Texture Modeling for Synthetic Fingerprint Generation

Peter Johnson
Clarkson University
Potsdam, NY, USA
johnsopa@clarkson.edu

Fang Hua
Clarkson University
Potsdam, NY, USA
huaf@clarkson.edu

Stephanie Schuckers
Clarkson University
Potsdam, NY, USA
sschucke@clarkson.edu

Abstract

The development of biometric recognition technologies often requires large sets of biometric data for training and evaluation purposes. The use of synthetically generated biometric samples has been explored as a means of avoiding the challenges of large scale data collection. Our paper builds on previous work in synthetic fingerprint generation research through the modeling and synthesis of texture characteristics for synthetic fingerprint generation. The proposed texture characterizing features can be modeled from real fingerprint images to generate synthetic fingerprint texture statistically representative of a particular real fingerprint database. The texture characterizing features include ridge intensity along the ridge center-lines with seven frequency components, ridge width, ridge cross-sectional slope, ridge noise, and valley noise. A comparison of these feature densities from real and synthetic fingerprints is shown, which demonstrates the effectiveness of this method of modeling and generating synthetic fingerprint textures.

1. Introduction

Fingerprint ridge patterns are one of the most widely utilized biometric indicators of identity. Though fingerprint recognition systems have displayed low error rates in matching performance, this technology is far from reaching the theoretical limits of fingerprint individuality [1]. Furthermore, some applications require 1 to N matching, where N may be upwards on the order of 10^9. As new matching algorithms are introduced and refined over time, large numbers of biometric samples must be collected, firstly, for training purposes and secondly, for performance evaluation. As fingerprint databases grow quite large, the scalability of fingerprint matching systems becomes an increasing concern.

The concept of using synthetically generated fingerprints to create large databases has been explored. Synthetic biometrics is defined as artificially generated biometric data, which exhibits meaningful biological characteristics as measured by an existing biometric system [2]. Typically, the generated biometric data represents unique virtual identities. These generated characteristics of an artificial fingerprint are typically defined by controlling parameters set to imitate the biological characteristics of real fingerprints. Introducing a certain amount of variation in the setting of these parameters allows for the generation of large numbers of unique fingerprint images. With the ability to generate large numbers of realistic synthetic fingerprint images, the time consuming and expensive process of collecting data from live subjects can be reduced. However, with a switch from real fingerprints to synthetic, there also comes a loss of empirical accuracy, which motivates further development in and testing of synthetic generation techniques.

Synthetic biometrics also allow for the protection of personal identity [3]. A synthetic biometric sample in an evaluation database is not associated with any real person's identity. Thus, synthetic biometric databases could be distributed among research organizations with less regulation.

Through modeling, one tries to represent aspects of the real world, but the models are only as good as what can be observed. Yanushkevich [2] defined the crucial point of biometric modeling as the analysis-by-synthesis paradigm, where analysis incorporates feature extraction and modeling. The model for analysis produces features from the biometric sample and the model for synthesis reproduces the biometric sample from the extracted features. The diagram in Figure 1 demonstrates the analysis-by-synthesis process for fingerprint modeling and generation, where features are extracted from real fingerprint images and modeled according to their densities. These densities are sampled to generate unique features, which are mapped to synthetic fingerprint images. After extracting and modeling the features from these synthetic images, the densities of the synthetic images can be compared back to the densities from the real images to validate the realistic nature of the synthetic images.

The first step in generating a synthetic fingerprint image is creating the master print. The master print is defined by the ridge orientation and ridge density and is represented by a binary image. Common methods of generating the

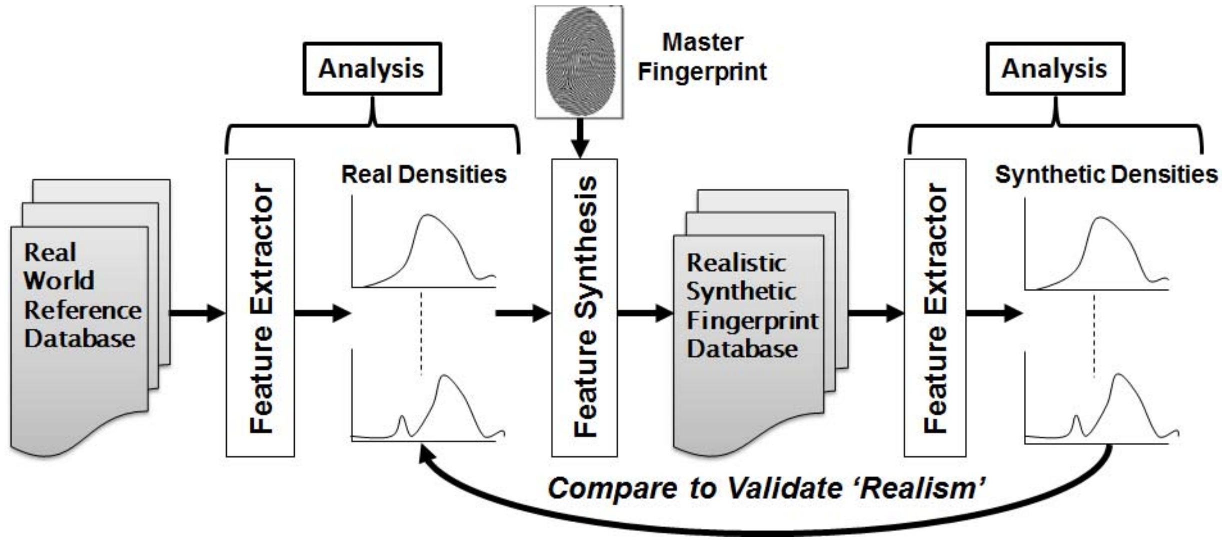

Figure 1: Analysis-by-synthesis process for synthetic fingerprint generation.

master print are described in [4], [5]. Given the master print, non-linear distortions are typically applied to simulate the effects of pressing the finger on the sensor [6]. The final step is texture rendering and adding noise. Previously, this step has been accomplished with uniform noise generation [4] and improved upon with the non-uniform Perlin noise function [7]. For these methods, the realistic nature of the texture was analyzed with visual inspection. In this paper, we propose a new approach for rendering synthetic fingerprint texture by modeling characteristics of real fingerprint ridges and valleys. Unlike previous approaches, our approach is objectively validated by comparing statistical distributions of the measured characteristics from the real and synthetic datasets.

Section 2 gives an overview of specific research in the generation of synthetic fingerprints. Section 3 describes the ridge and valley modeling approach for fingerprint synthesis proposed in this paper. Section 4 then evaluates and discusses the ridge and valley models. Finally, Section 5 gives the conclusions.

2. Synthetic Fingerprint Generation

The process of generating a synthetic fingerprint image begins with the creation of the master fingerprint image, which is a binary image not containing any texture or noise. The state of the art is the SFinGe technique developed by Cappelli et al. [4], [6], [7]. This technique begins with the Sherlock and Monro model [8] for generating the orientation image from randomly placed cores and deltas. Cappelli and Maltoni [9] modeled the spatial distributions of cores and deltas using Gaussian mixture models (GMM). Next, the ridge density image is defined. Then, beginning with a small spike placed in the center of the image, or multiple randomly places spikes, a Gabor filter, tuned to local orientation and frequency, is used to iteratively enhance the fingerprint image from those points. In order to produce a more realistic distribution of minutiae points, Zhao et al. [5] proposed modeling the realistic distribution of minutiae points and built on the technique of reconstructing fingerprints from minutiae templates of Feng and Jain [10] to generate unique synthetic fingerprints.

After the master fingerprint is generated, the next step is adding distortions to the image to simulate the effects of traction and torsion forces applied during placement of the finger on the sensor surface. For the modeling of the resulting non-linear distortions, a skin-distortion model that had been developed to aid in matching is implemented [6]. For this step, every pixel is re-mapped according to a specified distortion function.

The final step in the generation of a synthetic fingerprint image is rendering texture and adding noise. An initial attempt at creating a noisy fingerprint by Cappelli et al. [4] is to add small white circles of varying size at random positions on the image. After applying a smoothing filter and reapplying the valleys, a textured, somewhat realistic fingerprint is obtained. This method is then improved upon by Cappelli et al. [7] with the use of the Perlin Noise function [11]. In this method, random waves at multiple frequency ranges are added together to create a non-uniform noise function. The downside of these methods is that they do not necessaraly produce fingerprint texture statistically representative of real fingerprint texture.

Research has shown that image quality is influenced by various skin characteristics, such as moisture, natural oils, elasticity, and temperature, as well as different sensing technologies [12]. Our paper focuses on accurately representing the ridge and valley segments in fingerprint images, which can influence quality, particularly ridge clarity [13]. While previous methods produce fingerprint images that have texture and noise characteristics of the ridges and valleys visually similar to that of real fingerprint images, we propose a new approach for adding these characteristics to synthetic fingerprint images, where the texture and noise are modeled from real fingerprint images. The advantage of this approach is that the characteristics are measurable in a real database and can be mapped to a synthetic database. Through this approach, the synthetic database models the real database such that the characteristics of the synthetic database can vary depending on the real database it seeks to model.

3. Proposed Approach

A set of key features are identified and used to generate texture and noise for synthetic fingerprint images. By accurately modeling these features from a database of real fingerprint images, a database of synthetic fingerprint images can be generated, representative of the database from which it was modeled.

The proposed method of this paper is as follows: First, features from a database of real fingerprint images are extracted (Section 3.1) and modeled (Section 3.2). Next, the models are sampled to obtain the controlling parameters for the synthetic fingerprint generator, allowing the features to be effectively mapped from a real world database to a synthetic database (Section 3.3). Finally, the analysis by synthesis cycle is completed by extracting and verifying the features from the synthetic database (Section 4). This process is demonstrated through the modeling of ridge texture and valley noise features.

3.1. Feature Extraction

The process of extracting the ridge features begins with enhancement of the fingerprint image to obtain a binary ridge mask. The typical method of ridge enhancement is using a Gabor filter to bandpass filter the fingerprint image using the local ridge orientation and frequency. The estimation of the ridge orientation map is conducted using the gradient method, as described in [14], [15]. Given the estimated orientation, the ridge frequency can then be estimated for each pixel by counting peaks in the cross section signature. The ridge segments from the enhanced fingerprint are then thinned to a single pixel wide 8-connected segment. After removing all ridge bifurcation minutiae points, each ridge segment is individually extracted.

During the extraction of ridge segments, the cross section perpendicular to the ridge direction is analyzed and the ridge width and cross-sectional slope are measured. The edges of the ridge are identified by locating the points of maximum derivative on either side of the ridge center. After the successful tracking of each ridge segment, the 1-dimensional signal along the center of the ridge is analyzed, allowing for the measurement of the frequency components along the length of the ridge and the gray level intensities forming the ridge. A short-time Fourier transform is performed on this ridge signal, allowing for the modeling of the variation of ridge signal frequency along the length of each ridge. Frequency analysis of this ridge signal has been demonstrated to be a useful measure of perspiration patterns along ridge segments in fingerprint liveness detection research [16]. The ridge frequency is modeled using the absolute values of the first seven frequency components of a 32-point Fast Fourier Transform (FFT). The first seven frequency components were considered sufficient for capturing the variation of gray level from perspiration emanating from pores. By reconstructing the signal with these seven frequency components and subtracting the reconstructed signal from the original signal, the difference gives an approximation of the noise in the ridge. Finally, in a similar way, the valley segments are isolated and the gray level values are extracted for determination of the noise in the valleys.

This process results in five features as follows: ridge intensity along the ridge center-lines with seven frequency components, ridge width, ridge cross-sectional slope, ridge noise, and valley noise. These features are characterized as described in Section 3.2 and used when generating the texture for the synthetic images as described in Section 3.3.

3.2. Feature Modeling

The distribution of each feature is analyzed on a per image basis and an appropriate model is constructed. Since it is unknown whether the densities belong to a particular distribution, a non-parametric modeling approach is implemented. In this case, the simple histogram method is applied, where the model is defined as the histogram of the data. Features are extracted from a real fingerprint image and modeled. These models are then sampled to create texture for one or many synthetic fingerprint images, where the texture of each synthetic fingerprint is different, yet representative of the single real fingerprint.

3.3. Feature Mapping

For the mapping of texture features to a synthetically generated ridge pattern, the reverse of the feature extraction process must be implemented. The master synthetic fingerprint and the orientation map are required for this step. First, the ridges of the master fingerprint are thinned to one pixel wide 8-connected segments as described in Section 3.1. Then, each ridge segment is tracked, overlaying a synthetically generated ridge segment on the master print. The following procedure is repeated for each ridge segment.

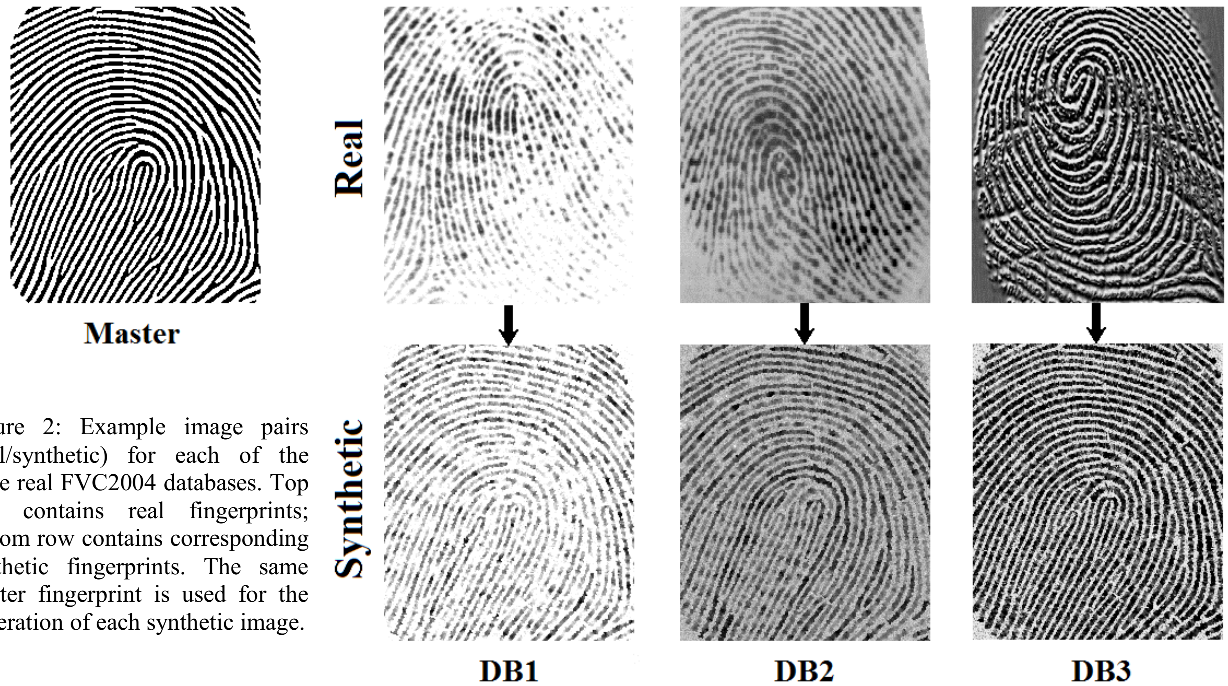

Figure 2: Example image pairs (real/synthetic) for each of the three real FVC2004 databases. Top row contains real fingerprints; bottom row contains corresponding synthetic fingerprints. The same master fingerprint is used for the generation of each synthetic image.

1. Determine length of segment.
2. Generate ridge center signal, sampling from ridge frequency, intensity, and noise models.
3. Generate a cross-section for each ridge center pixel, sampling from ridge width and cross-sectional slope models.
4. Travel along each ridge segment, placing a cross-section, rotated to be perpendicular to local ridge orientation.

The last step is to scan through the image and locate any unassigned ridge pixels and choose a value based on its neighborhood, here the median value was chosen.

Three examples of texture mapping (one for each database) from a real fingerprint image to a synthetic are shown in Figure 2. These examples visually demonstrate the capability of the proposed texture modeling approach in capturing particular image quality characteristics to be recreated in synthetic fingerprints.

4. Model Evaluation & Discussion

As a check to ensure the appropriate mapping of features, the feature extraction process is repeated for the synthetic images. The distributions are then compared to the extracted features of the real fingerprints. For this evaluation, 800 images are selected from each of the three real FVC2004 databases and the texture characteristics of each image are modeled. One synthetic image is created from each of the models, giving 800 synthetic images per database, each with unique, yet statistically representative texture. Ten unique master prints are used for this process. The comparisons of the three real FVC2004 databases and three corresponding synthetic databases created here are shown in Figure 3, demonstrating a close similarity between the feature distributions of the synthetic fingerprints and the real fingerprints from which they were modeled. For this analysis, the feature distributions from all 800 images for each database are combined into a single plot for each feature. Each pair of distributions is compared using the two-sample Kolmogorov-Smirnov (K-S) test, where the K-S statistic is a measure of the distance between the distribution functions of the two samples. Each K-S statistic is labeled in Figure 3. Figure 3 also demonstrates the differences in the feature distributions across different types of fingerprint data, e.g. the different sensing technologies used in the FVC2004 collection.

One issue encountered is in DB3, where significant portions of the ridges are saturated at a gray level of zero. During generation of the synthetic ridges, clipping caused the distribution of each frequency component magnitude to be shifted towards higher gray levels when adding in each component to the synthetic ridge signal. Consequently, the K-S statistics for the DB3 intensity distributions are significantly higher than all the others. This case demonstrates what can happen when the synthetic generation process is not tightly coupled with the feature extraction process. In order to solve this problem, a new synthetic generation method would need to be developed which includes modeling of this saturation effect. A contribution of this work is that the method "couples" analysis and synthesis for texture generation, particularly in DB1 and DB2.

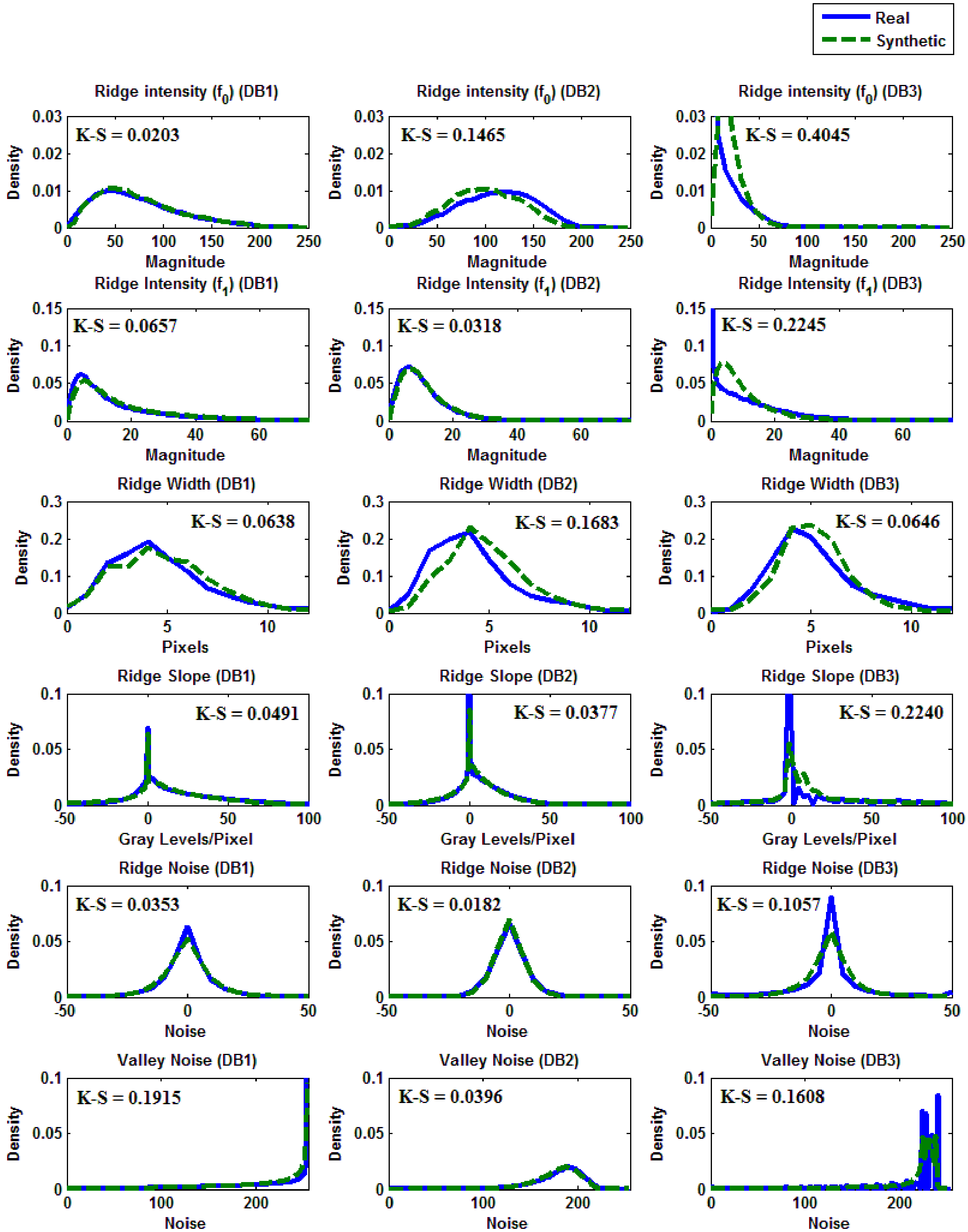

Figure 3: Comparison of real and synthetic feature distributions measured from 800 real fingerprint images from each of the three real FVC2004 databases and 800 synthetic fingerprint images representative of each of the real databases in terms of texture. Distributions are compared using the two-sample Kolmogorov-Smirnov (K-S) test and each K-S statistic is reported.

When comparing our texture generating approach to the commercial software SFinGe, some important differences should be pointed out. The input parameters for texture in SFinGe include ridge noise, ridge prominence, valley noise, and background noise. Our input parameters for texture include ridge intensity along the ridge center with seven frequency components, ridge width, ridge cross-sectional slope, ridge noise, and valley noise. While the commercial SFinGe includes texture analysis, our features expand upon this work and include additional components of the texture characteristics. However, the main contribution of our approach is in the modeling aspect. Our features are *measured and modeled directly from real fingerprint images and mapped to synthetic* fingerprint images. The SFinGe software allows the parameters to be selected from specified distributions for creating a database of images. However, the linkage between the distributions of a real database and controlling parameters of the synthetic features must be performed manually. Furthermore, the distributions are for selection of parameters at the database level and distributions at the image level can not be controlled. Currently our features are spatially independent, causing a slight lack in realism. Spatial dependence will be considered in future work.

5. Conclusion

Modeling of fingerprint features is an important aspect of synthetic fingerprint generation, in that it allows for generation of new fingerprint samples, which are unique from the fingerprint used for modeling, while still retaining the statistical properties of the database of real fingerprints. This process has been demonstrated by mapping the feature statistics from real to synthetic fingerprints. The synthesis process could be repeated any number of times by resampling the models to create many synthetic fingerprints with unique textures, representative of real fingerprint images.

A future direction to be taken in this work is to analyze the correlations between features to assess the possible need for joint-feature models. It is likely that at least some of the features are highly correlated and that correlation should be incorporated into the models. An additional step will be to analyze the spatial dependence of each feature. An example of this type of dependency could be the ridge intensity being greater towards the center of the fingerprint image than on the outside. Ultimately, these texture features should be combined with minutia distribution and distortion models to assess matching performance with synthetic fingerprint images.

References

[1] S. Pankanti, S. Prabhakar, and A. K. Jain, "On the Individuality of Fingerprints," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 8, pp. 1010–1025, 2002.

[2] S. N. Yanushkevich, "Synthetic Biometrics: A Survey," in *International Joint Conference on Neural Networks, 2006. IJCNN '06*, 0-0 0, pp. 676–683.

[3] K. Sumi and T. Matsuyama, "Privacy protection of biometric evaluation database--a preliminary study on synthetic biometric database," presented at the Japan-Korea Joint Workshop on Frontiers of Computer Vision, 2005, pp. 189–194.

[4] R. Cappelli, D. Maio, D. Maltoni, and A. Erol, "Synthetic Fingerprint-Image Generation," in *Pattern Recognition, International Conference on*, Los Alamitos, CA, USA, 2000, vol. 3, p. 3475.

[5] Q. Zhao, A. K. Jain, N. G. Paulter Jr, and M. Taylor, "Fingerprint Image Synthesis based on Statistical Feature Models," presented at the The IEEE Fifth International Conference on Biometrics: Theory, Applications and Systems, Washington D.C., Sept. 23-26.

[6] R. Cappelli, D. Maio, and D. Maltoni, "Synthetic Fingerprint-Database Generation," in *Pattern Recognition, International Conference on*, Los Alamitos, CA, USA, 2002, vol. 3, p. 30744.

[7] R. Cappelli, D. Maio, and D. Maltoni, "An improved noise model for the generation of synthetic fingerprints," in *Control, Automation, Robotics and Vision Conference, 2004. ICARCV 2004 8th*, 2004, vol. 2, pp. 1250– 1255 Vol. 2.

[8] B. G. Sherlock and D. M. Monro, "A model for interpreting fingerprint topology," *Pattern Recognition*, vol. 26, no. 7, pp. 1047–1055, Jul. 1993.

[9] R. Cappelli and D. Maltoni, "On the Spatial Distribution of Fingerprint Singularities," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 4, pp. 742 –448, Apr. 2009.

[10] J. Feng and A. K. Jain, "Fingerprint Reconstruction: From Minutiae to Phase," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 2, pp. 209 –223, Feb. 2011.

[11] K. Perlin, "An image synthesizer," *SIGGRAPH Comput. Graph.*, vol. 19, no. 3, pp. 287–296, Jul. 1985.

[12] C. R. Blomeke, S. K. Modi, and S. J. Elliott, "Investigating the relationship between fingerprint image quality and skin characteristics," in *42nd Annual IEEE International Carnahan Conference on Security Technology, 2008. ICCST 2008*, 2008, pp. 158 –161.

[13] T. P. Chen, X. Jiang, and W. Y. Yau, "Fingerprint image quality analysis," in *2004 International Conference on Image Processing, 2004. ICIP '04*, 2004, vol. 2, pp. 1253 – 1256 Vol.2.

[14] M. Kass and A. Witkin, "Analyzing oriented patterns," *Computer Vision, Graphics, and Image Processing*, vol. 37, no. 3, pp. 362–385, Mar. 1987.

[15] Lin Hong, Yifei Wan, and A. Jain, "Fingerprint image enhancement: algorithm and performance evaluation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 8, pp. 777–789, Aug. 1998.

[16] B. Tan and S. Schuckers, "Spoofing protection for fingerprint scanner by fusing ridge signal and valley noise," *Pattern Recognition*, vol. 43, no. 8, pp. 2845–2857, Aug. 2010.