The FBI Wavelet/Scalar Quantization Standard for gray-scale fingerprint image compression *

Jonathan N. Bradley and Christopher M. Brislawn
Los Alamos National Laboratory, Los Alamos, NM 87545

Tom Hopper
Federal Bureau of Investigation, Washington, D.C. 20537

ABSTRACT

The FBI has recently adopted a standard for the compression of digitized 8-bit gray-scale fingerprint images. The standard is based on scalar quantization of a 64-subband discrete wavelet transform decomposition of the images, followed by Huffman coding. Novel features of the algorithm include the use of symmetric boundary conditions for transforming finite-length signals and a subband decomposition tailored for fingerprint images scanned at 500 dpi. The standard is intended for use in conjunction with ANSI/NBS-CLS 1-1993, American National Standard Data Format for the Interchange of Fingerprint Information, and the FBI's Integrated Automated Fingerprint Identification System.

1 INTRODUCTION

Since the FBI began keeping fingerprint records in 1924, their collection has grown from an initial 810,000 cards to a present size of over 25 million cards. Archiving this information in the form of inked impressions on paper cards has obvious drawbacks when it comes to transmission, storage, and automated analysis of fingerprints. While bit-mapped (black/white) facsimile scans of inked impressions have been used to provide rapid transmission of "Post Office-grade" reproductions, the quality of facsimile digitization is not high enough to permit replacing the original cards with their facsimile scans. Nonetheless, there are many advantages to storing and transmitting fingerprint records in some type of digital format, and a number of municipal and state jurisdictions have been implementing different commercial digital imaging systems for recording fingerprint data. This has led to compatibility problems resulting from the use of multiple competing, proprietary hardware systems and data formats, a situation that has generated a demand for standardization in the criminal justice community. Another major factor behind the FBI's interest in fingerprint digitization is improving their response time to justice system inquiries regarding criminal histories or outstanding warrants for arrested suspects prior to arraignment. This will require both rapid transmission of arrest records, such as fingerprints, and automation of background checks, a task that will be facilitated by an Integrated Automated Fingerprint Identification System, currently undergoing development.

In response to these issues, the FBI's Criminal Justice Information Services Division (CJIS) has developed standards for fingerprint digitization in cooperation with the commercial vendor and criminal justice communities, with the assistance of the National Institute of Standards and Technology (NIST) and Los Alamos National Laboratory (LANL). The pertinent specifications for fingerprint digitization are contained in [1, 2]; the present article is devoted to providing an informal description of the digital fingerprint image compression algorithm specified in [2]. If there are any discrepancies between the description offered here and the official CJIS specification, the official specification naturally takes precedence.

*Reproduced from: SPIE Proceedings, vol. 1961, *Visual Information Processing II*, Orlando, FL, pp. 293–304, Apr. 1993.

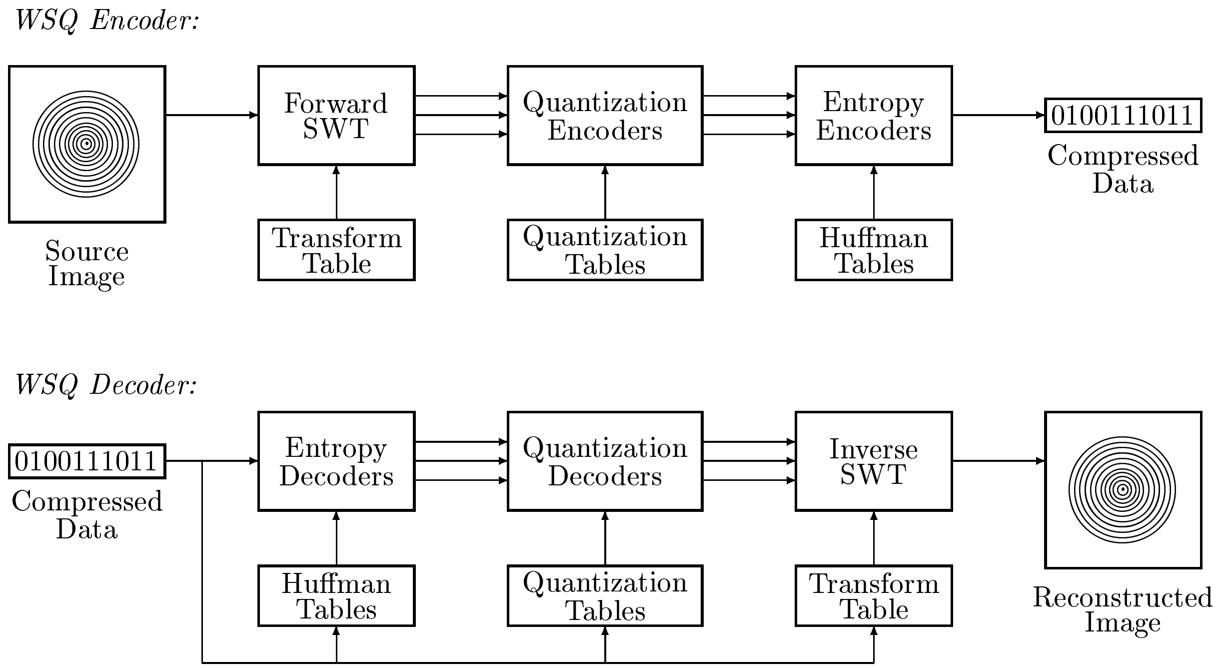

Figure 1: Simplified WSQ Encoder and Decoder Diagrams.

Fingerprint images are digitized at a resolution of 500 pixels/inch with 256 levels (8 bits) of gray-scale information; gray-scale images have a more "natural" appearance to human viewers than do bit-mapped images and allow a higher level of subjective discrimination by fingerprint examiners. The data storage and transmission requirements imposed by this level of resolution are considerable: a single 1.5 inch^2 fingerprint block is transformed into around 600 kilobytes of digital information; an entire card (10 rolled impressions, plain impressions of the thumbs and simultaneous impressions of both hands) produces about 10 megabytes of data. At this rate, digitizing the FBI's current holdings would result in some 250 terabytes of archival data. There are also significant transmission considerations: at conventional high-speed modem rates (9600 bits/second, 20% overhead), electronic transmission of a single 10MB card would take almost 3 hours. While this is still considerably faster than even overnight delivery services and eliminates the danger of an irreplaceable card being lost or damaged in the mail, the magnitude of the fingerprint database is such that the FBI has made data compression part of the digitization standard.

The compression algorithm selected by CJIS is based on adaptive uniform scalar quantization of a wavelet transform subband image decomposition and is referred to as the *wavelet/scalar quantization* (WSQ) standard. This particular approach was chosen on the basis of follow-up studies to the investigation reported in [3], and the WSQ algorithm's suitability for fingerprint image data has been verified in tests performed by FBI fingerprint examiners. Testing has shown that the algorithm's compressed image quality is high enough to be acceptable for archival purposes at compression ratios of around 20:1. The WSQ standard contains some elements of the LANL wavelet/vector quantization algorithm described in [3, 4] and has been developed jointly by CJIS and LANL.

2 OVERVIEW OF THE WSQ ALGORITHM

An overview of the WSQ algorithm is shown in Figure 1. The algorithm consists of three main processes: wavelet transform decomposition of the source fingerprint image, scalar quantization of the wavelet

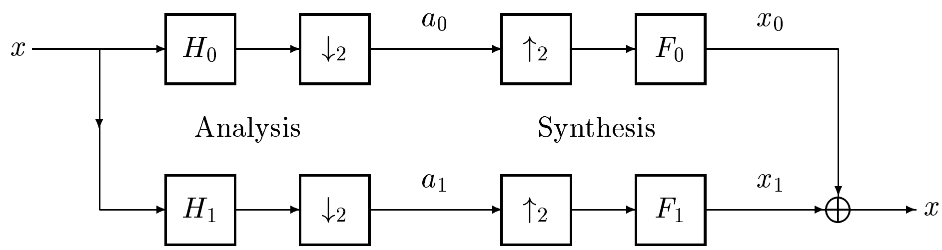

Figure 2: Two-Channel Perfect Reconstruction Subband Coder.

coefficients, and lossless entropy coding of the quantizer indices. The standard specifies a class of encoders and a single decoder with sufficient generality to decode compressed image data produced by any compliant encoder.

In a WSQ encoder, the digitized source image is decomposed into 64 subbands of floating-point wavelet coefficients by a two-dimensional symmetric wavelet transform (SWT). The SWT subbands are then passed to a bank of uniform scalar quantizers; parallel arrows indicate multiple information channels at this stage in the process. The integer indices output by the quantization encoders are entropy-encoded by run-length coding of zeros and Huffman coding. The compressed image data, a table of wavelet transform specifications, and tables for the scalar quantizers and Huffman coders are catenated into a single bit-stream of compressed data.

There are two principal formats specified for the compressed data: an interchange format containing all tables needed to decode an image after transmission between applications, and an abbreviated format for use within a single application in which the tabulated data is available from other sources. The syntax for compressed data is modelled closely on the syntax employed in the JPEG still image compression standard [5]. Two-byte markers are included to enable the decoder to parse the compressed bit-stream and locate side information—like table specifications—before decoding the compressed image data.

The WSQ decoder parses the compressed data and extracts the tables needed in the decoding process. An entropy decoder uses the Huffman tables to decode the compressed SWT subbands, and the scalar quantizer indices are then decoded to reconstruct quantized wavelet coefficients, which are approximations of the original wavelet coefficients. The quantized coefficients are run through an inverse SWT to produce the reconstructed image.

The standard allows for encoders using wavelet filters from either of two distinct classes of linear phase perfect reconstruction filter banks in conjunction with two different symmetric wavelet transform algorithms and image-specific scalar quantizers and entropy coders. Parameter settings for the first FBI-approved encoder, including the choice of filters, scalar quantizer parameters, and Huffman coding specifications, are given in [2] and will be described in Section 6. The CJIS specification also includes compliance tests for encoders, decoders, and the compressed data format; we refer the reader to [2] for details on compliance testing.

3 THE WAVELET TRANSFORM SUBBAND DECOMPOSITION

A one-dimensional, two-channel, perfect reconstruction multirate filter bank is depicted in Figure 2. The invertible linear transformation $x \to \{a_0, a_1\}$ induced by the analysis filter bank is called a *discrete wavelet transform* (DWT). For image processing applications, it is necessary to specify how boundary

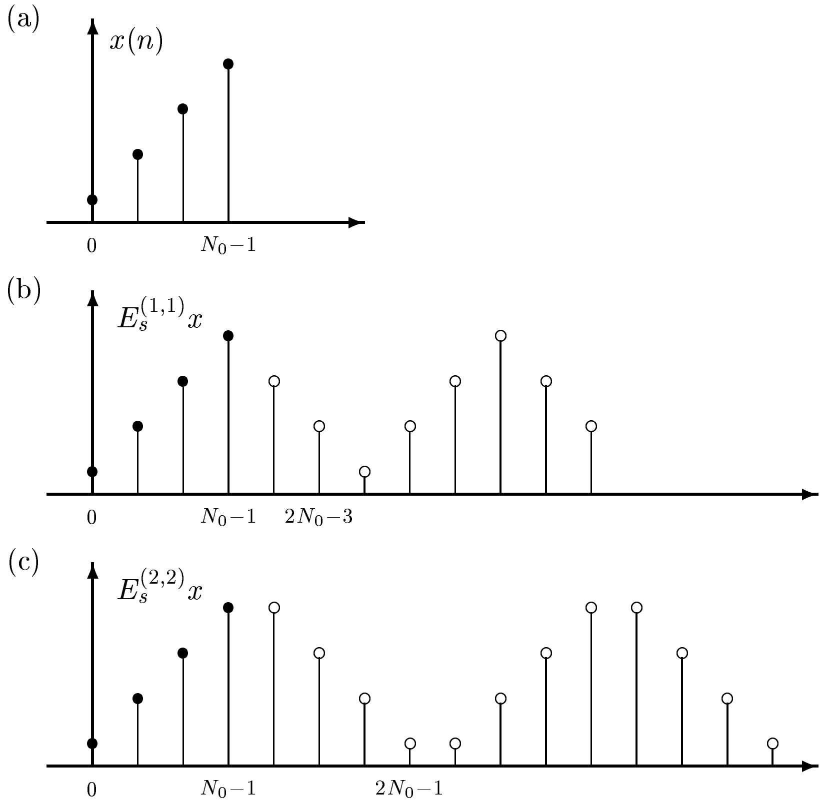

Figure 3: Two Types of Periodized Symmetric Signal Extension.

conditions are to be handled when the input, x, is a finite-duration signal, such as a row or column vector from a digitized image. In general, a single filter bank is capable of furnishing a number of different transforms depending on how it is applied to finite-duration inputs. The simplest way to handle an input of length N_0 is to apply finite impulse response filters by N_0-periodic circular convolution, followed by 2:1 circular downsampling (the *circular DWT*). There are two problems with this approach: first, periodization introduces a jump discontinuity in the signal, which adds variance to the highpass subbands, and second, 2:1 circular downsampling is only possible when N_0 is even. Added highband variance adversely affects quantizer performance, and the digitization standard imposes no constraints on image dimensions (e.g., dimensions need not be powers of 2, or even divisible by 2), so the WSQ standard addresses both of these problems at once by applying the filter bank to a periodized *symmetric extension* of the input. A transformation defined in this way is called a *symmetric wavelet transform* (SWT) [6].

Two different symmetric extensions are shown in Figure 3; note that both have even periods, regardless of whether N_0 is even or odd. Since the length (or period) of the input has been effectively doubled by the symmetric extension process, the crucial issue is ensuring that the transform does not expand the size of the signal being transformed. An SWT is *nonexpansive* if the original signal of length N_0 can be reconstructed perfectly from just N_0 transform coefficients. This is accomplished by using linear phase filters designed so that the downsampled SWT subbands will also be symmetric and can therefore be windowed (or truncated) with no loss of information. The standard [2] specifies two distinct classes of SWT's for the two nontrivial families of linear phase perfect reconstruction filter banks.

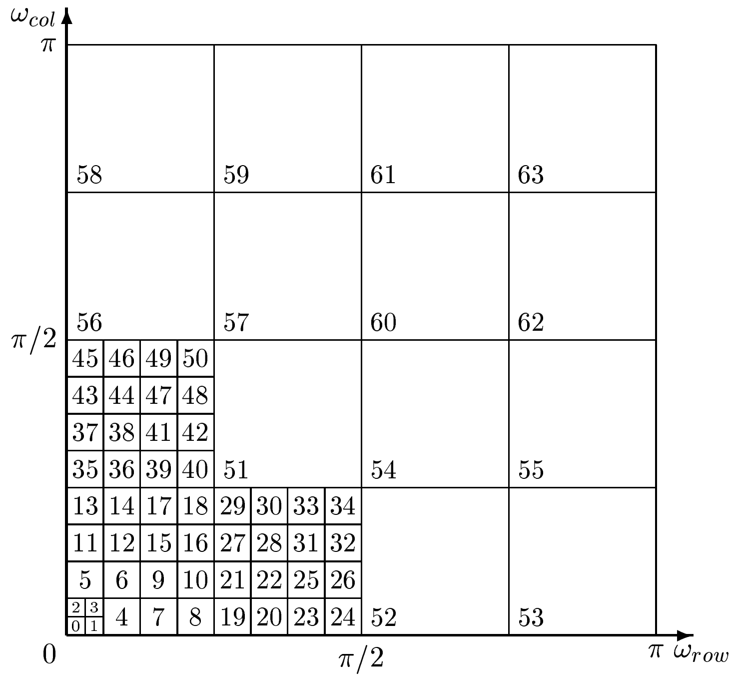

Figure 4: Frequency Support of Wavelet Transform Subbands in WSQ Standard.

The extension $y = E_s^{(1,1)}x$ shown in Figure 3(b) is used with odd-length linear phase filters, a lowpass filter, h_0, symmetric about $n = 0$ and a highpass filter, h_1, symmetric about $n = -1$. The extension $y = E_s^{(2,2)}x$ shown in Figure 3(c) is used with banks of even-length linear phase filters containing a symmetric lowpass filter and an antisymmetric highpass filter, both centered at $-1/2$. With these conventions, when N_0 is even it is necessary to transmit just $N_0/2$ coefficients in both the lowpass and highpass channels, for either SWT method. When N_0 is odd, perfect reconstruction can be assured by transmitting $(N_0 + 1)/2$ coefficients in the lowpass channel and $(N_0 - 1)/2$ in the highpass channel, again, for both of the two SWT methods. Thus, the SWT's specified by the WSQ standard are nonexpansive for input signals of either even or odd length, so the SWT accomodates the FBI requirement that there be no constraints on image dimensions. The encoder transmits the analysis filters, h_0 and h_1, along with the compressed image data, and the decoder is able to use this information to construct the synthesis filters, f_0 and f_1, via known *anti-aliasing relations*. A detailed treatment of symmetric wavelet transform methods is given in [6], including discussion of the specific implementations employed in [2].

As mentioned above, the SWT is applied to a two-dimensional digital input image by transforming first the rows and then the columns of the image, yielding a four-channel decomposition. The four subbands are then cascaded back through the two-dimensional analysis bank to produce a more refined 16-channel decomposition. The cascade is repeated several more times on some of the resulting lowpass subbands until a 64-band decomposition is achieved; see Figure 4 for a depiction of the approximate frequency passbands for the resulting channels. This decomposition was designed from an analysis of the power spectral density (PSD) of fingerprint images, the information-packing performance of different decompositions, and empirical studies of the effects of quantization of specific subbands on reconstructed image quality. Note that the PSD estimate made in [4] shows that the natural frequencies of the ridges in fingerprint images are in the portion of the spectrum contained roughly in subbands 7–18, that is, periods of about 8–16 pixels.

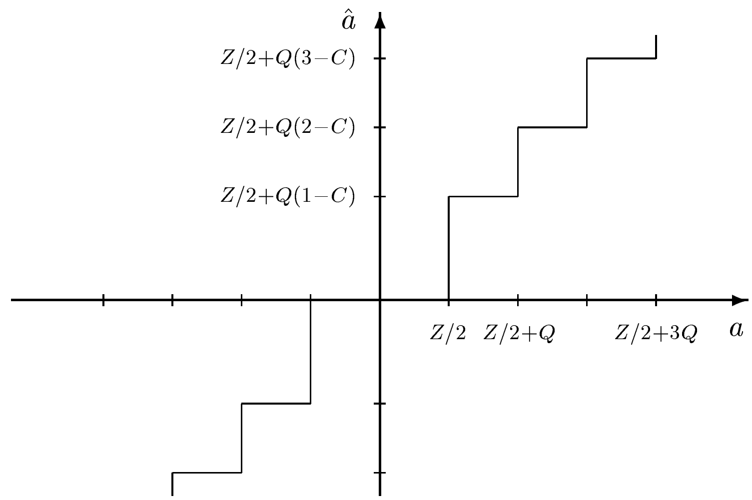

Figure 5: WSQ Scalar Quantizer Characteristic.

4 SUBBAND QUANTIZATION

Lossy compression in the WSQ algorithm is achieved by uniform scalar quantization of the SWT subbands. The source images with their 8-bit gray-scale resolution are *continuous-tone* images as far as the compression standard is concerned, and the resulting wavelet coefficients are regarded as analog input by the encoding process. The term *amplitude quantization*, or simply *quantization*, refers to the procedure that maps an analog floating-point wavelet coefficient, a, to one of finitely many *quantized* floating-point values, $\hat{a}$. This is done in two stages. In the WSQ encoder, a *scalar quantization encoder* maps a to the (integer) *quantizer index*, p, that points to the quantization bin in which a lies. A quantization encoder is described mathematically by a function $E : \mathbf{R} \to \mathbf{S}$, where $\mathbf{S}$ is a discrete set of quantizer indices. It is the quantizer indices that are entropy-encoded and transmitted in a compressed format. Since E is not invertible, the resulting compression is inherently lossy.

In the WSQ decoder, a *quantization decoder* maps quantizer indices, p, back to a discrete set of reconstructed floating-point values, $\hat{a}$, which are called *quantized wavelet coefficients*. A quantization decoder is described mathematically by a one-to-one function $D : \mathbf{S} \to \mathbf{R}$; note that the decoder, D, is *not* the inverse (in the mathematical sense) of the encoder, E. The composite function $F = D \circ E$, which maps analog floating-point input to quantized floating-point output, is known as the *quantizer characteristic* [7] (e.g., see Figure 5).

Within a single subband, the quantization intervals, or bins, are of equal width with the exception of the bin containing the origin (the *zero bin*), which is somewhat wider based on noise threshold estimates. Bin widths may vary from subband to subband; the bin width for the k^{th} subband is denoted by Q_k, and Z_k denotes the width of the k^{th} zero bin. A procedure for selecting bin widths Q_k and Z_k for the subbands in the WSQ decomposition will be discussed in Section 6. The WSQ quantizer characteristic in Figure 5 is defined mathematically by the following equations. Quantization encoding of the k^{th} two-dimensional 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. Decoding is 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 is the bin's midpoint.

Some remarks are in order concerning the above terminology. When [2] was drafted, the terminology used for scalar quantization was initially chosen to maintain consistency with the JPEG standard [5, 8], including the use of the terms "quantizer/dequantizer" for what have here been called the "quantization encoder/decoder." In much of the signal processing literature on amplitude quantization, however, the term "quantizer" is used for the composite quantizer characteristic function, F. Moreover, when authors consider amplitude quantization in isolation from other system components, such as entropy encoding, they also commonly refer to the functions E and D simply as the "encoder" and "decoder," respectively, which suggests the use of the terms "quantization encoder/decoder" when more specificity is required. Since the decoding process doesn't undo the quantization of the coefficients effected by the encoder but merely maps discrete quantizer indices to discrete floating-point values, the terms "quantization decoder" and "quantized signal" seem more appropriate for D and $\hat{a}$ than "dequantizer" and "dequantized signal."

For thorough treatments of quantization theory see [7, 9].

5 ENTROPY CODING OF QUANTIZER OUTPUT

Following scalar quantization of the image subbands, the indices $p_k(m,n)$ are mapped to a stream of symbols, given in Table 1, which are then Huffman coded. Table 1 specifies how the various indices and zero run-lengths are represented by a set of 254 symbols. Symbols 107–254 are used to transmit index values between -73 and 74. If a nonzero index from outside this range is encountered, the appropriate escape symbol is transmitted followed by the actual integer index value. For example, for a positive (resp., negative) index with absolute value less than 256 and greater than 74, symbol 101 (resp., 102) is transmitted followed by the absolute value of the index as an 8-bit integer. Similarly, for an index of absolute value less than 65536 and greater than or equal to 256, symbol 103 (resp., 104) is transmitted followed by the absolute value of the index as a 16-bit integer. No means for transmitting an index of absolute value greater than or equal to 65536 is provided. For realistic values of Q_k, far fewer than this number of quantization levels will actually be needed; this implies the absence of quantizer overload distortion [7]. Symbols 1–100

Table 1: Huffman Table Input Symbols.

Symbol	*Value*
1	zero run-length 1
2	zero run-length 2
3	zero run-length 3
⋮	⋮
100	zero run-length 100
101	escape for positive 8 bit index
102	escape for negative 8 bit index
103	escape for positive 16 bit index
104	escape for negative 16 bit index
105	escape for zero run—8 bits
106	escape for zero run—16 bits
107	index value −73
108	index value −72
109	index value −71
⋮	⋮
180	*Not used. Use symbol 1.*
⋮	⋮
253	index value 73
254	index value 74

in Table 1 are used for transmitting zero run-lengths. Zero run-lengths greater than 100 are coded by transmitting the escape symbol 105 or 106 followed by an integer specifying the length of the run.

The Huffman coding tables are image dependent and therefore must be contained in the coded data format. The standard specifies that subbands will be grouped into 3 to 8 blocks for Huffman coding to facilitate progressive transmission capabilities. All subbands within a block are coded using the same Huffman table. The occurrence of each symbol in Table 1 is counted for each block; the resulting counts are used to calculate the Huffman table codeword lengths, which in turn determine the codeword for each symbol in an unambiguous manner according to a prescribed procedure. Transmission of the Huffman tables is accomplished by transmitting the array of codeword lengths and a corresponding list of symbols. The method employed for Huffman coding was adopted from the JPEG specification [5], which contains many relevant suggestions that are helpful to the WSQ implementer. A good discussion of Huffman coding can also be found in [8].

6 THE FIRST APPROVED WSQ ENCODER

At present, only one encoder has been approved by CJIS for fingerprint image compression, although the standard allows for additional encoders in the future within the range specified in [2]. Improvements in quantizer performance are certainly desirable provided the resulting compressed data complies with the format specification expected by the decoder.

The digital filter bank in encoder #1 corresponds to a regular biorthogonal wavelet basis constructed by Cohen, Daubechies, and Feauveau [10]; details of the construction and pictures of the mother wavelets and scaling functions can be found in [11, 12]. Both the analysis and synthesis banks consist of pairs of symmetric filters with 7 and 9 impulse response taps. Exact expressions for the taps can be found in [2].

These filters were selected by trial and error based on quantizer performance in comparison with other perfect reconstruction filter banks; this particular filter bank produced quantized images superior to those generated by the other filters tested. The importance for digital image coding applications of the regularity of the associated continuous wavelets is not well understood at present, although regularity does imply that distortion in the quantized lowpass subbands will be smooth and slowly varying rather than abrupt or fractal-like. The fact that the support of these filters closely matches the natural frequency of fingerprint ridges is probably another factor in their excellent performance on fingerprint images.

Scalar quantization for encoder #1 is based on the following quantizer bin-width formula:

$$Q_k = \begin{cases} 1/q \quad , & 0 \le k \le 3 \quad , \\ 10/\left(qA_k \log_e \sigma_k^2\right) \quad , & 4 \le k \le 59 \quad , \\ 0 \quad , & 60 \le k \le 63 \quad . \end{cases}$$

The value $Q_k = 0$ for bands 60–63 is interpreted to mean that these subbands are discarded altogether by the encoder and not transmitted. The constants A_k are tabulated in [2]; these constants and the factors $\log_e \sigma_k^2$ give the relative widths of the quantizer bins for the 64 subbands. The logarithmic factor was determined empirically to give good qualitative results when quantizing fingerprint images. When $\log_e \sigma_k^2 \le 0$, the k^{th} subband is assigned a bit rate of 0 bits/pixel. The parameter q sets the overall lossy quantization rate for the encoder, which determines the distortion introduced by the compression process. The authors are presently working on verifying a formula for computing q to ensure that the lossy compression ratio (and therefore the quantization distortion) is consistent from image to image. Tests show that q can be set to obtain a prespecified compression ratio, modulo benign (distortion-less) compression due to variable amounts of zero run-length and Huffman coding gain. The reconstruction parameter, C, has the value $C = 0.44$, and the width of the zero bin is given by

$$Z_k = 1.2\, Q_k \quad .$$

Finally, encoder #1 constructs just three entropy coding blocks and only two Huffman coding tables, one for low- and mid-frequency subbands 0 through 18, and another for the highpass detail subbands 19–59, which are divided into two blocks with a block boundary between bands 51 and 52.

7 CONCLUSIONS AND FUTURE ACTIVITY

This paper has provided a brief overview of the WSQ fingerprint image compression standard. The algorithm has been approved by the FBI for compression ratios on the order of 20:1. The effectiveness of the WSQ standard can be readily judged from empirical results. Figure 6 is an original 768×768 8-bit gray-scale fingerprint image; Figure 7 shows the results of compressing this image 26.0 : 1 using WSQ encoder #1 as per Section 6. One can see that minutiae such as ridge endings and bifurcations are accurately preserved, as well as finer features like ridge textures.

An incomplete aspect of the encoder specification discussed in Section 6 is that at present a formula is given only for setting the *relative* widths of the quantization bins; calculation of the parameter q, which determines the overall compression ratio, remains to be specified. Methods of calculating q based on optimal bit allocation techniques are currently under investigation so that a specification for q can be included in the WSQ standard.

Future research efforts will involve the development of additional compliant encoders that produce high-quality compressed images. This work will investigate the use of different linear phase perfect reconstruction filter banks as well as improved methods for performing bit allocation for the subband quantizers.

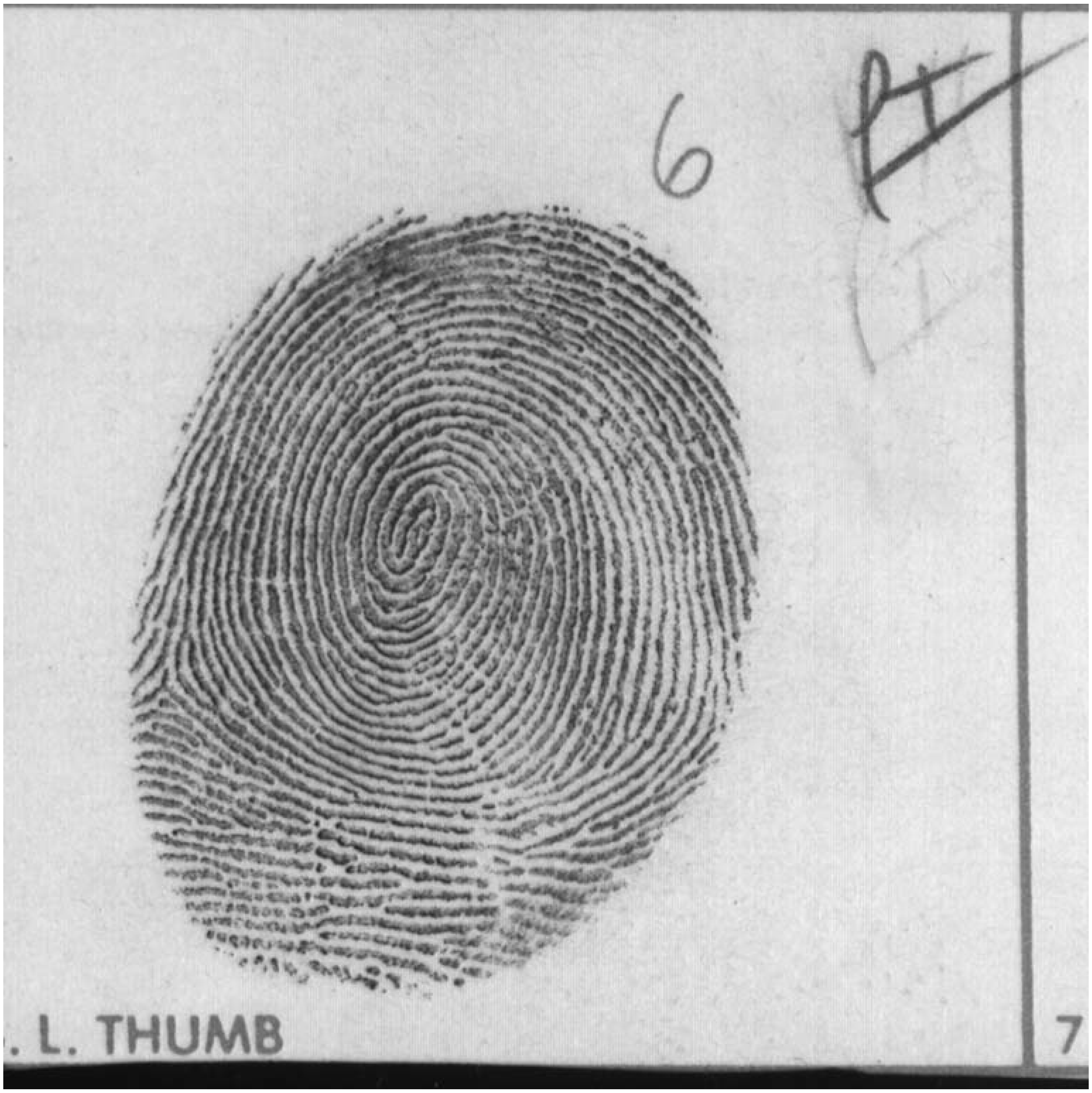

Figure 6: Original 768 × 768 8-bit Gray-Scale Fingerprint Image.

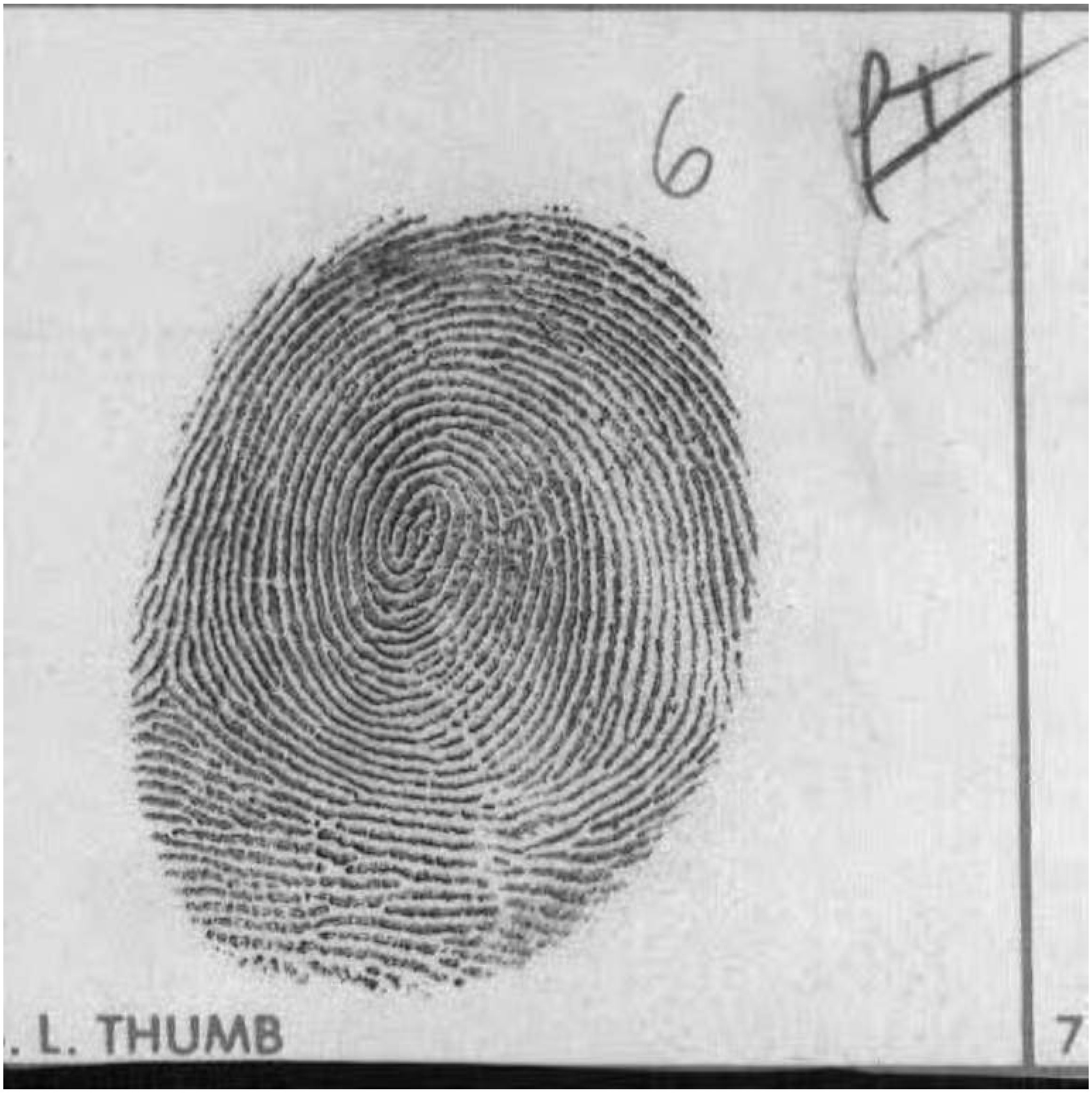

Figure 7: Fingerprint Image Compressed 26.0:1.

8 REFERENCES

[1] *American National Standard Data Format for the Interchange of Fingerprint Information*, ANSI/NBS-CLS 1-1993, (revised), American Nat'l. Standards Inst., 1993.

[2] *WSQ Gray-Scale Fingerprint Image Compression Specification*, (ver. 2.0), Criminal Justice Information Services, Federal Bureau of Investigation, Feb. 1993.

[3] T. Hopper and F. Preston, "Compression of grey-scale fingerprint images," in *Proc. Data Compress. Conf.*, (Snowbird, UT), pp. 309–318, IEEE Computer Soc., Mar. 1992.

[4] J. N. Bradley and C. M. Brislawn, "Compression of fingerprint data using the wavelet vector quantization image compression algorithm," Tech. Rep. LA-UR-92-1507, Los Alamos Nat'l. Lab., Apr. 1992. Progress report to the FBI.

[5] *Digital Compression and Coding of Continuous-Tone Still Images, Part 1, Requirements and Guidelines*, ISO Draft Int'l. Standard 10918-1, (a.k.a. "the JPEG Standard"), American Nat'l. Standards Inst., Feb. 1991.

[6] C. M. Brislawn, "Classification of symmetric wavelet transforms," Tech. Rep. LA-UR-92-2823, Los Alamos Nat'l. Lab., Mar. 1993. Revised draft.

[7] N. S. Jayant and P. Noll, *Digital Coding of Waveforms.* Prentice-Hall Signal Processing Series, Englewood Cliffs, NJ: Prentice Hall, 1984.

[8] W. B. Pennebaker and J. L. Mitchell, *JPEG Still Image Data Compression Standard.* New York, NY: Van Nostrand Reinhold, 1992.

[9] 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.

[10] A. Cohen, I. C. Daubechies, and J.-C. Feauveau, "Biorthogonal bases of compactly supported wavelets," Tech. Rep. 11217-900529-07 TM, AT&T Bell Labs, Murray Hill, NJ, May 1990.

[11] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform," *IEEE Trans. Image Process.*, vol. 1, pp. 205–220, Apr. 1992.

[12] I. C. Daubechies, *Ten Lectures on Wavelets.* No. 61 in CBMS-NSF Regional Conf. Series in Appl. Math., (Univ. of Lowell, Lowell, MA, June 1990), Philadelphia, PA: Soc. Indust. Appl. Math., 1992.