Encoding Strategies for Realizing MPEG-4 Universal Scalable Video Coding

Yi-Shin Tung, Jin-Hau Kuo, Ja-Ling Wu, Wen-Huang Cheng, Ting-Jian Pan
The Communications and Multimedia Laboratory,
Department of Computer Science and Information Engineering,
National Taiwan University, Taipei, Taiwan

ABSTRACT

The universal scalability, which integrates different types of scalabilities and consequently provides a large scaling range for each codec parameter, is of high interests to the applications in the current heterogeneous surroundings. In our previous work, an MPEG-4 universal scalable codec basing on a layered path-tree structure [1,2] has been addressed, in which a video layer and the coding order of two consecutive layers are interpreted as a node and the parent-to-child relationship in a path-tree, respectively. Since individual video layers can be coded separately using different coding tools, defined in MPEG-4 simple scalable profile (SSP) [3] and fine-granularity scalable profile (FGS) [4], the proposed scalable video coder may achieve spatial, temporal and SNR enhancements simultaneously. In this paper, based on some visual observations, we will address some encoding strategies applicable to the universal scalable coding, including the spatial-temporal quality tradeoff, the region sensitivity and the frequency weighting. Applying these strategies will take the content characteristics into consideration and can determine better coding parameters. As a result, the bit allocation of an encoder becomes more sensitive and correlative to those perceptually important parts of spatial, temporal and SNR enhancements. Next, a batch encoding process is conducted to generate universal scalable streams automatically, in which all the abovementioned encoding strategies are fully integrated together. The preliminary experiments show that better visual quality can be obtained within a large range of bitrates.

Keywords: Video scalability, video perceptibility, FGS, MPEG-4, universal scalability, region sensitivity, frequency weighting

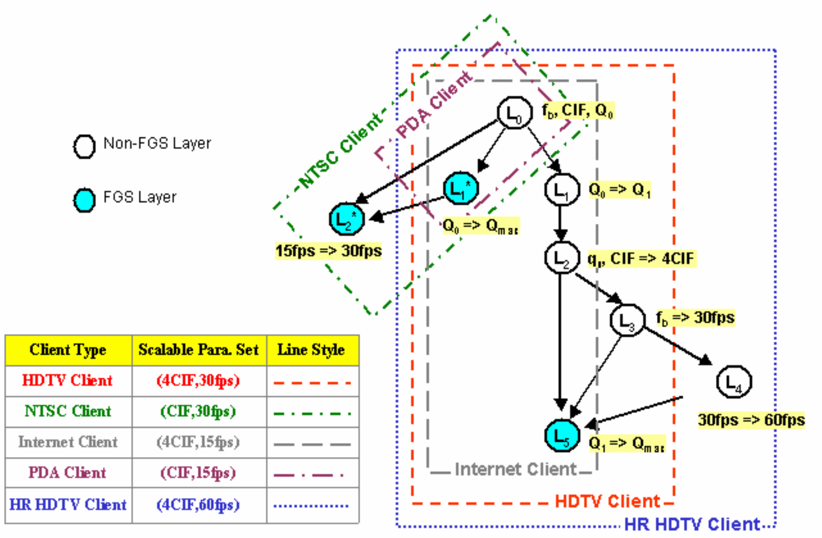
1. INTRODUCTION

Internet, the largest heterogeneous network in the world, becomes the worldwide interoperable medium for digital communication. Once any individuals or sub-networks connect to Internet, they become a legal member of the network and able to access any authorized services executed by some organizations. Different types of connections are adopted today. For examples, mobile devices uses 3G-H.324M, home devises connect through ADSL lines, and community devices choose wireless LAN solutions. All these physical connections have different networking characteristics, i.e. bandwidth, data lost rate, delay and network buffer size. On the other hand, the devices used to communicate in the heterogeneous networks are also varying in their available computation resources, input and output capabilities and etc. Cellular phone, PDA, notebook, and a lot of consumer products have different computation power and display resolution, but can access to Internet as well. These variant communications and clients' properties comprise the heterogeneous interconnected environment, and they stimulate the demand for a common video service, which can be practiced under different settings of bandwidth, computation, or output video resolution. To this end, a methodology for hybrid integration of different types of scalabilities is desirable such that a single encoding is enough to serve all clients coming from different networks or equipped with different processing capabilities. Thus, the universal scalability, which includes all practical types of scalabilities and fits to most of clients, deserves a lot of attentions today.

In our previous work, we have designed a universal scalable video coder whose configuration can be described based on a predefined path-tree structure. In our design, a traditional video layer is represented as a node in the tree, and the correlation between the base layer and each enhancement layer are considered as the parent-to-child relationship between the corresponding two nodes. Generally speaking, a single video enhancement (layer) can carry about one enhancement type (will be detailed later). To be universal scalable, multiple video layers should be simultaneously adopted, each of them improves spatial, temporal or SNR resolution, respectively, and combining all enhancements achieve the universal scalable capability. The scalability configuration (represented by the path-tree structure) of the proposed coder can be changed adaptively to fit for same major groups of clients. That is to say, the path-tree layered enhancement structure is constructed according to application requirements. Based on this design, we realized a universal scalable coding scheme for five types of clients (HDTV, NTSC, Internet, PDA, and High Resolution HDTV), each has different optimal quality requirements. The base layer and enhancement layers are all coded using tools defined in the MPEG-4 simple profile (SP), the simple scalable profile (SSP) and the streaming video profile. The decision on path-tree traversal conducts the layer selection process to achieve five desired types of scalabilities: spatial, temporal, quality, rate and computational scalabilities.

The path-tree structure of the resultant coding scheme is shown in Fig. 1. For users of different requirements, a sub-tree including nodes bring out positive effects is chosen statically when the service is requested. During the operating period of a service, one of the paths from root to leaf nodes in the truncated sub-tree is dynamically selected according to the time-varying environmental parameters, such as bitrate and computation capability. As a result, the visual quality is getting improved when the lower video layers are adopted first and then the higher enhancement ones.

In summary, our previous work addressed an MPEG-4 based universal scalable coding structure but left the decision of coding parameters for being determined during the later encoding optimization. Developing a good encoding mechanism is the key to demonstrate the superiority of the designed coding system. In this paper, we first investigate some perception-based encoding strategies and then apply them to decide the coding parameters of the proposed universal scalable encoding scheme. We further implement a batch encoding process, which integrates all devised strategies together, and therefore, can better approach to the optimal video quality that human can perceive.



Client Type	Scalable Para. Set	Line Style
HDTV Client	(4CIF,30fps)	– – – – –
NTSC Client	(CIF,30fps)	– · – · –
Internet Client	(4CIF,15fps)	— — —
PDA Client	(CIF,15fps)	— · — ·
HR HDTV Client	(4CIF,60fps)	··············

Figure 1: The realized MPEG-4 universal scalable codec and its path-tree structure.

2. ENCODING STRATEGIES

A coding performance comes from a better understanding of the characteristics of the content that being encoded. Therefore, conducting a precise video analysis and investigating the human visual perceptibility are keys to derive better coding parameters. Effective video analyzing tools and appropriate formulation for exploring content characteristics are the mandatory tasks when implementing an optimized encoder. Moreover, in the universal scalable coding the encoding strategies are no longer limited to be applied to a single layer. Instead, a top-down encoding strategies involved in harmonizing among different layers should be considered, carefully. Based on some visual observations we will address some encoding strategies in this section, which help us to find better coding parameters.

2.1. Video Smoothness, Frame Quality and Dynamic Frame Rate Encoding

It is observed that when seeing a high-activity video clip, people cannot see the individual frame in detail; oppositely, when seeing a smooth video clip, people cannot acknowledge a few frame-drops even if the presenter did it on purpose. Thus, when applying temporal scalability, the number of representative frames in the base layer should change depending on content's smoothness. A better choice of these representative frames would improve not only the visual perceptibility of the temporal base layer but also the rate utility with respect to the overall spatial and temporal quality. For example, if the number of selected frames per second is low, the bitrate should be allocated first to improve the quality of frames of the temporal base layer. This arrangement is preferred when coding smooth videos. On the other hand, if the percentage of base layer frames is high, more low-quality frames should be reconstructed in advance. This setting is more suitable for coding high motion-intensive videos.

In order to enable the adaptability to content smoothness, a dynamic-frame-rate scheme should be adopted for coding the temporal base layer. Several simple operations can be used to measure frame's smoothness. All of them can be classified into two classes: the pixel-based measurement and the histogram-based measurement. In general, the former compares the frames' signal samples by a certain operator and then sums up their results together. This class of measurement is suitable for assessing the motion-activity of a video. On the other hand, the histogram-based one gets the intensity histogram of all pixels first, and then calculates the histogram difference of neighboring frames. However, since histograms do not contain the spatial-related information, it is more suitable to be applied to shot-change detection based applications. So, the pixel-based measurement is adopted to evaluate the frames' smoothness, in this work.

In our design, the smoothness evaluation of the current frame to the last selected frame is made as the basis of choosing the next representative frame. The result of the smoothness measurement will increase when the frames' distance is enlarged, because the object motion within a shot is accumulated. Finally, the accumulated frame difference will exceed a certain threshold after a few frames (e.g. 2 to 4 frames). The frame at this instance is selected and coded into the base layer. All other frames are thought to be of less importance and are coded into the temporal enhancement layer. As a result, by using the prescribed configuration, the video smoothness will affect the frame rate of the base layer as what we want. Fig. 2 depicts the flowchart of the frame classification process.

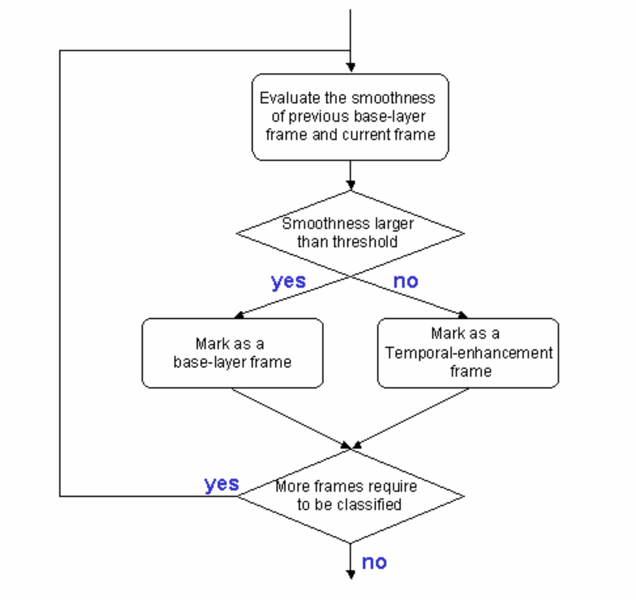


Figure 2: The flowchart of selecting the representative frames for the temporal base layer.

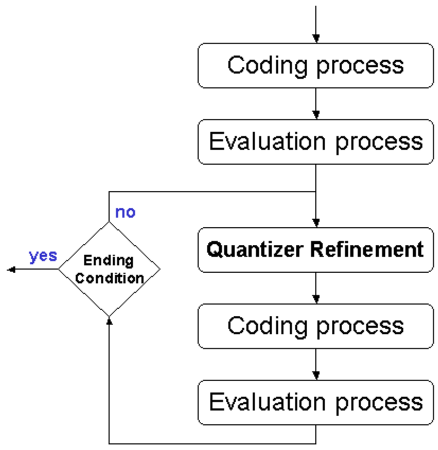


Figure 4: The flowchart of the iterative process to obtain a better quantizer map.

Moreover, the sensitivity to content smoothness and the individual frame's quality behaves contrarily. Taking this fact into consideration, when the coded frame rate in the base layer changes from one-half to one-third, the corresponding average quantization scale also changes from 10 to 15. The formula we used to reflect this phenomenon is

$$q \times f = C, \tag{1}$$

where q and f respectively represent the average quantizer of one frame and the average frame subsampling ratio within the window centered at that frame, and C is a predefined constant.

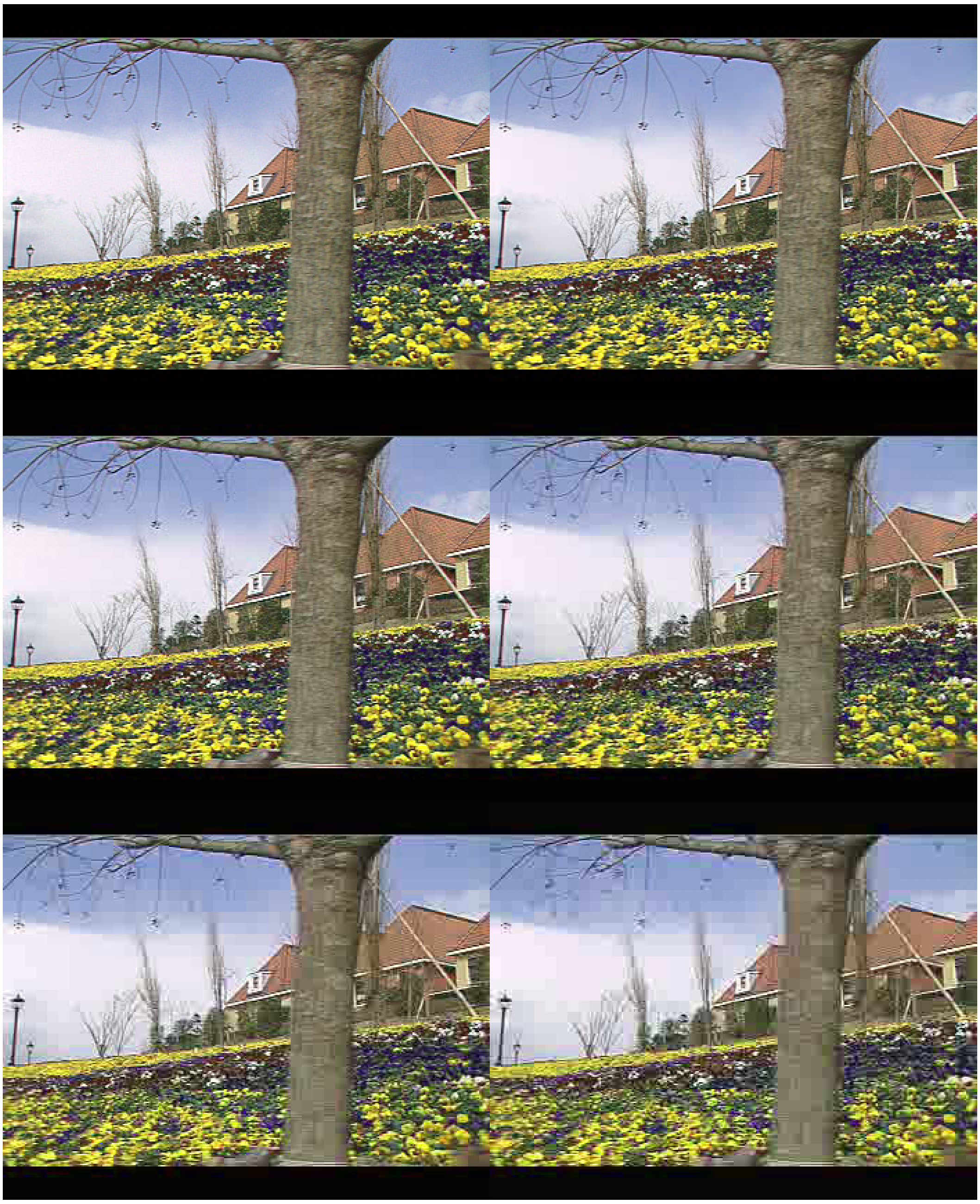


Figure 3: The first frame of the garden sequence coded by different quantizers: They are original, q=4, q=8, q=16, q=23 and q=31, from top to down and from left to right.

2.2. Region Sensitivity and the Sarnoff's JNDmetrix

Region sensitivity depends on the characteristics of the inspection region. In general, distortion occurred in the smooth area is most perceptible; the next is the edge area, and finally the texture area [5]. The less sensitive the region is, the higher masking property it possesses. An example is shown in Fig. 3. Different levels of quantization distortions are added to the test image. Usually, an observer detects the main truck of the tree to be the most annoying. The next may be the thin branch of the tree in front of the blue sky or the intersection of the blue sky and the white cloud. Oppositely, the distortions of the flowers and the house roof are nearly invisible. When encoding videos with constrained bitrates, some signal distortions are unavoidable. In such cases, making all inspection regions with nearly the same visible distortion will bring out the maximal overall quality. A mechanism to detect region characteristic is desired for video coding. Research on region perceptibility has been developed for years. Several famous quality assessment metrics have been devised, such as Sarnoff's JNDMetrix [6], NASA's DVQ [7], and EPFL's model [8]. Their performances are statistically equivalent [9], and no one can absolutely outperform the others. Instead of re-developing a similar metric by exhausting our vitality, we decide to adopt the famous assessment metric, the Sarnoff's JNDMetrix, to measure the region sensitivity.

In our work, finding a better quantizer map under the constraint of a given average quantizer value, such that the visible distortion is equally distributed, is our goal. Taking advantage of the pixel-based distortion measurement (called JND map) exported by the JNDmetrix, a simple Macroblock(MB)-based distortion, d_{mb}, can be obtained by summing up the quantitative values of all pixels within the MB. The derived MB-based measurements are used to adjust the corresponding quantizer to that MB. We designed a coding-result-driven process, in which the following three processes are operated in turn to iteratively approach to the final quantizer map.

(i) The encoding process: The current quantizer map is used to code the input frame.

(ii) The evaluation process: The coded frame and the original frame are fed to the JNDmetrix to obtain its pixel-based JND measurement, and then converted to MB-based one. The root of the variance (d_{var}) of all MB-based distortion values multiplied by their mean (d_{mean}) for each frame is used as the objective function to be minimized.

(iii) The quantizer map refinement: The MB-based measurement, d_{mb}, is then adopted to refine the corresponding quantizer for the next evolution. The ratio of the MB's distortion to its mean (γ) are used to derive the new quantizer, as shown in Eqns. (2) and (3). If the ratio falls within a certain range near to 1, we regard that quantizer feasible, and do not alternate it. Otherwise, additive-increase and multiplicative-decrease scheme is applied to derive the new quantizer. The symbols α and β are both constants, which represent the allowable stable range and the adaptive increasing step, respectively. The formulas we used to conduct this refinement process are:

$$\gamma = {}^{d_{mb}}\!\big/_{d_{mean}} \tag{2}$$

and

$$q_{new} = \begin{cases} q_{old} \times (1-\gamma) \times \beta, & \text{if } \gamma < (1-\alpha) \\ q_{old}, & \text{if } (1\text{-}\alpha) \le \gamma \le (1+\alpha) \\ q_{old} \times \gamma, & \text{if } (1+\alpha) < \gamma \end{cases} \tag{3}$$

The flowchart of this process is illustrated in Fig. 4. In each iteration, the new evaluation result is compared to the one obtained in the last iteration. If the objective function value becomes better, the iterative process continues; otherwise, we stop the whole process. And the quantizer map in the last run is adopted as the final quantizer map. A resultant evolution of this refinement process for the first frame of the garden sequence is shown in Fig. 5. The intensity in this figure indicates the quantitative distortion from user perspective (reported from JND tools). As we expected, the refinement process will quickly make the distortion equally distributed (all regions have similar intensity after a few runs of iterations).

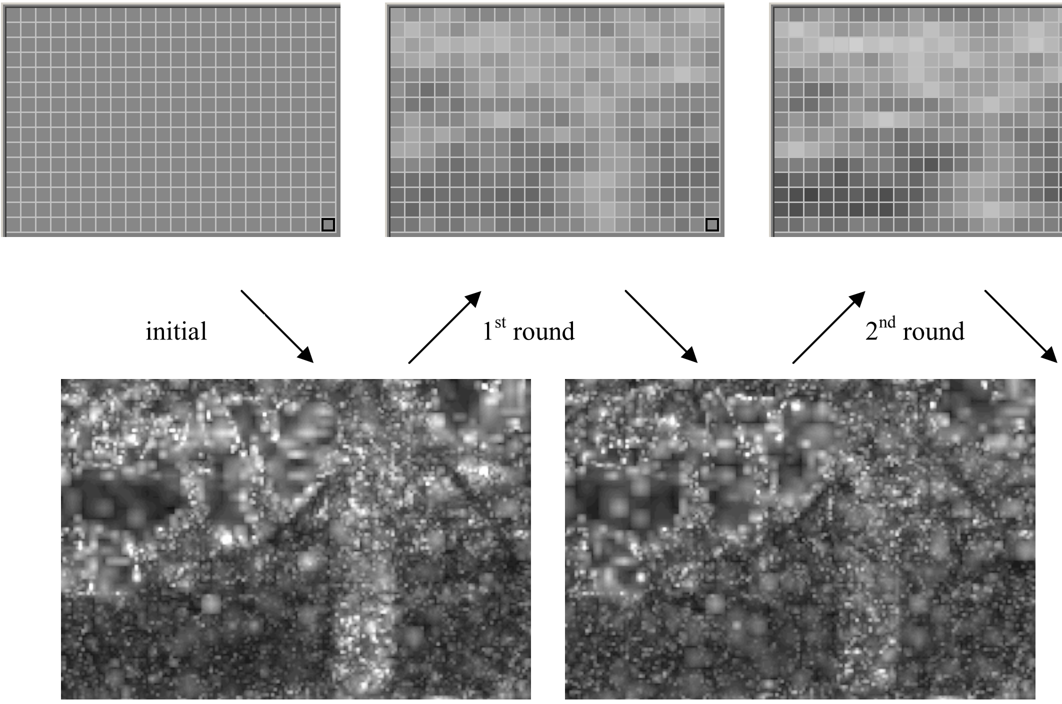


Figure 5: The iterative process to get better quantization map. Each macroblock is depicted as a grid in the map. A brightly colored grid indicates a low quantizer value, while a dark grid indicates a large one.

2.3. FGS Coding Strategies

MPEG-4 FGS is a new advance to scalable coding technology in which the bitplane coding of DCT coefficients is used such that the enhancement bitstream can be truncated at any location while still keep the obtainable best perceptual quality. It helps to achieve fine-rate adaptability. In our universal scalable system, FGS coding tool is always used in the last enhancement layers (i.e. all leaf nodes in the path-tree) to improve the SNR resolution. Through this setting, the whole system can have fine rate and quality adjustability if we choose the scheduling path wisely.

MPEG provides quantization matrices and MB quantizers, which allow the user to adjust the perceptibility of DCT subbands and MB regions, respectively. However, in the default processing of MPEG-4 FGS coding, it did not take these factors into consideration. The connections between the quantization matrices and MB quantizers of the base layer and the frequency weighting and selective enhancement tools of the FGS layer will be addressed in the following.

2.3.1. Strategy I: Applying Frequency Weighting Tool to Represent Different Sensitivity of DCT Subbands

MC-DCT coders convert each frame to frequency domain by performing 8x8 DCT transforms. Each frequency in the DCT domain has different weighting on the human's perceptibility. The quantization matrices defined by MPEG have taken this phenomenon into consideration, and therefore, the visual perceptibility to different frequencies can be reflected by varying these matrices. The different visual sensitivities of DCT subbands also have an impact on FGS layers. In the FGS coding, residues of the base layer are further transformed to the DCT domain, and then coded bitplane by bitplane. All the significant bitplanes of all MBs are coded first, and then the less significant bitplanes. Fortunately, FGS does provide frequency-weighting tools, which can lift bitplanes up for different frequencies before they are coded. To correlate human sensitivity with frequencies, the quantization matrices used in the base layer are exploited again to get the corresponding frequency-weighting matrices. That is

$$BitDiff_i = \log_2 \frac{Q_i}{Q_{\min}} \tag{4}$$

$$freq\_weight_i = \underset{j}{Max}(BitDiff_j) - BitDiff_i \tag{5}$$

where Q_i is the value of the quantization matrix for the i-th DCT coefficient, and Q_{min} is the smallest value among all bins. Table 1 presents the instances of MPEG-4 default quantization matrices and their corresponding frequency weighting tables.

2.3.2. Strategy II: Applying Selective Enhancement Tool to Make Difference of Region Sensitivity

Different regions may have different characteristics, and therefore, are of distinct perceptibility. As prescribed, a quantizer map can be treated as the description of relative importance of all MBs, and the effect of this map should also be considered when applying FGS coding. The selective enhancement tools provided by FGS help us to achieve different sensitivities to different regions. Similar to the frequency weighting tool, the DCT residues of different MBs can be lifted before coded. The quantizer based selective-enhancement map can be formulated as follows.

$$BitDiff_i = \log_2 \frac{q_i}{q_{\min}} \tag{6}$$

$$sel\_enhance_i = \underset{j}{Max}(BitDiff_j) - BitDiff_i \tag{7}$$

where q_i is the quantizer of the i-th macroblock used in the base layer, and q_{min} is the smallest value among all quantizers.

8	17	18	19	21	23	25	27
17	18	19	21	23	25	27	28
20	21	22	23	24	26	28	30
21	22	23	24	26	28	30	32
22	23	24	26	28	30	32	35
23	24	26	28	30	32	35	38
25	26	28	30	32	35	38	41
27	28	30	32	35	38	41	45

Default MPEG-4 Intra-Matrix

16	17	18	19	20	21	22	23
17	18	19	20	21	22	23	24
18	19	20	21	22	23	24	25
19	20	21	22	23	24	26	27
20	21	22	23	25	26	27	28
21	22	23	24	26	27	28	30
22	23	24	26	27	28	30	31
23	24	25	27	28	30	31	33

Default MPEG-4 Non-intra-Matrix

2	1	1	1	1	0	0	0
1	1	1	1	0	0	0	0
1	1	1	0	0	0	0	0
1	1	0	0	0	0	0	0
1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Designed Intra-Frequency-Weighting

1	1	1	1	1	1	1	0
1	1	1	1	1	1	0	0
1	1	1	1	1	0	0	0
1	1	1	1	0	0	0	0
1	1	1	0	0	0	0	0
1	1	0	0	0	0	0	0
1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Designed Non-intra-Frequency-Weighting

Table 1: Default Quantization Matrices of MPEG-4, and the designed frequency weighting matrices.

3. THE REALIZED BATCH ENCODING PROCESS

By integrating all the prescribed strategies together, a batch trancoding system is implemented, which can automatically convert MPEG-1/2 videos to the universal scalable bitstreams without too much user intervention. The inputs to this batch process are simply the decisions of whether to reuse the available region perceptibility implicitly embedded in the input bitstreams (MPEG-1 or MPEG-2) and which path-tree structure is adopted in the scalable encoding process. The flowchart of the realized batch process is depicted in Fig. 6. There are seven stages in the process. The first five stages are trying to generate the content descriptions or encoding parameters in the form of descriptors [10, 11], and they are used to conduct the final scalable encoding process. The following paragraphs describe each of the seven stages briefly.

Stage-1: Video import and descriptions extraction: All input sources are classified based on whether they are MC-DCT coded streams. If they are, the usable information (i.e. the quantizer map and coding mode) is extracted and formatted to generate the corresponding descriptions, which can be utilized by latter encoding process.

Stage-2: Shot change detection: After video is imported, the shot-change detection process is activated and all frames are marked on with respect to their corresponding shots. Histogram-based operation is used to judge the occurrence of shot change. Since few correlations are existed between neighboring shots, processing stages from the next one to the fifth are executed shot-by-shot.

Stage-3: Frame classification and quality decision: If temporal scalability is indicated in the input path-tree structure, the frame classification and average quality decision step (as described in Section 2.1) is applied to select the representative frames for the temporal base layer. Since nearly no informative priority is available from input video, this process is always operated in the encoding process. Once the frame rate of the base layer is decided, the trade-off between spatial and temporal qualities is also utilized to obtain the average quality requirement for each frame.

Stage-4: Coding pattern decision: After deciding which frames belong to the base layer, the B-frame partitioning scheme [12] is adopted. As a result, the coding patterns of the base and enhancement layers are set to “IPPP…” and “BBBB…”, respectively, which can eliminate the temporal redundancy as much as possible.

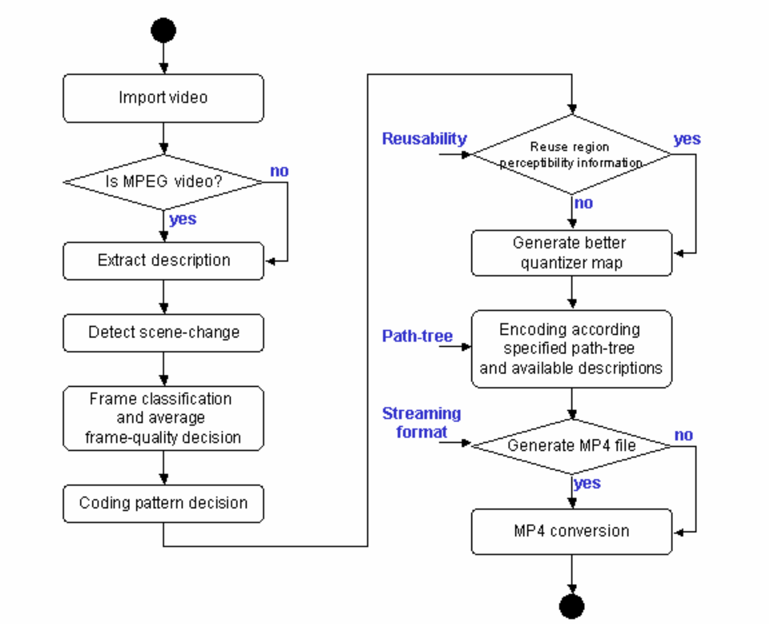


Figure 6: The flowchart of the realized batch encoding process.

Stage-5: Region perceptibility assessment: Well-compressed streams will deliver a good assessment of region perceptibility, such as the highly performed MPEG-2 bitstreams generated by professional encoders. In our batch process, the region perceptibility depends on user's indication of acknowledging if the input stream is good enough. If a user indicates the region perceptibility is worthy to be reused, the time-consuming processes of video analysis and region sensitivity judgment can be eliminated. Otherwise, the refinement process, as described in Section 2.2, has to be conducted.

Stage-6: Universal scalable encoding: Up to now, all required descriptions are generated and refined if necessary. These descriptions and the specified path-tree structure elaborated to fit major types of clients are fed to the universal scalable coder, together. As a result, the desired universal scalable streams are produced according to the so-obtained content smoothness and region perceptibility. In addition, the different weightings of frequency and region sensitivities, as described in Section 2.3, are also applied to MPEG-4 FGS coding in this stage.

Stage-7: MP4 file conversion: Finally, the scalable streams and the corresponding metadata for dynamic delivery of multiple video layers are then converted to MPEG-4 file format (MP4) [3], such that efficient scheduling algorithms can be easily achieved and deliver different quality of video services without involving in codec's implementation details too much.

Based on the above batch process, the universal scalable coding system is implemented. The content descriptions and encoding parameters can be visually represented such that we can make sure if it meets user's feeling well. The operational snapshot of the system is shown in Fig. 7. Up to now, even though a lot of researches focused on understanding the human visual system, it is still not so precise for writing a formula to represent the actual feeling of human beings. Thus, to make the encoding system more subject-oriented, the interactive interface is also provided in our system. Users can encode arbitrary frames several times in any allowable orders, and adjust the coding parameters after subjectively assessing the reconstruction in the last turn. Through this way, we believe that from the interactive encoding process, developers can acquire some tuning experiences, and finally get some refined rules to improve the batch encoding process.

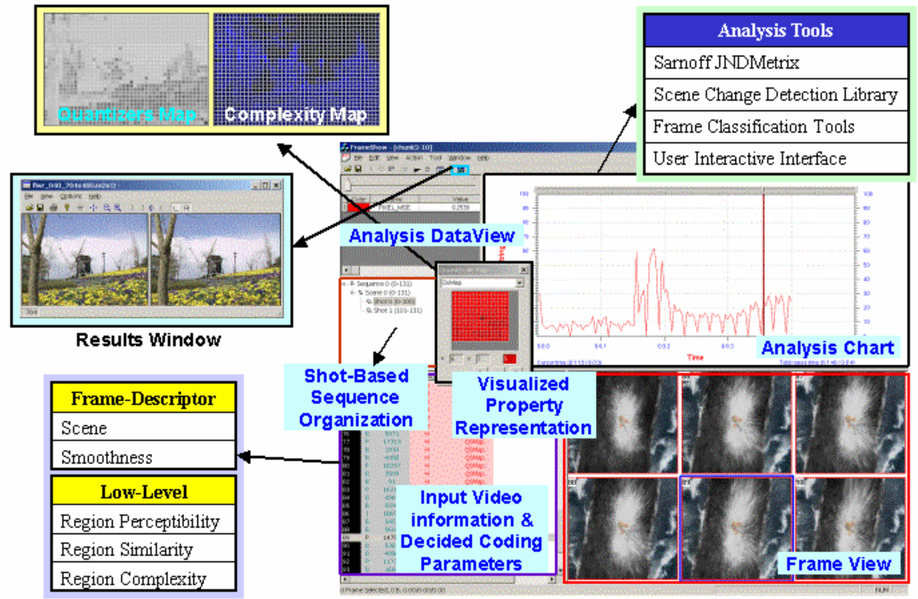


Figure 7: The operational snapshot of the universal scalable encoding system and some visualizations of coding parameters.

4. EXPERIMENTAL RESULTS AND CONCLUSIONS

In our experiments, we merge several test sequences together and treat them as the input sequence. The path-tree structure, as shown in Fig. 1, is input to the encoding system and all prescribed strategies are activated. The dynamic frame rate strategy successfully chooses both the lower frame rate for smooth shots and higher quality setting to those frames in the temporal base layer. A comparison of PSNR and JND before and after applying the strategies under the same bitrate constraint is shown in Fig. 8. We can see that the PSNR becomes slightly worse but JND gets better up to one JND unit improvement (A JND value of 1 indicates the threshold amount of difference required between two stimuli in order for 75% of subjective observers to be able to just barely notice that difference [13]).

According to our preliminary results, the coded streams produced by the batch encoding process possess the advantages of multiple-scalability, better coding efficiency and well-accepted perceptibility. In other word, with the aid of the prescribed encoding strategies, the proposed coding scheme can adapt to the content smoothness and region perceptibility well. The experimental results show that the subjective visual quality and objective based visual measurement become well in spite of the decrease in PSNR.

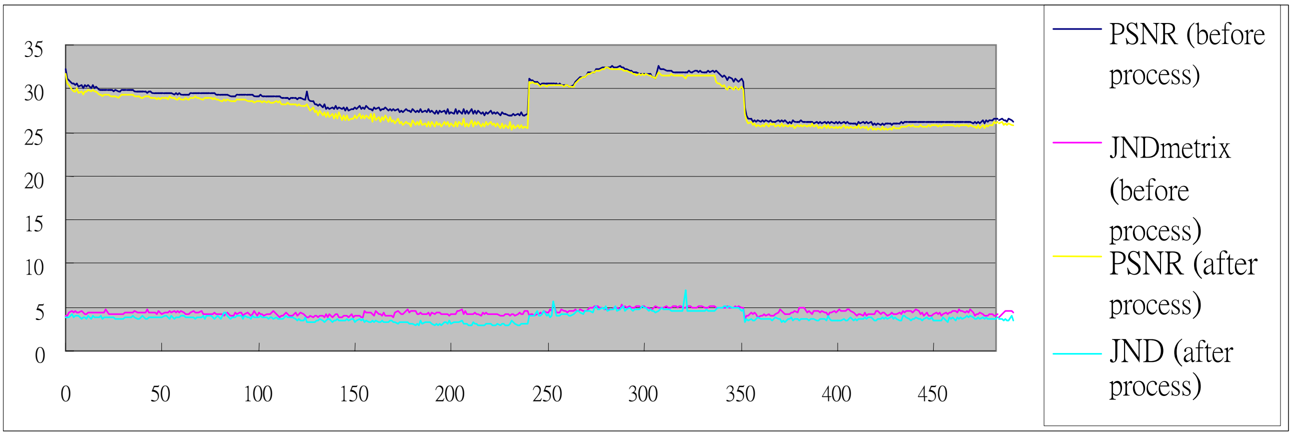


Figure 8: The resultant PSNR and JND of our experimental results before and after the region perceptibility is considered. The input videos ,in sequence, are football, garden, table tennis and mobile & calendar. It shows that the visual quality will get better after the refinement process even if the PSNR becomes worse.

REFERENCES

1. Yi-Shin Tung, Po-Kang Hsiao, Ja-Ling Wu, and Chi-Hui Huang, “A Highly Scalable Video Service based on MPEG-4 Spatial-Temporal-FGS Scalable Codec and Hybrid TCP/UDP Transmission,” *Proc. of 3rd Workshop and Exhibition on MPEG-4*, 2002.
2. Yi-Shin Tung, Ja-Ling Wu and Po-Kang Hsiao, “The Design and Implementation of an MPEG-4 Universal Scalable Video Codec Based on Configurable Layered Path-Tree Structure,” submitted to *IEEE Trans. on Circuits and Systems for Video Technology*.
3. “Information Technology – Coding of Audio-Visual Objects,” *ISO/IEC 14496-2:2001*, Second Edition, International Organization for Standardization, 2001.
4. “Streaming Video Profile,” *ISO/IEC 14496-2:2001/Amd 2:2002*, International Organization for Standardization, 2002.
5. Marcia G. Ramos, Sheila S. Hemami, “Perceptually based scalable image coding for packet networks,” *Journal of Electronic Imaging*, Vol. 7, No. 3, pp. 453-463, July 1998.

6. J. Lubin, “A HUMAN VISION SYSTEM MODEL FOR OBJECTIVE PICTURE QUALITY MEASUREMENTS,” *Proc. of IEEE International Broadcasting Convention*, pp. 12-16, September 1997.
7. Andrew B. Watson, James Hu and John F. McGowan III, “Digital Video Quality Metric Based on Human Vision,” *Journal of Electronic Imaging*, Vol. 10, No. 1, pp. 20-29, January 2001.
8. Stefan WINKLER, “Vision Models and Quality Metrics for Image Processing Applications,” *EPFL Thesis No. 2313*, 2000.
9. Ann Marie Rohaly, John Libert, Philip Corriveau, Arthur Webster, “Final Report from the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment,” available through http://www.vqeg.org, March 2000.
10. “Information Technology – Multimedia Content Description Interface -Visual,” *ISO/IEC 15938-3:2002*, International Organization for Standardization, 2002.
11. Kan-Li Huang, Yi-Shin Tung, Ja-Ling Wu, Po-Kang Hsiao, and Hsien-Shuo Chen, “Characteristic Extraction from MPEG-Compressed Bitstream and Its Applications,” *IEEE Transaction on Consumer Electronics*, pp. 522-532, Aug 2002.
12. Yi-Shin Tung, Ja-Ling Wu, Hsi-Kang Tsao, and Chi-Hui Huang, “A Highly Rate-Scalable System Based on MPEG-4 Spatial-Temporal-FGS Video Coding,” *Proc. of Advances in Automation, Multimedia and Video Systems, and Modern Computer Science*, ISBN 960-8052-44-0, WSES Press, pp. 320-325, 2001.
13. “JNDmetrix-IQTM Reference Guide, Version 7.6”, Sarnoff Corporation, May 2001.