Subject: WSQ Compression Change in Subband Variance Computation.

This announcement proposes a change to the wavelet scalar quantization (WSQ) compression algorithm addressing an issue which can cause degradation of the compressed image. The problem occurs with the two-thumb images from live-scan identification flats (ID Flats) captured with a 3-inch high platen device. If the thumbs are placed sufficiently close to the bottom edge of the platen during capture, the compression/decompression of the image can result in a severely degraded fingerprint image (See the included images). In extreme cases all prints in the image are lost.

The degradation occurs because the variance computation is based on a subregion of the Discrete Wavelet Transform (DWT). If the thumbs are placed near the bottom of a 3-inch high platen, they can be outside of the variance computation subregion resulting in a variance computation that doesn't properly represent the fingerprint data in the image. The bad variance computation leads to poorly informed quantization decisions that cause irreversible corruption to the compressed fingerprint data.

This new method for subband variance computation will make a decision to use either the subregion or the full region of the DWT subband. The decision on which region depends on the sum of the subregion variance estimates for DWT subbands 0-3. If the sum of these variances is >20,000, the variance computation continues using the subregion of each DWT subband, otherwise the variance computation uses the full region of each DWT subband. A paper with details of this analysis is being prepared for public release. The Appendix shows the proposed modifications to Part 3-section 3.1 of the WSQ specification that are being reviewed.

The algorithm modification only alters the variance computation on an as needed basis and mainly for larger platen live-scan capture of two-thumb images. It is slightly possible that single finger captures and slap images captured on platens smaller than 3 inches in height will use this new variance computation when compressed. Studies done on existing NIST single finger live-scan, rolled image, and Identification Flat datasets showed that a small percentage of non two-thumb captures were affected by this change. Compression analysis of 500,000 single finger live-scan capture impressions and 100,000 four-finger slap capture impressions showed that <0.1% and <0.5% of each image type respectively did undergo full-region variance estimation as a result of this modification. Visual inspection of both types showed that the affected images were generally low quality with very little visible ridge structure. Based on compression analysis of 330,000 two-thumb impressions this modification affected approximately 12% of these impressions captured on a large platen live-scan device. The code modification specifically targeted cases where the subregion variance computation showed little to no variance. The modification also protects other applications that use a different capture device which allow most of the image to be captured outside the subregion variance computation area.

Operationally, very little change occurs for images other than two-thumb captures. Since the variance computation decision is based on the subregion variances for subbands 0-3, the compression time for the subregion variance method is practically identical to the previous algorithm. There is approximately a 20% increase in compression time for the 12% of two-thumb captures that used full region variance computation.

This modification changes only the variance computation in the compression algorithm. The decompression algorithm remains the same and can decompress images from either version of the compression algorithm.

The change has been made to the NIST WSQ algorithm that is part of NBIS release 3.3.1 and is available at http://www.nist.gov/itl/iad/ig/nigos.cfm

The WSQ specification will be modified to reflect this modification and when ready be made available for downloaded at https://www.fbibiospecs.org/biometric_specs.html.

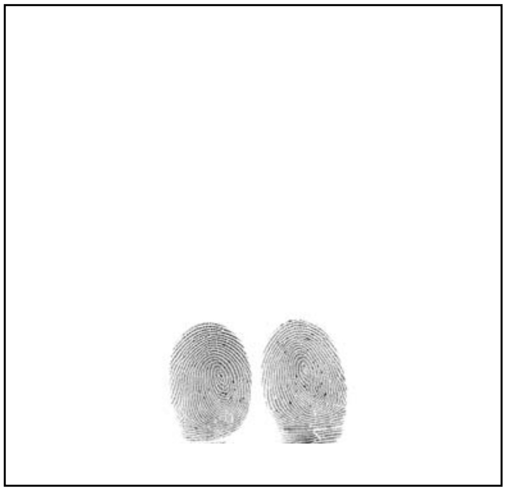

Figure 1: original

Figure 2: decompressed

$$B = \left(\frac{-14\sqrt{15} - 63}{1080\sqrt{15}}\right)^{1/3}$$

Table 1. Analysis Wavelet Filters

3 Adaptive Quantization of DWT Output

This section defines the parameters Q_k, Z_k, and C from Annex A.3 for WSQ encoder number two. The dequantization output level parameter, C, used in the quantization decoder shall be set to the value

$$C = 0.44$$

The bin widths, Q_k and Z_k, are determined from the variances of the DWT subbands as follows.

3.1 Subband variance computation

A subband variance computation is made based on either a subregion estimate or the entire region of each DWT subband. The decision on which region depends on the sum of the subregion variance estimates for DWT subbands 0-3. If $\sum_{k=0}^{3} \sigma_k^2 > 20{,}000$, the variance estimate is based on the subregion of each DWT subband, otherwise the variance computation uses the full region of each DWT subband.

For the variance estimate based on a subregion of each DWT, let $a_k(m,n)$ denote the floating point array of which X_k and height Y_k comprising the k^{th} subband, indexed as $0 \le m < Y_k$ and $0 \le n < X_k$ with (0,0) referring to the upper left corner of the subband. The width and height of the subregion used for the variance estimate are, respectively,

$$X'_k = \left\lfloor \frac{3X_k}{4} \right\rfloor \ and \ \ Y'_k = \left\lfloor \frac{7Y_k}{16} \right\rfloor$$

or for the variance computed on the entire region of each DWT the width and height are,

$$X'_k = X_k \ and \ Y'_k = Y$$

The variance shall be calculated with the unbiased estimator

$$\sigma_k^2 = \frac{1}{X'_k Y'_k - 1} \sum_{n=x_{0,k}}^{x_{1,k}} \sum_{m=y_{0,k}}^{y_{1,k}} \left(a_k(m,n) - \mu_k\right)^2$$

where μ_k denotes the mean of a_k.

The horizontal and vertical offsets for the subregions ($x_{i,j}$ and $y_{i,k}$, respectively), relative to the upper left corner, are

$$x_{0,k} = \left\lfloor \frac{X_k}{8} \right\rfloor$$

$$x_{1,k} = x_{0,k} + X'_k - 1$$

$$y_{0,k} = \left\lfloor \frac{9Y_k}{32} \right\rfloor$$

$$y_{1,k} = y_{0,k} + Y'_k - 1$$

The horizontal and vertical offsets for the full region ($x_{i,j}$ and $y_{i,k}$, respectively), relative to the upper left corner, are

$$x_{0,k} = 0$$

$$x_{1,k} = X_k - 1$$

$$y_{0,k} = 0$$

$$y_{1,k} = Y_k - 1$$

3.2 Bin width computation

The formula for the relative bin widths, Q'_k, used in encoder number two is:

$$Q'_k = qQ_k = \begin{Bmatrix} 1 & k = 0-3 \\ 10/(A_k \log_e(\sigma_k^2)) & k = 4-59 \text{ and } \sigma_k^2 \geq 1.01 \\ 0 & k = 60-63 \text{ or } \sigma_k^2 < 1.01 \end{Bmatrix} \qquad A_k = \begin{cases} 1.32 & k = 52,56 \\ 1.08 & k = 53,58 \\ 1.42 & k = 54,57 \\ 1.08 & k = 55,59 \\ 1.00 & otherwise \end{cases}$$

The proportionality constant, q, controls the absolute bin widths Q_k, and the overall level of compression. Zero bin widths, Z_k, shall be computed in terms of Q_k by the formula $Z_k = 1.2Q_k$.

We now specify the procedure for computing the parameter q that determines the bin widths Q_k as specified above. For the k^{th} DWT subband, let σ_k^2 denote the subband variance estimate, computed according to the above specification. Let m_k be the downsample factor, which is defined to be the ratio of image size to subband size; e.g., $m_{63} = 16$ and $m_4 = 256$. The bit rate to be assigned to the k^{th} DWT subband will be denoted r_k, and r is the targeted overall bit rate, which imposes a constraint on the subband bit rates via the relation

$$r = \sum_k \frac{r_k}{m_k}$$

As explained in Annex A.3, the standard allows the decoder to discard some subbands and transmit a bin width of zero ($Q_k = 0$) to signify that no compressed image data is being transmitted for subband k. For instance, this is always done for $60 \leq k \leq 63$ in encoder number two, and may be done for other subbands as well on an image by image basis if the encoder determines that a certain subband contains so little information that it should be discarded altogether. To keep track of the subband bit allocateion, let K denote the set of all subbands assigned positive bit rates (in particular, for encoder number two, $K \subset \{0, 1, ..., 59\}$). The fraction of DWT coefficients being coded at a positive bit rate will be denoted by S, where

$$S = \sum_{k \in K} \frac{1}{m_k}$$

To relate bit rates to quantizer bin widths, we model the data in each subband as lying in some interval of finite extent, specifically, as being contained within an interval spanning 5 standard deviations. This assumption may not be valid in general, but we will not incur overload distortion due to outliers because outliers are coded using escape sequences in the Huffman coding model. Therefore, for the sake of quantizer design we assume that the data lies in the interval $[\mu_k - \gamma\sigma_k, \ \mu_k + \gamma\sigma_k]$; this implies that