simulation of Computer and Telecommunications Systems, (pp. 119-126).

Baturone, I., et al. (2000). Microelectronic design of fuzzy logic-based systems. Boca Raton, FL: CRC Press.

Böröczy, L., Ngai, A.Y., & Westermann, E.F. (1999). Statistical multiplexing using MPEG-2 video encoders. IBM Journal of Research and Development, 43(4), 511:520.

Dravida, S. et al. (2002). Broadband access over cable for next-generation services: A distributed switch architecture. IEEE Communications Magazine, 40(8), 116:124.

Grant, P.M., Saw, Y.-S., & Hannah, J.M. (1997). Fuzzy rule based MPEG video rate prediction and control. In EURASIP ECASP Conference, (pp. 211–214).

Guha, A., & Reininger, D.J. (1994). Multichannel joint rate control of VBR MPEG encoded video for DBS Applications. IEEE Transactions on Consumer Electronics, 40(3), 616:623.

Haskell, B.G. (1972). Buffer and channel sharing by several interframe picture phone coders. Bell System Technical Journal, 51(1), 261:289.

Heeke, H. (1993). A traffic control algorithm for ATM networks. IEEE Transactions on Circuits and Systems for

Video Technology, 3(3), 183:189.

He, Z., & Wu, D.O. (2008). Linear rate control and optimum statistical multiplexing for H.264 video broadcast. IEEE Transactions on Multimedia, 10(7), 1237:1249.

Hermsmeyer, C., Hernandez-Valencia, E., Stoll, D., & Tamm, O. (2007). Ethernet aggregation and core network models for efficient and reliable IPTV services. Bell Labs Technical Journal, 12(1), 57:76.

Horowitz, M., Joch, A., Kossentini, F. & Hallapuro, A. (2003). H.264/AVC Baseline Profile decoder complexity analysis. IEEE Transaction on Circuits and Systems for Video Technology, 13(7), 704:716.

Huynh-Thu, Q., & Ghanbari, M. (2008). The scope of validity of PSNR in image/video quality assessment. Electronics Letters, 44(13), 800:801.

ITU-T. (1992). Draft recommendation L371: Traffic control and congestion control in B-ISDN.

Jain, R., Kalyanaraman, K., Goyal, R., Fahmy, S., & Viswanathan, R. (1996). ERICA switch algorithm: a complete description. ATM Forum Document number /96-1172.

Jammeh, E.A., Fleury, M., & Ghanbari, M. (2008). Fuzzy logic congestion control of transcoded video streaming without packet loss feedback. IEEE Transactions on Circuits and Systems for Video Technology, 18(3), 387:393.

Kasai, H., Nilsson, M., Jebb, T., Whybray, M., & Tominaga, H. (2002).The development of a multimedia transcoding system for mobile access to video conferencing. IEICE Transactions on Communications, 10(2), 2171:2181.

Lee, J.M., Park, H.-J., Choi, S.G., Choi, J.K. (2009). Adaptive hybrid transmission mechanism for on-demand mobile IPTV over WiMAX.IEEE Transcations on Broadcasting, 55(2), 468:477.

Luby, M., Stockhammer, T., & Watson, M. (2008). Application FEC in IPTV systems. IEEE Communications Magazine, 46(5), 94:102.

Ma, H., & El Zarki, M. (1999). Bandwidth reservation for the provision of VoD services over wireless access networks using hybrid ARQ schemes. In IEEE Wireless Communications and Networking Conference, (pp. 114-

118).

Maisonneuve, J., Deschanel, M., Heiles, J., Li, W., Liu, H., Sharpe, R., & Wu, Y. (2009). An overview of IPTV standards development. IEEE Transactions on Broadcasting, 52(2), 315:328.

Mamdani, E.H., & Assilian, S. (1975). An experiment in linguistic synthesis with a fuzzy logic controller. International Journal of Man-Machine Studies, 7(1), 1:13.

Moiron, S., Razavi, R., Fleury, M., & Ghanbari, M. (2010). Statistical Multiplexing of Transcoded IPTV Streams based on Content Complexity. In 7th International ICST Mobile Multimedia Communications.

Muntean, G. (2006). Efficient delivery of multimedia streams over broadband networks using QOAS. IEEE Transactions on Broadcasting, 52(2), 230:235.

Nam, H.-M., Dan, B.-K., Kim, H.-S., Jeong, J.-Y., & Ko, S.-J. (2006). Low complexity H.264 transcoder for bitrate reduction. In International Symposium on Communications and Information Technologies, (pp. 679–682).

Nim H.-M. et al. (2006). Low complexity H.264 transcoder for bitrate reduction. In International Symposium on Communications and Information Technologies, (pp. 679-682).

Rezaei, M., Hannuksela, M.M., & Gabbouj, M. (2008). Semi-fuzzy rate controller for variable bit rate video. IEEE Transactions on Circuits and Systems for Video Technology, 18(5), 633:645.

Rosdiani, E., & Ghanbari, M. (2000). Picture complexity based rate allocation algorithm for transcoded video over ABR networks. Electronics Letters, 36(6), 521:522.

Seeling, P., & Reisslein, M. (2005). The rate variability-distortion (VD) curve of encoded video and its impact on statistical multiplexing. IEEE Transactions on Broadcasting, 51(4), 473:492.

She, J., Hou, F., Ho, P.-H., Xie, & L.-L. (2007). IPTV over WiMAX: Key success factors, challenges, and solutions. IEEE Communications Magazine, 45(8), 87:93.

Singh, H., Kvvon, C., Kim, S.S., & Ngo, C. (2008). IPTV over wireless LAN: Promises and challenges. In IEEE Consumer Communications and Networking Conference, (pp. 626-635).

Vukadinovic, V., & Huschke, J. (2008). Statistical multiplexing gains of H.264/AVC video in E-MBMS. In 3rd International Symposium on Wireless Pervasive Computing, (pp. 468–474).

Wagner, M., Kopilovic, I., Döhla, S., Färber, N., & Fuchs, H. (2009). Towards an RTP profile for IPTV. IEEE Transactions on Broadcasting, 55(2), 343:352.

Wang, L., & Vincent, A. (1996). Joint rate control for multi-program video coding. IEEE Transactions on Consumer Electronics, 42(3), 300:305.

Zheng, H., & Lin, K. J. R. (2000). Multimedia services over Digital Subscriber Line. IEEE Signal Processing, 17(4), 44:58.