

Impact of Age Groups on Fingerprint Recognition Performance

Shimon K. Modi, Prof. Stephen J. Elliott, Ph.D., Jeff Whetsone
Industrial Technology, College of Technology
Purdue University
West Lafayette, U.S.A.
{shimon,Elliott,jwhesto}@purdue.edu

Prof. Hakil Kim, Ph.D.
INHA University
Incheon, Korea
hikim@inha.ac.kr

Abstract—Ever since introduction of automated fingerprint recognition in law enforcement in the 1970s it has been utilized in applications ranging from personal authentication to civilian border control. The increasing use of automated fingerprint recognition puts on it a challenge of processing a diverse range of fingerprints. The quality control module is important to this process because it supports consistent fingerprint detail extraction which helps in identification / verification. Inherent feature issues, such as poor ridge flow, and interaction issues, such as inconsistent finger placement, have an impact on captured fingerprint quality, which eventually affects overall system performance. Aging results in loss of collagen; compared to younger skin, aging skin is loose and dry. Decreased skin firmness directly affects the quality of fingerprints acquired by sensors. Medical conditions such as arthritis may affect the user's ability to interact with the sensor, further reducing fingerprint quality. Because quality of fingerprints varies according to the user population's ages and fingerprint quality has an impact on overall system performance, it is important to understand the significance of fingerprint samples from different age groups. This research examines the effects of fingerprints from different age groups on quality levels, minutiae count, and performance of a minutiae-based matcher. The results show a difference in fingerprint image quality across age groups, most pronounced in the 62-and-older age group, confirming the work of [7].

Keywords—*fingerprint quality; fingerprint performance; impact of age; aging effect; quality assessment*

I. INTRODUCTION

Assessment of biometric sample quality has captured considerable interest and triggered many research efforts. Research on fingerprint recognition, as the oldest scientifically recognized biometric modality, has typically focused on assessment of image quality. Differing methodologies exist; [2] use orientation certainty and strength of dominant frequency to calculate fingerprint image quality. The fingerprint image is divided into sub-blocks and image quality is computed for each sub-block; these individual computations are then used to calculate an overall image quality score ([2]). [3] describe a quality assessment scheme that first counts the foreground blocks in an image, and then identifies the dominant direction of those blocks. Blocks close to the

foreground centroid are given more weight. A ratio of the weighted sum of dominant-direction blocks and the weighted sum of foreground blocks is used to compute image quality ([3]). [4] perform image quality assessment using the fingerprint's global structure: a 2D Discrete Fourier Transform is calculated, and the measure of energy concentration in regions of interest is used as a determinant of quality; higher energy concentrations yield better image quality. Tabassi and Wilson describe an approach to classifying image quality assessment whereby the quality of fingerprint features used for matching operations is computed and defines the degree of separation between match and non-match distributions ([6]). In their work, a neural network is trained to map this degree of separation according to levels of quality, thus making fingerprint image quality an indicator of matcher performance. The impact of image quality degradation on performance of fingerprint matching systems shows that ridge-based matchers outperform minutiae-based matchers on lesser-quality images ([1]). This also holds true when examining different populations; in one such case, fingerprint image quality for two groups (18-25, ≥ 62) are significantly different ([7]). Further, there was a negative impact on fingerprint algorithm performance on these two age groups ([5]). The goal of this paper is to expand the works of [5] and [7], to evaluate the impact of age on image quality and performance by examining four age groups (18-25, 26-39, 40-62, and ≥ 62).

II. BACKGROUND WORK

Individuals aged 18 and older participated in the study in one of four distinct age groups: 18-25 years, 26-39 years, 40-62 years, and 62 years and above. Data collection occurred in two phases, the first of which collected fingerprints from the 18-25 and 62+ populations ([7]), and another data collection period in 2006, which augmented the database. In the latter data collection effort, fingerprints were collected regardless of age. These age groups were selected so that fingerprints could be obtained from at least 30 different fingers for each age group and divided among the range between 26 and 62+ years as closely as possible to the midpoint. In both phases, fingerprints were collected using an DigitalPersona® U.are.U® 4000 optical sensor; the resulting database

contained three fingerprint images from both the right and the left index fingers of each participant (total of six fingerprints per subject) Table I and Figure 1 summarize the contents of this study's database.

TABLE I. DATASET SUMMARY

Age Grp.	Num of subjects	Fingers per Subject	Total Distinct Fingers	Sam- ples per Finger	Total Samples Per Age Group
18-25	79	2 (right index, left index)	158	3	948
26-39	24	2	48	3	288
40-62	26	2	52	3	312
62+	60	2	120	3	720

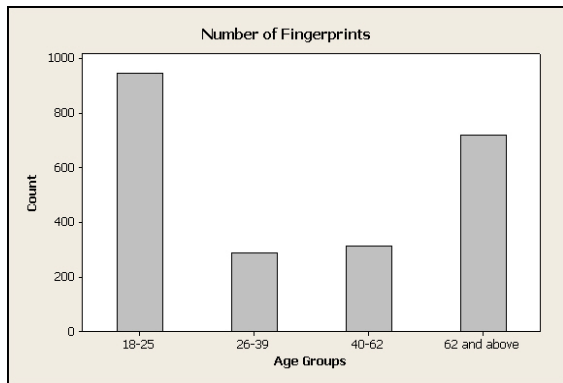


Figure 1. Fingerprint count per age group

A commercially available fingerprint image quality assessment tool extracted the quality scores and minutiae count for fingerprints from the four different age groups. This study's hypotheses were to establish whether the population means of quality scores for the four different age groups were equal. Subsequently, the impact of low-quality image scores was calculated using the following protocol:

1. Combine all the fingerprints into one data set and obtain a Detection Error Tradeoff (DET) curve for the data set.
2. Remove the lowest 10 percent quality images from the 18-25 age group in the master data set and obtain a DET curve for the modified data set.
3. Remove the lowest 10 percent quality images from the 26-39 age group in the master data set and obtain a DET curve for the modified data set.
4. Remove the lowest 10 percent quality images from the 40-62 age group in the master data set and obtain a DET curve for the modified data set.
5. Remove the lowest 10 percent quality images from the 62+ age group in the master data set and obtain a DET curve for the modified data set.

III. RESULTS AND ANALYSIS

Figure 2 illustrates the distribution of quality scores, as calculated by the image quality algorithm. The spread of quality scores for the 18-25 age group was more compact than that compared to the scores of the 62+ age group. Largely, the scores of the middle groups (26-39 and 40-62 age groups) overlap.

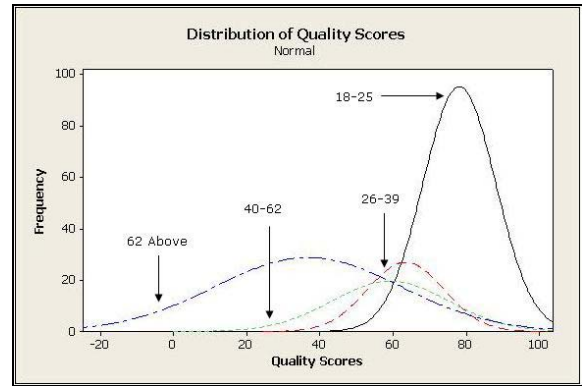


Figure 2. Distribution of quality scores

Figure 3 shows representative samples of high and low quality for each of the four groups, as determined by the image quality tool.



Figure 3. Representative high quality (top row) and low quality (bottom row) fingerprints

Preliminary visual analysis (Figure 4) illustrates a clear variation in quality scores for fingerprints among the four different age groups. The data set was determined to be non-parametric; therefore, a Kruskal Wallis test was performed to examine whether population means of quality scores for the four different age groups are equal. The following hypotheses were determined:

$$\text{Null Hypothesis (H}_0\text{)} : \mu_{18-25} = \mu_{26-39} = \mu_{40-62} = \mu_{62\text{above}} \quad (1)$$

$$\text{Alternate Hypothesis (H}_a\text{)} : \text{Not all } \mu \text{ are equal.} \quad (2)$$

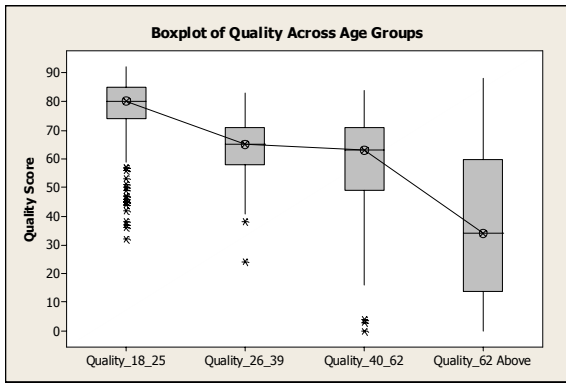


Figure 4. Box-plot of quality scores

The Kruskal Wallis test is a non-parametric test when assumptions of data normality do not apply ([8]). The p -value indicates the probability of observing a test statistic as extreme as the one observed from the test if the null hypothesis is true. A simple interpretation of the p -value is that, if it is greater than the significance level, then the null hypothesis is accepted, else the alternate hypothesis is concluded. The results, shown in Table II, led the researchers to conclude that the quality score means for all age groups are not the same, using a significance level (α) of .05. Table II shows the results of the test.

TABLE II. KRUSKAL-WALLIS TEST RESULTS FOR QUALITY SCORES

Age Groups	18-25	26-39	40-62	62+
Median	37	47	46	40
p -value	< 0.0001			

The Kruskal Wallis test examines whether all the groups being compared are similar, but will not test the difference between two particular groups. To do this, a multiple paired test for equality between each pair of age groups was performed. Since there were four age groups, there were six possible paired comparisons. Using the Bonferroni adjustment procedure, the new significance level was set to be at .0083 for these six paired comparisons to ensure that the original significance level of .05 is maintained ([8]). Table III shows that the p -value from the Mann Whitney test for comparisons of quality scores between age groups 26-39 and 40-62 was not significantly different. The remainder of the comparisons shows statistically significant differences in quality scores.

TABLE III. PAIRWISE COMPARISONS FOR QUALITY SCORES

Age Groups	26-39	40-62	62+
18-25	$p < 0.001$	$p < 0.001$	$p < 0.001$
26-39		$p = 0.08$	$p < 0.001$
40-62			$p < 0.001$

While a difference in image quality across the groups was observed, further analysis was undertaken to establish whether there was also a difference in minutiae count for fingerprints collected from the four different age groups. The Kruskal

Wallis test was used on the four different age groups, with the following hypotheses:

Null Hypothesis (H_0) :

$$\mu_{18-25} = \mu_{26-39} = \mu_{40-62} = \mu_{62+above} \quad (3)$$

Alternate Hypothesis (H_a) :

$$\text{Not all } \mu \text{ are equal.} \quad (4)$$

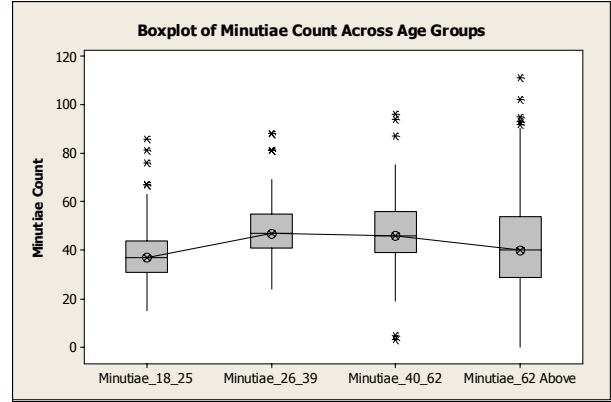


Figure 5. Box-plot of minutiae count

The Kruskal Wallis test with a significance level of .05 examined the difference in minutiae across the groups. Figure 5 shows the resulting box plot of the minutiae across the four age groups, while the detailed results of this Kruskal Wallis test are identified in Table IV.

TABLE IV. KRUSKAL-WALLIS TEST RESULTS FOR MINUTIAE COUNT

Age Groups	18-25	26-39	40-62	62+
Median	37	47	46	40
p -value	< 0.0001			

TABLE V. PAIRWISE COMPARISONS FOR MINUTIAE COUNT

Age Groups	26-39	40-62	62+
18-25	$p < 0.001$	$p < 0.001$	$p < 0.001$
26-39		$p = 0.382$	$p < 0.001$
40-62			$p < 0.001$

The multiple paired comparison procedure was used to compare the magnitude of differences between every possible pair of age groups. Using the Bonferroni adjustment procedure, the new significance level is set at .0083 for these six paired comparisons to ensure that the original significance level of .05 is maintained. Table V shows that the p value from the Mann Whitney test for comparisons of minutiae count between age groups 26-39 and 40-62 was not significantly different and the rest of the comparisons show a statistically significant difference in minutiae count. As the comparison of quality scores for these age groups does not show a statistically significant difference, it follows that minutiae count should not differ either.

A scatter plot of minutiae count against quality scores for all the age groups combined shows a curvilinear relationship

between the two, as seen in Figure 6. This should not be mistaken for a causal relationship; this only indicates the direction of relationship between the two. The graph indicates that fingerprints with a minutiae count at the extreme ends of the range are assigned a lower quality score, while fingerprints with minutiae count in the mid-range were consistently given a higher quality score. Fingerprint samples with a minutiae count between 40 and 60 are assigned low quality scores, indicating that there are other factors taken into account when determining quality scores.

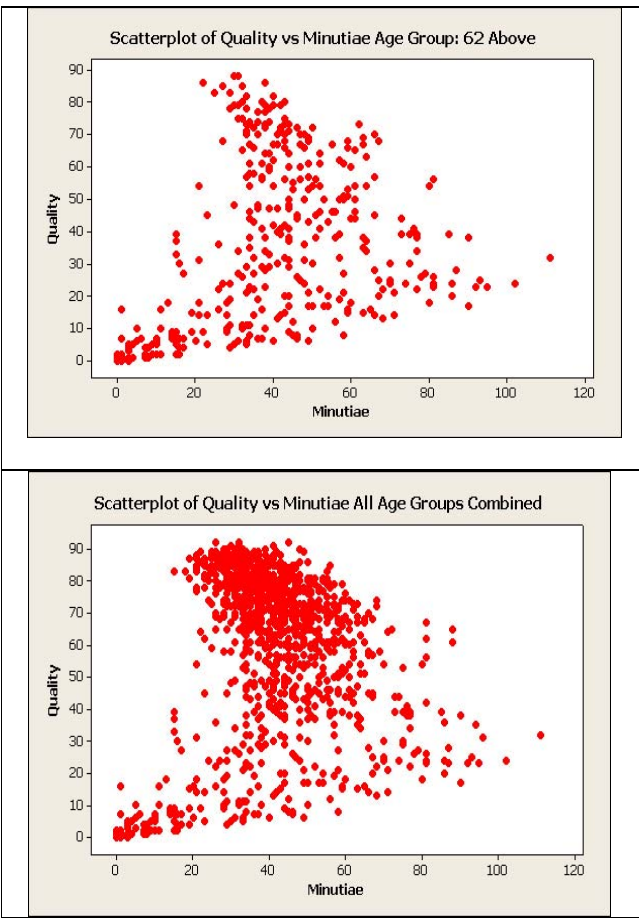
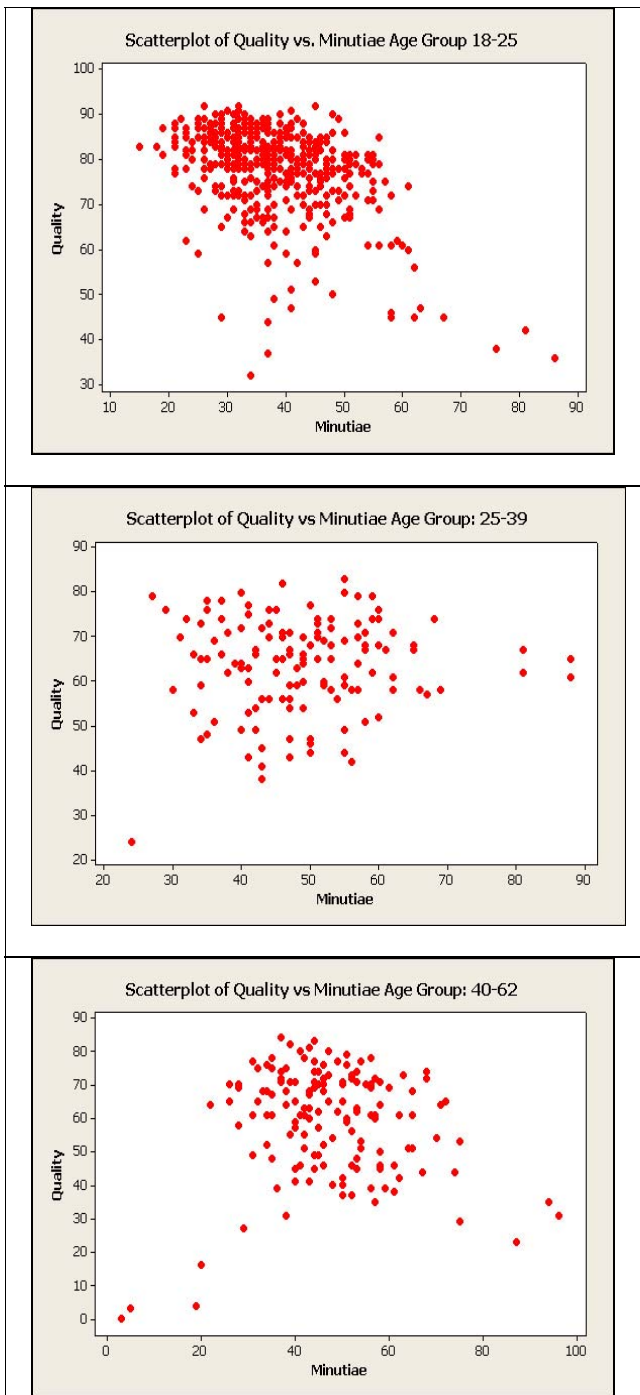


Figure 6. Scatter-plot of quality vs. minutiae count

Following the determination that there was a statistical difference in both image quality and minutiae count for fingerprints from the different age groups, the next step was to study the contribution of low quality images to the performance of the minutiae-based fingerprint matcher. According to the methodology outlined in the previous section, five DET curves were calculated, as shown in Figure 7. The most significant shift in the DET curve was observed when the lowest 10 percent quality images from the 62+ age group were pruned from the data set. Removing 10 percent of the lowest quality images from the other age groups did not show a significant shift in the DET curves. DET curves for all the other age groups are clustered along the same path showing a very similar change in performance on removing the lowest 10 percent quality images. The greatest impact on performance can be attributed to the 62+ age group.

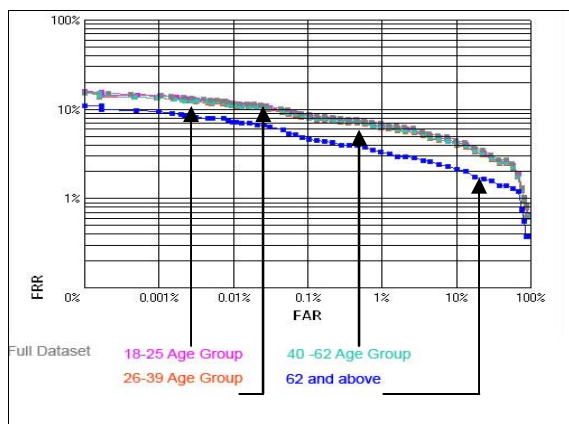


Figure 7. Full dataset DET curves and effect of pruning lowest 10% quality images from each dataset

IV. CONCLUSIONS AND FUTURE WORK

The results confirm a difference in fingerprint image quality across age groups, although it is most pronounced with the 62+ age group. This confirms the work done previously by [7] with regard to image quality. Furthermore, the performance of the groups is also different, as shown by the DET curves. What we have learned from the statistical results is that fingerprint image quality is not similar between age groups because the quality score were not within a reasonable tolerance to be similar. This work also points to the importance of automated quality control of fingerprints: the results clearly show an increase in error rates for fingerprints of older individuals. Error rates are amplified when a fingerprint recognition system is deployed for large-scale use by subjects from different age groups. There is a clear need to understand issues with fingerprints representing different age groups, because the pervasiveness of this technology in the near future will require it to handle fingerprint images from different age groups and differing levels of quality. Based on Figure 2 and Figure 4, the overall quality score decreases with increasing variances as the age increases, which can be easily presumed. The average minutiae counts, however, are within a similar range for all the age groups, with only the variance increasing with age groups. A visual analysis of outliers in the youngest age group showed temporary skin changes caused their minutiae count to be significantly different than the rest of the age group. Meanwhile, for the most senior age group,

the minutiae count contains a significant portion of false minutiae and the ratio of fingerprints with low quality is not negligible. The effects of these can be seen in error rates shown in Figure 7. Figure 7 also indicates that, for the most senior group having a significant portion of false minutiae and low-quality fingerprints, the lower performance is related to the quality score. For the rest of the age groups, however, quality is not the only contributing factor leading to a change in performance rates. This indicates a need to consider other factors that could cause a difference in performance rates. Further research is needed in order to investigate factors other than quality that might affect the performance of fingerprint recognition systems. Emphasis on ridge structure for matching operations is another research topic warranting further study on the effects of aging and fingerprint recognition. The results obtained from this research indicate a need for a framework for data modeling of different age groups in order to improve performance rates of fingerprint matching systems.

REFERENCES

- [1] J. Aguilar-Fuierrez, L. Munoz-Serrano, F. Alonso-Fernandez, and J. Ortega-Garcia, "On the effects of image quality degradation on minutiae- and ridge-based automatic fingerprint recognition," paper presented at the 39th Annual International Carnahan Conference on Security Technology, Albuquerque, NM, October 12-14, 2005.
- [2] T. P. Chen, X. Jiang, and W. Y. Yau, "Fingerprint image quality analysis," paper presented at the International Conference on Image Processing, Singapore, October 24-27, 2004.
- [3] N. Haas, S. Pankanti, and M. Yao, *Fingerprint Quality Assessment*. NY, NY: Springer-Verlag, 2004, pp. 55-66.
- [4] A. Jain, Y. Chen, and S. Dass, "Fingerprint quality indices for predicting authentication performance," paper presented at the 5th International Conference on Audio- and Video-Based Biometric Person Authentication, Rye Brook, NY, July 20-22, 2005.
- [5] S. K. Modi, and S. J. Elliott, "Impact of image quality on performance: Comparison of young and elderly fingerprints," paper presented at the 6th International Conference on Recent Advances in Soft Computing (RASC), Canterbury, UK, July 10-12, 2006.
- [6] E. Tabassi, and C. L. Wilson, "A novel approach to fingerprint image quality," paper presented at the International Conference on Image Processing, Genoa, Italy, September 11-14, 2005.
- [7] S. J. Elliott, and N. C. Sickler, "An evaluation of fingerprint image quality across an elderly population vis-a-vis an 18-25-year-old population," paper presented at the International Carnahan Conference on Security Technology, Las Palmas, Gran Canaria, October 12-14, 2005.
- [8] J. Neter, M. Kutner, C. Nachtsheim, and W. Wasserman *Applied Linear Statistical Models* (4th ed.). Chicago: Irwin, 1996, pp.1080.