

A Matcher-Independent Approach to Quality Normalization

James Cambier

Cross Match Technologies

3 April 2014

Introduction

This paper is motivated by the need to develop a set of quality metrics that can be calculated from iris images and will quantify specific attributes of iris images and overall quality. Metrics that have been implemented thus far provide indications of sharpness, usable iris area, and pupil dilation. Additional metrics can be added, including contrast, gaze angle, motion blur, etc.

The metrics described here meet the following criteria:

- Sample-based – the range of values for each metric is derived from a population of images captured by one or more iris capture devices; this implies that all such images met some capture criteria and are suitable to some degree for template generation and matching
- Uniform scale – each individual metric is expressed on a scale of 1-100 with 100 representing the highest level of observed quality and 1 representing the lowest level. An “average” image has a score of about 50, and each image has a score that indicates its relative rank within the entire population
- Matcher independence – the quality metrics are not intended to predict the performance of any particular template generation and/or matching algorithm
- Standards compliance – the metrics are consistent with ISO CD 29794-6, a current draft standard titled “Biometric sample quality – Part 6: Iris image”[1], which lists a total of 16 iris image metrics plus a “unified quality score” that seeks to merge the individual metrics into a single overall quality value.

Technical Approach

General Methodology

The analysis described here focused on three specific quality metrics: a **sharpness** measure that indicates how well-focused the image is; **visible iris area**, which measures how much of the iris is not occluded by eyelids, eyelashes, reflections, etc. and is therefore available for encoding and matching; and **iris compression**, which indicates the extent to which the iris is compressed by dilation of the pupil, and is closely related to the pupil-iris diameter ratio. Details on the computation of these metrics are provided below. The three metrics were computed from images captured with three Cross Match Technologies production iris cameras: IScan2, a USB tethered two-eye imager, SEEK2, a handheld multibiometric enrollment and matching device, and SEEK Avenger, a newer version of SEEK that offers enhanced iris capture capabilities. The complete gallery consisted of 3048 images: 1251 from IScan2, 780 from SEEK2, and 1017 from SEEK Avenger. The subject population consisted entirely of company employees, ranging in age from approximately 20 to 65 years.

For each metric a histogram of the observed values, which we will call the raw metric value, was constructed and rescaled such that it summed to 1.0; the result is a sample pdf of the metric. A cumulative distribution function (cdf) was then constructed and the cdf values were multiplied by 100 to generate a normalized quality metric, so that the lowest observed raw metric value would have a normalized value of 1 and the highest observed raw value would have a normalized value of 100. The normalized value can thus be interpreted as the relative rank of the raw metric value within the entire

population. Note that the median raw value within the population would have a normalized quality metric of 50, since half the values would be greater than the median and half less. For all three of the metrics analyzed, the average value of the distribution was within 2 integer values of the median, indicating that the distributions were reasonably symmetrical.

Figure 1 shows a histogram of observed values for the raw sharpness metric, which has a minimum of about 20 and a maximum of 60 for this population. This example is compiled from the entire set of 3048 images. Figure 2 is the CDF for this histogram, scaled to a range of 1 to 100, which constitutes the mapping from raw sharpness values to the normalized sharpness metric.

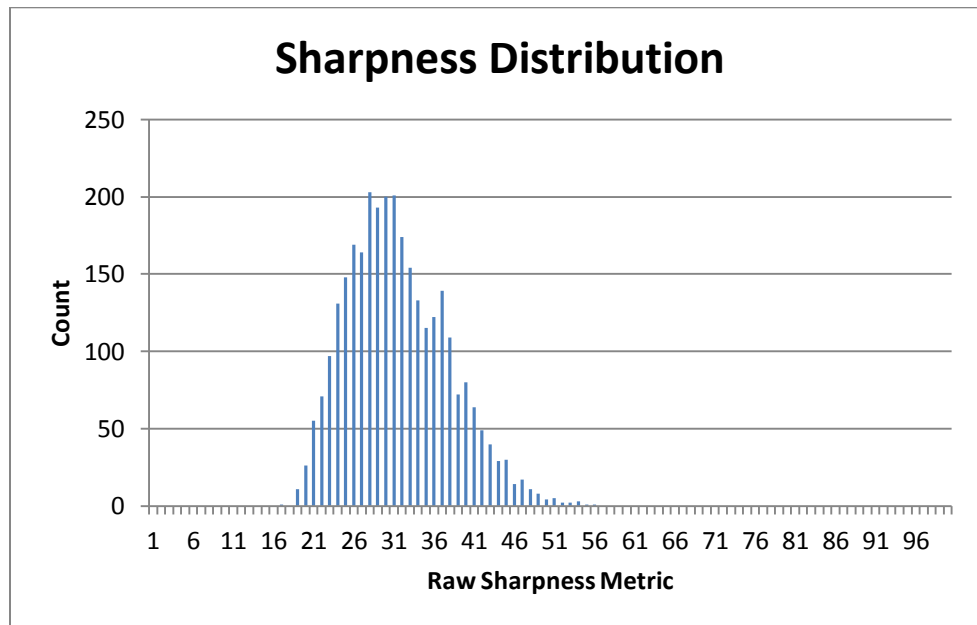


Figure 1 Raw sharpness distribution

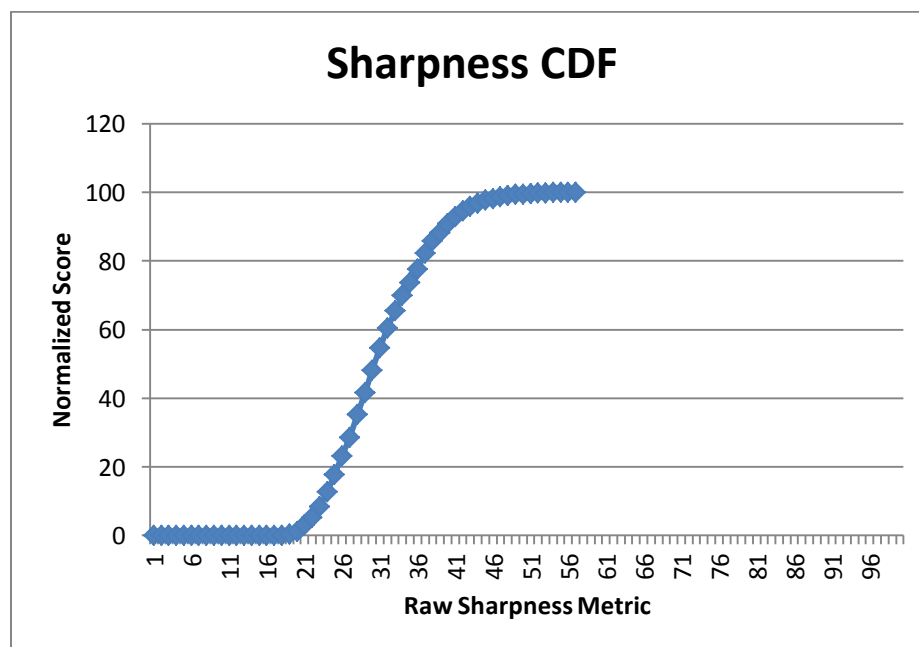


Figure 2 Mapping from raw to normalized sharpness

While the mapping displayed in Figure 2 is based on empirical data, what we need for determining the normalized sharpness value for a given raw sharpness value is a closed-form expression that gives the normalized sharpness as a function of the raw sharpness. We obtain such an expression by fitting a 3rd order polynomial to the points displayed in Figure 2. Details on this process are given below.

Note that since the quality mapping is based on empirical data from one image gallery, it is entirely possible that images from other cameras and other capture conditions might have raw sharpness values outside the range observed from the Cross Match images. One approach to handling such samples is to simply assign a normalized sharpness of 100 if the raw value is greater than 60 or 1 if it is less than 20. These out-of-range samples could also be added to the original data set so that from time to time the distribution and polynomial fit can be recalculated to provide updated parameters for calculating the normalized metric. These parameters would be incorporated into periodic software releases when convenient.

Sharpness

The sharpness measure used is calculated by computing the discrete cosine transform (DCT) of a series of non-overlapping 8 x 8 pixel patches, computing a sharpness value for each patch, and averaging the sharpness values over all the patches. The metric is calculated for a local subimage that encompasses the iris and extends slightly beyond the iris into the surrounding area. The sharpness measure, described by Kristan *et al* [2] is obtained by computing a Bayes entropy function for the normalized DCT spectrum; the entropy function yields a maximum when the spectrum tends to a uniform distribution, which happens when the image reaches best focus and its high-frequency content is maximized. When the image is defocused its energy becomes concentrated in the low frequencies and its entropy drops.

As described earlier, we obtained a distribution of the sharpness values for the test image population, computed its cumulative distribution function, and rescaled the CDF values to the 0-100 range to obtain the normalized quality metric. A 3rd-order polynomial fit was applied to the normalized values using Excel. The raw- to normalized- quality mapping is shown in Figure 3 along with values calculated using the best-fit polynomial. Although the calculated values exceed the 0-100 range at the extremes of the raw metric's range, this can be easily handled in software by constraining the normalized quality value to 0-100. The polynomial for mapping raw sharpness to normalized quality metric is

$$Q_s = S_3x^3 + S_2x^2 + S_1x + S_0 \quad (1)$$

where

$$S_i = \{-2.0778, -4.78202, 0.325824, -0.00375\} \quad (2)$$

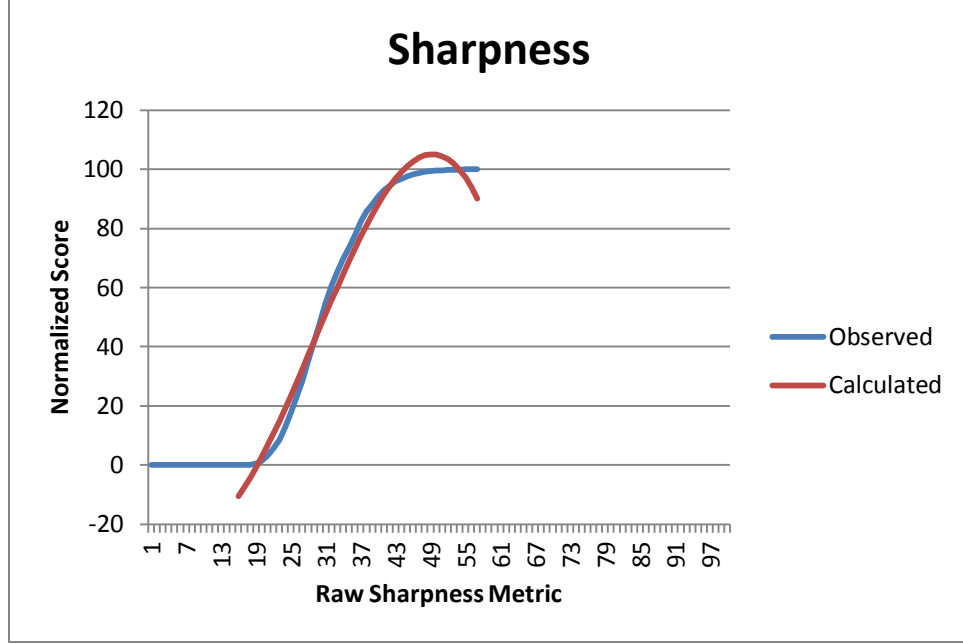


Figure 3 Sharpness quality normalization

Visible Iris

The visible iris area is calculated after image segmentation, which locates the iris and pupil boundaries, including eyelids, and occluded areas due to eyelashes, reflections, etc. The raw quality metric is $100 \times \text{visible iris} / \text{total iris}$ where the total iris area is the area within the circular iris boundary less the area within the (approximately) circular pupil boundary. In the test image population, the raw visible iris ranged from about 50% to 100% of the iris area. This range is mapped to the 0-100 normalized range using the empirical mapping and polynomial fit shown in Figure 4. The normalized visible iris area metric Q_A is given by

$$Q_A = A_3 x^3 + A_2 x^2 + A_1 x + A_0 \quad (3)$$

where x is the raw quality metric and the coefficients are

$$A_i = \{4.174934, 1.357796, -0.05514, -0.000524\} \quad (4)$$

As is the case with the normalized sharpness metric, the normalized visible iris area should be constrained to the 1-100 range.

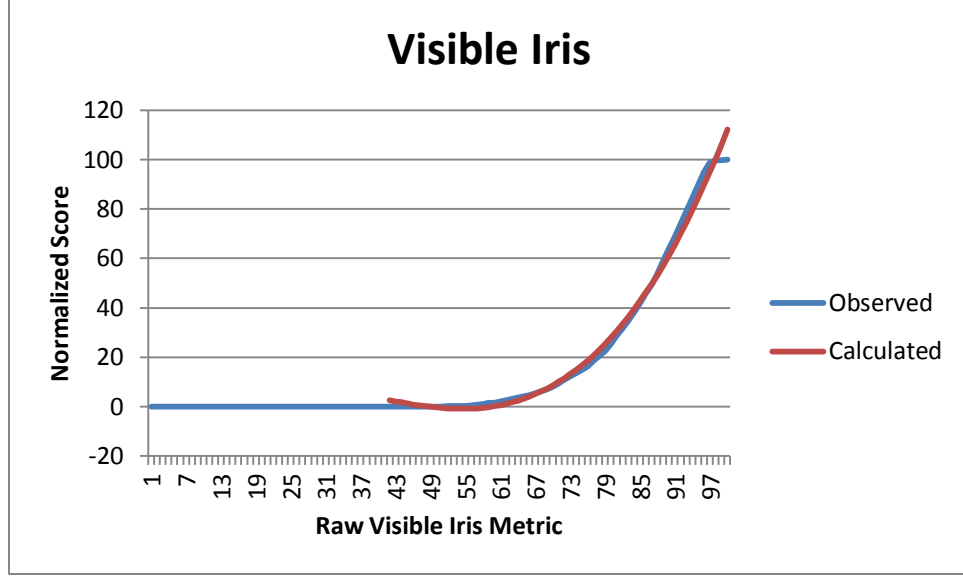


Figure 4 Visible iris quality normalization

Iris Compression

Iris compression (IC) is related to Pupil-Iris-Ratio (PIR), the ratio of pupil diameter to iris diameter, by

$$IC = 100 \cdot \left(1 - \frac{D_{pupil}}{D_{iris}} \right) \quad (5)$$

The use of iris compression as the raw quality metric is motivated by the definition of “Dilation Constancy” (DC) in Section 6.4.2.3 of ISO/IEC JTC1/SC37 DIS 29794-6[1]. This parameter is defined as

$$DC = 100 \cdot \frac{100 - \max(D_1, D_2)}{100 - \min(D_1, D_2)} \quad (6)$$

where D_1 and D_2 are the pupil-iris ratios (on a scale of 0-100) of the first and second iris images used in a matching operation. For most iris matchers, the best match performance is achieved if both images have similar pupil-iris ratios, so DC approaches 100. We have defined iris compression so that DC becomes

$$DC = 100 \cdot \frac{\min(IC_1, IC_2)}{\max(IC_1, IC_2)} \quad (7)$$

where IC_1 and IC_2 are just the iris compression values of the first and second images. Dilation Constancy is maximized if IC_1 and IC_2 are identical.

Of course, when we calculate a normalized quality value for iris compression for a single image we only have one pupil-iris ratio available to us. In order to handle this situation we note that the pupil-iris ratio observed in the test image set ranged from 14 to 55, with an average 28.2, which corresponds to an iris compression value of 71.8. Reasoning that the farther our sample’s iris compression is from the average

value, the more likely it is to result in a low value of Dilation Constancy, we calculated for our dilation metric a modified compression metric (MCM) as

$$MCM(IC) = 100 \cdot \left(1 - \frac{|IC - 72|}{72} \right) \quad (8)$$

where IC is the raw iris compression value.

MCM will be 100 if the IC is equal to 72 and will drop as it becomes larger or smaller than 72. The minimum and maximum observed values for IC in the population were 45 and 86, respectively, so MCM ranges from 62.5 to 100. Figure 5 shows the mapping from MCM to its normalized quality value. As for visible iris area, the mapping for MCM is a 3rd order polynomial

$$Q_M = M_3x^3 + M_2x^2 + M_1x + M_0 \quad (9)$$

where

$$M_i = \{7.499210, 7.492761, -0.210680, 0.001452\} \quad (10)$$

As is the case with the normalized sharpness metric, the normalized dilation metric should be constrained to the 1-100 range.

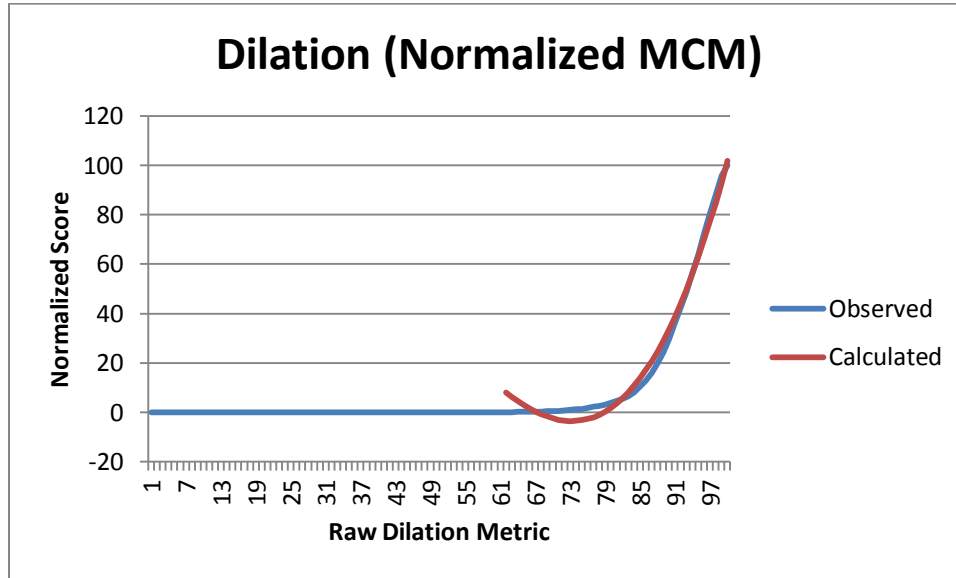


Figure 5 Dilation Normalization

It is interesting to compare our dilation metric to the Dilation Constancy described in the draft iris quality standard, which we can do if we assume that we are analyzing a probe image that is being matched against an enrollment image having the average pupil-iris ratio of 0.28 discussed above. Figure 6 is a plot of the normalized dilation metric and the Dilation Constancy as a function of the P-I ratio of the probe image. Note that the two plots are identical for P-I ratios above the average and nearly identical below the average P-I ratio except for very low P-I ratios.

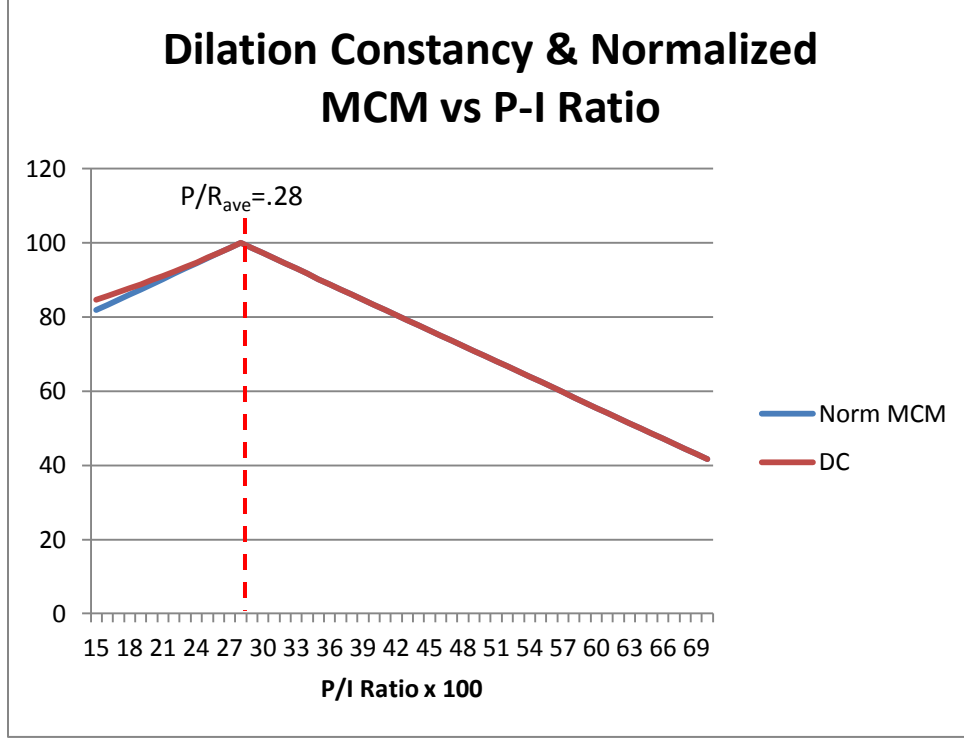


Figure 6 Normalized dilation metric and dilation constancy vs P-I ratio

Scalar Quality Metric

The final step is to combine the multiple quality values into a single overall scalar value. There are a number of ways to do this. Belcher and Du [3] simply multiply the normalized quality metrics together so that the overall quality will always be less than or equal to that of the lowest individual metric. If all the metrics are equal but less than one, their product will result in a value that may be considerably less than the individual metrics, whereas it seems intuitive that if the metrics are equal the resulting scalar value should be equal to the individual values.

Kalka *et al* [4] uses a Dempster-Shafer theory approach to evidential reasoning [5]. In this approach each individual quality metric is considered sequentially, and its value adds to the overall evidence that the image is good or bad. This approach is more complex, but probably merits further study.

A third approach is to compute the geometric mean of the normalized quality values, which indicates the typical value of a set of numbers. For normalized quality metrics q_1, q_2, \dots, q_n , the overall quality would be calculated as

$$Q = \sqrt[n]{\prod_{i=1}^n q_i} \quad (11)$$

where $\{q_i\}$ are the set of quality metrics. It can also be calculated using logarithms:

$$Q = \log^{-1} \frac{1}{n} \sum_{i=1}^n \log(q_i) \quad (12)$$

Finally, one could adopt a maximum likelihood approach, in which we would generate distributions of each quality metric for “good” and “bad” images, fit continuous functions to these distributions, and scale them to represent pdf’s for each metric under the hypothesis of “good” or “bad”. This would provide only a binary outcome whereas our intent is to provide a continuous quality metric.

We have used the geometric mean approach. Note that by constraining the minimum normalized quality value to 1 rather than 0, we assure that a minimal value in one metric does not cause the overall scalar quality to go to zero.

Test and Evaluation

Although the intent of this development was to create a method of normalization that does not seek to predict the performance of particular iris matcher, we performed some limited comparisons of the normalized metrics with match scores generated by a commonly used iris encoding and matching algorithm, using binary phase encoding of Gabor wavelet filters as described by Daugman[6]. We also elected to use a different image gallery, the ICE 2006 image set collected by Notre Dame University using an LG2200 imager, to see how the quality metrics compared to those generated by the Cross Match imagers.

Figure 7 shows the observed ranges ($\mu \pm \sigma$) of the three normalized metrics for the four imagers. As expected the means for the three cameras used to generate the training data cluster around a score of 50. The LG2200 data shows an interesting departure in that the dilation metric is quite low; we believe this is due to the fact that virtually all the test subjects were college-age students, and studies have found that younger people tend to have larger pupils[7].

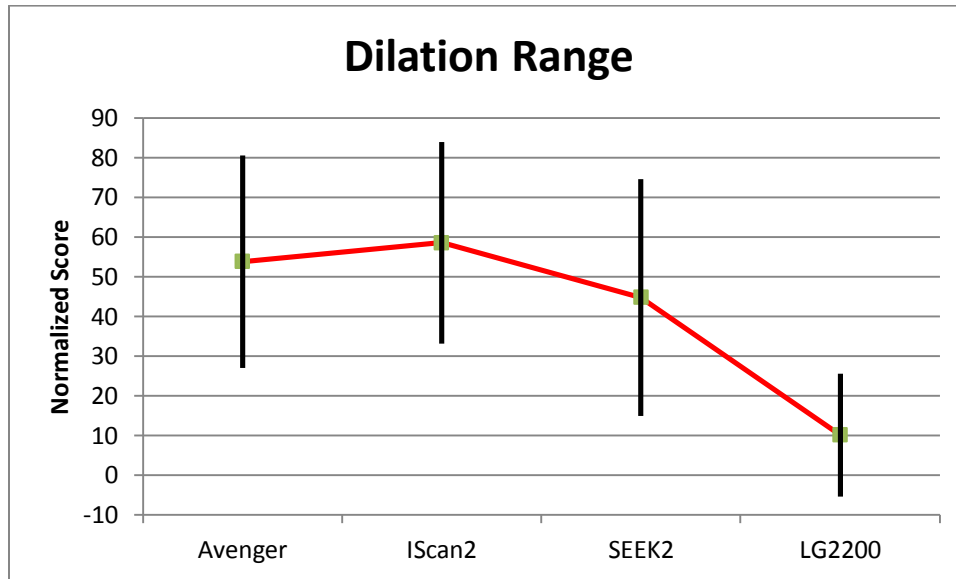


Figure 7a Normalized dilation range

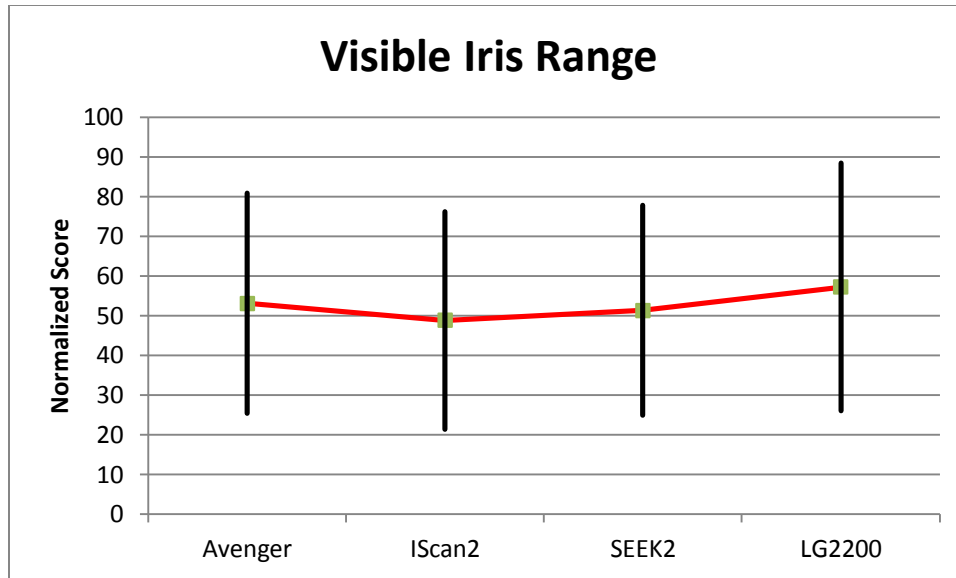


Figure 7b Normalized visible iris range

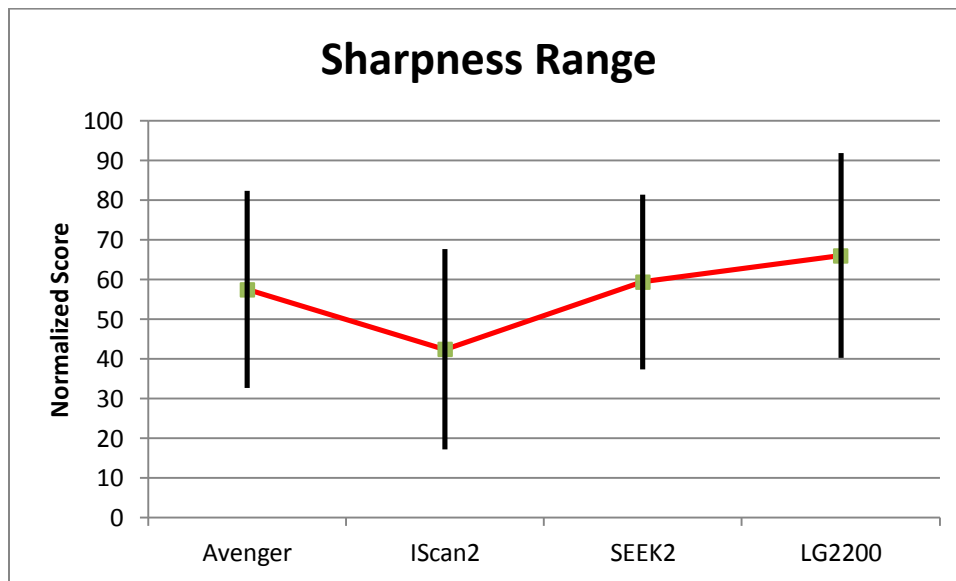


Figure 7c Normalized sharpness range

In order to evaluate the correlation of the normalized quality metrics with match performance we performed a total of 26743 genuine matches using 2946 images from the NIST ICE gallery. We recorded the match scores plus values of P-I ratio and the three normalized metrics for both images corresponding to each match. We then generated scatter plots of match scores vs various quality metrics and plotted linear regression curves

Figure 8 shows plots for two dilation metrics. Figure 8a is a plot of match score vs. dilation constancy computed using equation (7) above. In Figure 8b we have plotted match score vs. the minimum normalized dilation metric for the two images. Both show a clear trend toward higher scores with higher values of dilation constancy and minimum dilation.

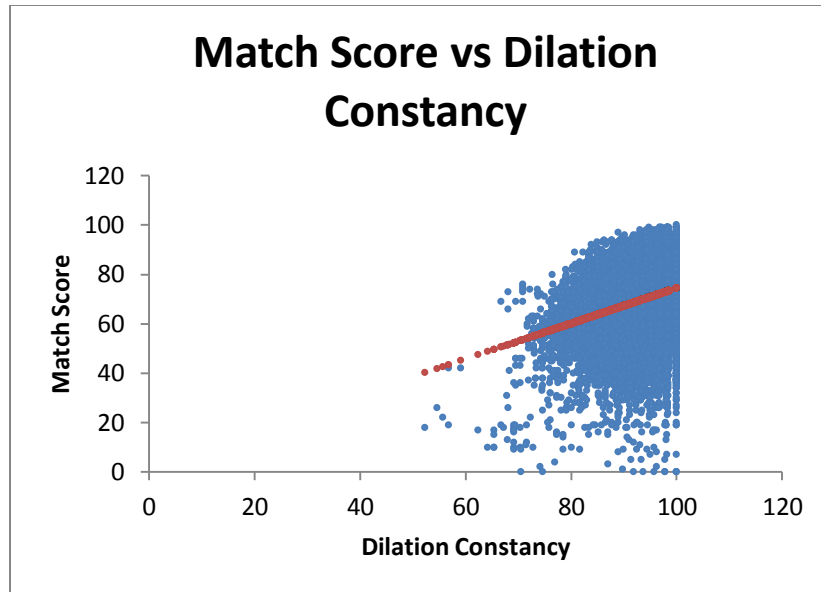


Figure 8a Match score vs dilation constancy

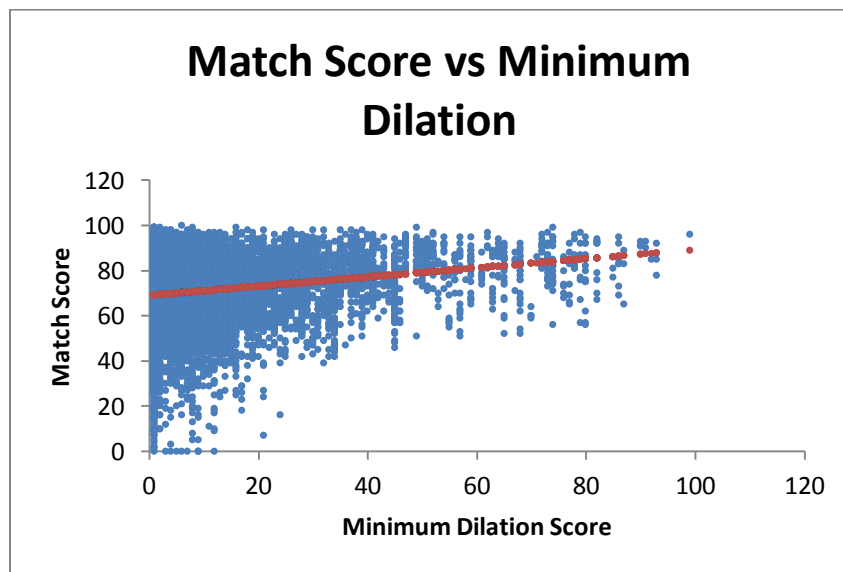


Figure 8b Match score vs minimum dilation metric

Figure 9 is a plot of match score vs the minimum normalized visible iris metric. We reasoned that the minimum of the two visible iris values would provide the best indicator of the iris area available for matching. The regression line indicates a clear trend toward higher scores for higher values of the minimum visible iris metric.

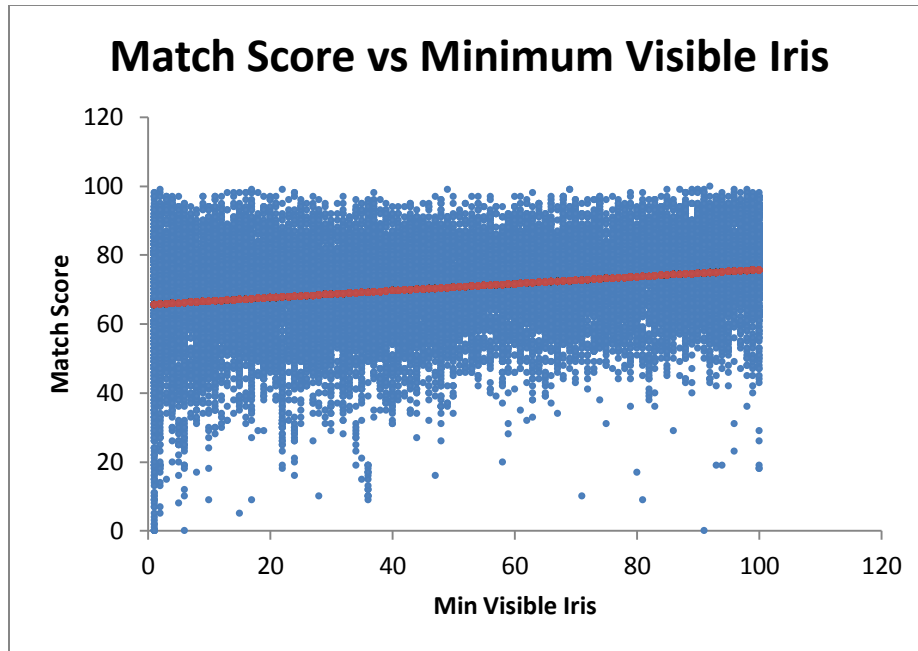


Figure 9 Match score vs minimum visible iris metric

In considering sharpness we reasoned that higher match scores should occur when the two images are most similar in appearance, i.e. their sharpness scores are close to each other, since this would assure that the same texture features are being encoded in both irises. To measure this we computed a “sharpness constancy” as $100 * \min(\text{sharpness}) / \max(\text{sharpness})$ and plotted the match score against this value. As shown in Figure 10, there is little if any correlation between the match score and this sharpness constancy value. Our conclusion is that the version of wavelet encoding we are using must be remarkably robust to variations in sharpness, and even to differences in sharpness in the two images being compared.

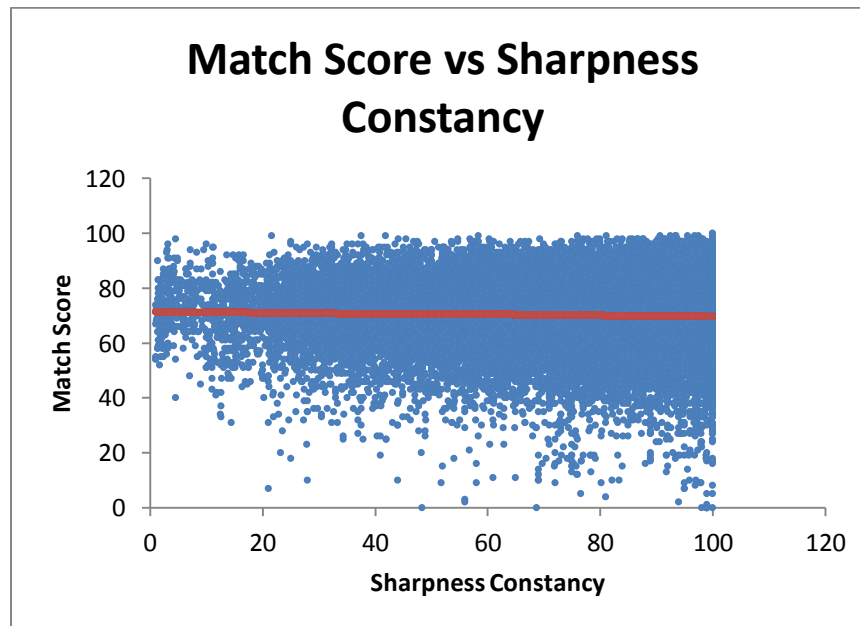


Figure 10 Match score vs sharpness constancy

References

1. Text of DIS 29794-6, Biometric sample quality – Part 6: Iris image data, ISO/IEC JTC 1/SC37 “Biometrics”, Document N5554, 2013.
2. Kristan, M. *et al* , “A Bayes-spectral-entropy-based measure of camera focus using a discrete cosine transform,” Pattern Recognition Letters 27(13):1431-1439, 2006.
3. Belcher, C. and Du, Y. “A selective feature information approach for iris image-quality measure,” IEEE Trans. Information Forensics and Security 3(3):572-577, 2008.
4. Kalka, N., Zuo, J., Schmid, N., and Cukic, B., “Estimating and fusing quality factors for iris biometric images,” IEEE Trans. Syst., Man, Cybern., 40(3):509-524, 2010.
5. Murphy, R., “Dempster-Shafer theory for sensor fusion in autonomous mobile robots,” IEEE Trans. Robot. Auto. 14(2):197:206, 1998.
6. Daugman, J., “High confidence visual recognition of persons by a test of statistical independence,” IEEE Trans. Pattern Anal. Mach. Intell., 15(11):1148-1161, 1993.
7. Bradley, J.C. *et al*, “Dark-adapted pupil diameter as a function of age measured with the NeuroOptics pupillometer,” J. Refract. Surg. 27(3):202-207, 2011.