CASE 82

Optimizing Video Compression

Abstract: Recent advances in Internet telephony and high-definition television (HDTV) have made real-time video compression a reality. The H.263 video compression standard investigated by this project takes advantage of image redundancies in both space and time. However, this standard was designed for a benign commercial environment. Using it effectively in a military environment requires additional parameter optimization. Through a series of L_{36} parameter design experiments, a strategy was developed which identified the H.263 video codec parameters that had the greatest impact on both video quality and transmitted bit rate. Using Taguchi experiment techniques, an optimal arrangement of control parameters were developed and tested.

1. Introduction

ITT Industries Aerospace/Communications Division, (IIN/ACD) located in Clifton, New Jersey, is a leader in the design and manufacturing of secure communications systems for military and government applications. The video codec investigated in this study is currently in use for the HMT and SUO programs.

Source Coding Algorithm

The H.263 video code is the ITU standard for very low bit rate video compression that makes it a prime candidate for use in a bandlimited military environment. Currently, it is used primarily for videoconferencing over analog telephone lines. Tests performed by the MPEG-4 committee in February 1996 determined the H.263 video quality outperformed all other algorithms optimized for very low bit rates.

The basic form of the H.263 encoder is shown in Figure 1. The video input to the codec consists of series of noninterlaced video frames occurring approximately 30 times per second. Images are encoded using a single luminance (intensity) and two chrominance (color difference) components. This format is commonly known as the Y, U, V format. Each pixel is sampled using an 8-bit linear quantizer, as defined in CCIR Recommendation 601. According to this standard, four pixels of the Y component are sampled and transmitted for every pixel of U and V (e.g., 4:1:1 format). This format was chosen by the ITU because human vision is more sensitive to luminance information than color information. Five standard image formats are supported by the H.263 codec: sub-QCIF, QCIF, CIF, 4CIF and 16CIF. For this investigation, the codec was operated using the QCIF format (e.g., 176 $\times$ 144 pixels/picture).

Images in the compressed video sequence are classified as either I-frames or P-frames. Intraframes (I-frames) do not rely on any previous spatial information and are encoded in a manner similar to still image compression. This processing results in a moderate amount of compression. Predicted frames (P-frames) are encoded based on spatial differences from the previous frame and allow a greater amount of compression to take place.

Both I and P frames are divided into 16 $\times$ 16 pixel luminance and 8 $\times$ 8 pixel chrominance macroblocks. H.263 uses motion estimation techniques to predict the pixel values of the target macroblock from the values of a reference macroblock in the

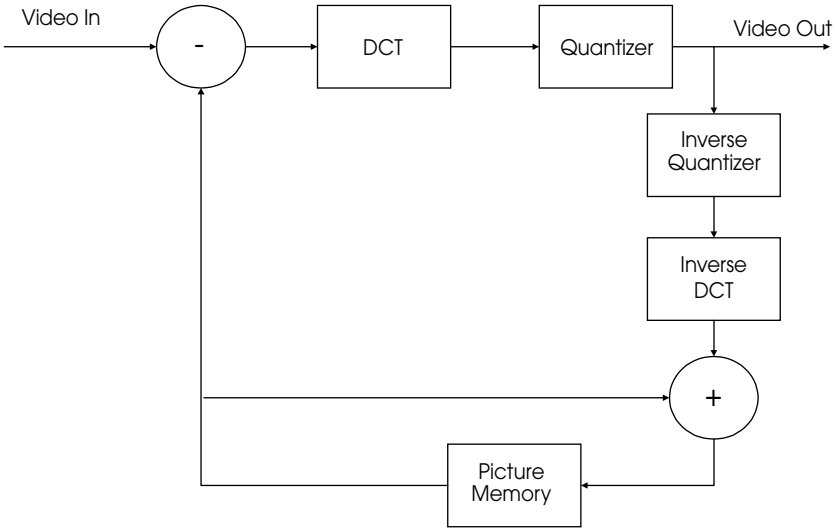

Figure 1
H.263 video codec system

previous frame. The best-fit prediction is found, usually as the result of an optimized search algorithm, and its location is transmitted as a motion vector that indicates its change from the target macroblock location. The macroblock is therefore encoded using two pieces of information, its motion vector and a prediction error.

Each macroblock undergoes a frequency-domain transform by using a discrete cosine transform (DCT). The DCT is a two-dimensional transform that maps values from the spatial domain into the frequency domain. For interframe macroblocks, the prediction error is calculated and transformed, while for intraframe macroblocks the pixel values themselves are transformed. The DCT coefficients are then quantized, mapping them into discrete steps. The step size, or *Q*-factor, remains constant throughout the video transmission and can range from 1 to 31. Because quantization introduces loss, higher *Q* values introduce more loss than lower *Q* values. Finally, the quantized coefficients are entropy encoded and transmitted to the decoder along with the motion vectors and other relevant formatting information.

The main purpose of the picture memory is to provide a reference image for the construction of P-frames. It accomplishes this by decoding the residuals of the transmitted video image through an inverse quantizer and DCT. It then adds these residuals with the previously stored image in the picture memory to construct the updated reference image necessary for P-frame difference processing.

In addition to I-frames and P-frames, four negotiable coding options are available for improved video performance (can be used together or separately):

1. *Unrestricted motion vectors.* For the default operation, motion vectors are restricted so that all pixels fall within the coded frame area. In unrestricted motion vector operation this restriction is removed and motion vectors are allowed to point outside the frame area.
2. *Syntax-based arithmetic coding.* For default operation, the H.263 codec using a variable-length encoding/decoding scheme for packing the compressed video into the bit stream. In this mode all variable-length encoding/decoding operations are replaced with arithmetic encoding/decoding techniques.
3. *Advanced prediction.* For the default H.263 codec, one motion vector is allocated per macroblock on a 16 × 16 pixel block size. As an

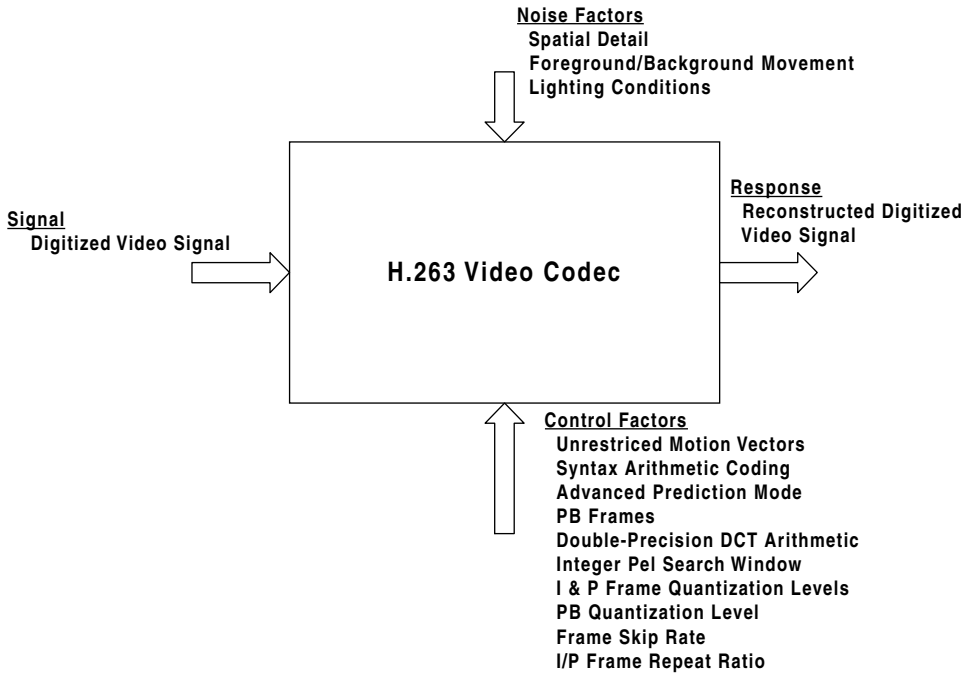

Figure 2
Parameter diagram

option, under advanced prediction mode, 8 × 8 vectors may be used.

4. *PB-frames.* Default operation of the H.263 video code only allows I and P frames. Under this option a PB frame is allowed, which consists of two pictures being coded as one unit. This PB-frame is predicted from a previously decode P-picture and a B-picture that is bi-directionally decoded.

2. Objectives

The objective of this experiment is to determine the sensitivity of the H.263 video codec parameters with respect to overall video quality and to identify the parameters that will best increase video quality without sacrificing the compressed video bit rate.

Parameter Diagram

To facilitate the robust design process, a parameter diagram of the system was generated. The parameter diagram is shown in Figure 2. The system noise and control parameters are as follows:

Noise Factors

Spatial detail: The amount of spatial detail in the image

Foreground motion: the amount of motion present in the primary or foreground object in the video scene

Background motion: the amount of motion present in the secondary or background objects in the video scene

Lighting conditions: the quality and direction of lighting in the video scene

Control Factors

Unrestricted motion vectors. Normal operation of the H.263 codec does not allow motion vectors to point "outside" the image areas. This option relieves this restriction and allows unrestricted motion outside the image area.

Syntax arithmetic coding. Normal operation of the H.263 codec uses variable-length coding/decoding for image transmission. Under this option, an arithmetic coding technique replaces all of the variable-length encoding/decoding operations.

Advanced prediction mode. For the H.263 encoder, one motion vector is allocated per macroblock on a 16 × 16 pixel block size. Under this option, 8 × 8 vectors are allowed.

PB frame. This allows the H.263 codec to encode pictures as PB-frames. A PB-frame consists of two pictures being codes as one unit. The PB-frame consists of a P-picture, which is predicted from a previously decoded P-picture, and a B-picture, which is bidirectionally decoded from past and future P-pictures.

Double-precision DCT. This option forces the decoder to use double-precision floating-point arithmetic in calculating the inverse DCT.

Integer pel search. This parameter controls the limit of the macroblock search window when calculating the motion vectors. The window size can range from 2 through 15.

I and P frame quantization. This parameter controls the quantization level for I-frames and P-frames. These parameters are set independently and do not vary during the course of a transmission. The quantization levels range from 2 through 30, with lower values indicating a greater number of quantization steps.

Frame skip rate. This parameter controls the amount of frames dropped prior to entering the encoder. These dropped frames are not encoded or used in any of the subsequent video analysis. This parameter can range from 1 thorugh 3.

PB quantization. This parameter controls the quantization level for the PB frames. It can range from 2 through 5.

I/P ratio. This parameter controls the number of P-frames transmitted between successive I-frames. It can range from 5 through 40.

Signal Factor

Digitized video image. The signal factor is the pixel valus for the luminance and two chrominance components of the video signal. The signal is measured on a frame-by-frame basis, and a composite SNR is calculated from these measurements.

Response Variable

Reconstructed video image. The response variable is thed pixel values for the luminance and two chrominance components of the reconstructed video signal.

Ideal Function

The effects of the control parameters on video performance were studied using an ideal function that measured the intensities of the reference video image pixels versus the uncompressed, reconstructed H.263 video image pixels for each of the video components (e.g., luminance and the two chrominance components; Figure 3). This ideal function was developed under this project and measures the overall video quality in an objective manner. Previously developed image quality metrics were based largely on the subjective techniques of a juried evaluation system. The ideal function has many advantages over a perceptive evaluation system. First, the measurement is repeatable. Given an image, the ideal function will always return the same measurement for image quality. Under a juried system, the image quality metric can vary considerably, depending on the makeup and composition of the jury. Second, the ideal function takes less time to score than the juried system. Typically, a video sequence of over 1000 frames took only 10 minutes of postprocessing analysis to calculate the quality metric. Third, the ideal function examines and scores each image in the sequence independent of the preceding image. This is an important distinction from a juried system, in which the human jurors' vision can "integrate" variations from frame to frame and miss some of the losses in image quality.

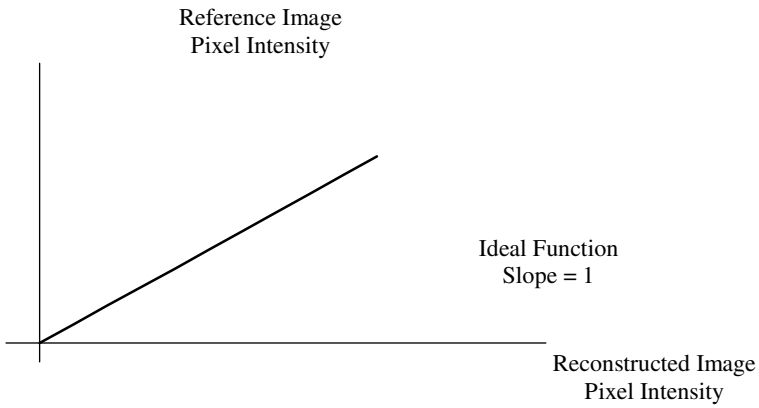

Figure 3
Image quality ideal function

For the ideal function, a perfect match between the input reference and the reconstructed video signals would result in a 45-degree straight line originating at the intersection of the *x*–*y* axis. However, since H.263 is a lossy compression scheme, the reconstructed video will deviate somewhat from the reference video, resulting in a degradation of video quality. This deviation increases with increasing video compression.

The SNR is calculated using the incoming video sequence as a reference. Since the incoming video is sampled at 30 frames/s and the video codec drops some of these video frames, the reconstructed video will be replayed at a much lower frame rate. Typically, the image decoder will decompress and display only 10 frames/second, resulting in a loss of video images. To account for the missing video images, the decoder was modified to update the display at a rate of 30 frames/s. "Missing" uncompressed images were displayed using the previous frames decompressed image. This resulted in an effective video display rate equal to the incoming video frame rate. The SN ratios in the experiments were calculated using the effective video signal against the reference video signal at a rate of 30 frames/s.

Table 1
Noise factors and levels

	Level	
Noise Factor	**1**	**2**
Spatial detail	Low	High
Lighting variation	Low	Medium
Movement	Low	High

3. Experimental Layout

Through a series of brainstorming sessions, noise factors and control factors were identified and levels assigned to these parameters, as described below.

Noise Factor Levels

Levels for the noise factors are shown in Table 1. Since there were a limited number of noise factors, a full factorial noise source was created, which used two video sequences. The first video sequence, called *claire,* is a "talking head" type of video with a low amount of both foreground and background movement and spatial details. The lighting condition for this sequence is low and does not vary. The second sequence, called *car phone,* contains a high amount of spatial detail and movement.

Control Factor Levels

Table 2 indicates the control factors, their mnemonics, and the levels used in this experiment.

Table 2
Control factors and levels

	Control Factor	Mnemonic	Level 1	Level 2	Level 3
A:	unrestricted motion vector	UVecFlag	0(off)	1(on)	
B:	syntax arithmetic coding	SCodFlag	0(off)	1(on)	
C:	advanced prediction	AFlag	0(off)	1(on)	
D:	PB frames	PBFlag	0(off)	1(on)	
E:	double precision DCT	DCTFlag	0(off)	1(on)	
F:	pel search window size	IPelWIn	2	7	15
G:	I-frame quantization	IQP	2	10	30
H:	P-frame quantization	PQP	2	10	30
I:	frame skip rate	FSkip	2	3	4
J:	PB quantization	DBQuant	0	1	3
K:	I/P ratio	IPRatio	4/1	19/1	39/1

Experimental Design

Based on the identified control factors, an L_{36} orthogonal array was chosen for the test configuration for the Taguchi experiments. Since there are five two-level factors and six three-level factors, a $2^{11} \times 3^{12}$, an L_{36} array was chosen to hold the parameters and was modified as follows. The first five columns in the standard L_{36} array were used to house the two-level control factors, and columns 12 through 17 were used to house the three-level control factors. Columns 6 through 11 and 18 through 23 were not used in the experimental design. All subsequent calculations were modified as appropriate to take into account the missing columns in the experiment design.

Each set of L_{36} Taguchi experiments was repeated six times, once for each of three video components (*Y*, *U*, *V*) and once for each of the two video scenes (car phone and claire), for a grand total of 216 experiments. In each experiment, the pixel intensity levels and the compression bit rates were measured and recorded. For each set of experiments, a data reduction and slope and SN ratio analysis was performed. Finally, a composite SN ratio analysis was performed for overall video quality and for the bit rate.

The final orthogonal array with the control factor levels is shown in the Table 3.

4. Experimental Results and Analyses

Derivation of Best-Fit Line (Sensitivity)

A data reduction analysis effort was performed on each set of collected data to obtain insight into the relationship between the reference video signal and the reconstructed video signal. The data from each of the Taguchi experiments was fitted to a linear relationship, which passed through the origin, using standard regression analysis. This regression produced a sensitivity (slope) analysis for the design.

Derivation of Composite SN Ratios

After deriving the best-fit sensitivity for each of the experiments, a composite SN ratio was calculated. In forming the composite SN ratio, we have taken into account two observations:

1. The sensitivities of the *Y*, *U*, and *V* components can be adjusted independently.

Table 3
Control factors L_{36} orthogonal array

No.	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*	*K*
1	0	0	0	0	0	1	2	2	2	0	5
2	0	0	0	0	0	7	10	10	3	1	20
3	0	0	0	0	0	15	30	30	4	3	40
4	0	0	0	0	0	1	2	2	2	1	20
5	0	0	0	0	0	7	10	10	3	3	40
6	0	0	0	0	0	15	30	30	4	0	5
7	0	0	1	1	1	1	2	10	4	0	20
8	0	0	1	1	1	7	10	30	2	1	40
9	0	0	1	1	1	15	30	2	3	3	5
10	0	1	0	1	1	1	2	30	3	0	40
11	0	1	0	1	1	7	10	2	4	1	5
12	0	1	0	1	1	15	30	10	2	3	20
13	0	1	1	0	1	1	10	30	2	3	20
14	0	1	1	0	1	7	30	2	3	0	40
15	0	1	1	0	1	15	2	10	4	1	5
16	0	1	1	1	0	1	10	30	3	0	5
17	0	1	1	1	0	7	30	2	4	1	20
18	0	1	1	1	0	15	2	10	2	3	40
19	1	0	1	1	0	1	10	2	4	3	40
20	1	0	1	1	0	7	30	10	2	0	5
21	1	0	1	1	0	15	2	30	3	1	20
22	1	0	1	0	1	1	10	10	4	3	5
23	1	0	1	0	1	7	30	30	2	0	20
24	1	0	1	0	1	15	2	2	3	1	40
25	1	0	0	1	1	1	30	10	2	1	40
26	1	0	0	1	1	7	2	30	3	3	5
27	1	0	0	1	1	15	10	2	4	0	20
28	1	1	1	0	0	1	30	10	3	1	5
29	1	1	1	0	0	7	2	30	4	3	20
30	1	1	1	0	0	15	10	2	2	0	40
31	1	1	0	1	0	1	30	30	4	1	40
32	1	1	0	1	0	7	2	2	2	3	5
33	1	1	0	1	0	15	10	10	3	0	20
34	1	1	0	0	1	1	30	2	3	3	20
35	1	1	0	0	1	7	2	10	4	0	40
36	1	1	0	0	1	15	10	30	2	1	5

Table 4
Y, *U*, *V*, composite SN ratio, and BPS results

	SNR (dB)				
No.	***Y***	***U***	***V***	**Composite**	**BPS**
1	−17.4	−6.4	−3.8	−15.7	55.0
2	−18.9	−11.1	−9.5	−17.4	44.7
3	−21.0	−14.8	−12.4	−19.7	38.3
4	−17.3	−6.4	−3.9	−15.7	54.1
5	−17.6	−10.4	−9.4	−16.2	43.4
6	−21.6	−14.9	−12.2	−20.2	41.9
7	−19.9	−10.4	−8.4	−18.3	46.1
8	−19.5	−12.5	−10.5	−18.0	41.1
9	−19.2	−10.2	−7.5	−17.6	52.8
10	−19.5	−10.7	−10.0	−17.9	41.9
11	−19.6	−8.7	−6.4	−18.0	51.9
12	−18.6	−11.8	−9.7	−17.2	45.2
13	−19.8	−12.1	−10.2	−18.3	42.2
14	−16.5	−5.3	−4.6	−14.8	51.4
15	−19.7	−8.8	−6.5	−18.1	48.9
16	−19.8	−11.5	−9.6	−18.3	44.8
17	−19.6	−8.6	−6.0	−18.0	51.4
18	−17.8	−9.7	−8.1	−16.3	45.6
19	−18.5	−6.6	−5.2	−16.8	51.0
20	−19.1	−12.5	−10.3	−17.7	47.1
21	−19.9	−11.0	−9.0	−18.3	44.7
22	−20.0	−11.3	−9.3	−18.4	46.0
23	−20.6	−14.8	−12.1	−19.2	41.3
24	−16.3	−4.2	−4.0	−14.7	51.6
25	−18.2	−11.4	−9.7	−16.8	44.9
26	−19.5	−8.8	−6.2	−17.9	50.0
27	−19.5	−8.0	−5.5	−17.8	51.7
28	−19.8	−12.6	−10.3	−18.4	45.9
29	−20.7	−11.3	−9.0	−19.1	43.5
30	−16.9	−6.2	−4.2	−15.2	53.0
31	−21.1	−14.8	−12.3	−19.7	38.1
32	−17.4	−6.4	−3.8	−15.7	54.3
33	−18.9	−11.1	−9.5	−17.4	44.3

Table 4
(*Continued*)

	SNR (dB)				
No.	***Y***	***U***	***V***	**Composite**	**BPS**
34	−18.3	−7.7	−5.4	−16.6	52.4
35	−19.1	−9.9	−8.6	−17.5	43.9
36	−19.0	−11.2	−9.2	−17.5	45.9
Avg.	−19.1	−10.1	−8.1	−17.5	47.0
Min.	−21.6	−14.9	−12.4	−14.7	55.0
Max.	−16.3	−4.2	−3.8	−20.2	38.1
Range	5.3	10.6	8.5	5.5	16.9

2. The sensitivities are static and cannot be changed, depending on the type of image being compressed because the image type is a noise factor.

For each component (*Y*, *U*, and *V*), a straight line with zero intercept was fitted to the scatter plot of the pixel values of the reconstructed image versus the original image. This slope is referred to as β_i, where $i = Y, U$, or V. Similarly, the variance around this fitted line is σ_i, where $i = Y, U$, or V.

The composite variance was defined as the weighted average of the *Y*, *U*, and *V* component variances, which were normalized for the individual slopes (e.g., make each slope equal to 1) and the number of data points for each component. The composite variance is given by

$$\sigma^2_{\text{comp}} = \frac{4\sigma^2_Y/\beta^2_Y + \sigma^2_U/\beta^2_U + \sigma^2_V/\beta^2_V}{6}$$

and the composite SN ratio is given by

$$\eta = 10 \log \frac{1}{\sigma^2_{\text{comp}}}$$

where η is the SN ratio and σ_{comp} is the composite variance.

Composite Bit-Rate SN Ratio

The second component measured was the bit rate of the compressed video bit stream (BPS). For each experiment the bit rate of the compressed signal was measured and recorded. A composite bit-rate SN ratio was calculated using the relationship

$$\text{BPS} = 10 \log \Gamma$$

where BPS is the composite bit rate SN ratio and Γ is the compressed video stream bit-rate (bits/s).

Experimental Results

The results of the Taguchi experiments are shown in Table 4, which lists the results of the SN calculations for the *Y*, *U*, *V* and composite video quality metric and the compressed video stream bit rate (BPS).

Factor Effects Plots

Based on the results of the Taguchi experiments and data reduction efforts, factor effects plots were generated for the composite image quality and BPS SNR. These results are shown in Figures 4 and 5.

5. Factor Effects Summary

Since the main objective of this project was to optimize video quality, an examination of the composite video quality and BPS factor effects plot is appropriate. An examination of Figures 4 and 2 indicates that only four of the 11 control factors significantly affect image quality and compression: I-frame quantization level (IQP), P-frame quantization level (PQP), frame skip rate (FSkip), and the I/P frame ratio (I/P ratio). The results of an ANOVA analysis (Table 5) reveals that these four factors account for over 92% of the total sum of

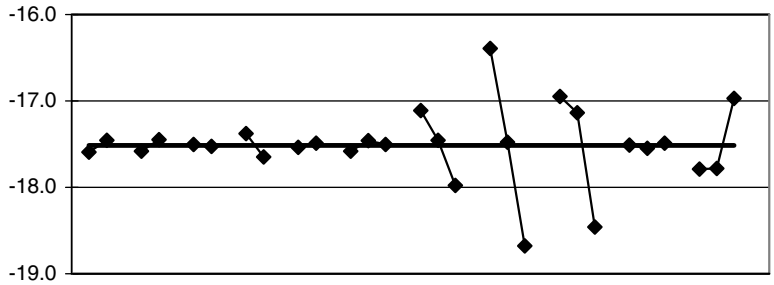

Figure 4
Composite image quality factor effects

squares for video quality and 94% of the total sum of squares for the BPS data.

The effect of the P-frame quantization parameter, PQP, is the strongest among all the control factors. For I-frame and P-frame quantization, the factor effects plot indicate that the smaller index values signify more levels of quantization, thereby increasing image quality. However, this increase in image quality does not come without an addition cost, as indicated from the BPS factor effects plot. That is, an increase in the number of quantization levels increases both the image quality and the compressed video bit stream. That is, for these parameters, an increase in image quality leads to a significant increase in bandwidth. A cost–benefit analysis is needed to arrive at the best values for these parameters. The factor effects plot shows that PQP has a greater effect on image quality than IQP. This is because the P-frame quantizer operates on the residuals of pixel intensities, while the I-frame quantizer operates on absolute pixel intensities.

Frame skip (FSkip) has the second-largest effect on image quality, with level 1 being the most desirable. This level implies that only alternating frames should be compressed and transmitted. Level 2 provides only a small degradation in image quality, while level 3, which corresponds to transmitting every fourth frame, causes a significant drop in image quality. An examination of the BPS factor effects plot indicates that the effect of FSkip on BPS is small, signifying that very little needs to be spent, in terms of bandwidth, to achieve significant gains in image quality.

A further examination of the slope of the FSkip parameter suggests an alternative scheme in which no frames are skipped, and the video codec compresses every image. If the factor effects plots were extrapolated from their present values, the image quality factor effects plot suggests a significant increase in image quality, while the BPS factor effects plot suggests a small increase in bandwidth because of it low slope.

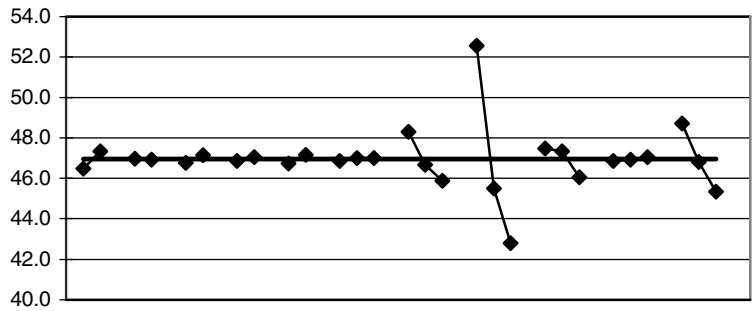

Figure 5
BPS factor effects

The data also suggest that larger values of I/P ratio are more desirable. It should be noted that this is the only win–win factor (i.e., larger values of the I/P ratio reduce BPS and at the same time increase image quality). Larger values of the I/P ratio should be investigated to confirm that this trend is accurate. However, I believe that there comes a point of diminishing returns, especially in noisy environments. Transmission errors may have a major interaction with the I/P ratio. Therefore, this factor should be studied when a simulation with transmission errors becomes available.

6. Video Parameter Design and Confirmation

To study the effects of the four major contributors to the SN ratio performance (e.g., IQP, PQP, FSkip, and I/P ratio), an L_9 confirmation experiment was run. In this experiment, the following factors were set for all nine experiments: A_1, B_1 C_1, D_1, E_1, F_2, and J_2. Also, the confirmation experiment used the same video sequences as the L_{36} experiment. In addition, the control factors were set as shown in Table 6.

Table 7 shows the results for each of the nine experiments plus the baseline experiment. This table shows both experimental results and the forecast for all nine experiments, plus the baseline configuration. Also included is an "optimal" image quality experiment, which consisted of factors A_1, B_1, C_1, D_1, E_1, F_2, G_1, H_1, I_1, J_2, and K_3.

7. Discussion

The factor effects plots for BPS are shown in Figure 6. This plot shows a similar factor effect sensitivity behavior, as was exhibited by the L_{36} experiment. The predictions for image quality and BPS SN ratio were confirmed by the actual experimental results. The experimental results confirm that the users can expect a 2.5-dB improvement in image quality when using the optimum factor effects settings. The image quality and BPS SNR forecasts correlate well with the experimental results in all cases.

Finally, the difference in image quality can readily be seen in Figure 7. This figure shows one image

Table 5
ANOVA results

Factor	Composite SN Ratio as % Sum of Squares	BPS SN Ratio as % Sum of Squares
UVecFlag	0.0	0.3
SCodFlag	0.3	0.0
AFlag	0.0	0.2
PBFlag	1.1	0.1
DCTFlag	0.0	0.2
IPelWin	0.1	0.0
IQP	7.3	4.7
PQP	50.3	78.8
FSkip	26.0	1.9
DBQuant	0.0	0.0
IPRatio	8.5	8.9

Table 6
Confirmation experiment orthogonal array

No.	*A* (IQP)	*B* (PQP)	*C* (FSkip)	*D* (IPRatio)
1	2	2	2	5
2	2	10	3	20
3	2	30	4	40
4	10	2	3	40
5	10	10	4	5
6	10	30	2	20
7	30	2	4	20
8	30	10	2	40
9	30	30	3	5
B	10	10	3	20

Table 7
Confirmation experiment results summary

	Composite SNR (dB)		BPS SNR (dB)	
No.	**Measured**	**Forecast**	**Measured**	**Forecast**
1	−15.7	−15.7	55.9	55.3
2	−17.2	−17.0	46.6	46.2
3	−18.9	−18.7	41.0	40.8
4	−14.7	−15.4	50.1	50.2
5	−18.5	−18.6	47.0	45.2
6	−18.3	−18.3	43.4	42.0
7	−17.9	−18.1	52.9	49.6
8	−16.8	−16.8	45.1	42.5
9	−19.8	−19.0	43.7	43.0
Optimum	−14.5	−14.9	53.0	52.8
Baseline	−17.6	−17.3	44.8	44.6

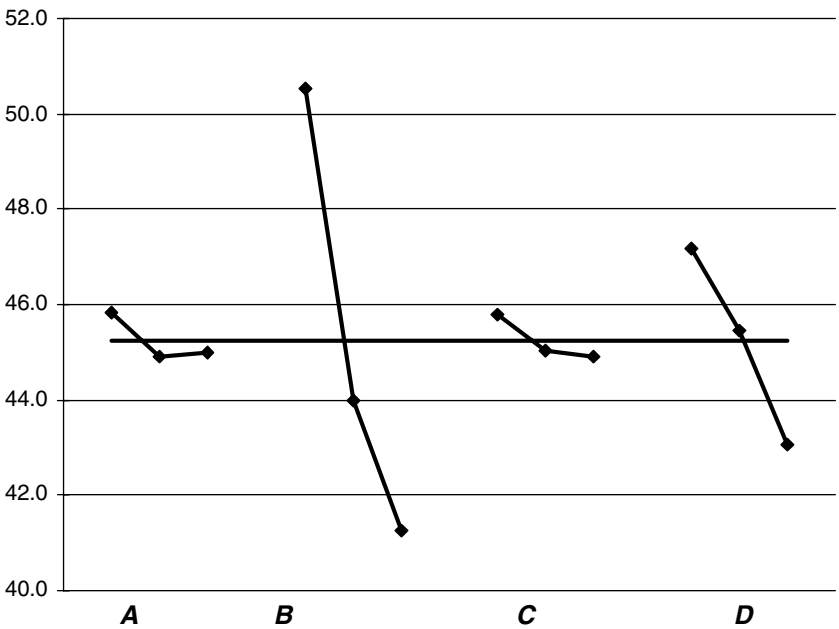

Figure 6
BPS factors effects

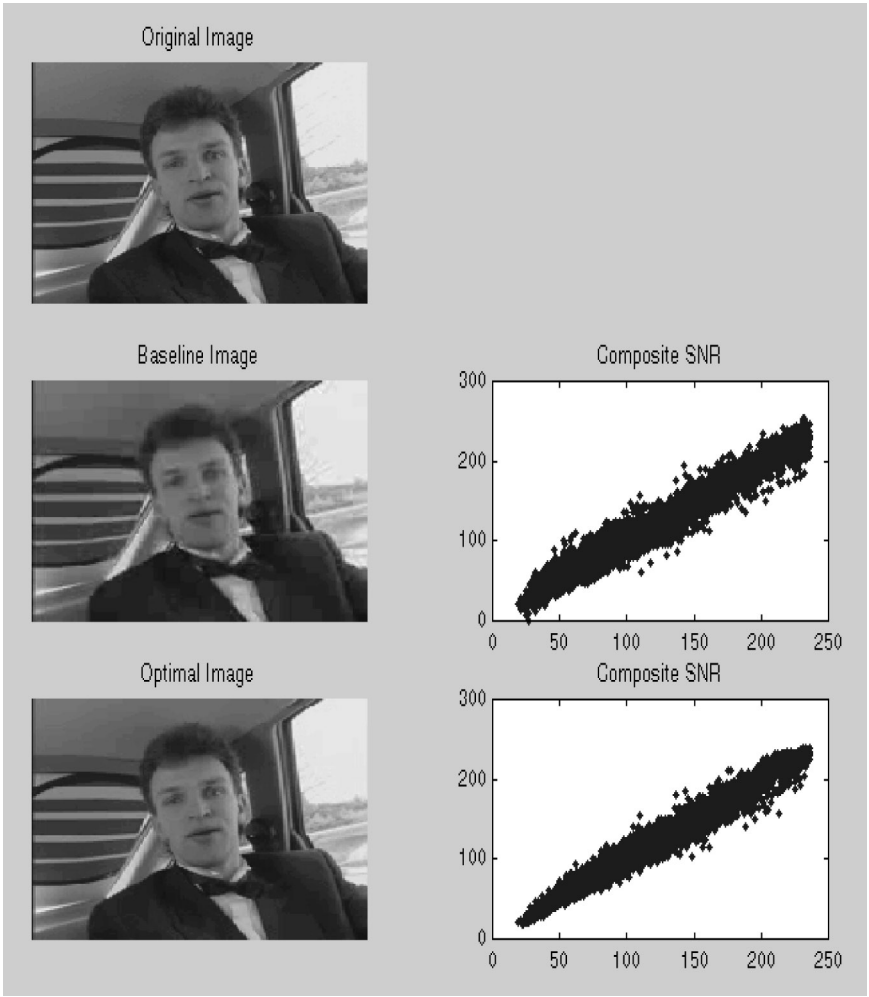

Figure 7
Verification experiment images

from the original video sequence and the reconstructed images from two of the validation experiments (e.g., the baseline and the optimal experiments). This figure shows that the baseline image suffers a considerable image quality loss. This can readily be seen in the facial details in the baseline image versus the original image. The baseline image suffers quality loss primarily in loss of facial details. In the optimal image sequence, the facial details are much improved and are very similar to the original image. Loss of spatial details can result in a significant quality loss, especially in combat identification systems. The figure also shows the ideal function for both the baseline and optimal images. Once again, it is apparent that there is a signal quality loss in the baseline image versus the optimal image. This shows up as a larger signal variance on the composite SNR graphs for these images.

8. Conclusions

In conclusion, the objective video scoring methods developed under this project, in conjunction with robust design techniques and Taguchi methods, have shown to be an unbiased, cost-effective approach to parameter design for the H.263 video codec. Using these techniques, an improvement of 2.3 dB was observed over the baseline approach in terms of the image quality metric derived. The analysis in this paper demonstrates the importance of applying robust design techniques early in the design process.

Additionally, the image quality SN ratio metric developed in this project was independent of any compression technologies and can be applied to a wide range of still and motion video codecs. Through the use of an objective quality metric, the time and costs involved with evaluating and quantifying video codecs have been reduced significantly.

Acknowledgments The author expresses his sincere thanks to Sessan Pejan of Sarnoff Research Labs, Princeton, New Jersey, for his help and expertise in video processing.

Reference

ITU-T, 1996. *Video Coding for Low Bit Rate Communications,* ITU-T Recommendation H.263.

This case study is contributed by Edward Wojciechowski and Madhav S. Phadke.