QOS MAPPING USING HUMAN FACTORS FOR FINE GRANULAR SCALABILITY

Masayuki Inoue, Shigetaro Iwatsu, and Yoshimitsu Ohtani

NTT Cyber Space Laboratories, NTT Corporation
1-1 Hikari-no-oka, Yokosuka-shi, Kanagawa, 239-0847 Japan
E-mail:{inoue.m, iwatsu.shigetaro, ohtani.yoshimitsu}@lab.ntt.co.jp

ABSTRACT

To offer successful streaming video services, the Internet must have some kind of QoS control at each level of the protocol stack. This paper uses Fine Granular Scalability for application-level QoS, because of its ability to adjust to a wide range of network channel capacities. We conduct an experiment using the DSIS method to subjectively assess Fine Granular Scalability. Our results show that Fine Granular Scalability provides only SNR scalability, not subjective image quality scalability. A multiple regression analysis shows that we can establish QoS mapping between user and application-levels by using both application-level QoS parameters and the human factor. Furthermore, we show that FGS can adjust to not only a wide range of channel capacities but also a wide variety of users by using the indicated QoS mapping.

***Index Terms*—** Video coding, Human factors, Quality factor, Quality control

1. INTRODUCTION

Many different types of networks are now converging on the all-IP network and the Internet. Moreover, network capacities are increasing radically. Therefore streaming video over the Internet has been rapidly adopted over the past few years. To offer successful streaming video services, the Internet must have some kind of QoS control at each level of the protocol stack and we have to remember that users have a wide variety of profiles. Tasaka considered the QoS at each level of the protocol stack. He identified six kinds of QoS: (1) physical-level QoS, (2) node-level QoS, (3) network-level QoS, (4) end-to-end-level QoS, (5) application-level QoS, and (6) user-level QoS (or perceptual QoS)[1].

This paper uses MPEG4 FGS (Fine Granular Scalability)[2] for application-level QoS, because of its ability to adjust to a wide range of network channel capacities. However, FGS provides only SNR scalability, not subjective image quality scalability. This is unfortunate since the latter is more important than the former; the users are the true determiners of service acceptance. Furthermore, the control of bitrate over FGS bitplane is not user friendly, because users can not imagine the subjective image quality from the bitrate.

No paper has considered in detail how to map the application-based QoS level to user-satisfaction level when using FGS. We rectify this omission by conducting an experiment in which the DSIS(double-stimulus impairment scale) method[3] is used to subjectively assess FGS. We consider QoS mapping relations between application-level parameters, such as FGS codec parameters, and user-satisfaction level parameters, such as subjective image quality grading points[4]. Furthermore, this paper describes how to control FGS bitplanes to support a wide variety of users through the human factors, i.e. user profiles.

2. EXPERIMENT METHODOLOGY

We examined the relation between user and application-levels by assessing the quality of decoded videos in terms of the QoS parameters at the application-level.

2.1. Experimental Model

In this paper, we examine the effect of subjective image quality using the simplified model. For QoS mapping between user-level and application-level, we consider that subjective image quality consists of four elements: 'image', 'encoder', 'decoder', and 'user'. In this case, we use 'SAD(Sum of Absolute Difference)' as the image-element parameter, 'QP value' and 'Base bitrate(Base Layer bitrate)' as the parameters for encoder-element, 'FGS bitplane' and 'FGS bitrate (Enhancement Layer bitrate)' as the parameters for decoder-element, and 'Gender', 'Age', 'Imagery', and 'Interest' as the user-element parameters, which are human factors as is shown in Table 1. For application-level QoS, we adopt 5 parameters: 'SAD', 'QP value', 'Base bitrate', 'FGS bitplane' and 'FGS bitrate'. For user-level QoS, we adopt the single parameter of 'grading point'[3].

2.2. Subjects

The 24 subjects, 12 males and 12 females, had no previous experience evaluating images. To obtain information on human factors, we used the questionnaire shown in Table 1.

Table 1. Questionnaire.

human factors	questions
gender	male or female?
age	twenties or thirties?
imagery	Please imagine yourself seeing an image. Which type of image do you imagine, TV, MOVIE, DVD, or others?
interest	Do you have an interest in image quality?

Table 2. Coding specifications.

item	mode
Base Layer	MPEG4 ASP(YUV)
Enhancement Layer	MPEG4 FGS(Y)
I-VOP	1 in every 30 frames
Quantization	H.263
Search range	$(\pm 8)\times(\pm 8)$[pels]

2.3. Methodology

To conduct the assessment from the user-level QoS point of view, we used the DSIS method. The reference video (the original video), and the test decoded-video were presented only once. After each presentation, the subject pushed a button to identify the grading scale. Grades 5, 4, 3, 2, and 1 mean 'imperceptible', 'perceptible, but not annoying', 'slightly annoying', 'annoying', and 'very annoying', respectively. We used three original videos (CIF size): 'Susie(frame no.:1-150)', 'Foreman(frame no.:91-240)', and 'Carphone(frame no.:91-240)'. Table 2 shows the coding specifications. Each original video was coded as per Table 2. Table 3 shows decoding parameter sets for the test decoded-videos. Decoding was performed using the parameter sets of QP values and FGS bitplanes. The three original videos were multiplied by the 30 parameter sets to yield 90 test decoded-videos. In this paper, an FGS bitplane value of 1 means that only the MS bitplane of the coded video was decoded. A QP value of 1 and an FGS bitplane value of 3 means that all data up to and including the LS bitplane was decoded.

3. RESULTS AND DISCUSSION

The experiment yielded 2160 results from the 24 subjects and 90 test decoded-videos.

3.1. Fine Granular Scalability Performance

Figure 1 shows the mean grading point of each set of QP values and FGS bitplanes. ANOVA(analysis of variance) showed that there were significant differences

Table 3. Parameter sets for test decoded-videos.

QP values	FGS bitplanes
1	1(M),2,3(L)
3	1(M),2,3(L)
5	1(M),2,3,4(L)
10	1(M),2,3,4,5(L)
15	1(M),2,3,4,6(L)
20	1(M),2,3,4,6(L)
31	1(M),2,3,5,7(L)

(M):MS bitplane
(L):LS bitplane

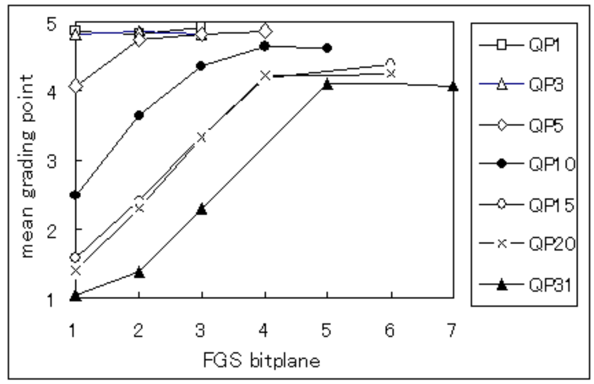

Fig. 1. Comparison of grading point for all videos.

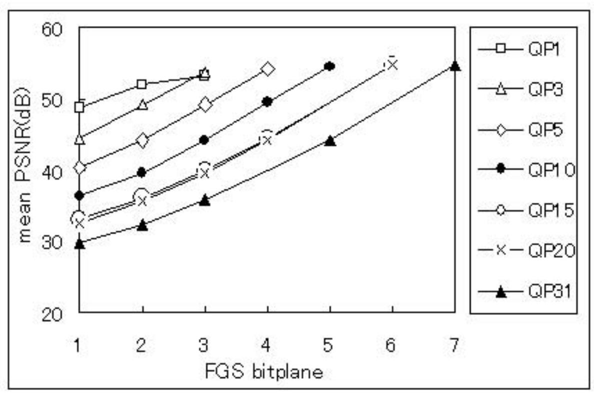

Fig. 2. Comparison of PSNR for all videos.

in the QP value($F(6,2130)=857.75$, $p < 0.05$) and FGS bitplanes($F(6,2130)=546.25$, $p < 0.05$), where F is F-ratio. We observe that the larger the base-layer QP value is, the smaller the maximum of the mean grading point is.

Figure 2 shows the mean PSNR of each set of QP values and FGS bitplanes. ANOVA showed that there were significant differences in the QP value($F(6,2130)=30273.62$, p

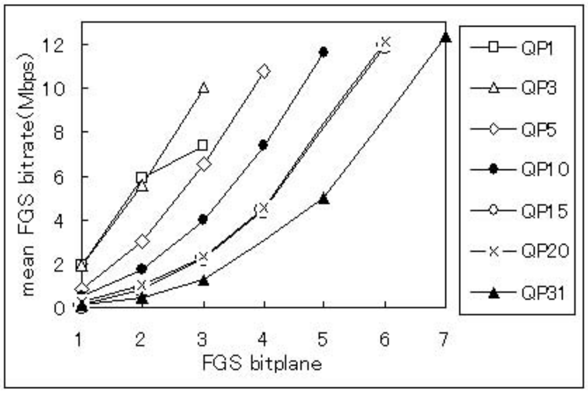

Fig. 3. Comparison of FGS bitrate for all videos.

Table 4. Principal component loading.

	loading	
Image quality factor	first	second
SAD	0.139	0.271
QP value	0.103	-0.864
FGS bitplane	0.951	-0.232
Base bitrate	-0.086	0.878
FGS bitrate	0.951	0.230

< 0.05) and that in FGS bitplanes($F(6,2130)=40088.29$, $p < 0.05$). The tendency is that the bigger the FGS bitplane is, the bigger the mean PSNR is, which is different from the trend in Figure 1.

Figure 3 shows the mean FGS bitrate of each set of QP values and FGS bitplanes. ANOVA showed that there were significant differences in the QP value($F(6,2130)=2907.47$, $p < 0.05$) and FGS bitplanes($F(6,2130)=11721.06$, $p < 0.05$). We observe the tendency in which the bigger the FGS bitplane is, the bigger the mean FGS bitrate is. Comparing FGS bitplane 5 with FGS bitplane 7 at QP value of 31, 7's bitrate is more than twice that of 5. However, Figure 1 shows that 5 and 7 have the same mean grading point. This means that the FGS bitplane value impacts SNR but not subjective image quality. Therefore, it is important to determine the QoS relation between application-level and user-level.

3.2. QoS mapping Application-level to User-level

We carried out principal component analysis using the 'SAD', 'QP value', 'FGS bitplane', 'Base bitrate', and 'FGS bitrate'. The result was that the first two principal components had characteristic-values of more than one, so we adopted the first and second principal components. Table 4 shows the principal component loading. According to the result, we adopted two sets of QoS parameters as predictor variables in multiple regression analysis. Owing to multi-colinearity, we selected QP value and FGS bitplane for one and Base bitrate and FGS bitrate for the other. This yielded

$$GP = 3.761 - 0.106QP + 0.492FP \quad (1)$$

$$GP = 2.607 + 0.234BB + 0.181FB \quad (2)$$

where GP, QP, and FP stand for Grading point, QP value, and FGS bitplane, respectively, and BB and FB mean Base bitrate and FGS bitrate, respectively. The contribution rates adjusted for the degrees of freedom for (1) and (2), are 0.639 and 0.429, respectively. This means that regression equation (1) is a better model than regression equation (2). Therefore, we can accurately estimate subjective image quality using the two application-level QoS parameters of QP value and FGS bitplane.

Table 5. Questionnaire results.

Sub. No.	gender	age	imagery	interest
1	male	30s	TV	yes
2	male	20s	others	no
3	female	20s	TV	no
4	female	30s	TV	yes
5	male	30s	TV	no
6	male	30s	DVD	yes
7	female	20s	DVD	yes
8	female	20s	MOVIE	no
9	female	30s	TV	no
10	male	30s	TV	no
11	male	20s	TV	no
12	male	20s	TV	yes
13	female	30s	DVD	no
14	male	20s	MOVIE	yes
15	female	20s	TV	no
16	female	20s	TV	yes
17	female	30s	TV	yes
18	male	20s	TV	yes
19	female	20s	DVD	no
20	female	30s	MOVIE	yes
21	female	30s	TV	no
22	male	30s	TV	yes
23	male	20s	TV	yes
24	male	30s	TV	yes

3.3. Human Factor

The questionnaire results are shown in Table 5. Since these results are qualitative, we transformed them into quantitative values for multiple regression analysis, as is shown in Table 6. The human factors 'gender', 'age', 'imagery', and 'interest' were assigned dummy variables 'g', 'a', 's1,s2,s3', and 'k',

Table 6. Dummy variables.

Sub. No.	gender	age	imagery			interest
-	g	a	s1	s2	s3	k
1	1	0	1	0	0	1
2	1	1	0	0	0	0
3	0	1	1	0	0	0
4	0	0	1	0	0	1
5	1	0	1	0	0	0
6	1	0	0	0	1	1
7	0	1	0	0	1	1
8	0	1	0	1	0	0
9	0	0	1	0	0	0
10	1	0	1	0	0	0
11	1	1	1	0	0	0
12	1	1	1	0	0	1
13	0	0	0	0	1	0
14	1	1	0	1	0	1
15	0	1	1	0	0	0
16	0	1	1	0	0	1
17	0	0	1	0	0	1
18	1	1	1	0	0	1
19	0	1	0	0	1	0
20	0	0	0	1	0	1
21	0	0	1	0	0	0
22	1	0	1	0	0	1
23	1	1	1	0	0	1
24	1	0	1	0	0	1

respectively. We then performed multiple regression analysis using the dummy variables as the predictor variables. We took the human factors of 'gender', 'age', 'imagery', and 'interest' into account as follows:

The multiple regression equation that considers g, a, s1, s2, s3, and k became

$$\begin{aligned} GP &= 3.856 - 0.106QP + 0.492FP + 0.095g - 0.127a \\ &\quad + 0.020s1 + 0.217s2 + 0.058s3 - 0.239k \end{aligned} \quad (3)$$

The contribution rate adjusted for the degrees of freedom for (3) became 0.646. Regression equation (3) shows that the goodness of fit is slightly better than that of regression equation (1). We make the following observations from (3). First, GP decreases as QP increases and increases as FP increases. Next, the coefficients of s2 and k are slightly larger than those of the others. This means that the human factors of 'imagery' and 'interest' impact GP more than the others. Also, we can transform equation (3) as follows:

$$GP = BGP + FGP \quad (4)$$

$$BGP = 3.856 - 0.106QP \quad (5)$$

$$\begin{aligned} FGP &= 0.492FP + 0.095g - 0.127a + 0.020s1 \\ &\quad + 0.217s2 + 0.058s3 - 0.239k \end{aligned} \quad (6)$$

where BGP and FGP stand for Base grading point and FGS grading point, respectively. Furthermore, we can also transform equations (5) and (6) into equations (7) and (8), respectively.

$$QP = (BGP - 3.856)/(-0.106) \quad (7)$$

$$\begin{aligned} FP &= (FGP - (0.095g - 0.127a + 0.020s1 \\ &\quad + 0.217s2 + 0.058s3 - 0.239k))/0.492 \end{aligned} \quad (8)$$

Equation (7) means that QP values can be determined if you know the Base grading point required by the user. Equation (8) means that the FGS bitplanes can be adjusted through the FGS grading point required by the user and the user's human factors. This indicates that we can control QP values and FGS bitplanes in compliance with the users' specified subjective image quality.

Note that it is easier to control the subjective image quality control via the FGS bitplane than by bitrate, because users can not imagine the subjective image quality from a bitrate. Equation (8) also indicates that we can control the FGS bitplane term so as to adjust to a wide variety of users according to the human factors, i.e. user profiles. In consequence, FGS can be adjusted freely to support not only a wide range of channel capacities but also a wide variety of user demands.

4. CONCLUSIONS

We conducted an experiment using the DSIS method to subjectively assess Fine Granular Scalability. By using multiple regression analysis, we obtained a QoS mapping relationship between user and application-levels through the use of both application-level QoS parameters and the human factors. Finally, using the QoS mapping so identified, we showed that FGS is one of the best approaches to supporting not only a wide range of channel capacities but also a wide variety of user demands.

Fruitful discussion with Prof. Yoshinori Hatori of Tokyo Institute of Technology is greatly appreciated.

5. REFERENCES

[1] S. Tasaka and Y. Ishibashi, "Mutually compensatory ...," in Conf. Rec. IEEE ICC2002, pp.1105–1111, 2002.

[2] W. Li, "Overview of fine ...," IEEE Trans. Circuits Syst. Video Technol., vol.11, pp.301–317, 2001.

[3] ITU-R Recommendation BT.500-11:"Methodology for the Subjective Assessment ...," 2002.

[4] M. Inoue, K. Jinzenji, K. Fukazawa, S. Iwatsu, and Y. Yashima, "A Study on QoS Mapping for ...," in Proc. IWAIT2006, S03-7, pp.105–110, 2006.