Content Classification Based on Objective Video Quality Evaluation for MPEG4 Video Streaming over Wireless Networks

Asiya Khan, Lingfen Sun and Emmanuel Ifeachor

Abstract— **User's perceptive video quality differs greatly with video contents hence, it is of practical importance to classify videos. Videos are most commonly classified according to their spatial and temporal features. In this paper, we classify videos into groups based on objective video quality evaluation. Video quality measured in terms of the Mean Opinion Score (MOS), obtained by the application and network level parameters is classified into groups using cluster analysis with good prediction accuracy. The QoS (Quality of Service) parameters that affect video quality considered in this paper are send bitrate, frame rate in application and packet error rate in the network level. We then find the degree of influence of each of the QoS parameter and analyze the relationship between QoS parameters and content types by using principal component analysis for streaming MPEG4 video over wireless networks. We compare our classified contents to the spatio-temporal dynamics of the content and establish the relationship between video contents and video quality by equations obtained by multiple linear regression. Finally, we apply the results to rate control methods. The proposed scheme makes it possible to apply priority control to content delivery, based on content types with similar attributes.**

Index Terms— **MOS, content classification, cluster analysis, QoS, MPEG4**

I. INTRODUCTION

Multimedia services are becoming commonplace across different transmission platforms such as Wi-Max, IEEE 802.11 standards, 3G mobile etc. The current trends in the development and convergence of wireless internet IEEE 802.11 applications and mobile systems are seen as the next step in mobile/wireless broadband evolution. Users' demand of the quality of streaming service is very much content dependent. It is therefore important to choose or adapt both the application level i.e. the compression parameters as well as network settings so that maximum end-to-end user perceived video quality can be achieved. The prime criterion for the quality of multimedia applications is the user's perception of service quality [1]. The most widely used metric is the Mean Opinion Score (MOS). Streaming multimedia services delivered via wireless LANS and 3G networks are hot topics. In order to provide high quality video streaming, we have to consider video contents and users' preferences of video quality. In content delivery systems such as video-on-demand video content delivered is highly sensitive to network errors. Hence, it is important to determine the relationship between the users' perception of quality to the actual characteristic of the content and hence increase users' service by using priority control for delivery. In this paper we explore this relationship by considering QoS parameters both in the network and application levels. The current trends in the development of wireless internet applications (IEEE 802.11) and mobile systems indicate that the future internet architecture will need to support various applications with different QoS requirements. More recently the term QoE (Quality of Experience) has been used and defined as the users perceived QoS. It has been proposed in [2][3] that a better QoE can be achieved when the QoS is considered both in the network and application layers as a whole.

Feature extraction is the most commonly used method to classify video contents. In [4],[5] video content is classified based on the spatial (edges, colours, etc) and temporal (movement, direction, etc) feature extraction which were then used to predict video quality together with other application-level parameters such as send bitrate and frame rate. The limitations of using feature extraction it does not express the semantic scene importance but the formal features such as degree of motion compensation. Recent studies in [6],[7] have classified video content based on content characteristics obtained from users' subjective evaluation using cluster [8] and Principal Component Analysis (PCA) [8]. In [9],[10] authors have used a combination of PCA [8] and feature extraction to classify video contents. These techniques are at early stages and do not take into account users' perception. Subjective testing is an accurate way of measuring users' perception of quality. However, it is expensive and time consuming and hence, the need for objective testing. Objective measurements can be performed in an intrusive or non-intrusive way. Intrusive measurements require access to the source then compares the original and impaired videos. Full reference and reduced reference video quality measurements are both intrusive [11]. Quality metrics such as Peak-Signal-to-Noise-Ratio (PSNR), VQM [12] and PEVQ [13] are full reference metrics. VQM and PEVQ are commercially used and are not publicly available. Non-intrusive methods (reference-free), on the other hand do not require access to the source video. Non-intrusive methods are either signal or parameter based. Non-intrusive methods are preferred to intrusive analysis and they are more suitable for on-line quality prediction/control.

From these backdrops, in this paper, we aim to recognize the most significant video content types, classify them based on objective video quality evaluation into groups using cluster analysis [8]. Video quality is evaluated in terms of the Mean Opinion Score (MOS) for all content types combining both the application level (send bitrate, frame rate) and network level (packet error rate) parameters. We then found the degree of influence of each QoS parameter and analyze the relationship between QoS parameters and content types

Manuscript received February 11, 2009. The work reported here is supported in part by the EU FP7 ADAMANTIUM project (contract No. 214751).

A. Khan, L. Sun and E. Ifeachor are with the Center for Signal Processing and Multimedia Communication, School of Computing, Communications and Electronics, University of Plymouth, Plymouth PL4 8AA, UK. (email: asiya.khan,l.sun,e.ifeachor@plymouth.ac.uk)

using principal component analysis [8].We further, compare our classified contents to the spatio-temporal dynamics of the content and establish the relationship between video contents and video quality by equations obtained by multiple linear regression. Finally, we apply the results to rate control methods to contents within a group. As subjective tests are costly and time consuming, the proposed test bed is based on simulated network scenarios using a network simulator (NS2) [14] with an integrated tool Evalvid [15]. It gives a lot of flexibility for evaluating different topologies and parameter settings used in this study. Our focus ranges from low resolution and low send bitrate video streaming for 3G applications to higher video send bitrate for WLAN applications depending on type of content and network conditions.

The paper is organized as follows. Section 2 describes the experimental set-up, whereas, section 3 outlines the relationship between video content and objective video quality. In section 4 we evaluate the proposed rate control scheme and section 5 concludes the paper and highlights areas of future work.

II. Experimental Set-up

A. Simulation Set-up

For the tests we selected twelve different video sequences of qcif resolution (176x144) and encoded in MPEG4 format with an open source ffmpeg [16] encoder/decoder with a Group of Pictures (GOP) pattern of IBBPBBPBB. Each GOP encodes three types of frames - Intra (I) frames are encoded independently of any other type of frames, Predicted (P) frames are encoded using predictions from preceding I or P frames and Bi-directionally (B) frames are encoded using predictions from the preceding and succeeding I or P frames.

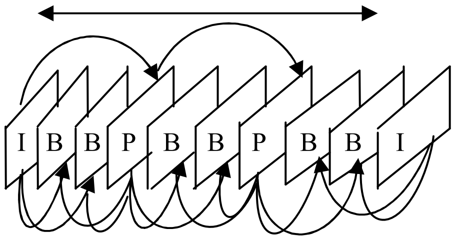

Figure 1. A sample of MPEG4 GOP (N=9, M=3)

A GOP pattern is characterized by two parameters, GOP(N,M) – where N is the I-to-I frame distance and M is the I-to-P frame distance. For example, as shown in Fig.1, G(9,3) means that the GOP includes one I frame two P frames, and six B frames. The second I frame marks the beginning of the next GOP. Also the arrows in Fig. 1 indicate that the B frames and P frames decoded are dependent on the preceding or succeeding I or P frames [17].

The experimental set up is given in Fig 2. There are two sender nodes as CBR background traffic and MPEG4 video source. Both the links pass traffic at 10Mbps, 1ms over the internet which in turn passes the traffic to another router over a variable link. The second router is connected to a wireless access point at 10Mbps, 1ms and further transmits this traffic to a mobile node at a transmission rate of 11Mbps 802.11b WLAN. No packet loss occurs in the wired segment of the video delivered path. The maximum transmission packet size is 1024 bytes. The video packets are delivered with the random uniform error model. The CBR rate is fixed to 1Mbps to give a more realistic scenario. The packet error rate is set in the range of 0.01 to 0.2 with 0.05 intervals. To account for different packet loss patterns, 10 different initial seeds for random number generation were chosen for each packet error rate. All results generated in the paper were obtained by averaging over these 10 runs.

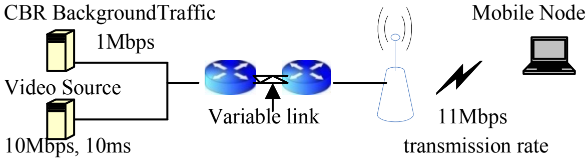

Figure 2. Simulation setup

B. Mean Opinion Score

To obtain a MOS (Mean Opinion Score) value we conducted experiments with an open source framework Evalvid [15] and network simulator tool NS2 [14]. Video quality is measured by taking the average PSNR (Peak-Signal-to-Noise-Ratio) over all the decoded frames. PSNR given by (1) computes the maximum possible signal energy to noise energy. PSNR measures the difference between the reconstructed video file and the original video file.

$$PSNR(s,d)]_{db} = 20 \log \frac{Vpeak}{MSE(s,d)} \qquad (1)$$

Mean Square Error (MSE) is the cumulative square between compressed and the original image.

MOS scores are calculated based on PSNR to MOS conversion from Evalvid [15] given in Table I below.

Table I.
PSNR to MOS Conversion

PSNR (dB)	MOS
>37	5
31 – 36.9	4
25 – 30.9	3
20 – 24.9	2
< 19.9	1

C. QoS Parameters

We considered the following quality affecting parameters both in the application level and the network level as follows:

Application Level parameters: The Frame Rate (FR): the number of frames per second. It takes one of three values as 10, 15 and 30fps. The Send Bitrate (SBR): the rate of the encoders output. It is chosen to take 18, 44, 80, 104 and 512kb/s.

Network Level Parameters: Packet Error Rate (PER): the simulator (NS-2) [14] drops packet at regular intervals using a random uniform error model. It takes five values as 1, 5, 10, 15 and 20%. It is widely accepted that a loss rate higher than 20% will drastically reduce the video quality.

III. Analysis Of Relationship Between Video Contents And Objective Video Quality

In this section, we analyze the relationship between video contents and objective video quality. We first classify video contents based on objective video quality evaluation (MOS scores), then find the degree of influence of each QoS parameter on video quality and finally, compare the classified contents to the spatio-temporal dynamics.

A. *Content Classification Model*

Video contents are classified based on the Mean Opinion Score obtained from parameters of SBR and FR in the application and PER network level. Content classification model is given in Fig. 3. A well known multivariate statistical analysis called cluster analysis [8] is used to classify the contents. This technique is used as it groups samples that have various characteristics into similar groups. Video MOS scores for all twelve video sequences obtained from objective video quality evaluation from the quality parameters of SBR, FR and PER are used as input to the statistical tool (cluster analysis) that classifies the video content into groups.

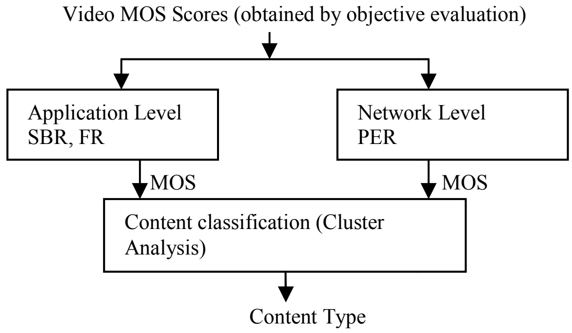

Figure 3. Content Classification Model

For our data, we used hierarchical cluster analysis in which samples that have the nearest Euclid distance are put together. Fig. 4 shows the obtained dendrogram (tree diagram) where the video sequences are grouped together on the basis of their mutual distances (nearest Euclid distance).

According to Sturge's rule ($k = 1 + 3.3\log N$), which for our data will be 5 groups. However because of the problems identified with this rule [18] we split the data (test sequences) at 62% from the maximum Euclid distance into three groups. (see the dotted line on Fig. 4) as the data contains a clear 'structure' in terms of clusters that are similar to each other at that point. From Fig. 4 the video contents are divided into three groups of content types of Slight Movement (SM), Gentle Walking (GW) and Rapid Movement (RM). The spearman correlation coefficient is 73.29%. The correlation coefficient should be very close to 100% for a high-quality solution.

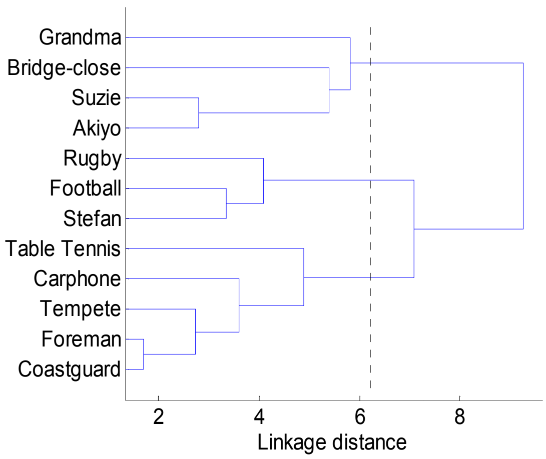

Figure 4. Tree diagram based on cluster analysis

To further verify the content classification from the tree diagram obtained (Fig. 4) we carried out K-means cluster analysis in which the data (video clips) is partitioned into *k* mutually exclusive clusters, and returns the index of the cluster to which it has assigned each observation. K-means computes cluster centroids differently for each measured distance, to minimize the sum with respect to the specified measure. We specified k to be three to define three distinct clusters. In Fig. 5 K-means cluster analysis is used to partition the data for the twelve content types. The result set of three clusters are as compact and well-separated as possible giving very different means for each cluster.

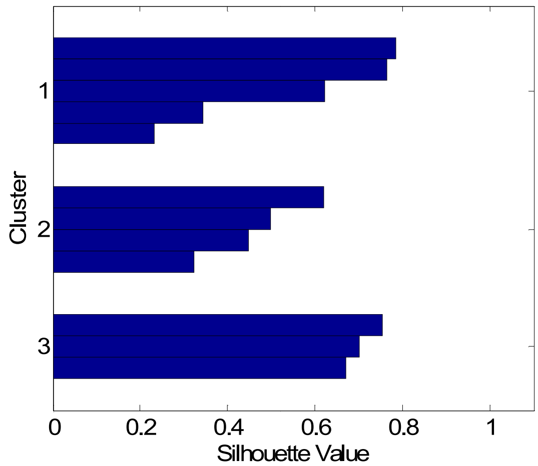

Figure 5. K-means cluster analysis

Cluster 1 in Fig. 5 is very compact for three video clips instead of five. Clips of Table tennis and Carphone are slightly out of the cluster. They can be within their own cluster and will be looked in much detail in future work. Both clusters 2 and 3 are very compact. All results were obtained using MATLAB™ 2008 functions. Based on both hierarchical and k-means cluster analysis the content types are divided into three groups as follows:

The three content types are defined for the most frequent contents for mobile video streaming as follows:

1. Content type 1 – Slight Movement (SM): includes sequences with a small moving region of interest (face) on a static background. See Fig. 6.

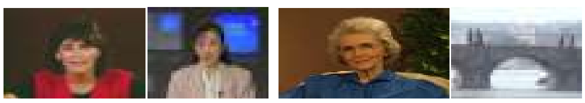

Figure 6. Snapshots of typical 'SM' content

2. Content type 2 – Gentle Walking (GW): includes sequences with a contiguous scene change at the end. They are typical of a video call scenario. See Fig. 7.

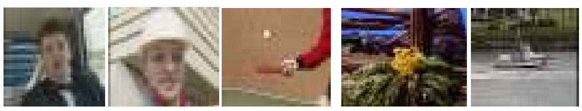

Figure 7. Snapshots of typical 'GW' content

3. Content type 3 – Rapid Movement (RM): includes a professional wide angled sequence where the entire picture is moving uniformly e.g sports type. See Fig. 8.

Figure 8. Snapshots of typical 'RM' content

We found that the 'news' type of video clips were clustered in one group, however, the sports clips were put in two different categories i.e. clips of ‘stefan’, ‘rugby’ and ‘football’ were clustered together, whereas, ‘table-tennis’ was clustered along with ‘foreman’ and ‘carphone’ which are both wide angle clips in which both the content and background are moving.

B. Degree of Influence of QoS Parameters

To find the degree of influence of each QoS parameter used in content classification (Fig. 3) we carried out Principal Component Analysis [8]. PCA is a method to reduce the dimensionality of a data set, in which there are a large number of inter-related variables. PCA uses a covariance matrix in the case where the same data has the same set of variables or correlation matrix in the case where data has a different set of variables. In this paper, we used a covariance matrix because of the same data set. The principal component scores for each content is shown in Table II.

TABLE II
PRINCIPAL COMPONENT SCORE TABLE

Content type	Content	Scores	SBR	FR	PER
SM	Akiyo	0.212	0.57	-0.58	-0.58
	Suzie	0.313	0.66	0.25	-0.71
	Grandma	0.147	-0.76	0.64	-0.05
	Bridge-close	0.092	0.41	-0.22	-0.89
GW	Table Tennis	0.287	0.08	-0.99	0.11
	Carphone	0.154	0.35	-0.93	0.10
	Tempete	0.231	0.25	-0.46	-0.85
	Foreman	0.204	0.56	0.45	-0.69
	Coastguard	0.221	0.62	-0.60	0.51
RM	Stefan	0.413	0.40	-0.72	0.58
	Football	0.448	0.62	-0.57	0.55
	Rugby	0.454	0.65	-0.59	0.48

Table II shows the influence of each QoS parameter on video quality. It shows that scores for sports video contents are higher than those of news type videos. Also in the category of RM higher packet loss have a greater impact on video quality compared to that of SBR and FR. Similarly, for SM content type PER does not have a bigger impact on video quality.

Total variation by first principal component was 72.7% and 10% by the second component. Hence only scores from the first component were chosen. From Table II the features of the three content types are described as follows:

Content type 1 – SM: The main factors degrading objective video quality are frame rate and send bitrate. However, for the sequence of Grandma SBR is a bigger degrading factor compared to frame rate. We know that the main purpose of this content type is to get some information instead of to view a video itself. Hence requirements of frame rate are higher than that of send bitrate.

Content type 2 – GW: The main factors degrading objective video quality are send bitrate and packet error rate. In this category packet loss has a much higher impact on quality compared to SM.

Content type 3 – RM: The main factor degrading the video quality are send bitrate and packet error rate. A video coded at low send bitrate and/ with high packet losses are very annoying for the user. As most people are interested in the processes of the sport and its circumstances, and are aware of a moments screen freeze.

The degree of influence of the QoS parameters is further given by the box and whiskers plot shown in Fig. 9. Fig. 9 shows the influence of each QoS parameter both in the application and network level on video quality for the three content types. The whiskers serve to show the extent of the rest of the data. For SM PER is not as important as for GW and RM. Similarly, FR has the least impact on RM compared to that of SM and GW.

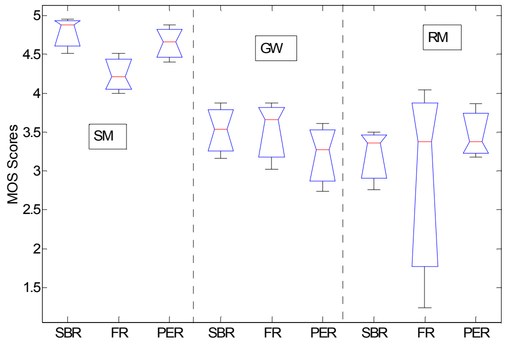

Figure 9. Significant effects of SBR, FR and PER

C. Comparison with the spatio-temporal dynamics

Video sequences are most commonly classified based on their spatio-temporal features. In order to classify video clip according to the spatial and temporal complexity of its content, a spatio-temporal grid [19] is considered and is depicted in Fig. 10.

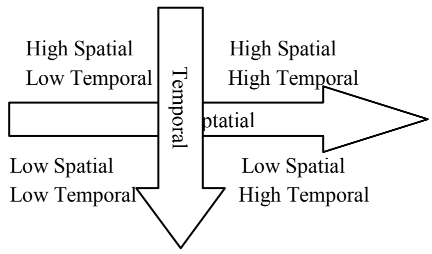

Figure 10. The spatio-temporal grid used for classifying a video sequence according to its content dynamics

From Fig. 10 the spatio-temporal grid divides each video into four categories based on its spatio-temporal features as follows:

- Low spatial – Low temporal activity: defined in the bottom left quarter in the grid.
- Low spatial – High temporal activity: defined in the bottom right quarter in the grid.
- High spatial – High temporal activity: defined in the top right quarter in the grid.
- High spatial – Low temporal activity: defined in the top left quarter in the grid.

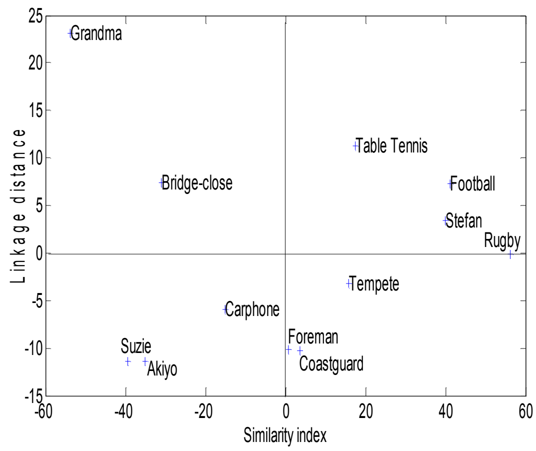

Figure 11. Principal co-ordinates analysis

Figure 11 shows the principal co-ordinates analysis also known as multidimensional scaling of the twelve content types. The function cmdscale in MATLAB™ is used to perform the principal co-ordinates analysis. cmdscale takes as an input a matrix of inter-point distances and creates a configuration of points. Ideally, those points are in two or three dimensions, and the Euclidean distances between them reproduce the original distance matrix. Thus, a scatter plot of the points created by cmdscale provides a visual representation of the original distances and produces representation of data in a small number of dimensions. In Fig. 11 the distance between each video sequence indicates the characteristics of the content, e.g. the closer they are the more similar they are in attributes.

Comparing Fig. 10 to Fig. 11 we can see that classifying contents from the MOS scores (through objective video quality evaluation in our case), contents of Football, Stefan Table-tennis and Rugby are high spatial and high temporal and would fit in the top right hand side, similarly contents of Suzie and Akiyo would fit in the bottom left hand side as they have low spatio-temporal features. Whereas, contents like Grandma and Bridge-close are in top left hand side indicating high spatial and low temporal features. Similarly, Foreman, Coastguard and Tempete are in the bottom right hand side with high temporal and low spatial features as expected. Only the video sequence of Carphone has been put in the bottom left hand side and will be investigated further.

This was also previously explained by Table II where the total scores as well as the QoS parameters scores for all content types are given (see sub-section B).

D. Relationship between video contents and objective video quality

In order to apply the results of the previous section to actual quality control, we investigate quantitative relationship between video contents and objective video quality. We carried out multiple linear regression analysis for the three content types. The quality parameters used are the Send Bitrate (SBR) and Frame Rate (FR) in the application level and Packet Error Rate (PER) in the network level. The results of the multiple regression analysis along with the coefficients of determination (R^2) are shown in Table III. The R^2 gives the goodness of fit of the proposed equations.

TABLE III
MOS BASED ON CONTENT TYPES

$$MOS_{SM} = 0.0075SBR - 0.014FR - 3.79PER + 3.4 \quad \text{Content type: SM } (R^2 = 85.72\%)$$

$$MOS_{GW} = 0.0065SBR - 0.0092FR - 5.76PER + 2.98 \quad \text{Content type: GW } (R^2 = 99.65\%)$$

$$MOS_{RM} = 0.002SBR - 0.0012FR - 9.53PER + 3.08 \quad \text{Content type: RM } (R^2 = 89.73\%)$$

Table III illustrates the derived classification functions in terms of the MOS for the three content types. The network operator can predict the users' perception of QoS based on the application-level parameters of SBR and FR and network-level parameters of PER then adjust the allocation of network resources accordingly. Similarly, the content provider can predict the video send bitrates for a given quality level as explained in the next section.

IV. EVALUATION

In this section, we compare the objective MOS values obtained from NS2 [14] and Evalvid [15] and apply the results to rate control methods.

We choose one video sequence from each content type and use the equations derived in the previous section to calculate the minimum send bitrate for a video sequence in a content type that will give minimum acceptable quality. Hence the content provider can specify the quality, video send bitrate can be reduced or increased according to the content type while keeping the same objective video quality.

TABLE IV
PREDICTED SEND BITRATE VALUES FOR SPECIFIC QUALITY LEVELS

Content type	FR	PER	MOS_{given}	SBR (Kbps) Predicted
SM	10	0	3.5	20
	15	0	3.6	55
	30	0/0.05	3.8	75/135
GW	10	0	3.7	125
	15	0	3.9	165
	30	0/0.02	4.1	215/235
RM	10	0	3.8	360
	15	0	4.1	500
	30	0/0.02	4.2	580/700

Table IV outlines the predicted send bitrates for the three content types. The content provider is able to identify the send bitrates that correspond to a given Quality of Service (in terms of MOS) level by simply using the equations from the previous section. Also we note that in the presence of packet loss (packet error rate) the send bitrates increases to try and compensate for the quality lost.

A. Comparison with existing work

In comparison to our previous work [20] where we classified video sequences based on their spatio-temporal features with good correlation (73.29%), in this paper contents are classified based on objective video quality evaluation. As shown in section IV sub-section C, comparing our classified contents with the spatio-temporal grid we achieved good correlation. The contents are classified using full reference i.e. would require access to the source. However, once the contents are classified then the model does not require access to the source (no-reference). The advantage of using this technique is that videos are classified as related to the users' satisfaction. Also we have shown that there is a clear link between the MOS scores (users' perceptive quality) to the types of content. This has also been explored in [21] where authors have evaluated the psychological factors that impact on quality. They have then linked the psychological factors to the objective quality parameters of packet loss, send bitrate, etc. Their opinion model has been standardized in ITU-T Rec J.247 [22]. Once video contents are classified based on users' perceptive quality (MOS scores), it is possible to apply priority control to content delivery.

V. CONCLUSIONS

In this paper, we analyzed the relationship between video contents and objective video quality and found the degree of influence of each of the QoS parameter using principal component analysis on video quality. We found the relationship between video contents and video quality using multiple linear regression analysis and finally, applied the results to rate control methods.

Video contents were classified into three groups based on MOS scores obtained from quality parameters both in the application and network level using cluster analysis with strong prediction accuracy. The purpose of content classification was not the re-organization of existing genres of video contents, but making new groups that had similar attributes. We also compared our content classification to the traditional method of spatio-temporal grid based on feature extraction. We found strong correlation between the spatio-temporal grid and our classification based on the MOS scores obtained by objective video evaluation. Future work will include MOS obtained from subjective testing to verify our classification model.

We further applied rate control methods to contents in the same group thus enabling content providers to identify the video send bitrates that correspond to various quality levels and hence provide high-quality streaming services over wireless networks. We also determined the influence of each QoS parameter on video quality for all content types, thus enabling network operators to allocate network resources according to users' perceptual quality and hence optimize network resources.

Our future work will focus on applying the same priority to video content within the same group and hence optimize users' perceptual quality based on both network resource and video codecs (e.g send bitrate, frame rate).

REFERENCES

[1] ITU-T. Rec P.800, "Methods for subjective determination of transmission quality," 1996.

[2] M. Siller and J. Woods, "Improving quality of experience for multimedia services by QoS arbitration on a QoE framework" in Proc. the 13th Packed Video Workshop 2003- polytech.univ-nantes.fr

[3] F. Agboma and A. Liotta, "QoE-aware QoS management", 6th Int. Conf. on Advances in Mobile computing and Multimedia, 24-26 Nov. 2008.

[4] M. Ries, O. Nemethova and M. Rupp, "Video quality estimation for mobile H.264/AVC video straming", *Journal of Communications*, Vol. 3, No.1, Jan. 2008, pp. 41-50.

[5] L. Yu-xin, K. Ragip, B. Udit, "Video classification for video quality prediction", *Journal of Zhejiang University Science A,* 2006 7(5), pp 919-926.

[6] Y. Kato, A. Honda and K. Hakozaki, "An analysis of relationship between video contents and subjective video quality for internet broadcasting", *Proc. Of the 19th Int. Conf. On Advanced Information Networking and Applications (AINA),* 2005.

[7] Y. Suda, K. Yamori and Y. Tanaka, "Content clustering based on users' subjective evaluation", *6th Asia-Pacific symposium on Information and Telecommunication Technologies ASPITT*, 2005, Volume , Issue , 09-10 Nov. 2005 pp. 177 – 182.

[8] W. J. Krzanowski, "Principles of multivariate analysis", Clarendon press, Oxford, 1988

[9] J. Wei, Video content classification based on 3-d Eigen analysis, *"IEEE transactions on image processing*", Vol. 14, No.5, May 2005.

[10] G. Gao, J, Jiang, J. Liang, S. Yang and Y. Qin, PCA-based approach for video scene change detection on compressed video, *Electronics letters*, Vol. 42, No.24, 23rd Nov. 2006.

[11] Video quality experts group, multimedia group test plan, Draft version 1.8, Dec 2005, http://www.vqeg.org

[12] Video quality Metric Software (VQM), http://www.its.bldrdoc.gov/vqm/

[13] Perceptual Evaluation of Video Quality (PEVQ), Opticom http://www.pevq.org

[14] Network Simulator (NS2), http://www.isi.edu/nsnam/ns/

[15] J. Klaue, B. Tathke, and A. Wolisz, "Evalvid – A framework for video transmission and quality evaluation", *In Proc. Of the 13th International Conference on Modelling Techniques and Tools for Computer Performance Evaluation,* Urbana, Illinois, USA, 2003, pp. 255-272.

[16]] Ffmpeg, http://sourceforge.net/projects/ffmpeg

[17] J. Mitchell and W. Pennebaker, "MPEG Video: Compression Standard, Chapman and Hall, 1996, ISBN 0412087715.

[18] R. J. Hyndman, "The problem with Sturges' rule for constructing histograms", Monash University 1995.

[19] N. Cranley and L.Murphy, "Incorporating user perception in adaptive video streaming services", in Digital Multimedia Perception and Design (Eds. G. Ghinea and S. Chen), published by Idea Group, Inc., May 2006. ISBN: 1-59140-860-1/1-59140-861-X.

[20] A. Khan, L. Sun and E. Ifeachor, "Content-clustering based video quality prediction model for MPEG4 video over wireless networks", Accepted to the IEEE ICC Conference, Dresden, Germany, 14-18 June, 2009.

[21] K. Yamagishi and T. Hayashi, "Opinion model using psychological factors for interactive multimodal services", IEICE Trans. Communication, Vol.E89-B, No. 2, Feb. 2006.

[22] ITU-T Recommendation J.247, "Objective perceptual multimedia quality measurement in the presence of a full reference", Aug. 2008.