Proposed First-Generation WSQ Bit Allocation Procedure

Jonathan N. Bradley and Christopher M. Brislawn
Group C–3, Computer Research and Applications
Los Alamos National Laboratory

September 8, 1993

I. Introduction.

The Wavelet/Scalar Quantization (WSQ) gray-scale fingerprint image compression algorithm [1] involves a symmetric wavelet transform (SWT) image decomposition followed by uniform scalar quantization of each subband. An overview of the standard can be found in [2]. The algorithm is adaptive insofar as the bin widths for the scalar quantizers are image-specific and are included in the compressed image format. Since the decoder requires only the actual bin width values—but not the method by which they were computed—the standard allows for future refinements of the WSQ algorithm by improving the method (referred to as a *bit allocation procedure*) used to select the scalar quantizer bin widths. This report proposes a bit allocation procedure for use with the first-generation WSQ encoder specified in Part 3 of [1].

In [1] a specific formula is provided for the *relative* sizes of the scalar quantizer bin widths in terms of the variances of the SWT subbands. An explicit specification for the constant of proportionality, q, that determines the absolute bin widths (and therefore the overall compression ratio) was not given in the February 1993 draft (version 2) of [1]. The actual compression ratio produced by the WSQ algorithm will generally vary from image to image depending on the amount of coding gain obtained by the run-length and Huffman coding stages of the algorithm, but testing performed by the FBI established that WSQ compression produces archival quality images at compression ratios of around 20 to 1. The bit allocation procedure described in this report possesses a control parameter, r (the *lossy bit rate*), that can be set by the user to achieve a predetermined amount of lossy compression, effectively giving the user control over the amount of distortion introduced by quantization noise. The variability observed in final compression ratios is thus due only to differences in lossless coding gain from image to image, chiefly a result of the varying amounts of blank background surrounding the print area in the images. Experimental results are presented that demonstrate the proposed method's effectiveness.

II. Scalar Quantization of SWT Subbands.

The WSQ standard specifies an image decomposition into 64 subbands with frequency supports as shown in Figure 1. Each subband is coded separately according to a scalar quantizer characteristic having uniform width bins with the exception of the zero bin, which is 20% wider (in the first-generation encoder), as specified in [1]. Each scalar quantizer is defined by an encoding and a decoding relationship. The quantization encoder maps a floating-point wavelet coefficient, a, to an integer quantizer index, p, that indicates the quantizer bin in which a lies. The quantization decoder maps the index to a prototypical ("quantized") real number, $\hat{a}$, representing all data values that lie within that bin. The standard specifies that the WSQ encoder transmit the values of Q_k (the bin width) and Z_k (the zero bin width) for each subband along with the Huffman coded quantizer indices. Quantization encoding of the k^{th} two-dimensional SWT subband, $a_k(m,n)$, is given by

$$p_k(m,n) = \begin{cases} \left\lfloor \frac{(a_k(m,n) - Z_k/2)}{Q_k} \right\rfloor + 1 & , \quad a_k(m,n) > Z_k/2 \\ 0 & , \quad -Z_k/2 \le a_k(m,n) \le Z_k/2 \\ \left\lceil \frac{(a_k(m,n) + Z_k/2)}{Q_k} \right\rceil - 1 & , \quad a_k(m,n) < -Z_k/2 \quad . \end{cases}$$

The notation $\lceil \cdot \rceil$ and $\lfloor \cdot \rfloor$ denotes the functions that round numbers to the next largest and next lowest integer, respectively. The quantized wavelet coefficients produced by the quantization decoder are given by

$$\hat{a}_k(m,n) = \begin{cases} (p_k(m,n) - C)Q_k + Z_k/2 & , \quad p_k(m,n) > 0 \\ 0 & , \quad p_k(m,n) = 0 \\ (p_k(m,n) + C)Q_k - Z_k/2 & , \quad p_k(m,n) < 0 \quad , \end{cases}$$

where C is a parameter between 0 and 1 that determines the reconstructed values. Note that if $C = 1/2$ then the reconstructed value corresponding to each quantization bin would be the bin's midpoint. The value specified in [1] for the first-generation quantizer is $C = 0.44$. The quantizer indices, $p_k(m,n)$, are transmitted losslessly by a combination of zero run-length and Huffman coding.

The formula for the bin widths, Q_k, used in the first-generation WSQ encoder is:

$$Q_k = \begin{cases} 1/q \, , & 0 \le k \le 3 \, , \\ \frac{10}{qA_k \log_e(\sigma_k^2)} \, , & 4 \le k \le 59 \, . \end{cases} \tag{1}$$

The constants A_k are empirically determined weights; $A_k = 1$ for $4 \le k \le 51$, and the remaining values are given in Table I. q is a constant that determines the overall compression

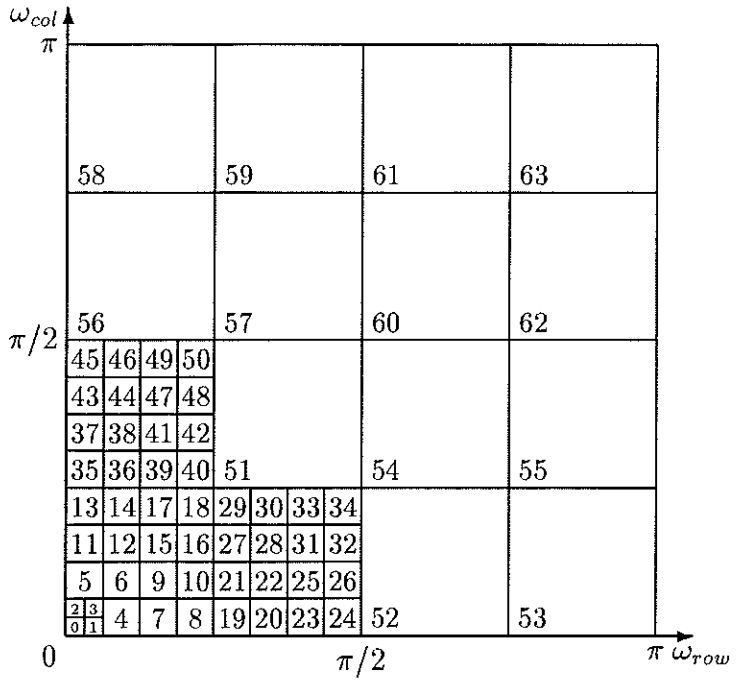

Figure 1: Frequency Support of Wavelet Transform Subbands in WSQ Standard.

Table I: Table of A_k.

Subband	A_k
52 & 56	1.32
53 & 58	1.08
54 & 57	1.42
55 & 59	1.08

ratio. Subbands 60–63 are not computed or transmitted in the first-generation version of the WSQ encoder. The variance computation is performed over a subregion in the center of each subband to adjust the coder to the wavelet coefficient statistics in the print detail area of the image.

III. Bit Allocation.

We now derive a formula for the parameter q in equation (1) that produces quantized subbands satisfying a user-imposed constraint on the overall lossy compression ratio. Although we do not do so here, it can be shown that this approach minimizes a weighted mean-square error distortion measure subject to a constraint on the overall compression ratio and hence can be regarded as an *optimal* bit allocation.

Since raw fingerprint image data has a bandwidth of 8 bits per pixel (bpp), a bit rate of r bpp for the compressed data corresponds to a compression ratio of $8/r$. For the k^{th} SWT subband, let r_k denote the subband bit rate, μ_k the subband mean, and σ_k^2 the subband variance. m_k denotes the downsample factor, which is defined to be the ratio of image size to subband size. For the decomposition pictured in Figure 1, all downsample factors are powers of 4; e.g., $m_{63} = 16$ and $m_4 = 256$.

Recall that the standard [1] allows the WSQ encoder to discard some subbands and transmit a bin width of zero ($Q_k = 0$) to inform the decoder that no compressed image data is being transmitted for subband k; i.e., that $r_k = 0$. For instance, this is always done for $60 \leq k \leq 63$ in the first-generation encoder, and may be done for other subbands as well on an image by image basis if the bit allocation routine determines that a certain subband contains so little information that it should be discarded altogether. To keep track of the non-discarded subbands, let K denote the set of all transmitted subbands (e.g., for the first-generation encoder, $K \subset \{0, 1, \ldots, 59\}$). The fraction of non-discarded SWT coefficients will be denoted by S, where

$$S = \sum_{k \in K} \frac{1}{m_k} \quad . \tag{2}$$

The targeted overall lossy bit rate, r, can be expressed as

$$r = \sum_{k \in K} \frac{r_k}{m_k} \quad , \tag{3}$$

since $r_k = 0$ if $k \notin K$.

To relate bit rates to quantizer bin widths, we need to assume that the data being quantized lies in some interval of finite extent. Accordingly, the assumption is made that the quantization bins cover the interval $[\mu_k - \gamma\sigma_k\,,\ \mu_k + \gamma\sigma_k]$; i.e., that

$$Q_k = \frac{2\gamma\sigma_k}{L_k} \quad , \tag{4}$$

where L_k is the number of bins in the quantizer, and the *loading factor*, γ, is a parameter that specifies the number of standard deviations of data that are being coded [3]. For the numerical experiments presented in the next section, the value $\gamma = 2.5$ was used. Although the WSQ algorithm eliminates the possibility of overload distortion by using escape sequences to code quantizer indices as large as ± 65535, it is still necessary to specify a realistic loading factor to model accurately the lossy bit rate of the quantizers. Thus, while a poor choice of loading factor will not result in overload distortion, as is the case with other quantization strategies, it *will* affect the extent to which the lossy bit rate constraint, r, models the actual observed compression ratio. We also assume that the average transmission bit rate for subband k is

$$r_k = \log_2 L_k \quad \text{bits/sample.} \tag{5}$$

This rate (5) actually models the "worst-case" scenario in which indices occur with equal probabilities and are coded with equal numbers of bits; if indices are not equally likely then Huffman coding will produce additional gain, resulting in a lower observed bit rate.

Now use the above model to determine q by applying the bit rate constraint to the bin widths. Isolate q in equation (1) by writing

$$Q_k = \frac{1}{q} Q'_k \tag{6}$$

where Q'_k is independent of q:

$$Q'_k = \begin{cases} 1 , & 0 \le k \le 3 , \\ \frac{10}{A_k \log_e(\sigma_k^2)} , & 4 \le k \le 59 . \end{cases}$$

Substituting (5) and (4) into (3) gives

$$r = \sum_{k \in K} \frac{1}{m_k} \log_2 \frac{2\gamma\sigma_k q}{Q'_k} \quad ,$$

and solving for q in terms of the constraint parameter, r, gives

$$q = \beta 2^{r/S-1} \left[\prod_{k \in K} \left(\frac{\sigma_k}{Q'_k} \right)^{1/m_k} \right]^{-1/S} \quad , \tag{7}$$

where $\beta = 1/\gamma$ is the *loading fraction* [3].

Two cases require special attention. First, in the limiting case $\log_e(\sigma_k^2) \searrow 0$ we have $Q_k \to \infty$, which corresponds to $r_k \to 0$, so we discard any subband whose variance is so small that $\log_e(\sigma_k^2) \le 0$. Second, if $Q_k \ge 2\gamma\sigma_k$ then (4) and (5) imply that $r_k \le 0$. Since this fouls up the relationship (3) that imposes the bit rate constraint, we use an iterative procedure to determine q, excluding bands that have a theoretically negative bit rate to ensure that the targeted bit rate, r, is met.

There are a couple reasonable ways of quantizing these excluded bands: 1) They could be discarded automatically by setting $Q_k = 0$, or 2) We could use the (large) bin widths given by (6) in the expectation that quantization of these bands will result in effectively zero bit rates. Since both of these options are equally straightforward to implement and both appear to be consistent with the admittedly incomplete specification provided in [1], we present both methods in this Report and compare the results of numerical experiments in the next section. (While it would save a few additional bytes in option 2—with no added distortion—to transmit $Q_k = 0$ if all values in band k did, in fact, get quantized to zero using the bin width given by (6), we have not tried implementing that refinement.)

Iterative Procedure for Computing Bin Widths.

1. Initialize; $j = 0$:

   (a) If $\log_e(\sigma_k^2) \leq 0$ then set $Q_k = 0$.

   (b) $K^{(0)} = \{k \mid \log_e(\sigma_k^2) > 0\} \subset \{0, 1, \ldots, 59\}$.

2. Iterate on (7) to calculate q:

   (a) $S^{(j)} = \sum_{k \in K^{(j)}} \frac{1}{m_k}$.

   (b) $q^{(j)} = \beta 2^{r/S^{(j)}-1} \left[\prod_{k \in K^{(j)}} \left(\frac{\sigma_k}{Q'_k} \right)^{1/m_k} \right]^{-1/S^{(j)}}$.

3. Exclude bands that would theoretically have negative bit rates:

   (a) $\Xi^{(j)} = \left\{ k \in K^{(j)} \mid Q'_k/q^{(j)} \geq 2\gamma\sigma_k \right\}$.

   (b) If $\Xi^{(j)} \neq \emptyset$ then

      i. $K^{(j+1)} = K^{(j)} \backslash \Xi^{(j)}$,

      ii. $j = j + 1$,

      iii. go to 2;

   else

      i. $K = K^{(j)}$, the bands with positive bit rates,

      ii. $\Xi = K^{(0)} \backslash K$, the excluded bands,

      iii. $q = q^{(j)}$,

      iv. continue.

4. Calculate bin widths:

   Option 1. Set $Q_k = Q'_k/q$ for $k \in K$ and $Q_k = 0$ for $k \in \Xi$.

   Option 2. Set $Q_k = Q'_k/q$ for all $k \in K^{(0)}$.

5. Exit.

Note that $\backslash$ denotes the set difference operator; i.e., $A \backslash B = A \cap B^c$. Given a reasonable choice of loading factor and negligible lossless coding gain, the bit rate for the compressed image will be constrained to r bpp. To demonstrate the validity of this claim, we now present the results of numerical experiments.

IV. Numerical Results.

This section presents numerical results on WSQ compression of fingerprint images using the bit allocation algorithm of the previous section. The purpose here is to establish the robustness of this method for controlling lossy quantization on fingerprint images. The independent parameters in the experiments are the image size, the exception-handling option, and the bit rate parameter, r. All images had 128 subtracted from the raw pixel values so that the input signals were between -128 and 127, in accordance with [1]. For each scenario considered, a set of 24 images is compressed and six compression ratio statistics are recorded: the theoretical compression ratio $\mathrm{CR}_{theo} = 8/r$, the mean observed compression ratio, the standard deviation, minimum and maximum observed compression ratios, and the average distortion as measured in terms of peak-signal-to-noise ratio (PSNR):

$$\mathrm{PSNR} = 20 \log_{10} \left(\frac{255}{\mathrm{RMSE}} \right) \ \mathrm{dB} \quad ,$$

where RMSE is the root mean square error. PSNR is a standard statistic widely used in image processing for reporting noise levels [4].

Results on 768×768 *Fingerprint Images.*

The first experiment was performed to insure that reasonable compression ratios are obtained on full-sized (768×768) images. The results of compressing a set of 24 images for various values of r using exception-handling option 1 are shown in Table II; results obtained using option 2 are shown in Table III. The observed compression ratios increase in a predictable fashion as r is decreased. Note that CR_{theo} provides a consistent lower bound on the minimum observed compression ratios. The actual range of observed values indicates varying amounts of additional lossless coding gain. Since option 2 discards fewer subbands than option 1, the mean observed compression ratios in Table III are, as expected, slightly lower than the mean observed compression ratios in Table II for the same target bit rate, r, with slightly higher PSNR measurements.

Results on 512×512 *Fingerprint Images.*

Next, the method was run on a set of 24 512×512 fingerprint images. The experimental results are shown in Tables IV and V. Since these cropped images have much less blank area than 768×768 images, there is much less lossless coding gain in these examples, and the final compression ratios come in closer to the predicted values. For instance, compare the mean compression ratios in Table IV with those in Table II for the same r values; this shows how little lossless coding gain there is in the absence of significant blank regions in the image. Zero run-length coding gain in the blank areas of typical 768×768 images therefore appears to be much more significant than Huffman coding gain in the print detail areas. This effect also accounts for the much smaller variability in the compression ratios seen in Table IV. Without additional lossless coding gain, the final compression ratio obtained using this bit allocation method comes in at a predictable value given by the lossy bit rate, r.

Table II: Compression Ratios Achieved on 768 × 768 Images, Option 1.

r bpp	CR_{theo}	CR_{avg}	STD	CR_{min}	CR_{max}	$PSNR_{avg}$
0.1	80.00	120.41	12.48	90.46	140.67	26.12
0.15	53.33	83.09	9.44	61.68	101.07	27.51
0.2	40.00	63.46	7.57	47.51	79.01	28.54
0.25	32.00	51.29	6.08	38.86	63.70	29.38
0.3	26.67	42.84	5.31	32.47	54.41	30.12
0.35	22.86	36.38	4.52	27.74	45.66	30.79
0.4	20.00	31.51	4.00	24.20	40.80	31.36
0.45	17.78	27.95	3.53	21.95	36.72	31.84
0.5	16.00	25.25	3.15	19.71	32.61	32.25
0.55	14.55	22.99	2.89	18.15	29.82	32.63
0.6	13.33	21.09	2.57	16.73	27.25	33.00
0.65	12.31	19.37	2.28	15.50	24.66	33.36
0.7	11.43	17.92	2.04	14.40	22.85	33.70
0.75	10.67	16.45	1.66	13.28	18.91	34.06
0.8	10.00	15.28	1.51	12.36	17.72	34.40
0.85	9.41	14.13	1.58	10.66	16.62	34.98
0.9	8.89	13.30	1.46	10.13	15.52	35.04
0.95	8.42	12.44	1.32	9.64	14.59	35.72
1.00	8.00	11.74	1.31	9.03	13.78	36.12

Table III: Compression Ratios Achieved on 768 × 768 Images, Option 2.

r bpp	CR_{theo}	CR_{avg}	STD	CR_{min}	CR_{max}	$PSNR_{avg}$
0.1	80.00	115.96	11.92	86.98	134.20	26.21
0.15	53.33	80.84	9.21	60.32	97.59	27.55
0.2	40.00	62.37	7.27	47.02	76.59	28.57
0.25	32.00	50.55	5.94	38.32	62.34	29.40
0.3	26.67	41.96	5.16	31.83	52.49	30.16
0.35	22.86	35.50	4.49	26.74	44.99	30.85
0.4	20.00	30.88	3.87	24.02	39.82	31.41
0.45	17.78	27.60	3.53	21.26	35.56	31.87
0.5	16.00	24.98	3.29	18.23	32.15	32.28
0.55	14.55	22.71	2.98	16.00	29.09	32.68
0.6	13.33	20.72	2.64	14.57	26.21	33.07
0.65	12.31	18.96	2.36	13.52	23.62	33.46
0.7	11.43	17.39	2.09	12.62	21.01	33.85
0.75	10.67	16.01	1.87	11.89	18.82	34.25
0.8	10.00	14.88	1.74	11.21	17.60	34.64
0.85	9.41	13.93	1.62	10.63	16.44	35.03
0.9	8.89	13.11	1.50	10.05	15.51	35.12
0.95	8.42	12.36	1.39	9.49	14.56	35.77
1.00	8.00	11.72	1.30	9.02	13.73	36.12

Table IV: Compression Ratios Achieved on 512×512 Images, Option 1.

r bpp	CR_{theo}	CR_{avg}	STD	CR_{min}	CR_{max}	$PSNR_{avg}$
0.1	80.00	80.34	2.87	74.98	85.81	23.72
0.15	53.33	54.73	1.72	51.25	58.19	25.15
0.2	40.00	41.68	1.16	39.31	43.52	26.25
0.25	32.00	33.72	0.84	32.00	35.39	27.14
0.3	26.67	28.10	0.64	26.85	29.25	27.96
0.35	22.86	23.80	0.60	22.58	25.02	28.72
0.4	20.00	20.68	0.49	19.60	21.75	29.36
0.45	17.78	18.46	0.36	17.84	19.16	29.89
0.5	16.00	16.71	0.33	16.26	17.41	30.37
0.55	14.55	15.27	0.28	14.87	15.91	30.82
0.6	13.33	14.05	0.28	13.61	14.64	31.25
0.65	12.31	12.99	0.25	12.54	13.52	31.66
0.7	11.43	12.07	0.25	11.56	12.57	32.07
0.75	10.67	11.25	0.21	10.80	11.70	32.46
0.8	10.00	10.47	0.34	9.11	10.86	32.92
0.85	9.41	9.79	0.39	8.72	10.25	33.34
0.9	8.89	9.23	0.34	8.36	9.68	33.71
0.95	8.42	8.69	0.36	7.90	9.17	34.12
1.00	8.00	8.18	0.38	7.53	8.70	34.58

Table V: Compression Ratios Achieved on 512×512 Images, Option 2.

r bpp	CR_{theo}	CR_{avg}	STD	CR_{min}	CR_{max}	$PSNR_{avg}$
0.1	80.00	77.53	3.00	71.08	83.09	23.76
0.15	53.33	53.33	1.51	49.79	55.69	25.18
0.2	40.00	41.06	1.08	38.59	42.64	26.23
0.25	32.00	33.30	0.78	31.49	34.64	27.15
0.3	26.67	27.58	0.63	26.12	28.59	27.99
0.35	22.86	23.30	0.55	22.04	24.34	28.76
0.4	20.00	20.36	0.42	19.54	21.35	29.38
0.45	17.78	18.27	0.34	17.74	19.05	29.91
0.5	16.00	16.58	0.34	15.86	17.36	30.39
0.55	14.55	15.13	0.39	13.87	15.83	30.84
0.6	13.33	13.89	0.44	12.21	14.54	31.28
0.65	12.31	12.81	0.45	11.03	13.41	31.71
0.7	11.43	11.84	0.44	10.22	12.41	32.15
0.75	10.67	10.99	0.43	9.61	11.51	32.58
0.8	10.00	10.23	0.44	9.09	10.80	33.00
0.85	9.41	9.60	0.44	8.69	10.22	33.41
0.9	8.89	9.05	0.43	8.32	9.68	33.82
0.95	8.42	8.56	0.42	7.90	9.16	34.22
1.00	8.00	8.14	0.39	7.52	8.68	34.62

Comparison of Exception-Handling Options.

Mean compression ratios are plotted alongside the predicted values corresponding to r bpp, for both exception-handling options 1 and 2, in Figures 2 and 3. The slight gap between the predicted curve and the curves for 512×512 images is due to the fact that we have evidently overestimated the loading factor; lowering the loading factor slightly would make the predicted and empirical curves coincide better but should have no effect on either option's rate-distortion performance. The important thing to note in these plots is that the observed compression ratios track the predicted values consistently across the entire range of bit rates tested and that the curves for 512×512 images match the predicted values quite closely. Moreover, the standard deviations reported for the observed compression ratios using both options 1 and 2 are comparable across the range of bit rates tested, indicating that the bit rate constraint is equally effective and robust at controlling lossy quantization with either exception-handling option.

Comparing the effects of the exception-handling options on quantifiable image distortion yields a similar story. While signal-to-noise ratio measurements are hardly the final word on an issue as subjective and application-dependent as image quality assessment, they at least provide a well-understood starting point for making quantitative comparisons. The plots of distortion (as measured in terms of mean PSNR) versus mean observed compression ratio in Figure 4 indicate that there is essentially no measurable difference in distortion between exception-handling options 1 and 2 for observed compression ratios spanning one full order of magnitude. The rate-distortion characteristics of options 1 and 2 both appear to be in line with theoretical rate-distortion modelling [3], which predicts a simple linear relationship between SNR and r at high bit rates. (Since we have plotted PSNR vs. compression ratio, and CR $\propto 1/r$, we expect a *reciprocal* relationship between PSNR and CR.) Thus, based on this numerical evidence, there is no quantitative reason for favoring either of these exception-handling options over the other. This issue will have to be settled on the basis of comparisons of subjective image quality, and we have not yet obtained conclusive qualitative assessments in this regard.

V. Conclusions and Recommendations.

We recommend adoption of the procedure outlined in Section III for the selection of quantizer bin widths, Q_k. The lossy bit rate constraint parameter, r, separates the lossy quantization level from the unpredictable but benign lossless coding gain achieved by zero run-length and Huffman coding. The extremely small variability in the observed compression ratios reported in Tables IV and V for 512×512 images shows how tightly the lossy bit rate constraint regulates the actual compression ratio in the absence of significant lossless coding gain, thereby acting as a "quality knob" for the WSQ algorithm. This demonstrates that the procedure is highly effective at preventing overcompression of images while still maintaining a consistently high level of lossy quantization, as specified by the user. It can also be shown that the bit allocation achieved by this method is mathematically equivalent to an optimal bit allocation based on minimizing a weighted mean-square error distortion measure, providing

a theoretical motivation for the procedure.

We suggest a value of about $r = 0.6$ bpp and a loading factor of $\gamma = 2.5$, which produces compression ratios of around 20:1 on full-sized fingerprint images. In light of the analysis presented in Section IV, we can endorse either exception-handling option and recommend that the choice be made on the basis of subjective image quality assessment.

References

[1] *WSQ Gray-Scale Fingerprint Image Compression Specification*, (version 2.0), Criminal Justice Information Services, Federal Bureau of Investigation, Wash. DC, Feb. 1993. Drafted by T. Hopper, C. Brislawn, and J. Bradley.

[2] J. Bradley, C. Brislawn, and T. Hopper, "The FBI wavelet/scalar quantization standard for gray-scale fingerprint image compression," in *Proc. Conf. Visual Info. Process. II*, vol. 1961 of *Proc. SPIE*, (Orlando, FL, Apr. 1993), Soc. Photo-Opt. Instrument. Engineers. To appear.

[3] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression.* No. 159 in Int'l. Series in Engineering & Computer Science, Norwell, MA: Kluwer Academic Publishers, 1992.

[4] M. Rabbani and P. W. Jones, *Digital Image Compression Techniques.* No. 7 in Tutorial Texts in Optical Engineering, Bellingham, WA: SPIE Optical Engineering Press, 1991.

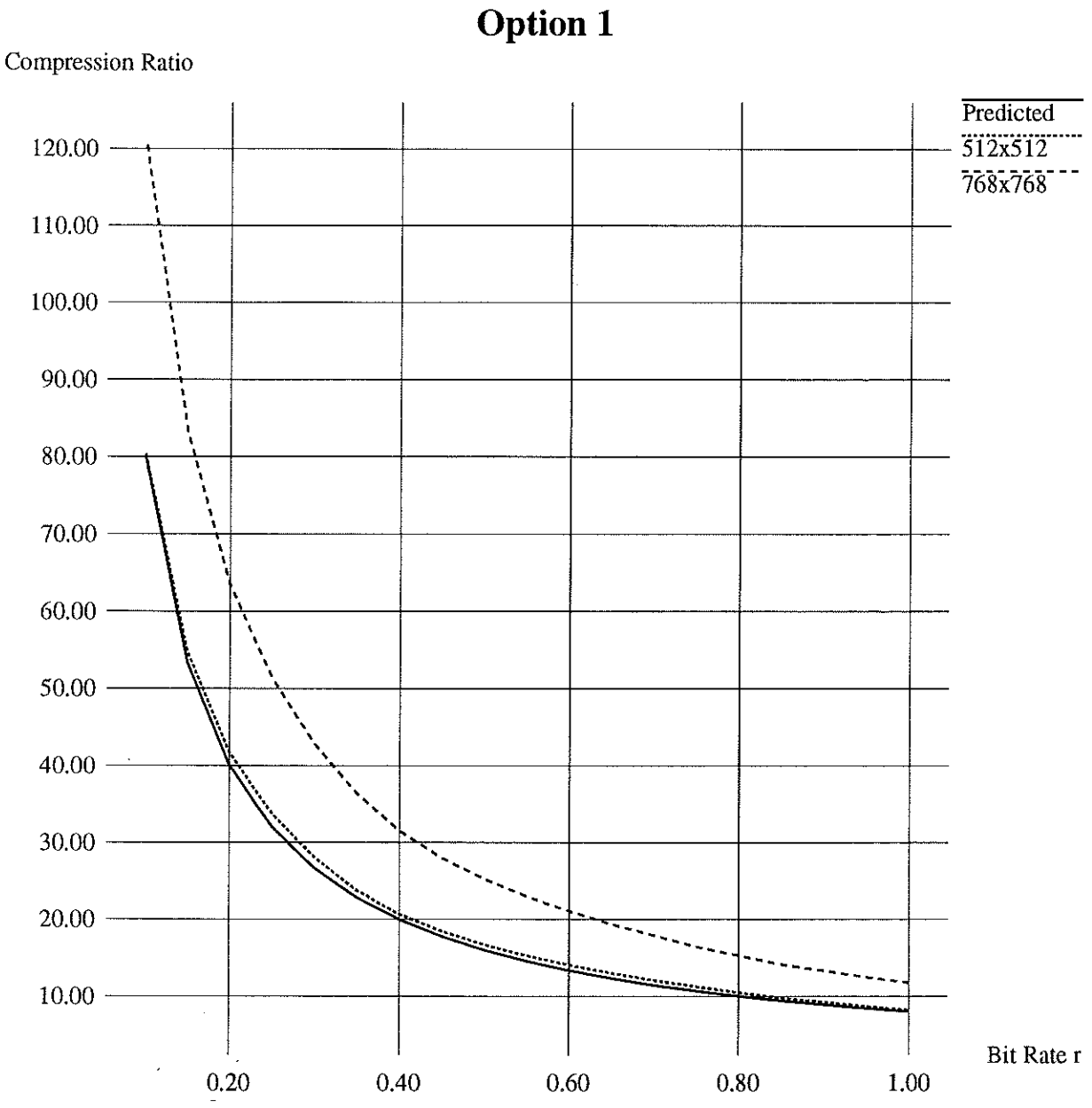

Figure 2: Compression Ratio vs. Bit Rate, Exception-Handling Option 1.

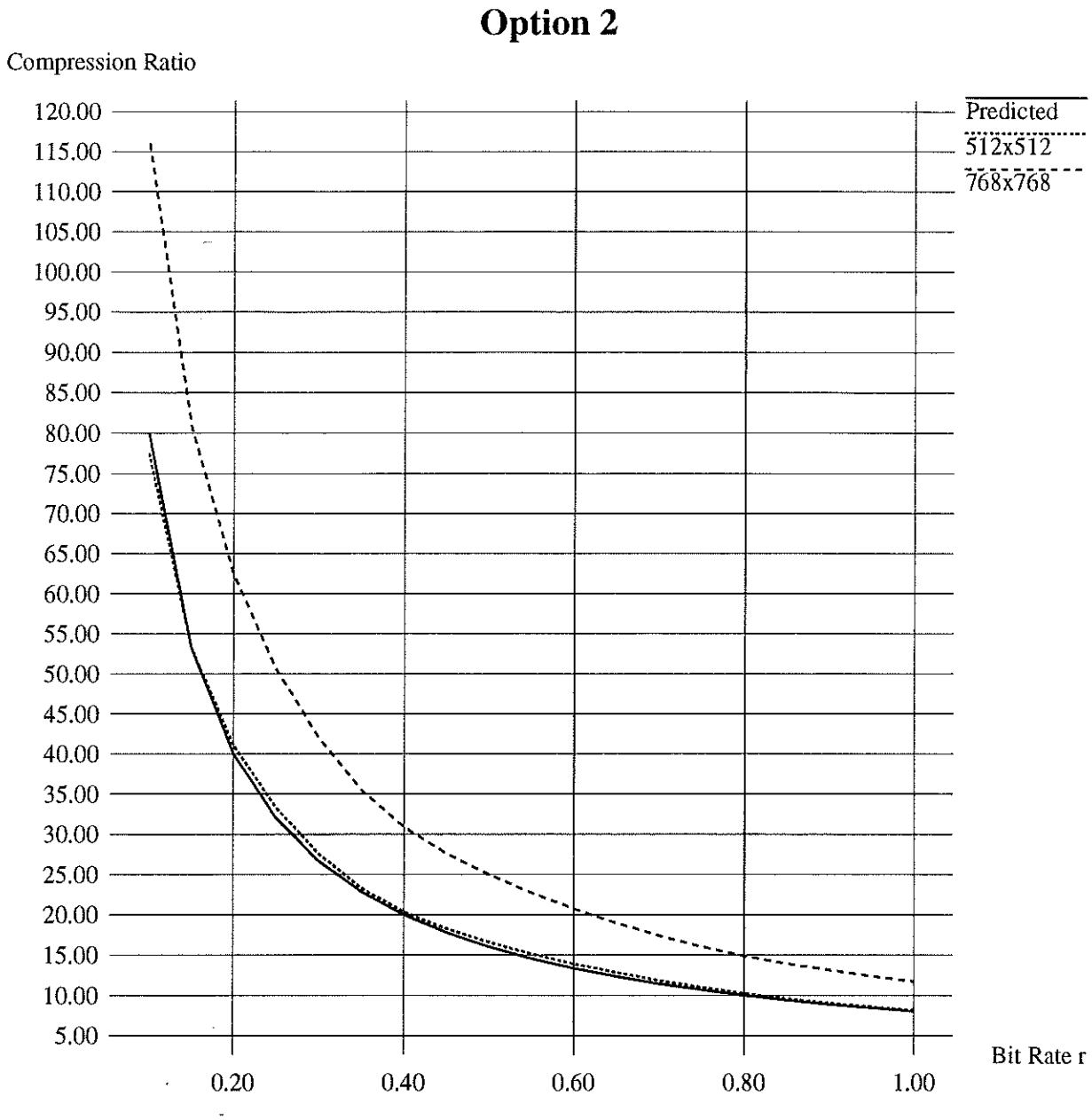

Figure 3: Compression Ratio vs. Bit Rate, Exception-Handling Option 2.

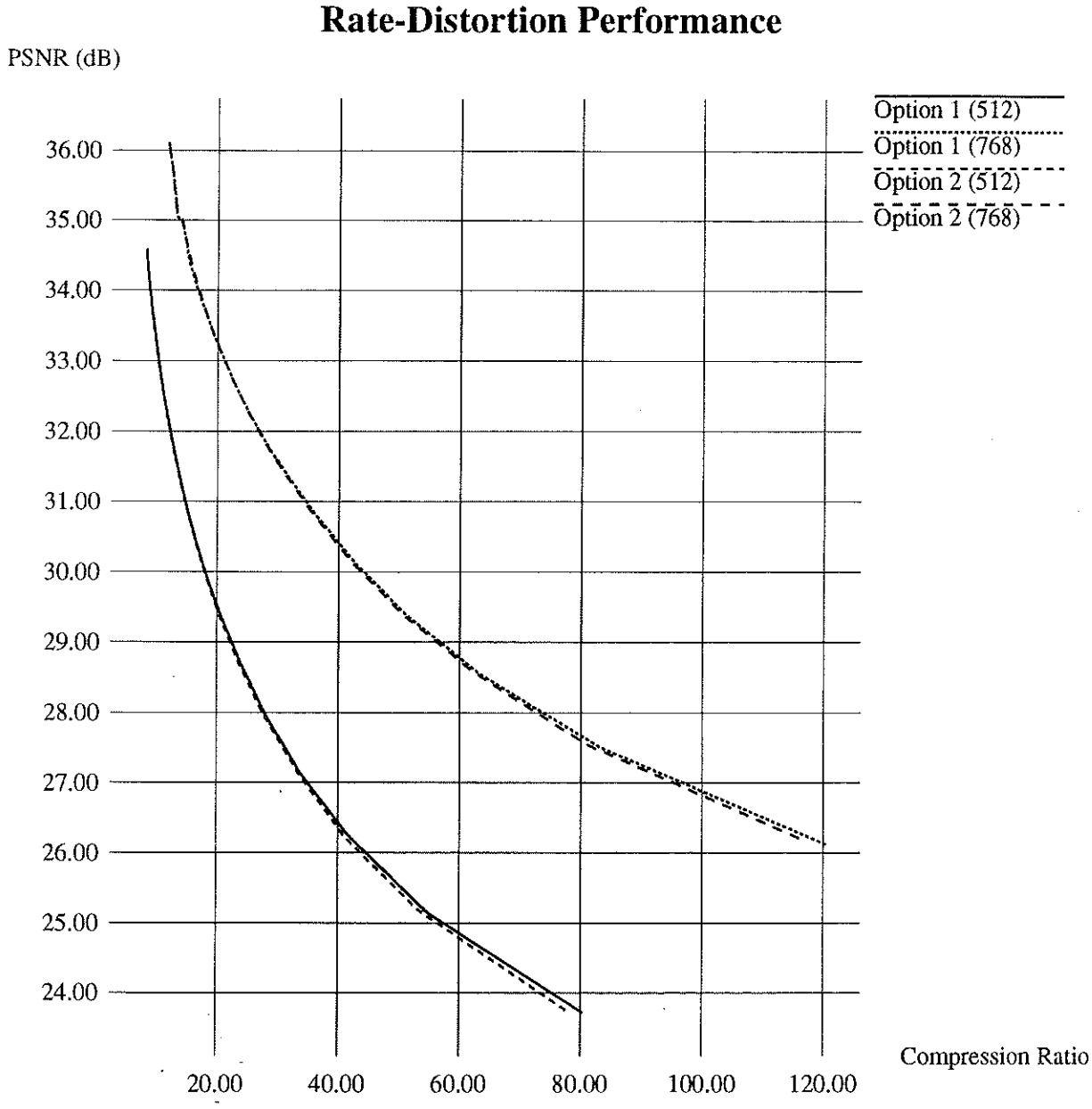

Figure 4: Distortion vs. Compression Ratio, Options 1 and 2.