Effect of Resolution and Image Quality on Combined Optical and Neural Network Fingerprint Matching

C L. Wilson, C. I. Watson, and E. G. Paek

Information Technology Laboratory

National Institute of Standards and Technology

Gaithersburg, MD 20899

Abstract

This paper presents results on direct optical matching, using Fourier transforms and neural networks for matching fingerprints for authentication. Direct optical correlations and hybrid optical neural network correlation are used in the matching system. The test samples used in the experiments are the fingerprints taken from NIST database SD-9. These images, in both binary and gray level forms, are stored in a VanderLugt correlator [1]. Tests of typical cross correlations and autocorrelation sensitivity for both binary and 8 bit gray images are presented. When Fourier Transform (FT) correlations are used to generate features that are localized to parts of each fingerprint and combined using a neural network classification network and separate class-by-class matching networks, 90.9% matching accuracy is obtained on a test set of 200,000 image pairs. These results are obtained on images using 512 pixel resolution. The effect of image quality and resolution are tested using 256 and 128 pixel images, and yield accuracy of 89.3% and 88.7%. The 128-pixel images show only ridge flow and have no reliably detectable ridge endings or bifurcations and are therefore not suitable for minutia matching. This demonstrates that Fourier transform matching and neural networks can be used to match fingerprints which have too low image quality to be matched using minutia based methods. Since more than 258,000 images were used to test each hybrid system, this is the largest test to date of FT matching for fingerprints.

1 Introduction

This paper presents data on inked fingerprint images matched with optical and hybrid optical neural network correlators. The matching method is tested on an authentication application. The inked fingerprint images are rolled prints scanned at 20 pixels/mm on a 4 cm by 4 cm area of a fingerprint card. We study matching of the inked fingerprints using global optical correlations [1] and partial optical correlation features and a system of neural classification and matching networks [2]. Images of three different resolutions are tested to determine the effect of image resolution and quality on matching accuracy.

For images of inked rolled fingerprints, even after core alignment and correction for rotation, optical matching of most prints is successful for matching the original image and rejecting other fingerprints, but fails on second copies of inked rolled images because plastic pressure distortions and image size variation are too large to allow global Fourier transform (FT) matching. Detailed computer simulations show that global optical matching uses the fine-grained phase-plane structure of the Fourier transform of the fingerprints to produce strong optical correlations. This fine-grained structure is very sensitive to pressure and plastic distortion effects which then dominate in correlations of static fingerprints.

The fine grained local variations in fingerprints can be compensated for by calculating optical correlations on smaller zones of the fingerprints. A training set was derived from disk two volume 1 of SD-9 and the testing set from disk one of volume one of SD-9 [3]. Since all fingerprints in disk one were tested against all each other, a total of 258,444 tests were performed in each experiment. This is the largest FT based matching experiment reported to date. In our experiments, two four by four matrices of correlations on zones of the fingerprint are used to produce a total of 32 features. One set of correlations is computed with the local zone grid centered on the core and one set is computed with the core in the center of the grid just above and to the left of grid center. Features were extracted from the correlation data using correlation peak height, correlation peak width, and correlation area.

These features are combined using two types of neural networks. The first network is used to classify the fingerprints [4, 5, 6]. This fingerprint classification network works directly with the fingerprint image. After each fingerprint is classified, class-by-class matching networks are trained for each class. These two networks function in a way similar to the binary decision networks discussed in [7]. For this particular problem, the network training is strongly dependent on regularization and pruning for accurate generalization [2].

The advantage of the combined optical neural network method is its insensitivity to image resolution and quality. The experiments presented in this paper were done with three different image sizes. Initial results were obtained with image samples with 512 by 512 pixels on each side, sampled at 20 pixels/mm. These images were down-sampled to 256 by 256 and 128 by 128 using averaging of the gray levels to achieve sampling rates of 10 pixels/mm and 5 pixels/mm. The full matching test was then performed for three combinations of extracted features and for images of each size. As we will discuss in section 3, analysis of ridge spacing data on the test fingerprints shows that the Nyquist sampling limit of two pixels for each ridge and valley occurs at the 256 pixel level. The 128 by 128 images were sampled at half the Nyquist level and are of low quality with few clear ridge endings and/or bifurcations. As we discuss in section 5, the accuracy of the hybrid matching method decreases with image resolution but remains usable even for 128 by 128 images.

In section 2 we describe the direct optical correlation experiment. In section 3 we present an analysis of ridge frequency data and its effect on image quality. In section 4 we discuss combining optical and neural network methods. In section 5 we present the results of the hybrid system and in section 6 we draw some conclusions about the difference in correlations of real time and rolled inked fingerprints.

2   Global Optical Correlations

In the global optical matching experiment, images from NIST Special Database 9 (SD-9) [3] are core aligned using the method discussed in [4] and cropped to fit the 640 by 480 pixel field of the pattern recognition system. Two hundred reference fingerprints and second rollings
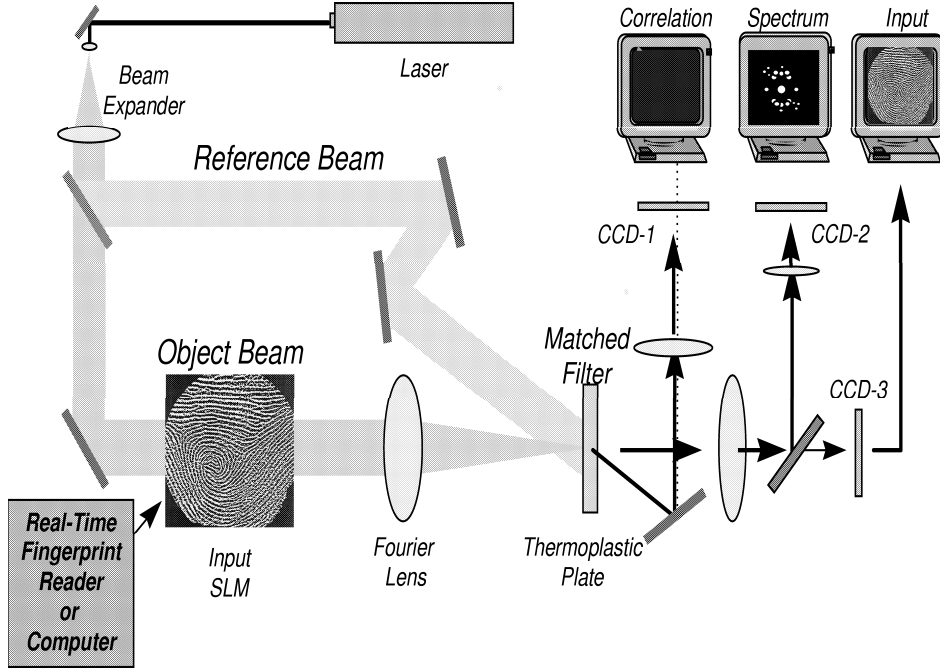
Figure 1: Diagram of the optical pattern recognition system

(inked images taken at a different time) are available for autocorrelation and cross correlation experiments. When binary finger prints are used the method used is based on that presented in [5].

Figure 1 shows a schematic diagram of the optical pattern recognition system. It is based on the conventional VanderLugt correlator [1]. The target fingerprint image is loaded on an spatial light modulator, (SLM) and is Fourier transformed by a lens. The resulting Fourier spectrum is interfered with a reference beam to record a Fourier transform hologram. After recording is finished, if an arbitrary input fingerprint is presented on the SLM the correlation of the input and the target appears in the correlation output plane.

Although the spatial heterodyning technique, often called joint transform correlator [8], has many advantages for real-time applications [9, 10] and was used in most of the recent fingerprint recognition experiments [11, 12, 13, 14, 15, 16], the VanderLugt correlator was adopted in this experiment. This is because the VanderLugt correlator does not require a fast SLM with high resolution and the large SBP (space bandwidth product) available from holographic recording materials provide a high degree of freedom to accommodate various distorted versions of a target that are simultaneously compared with an input. Also, since the information is recorded in the form of a diffraction pattern (hologram) instead of a direct image, it can be used on a credit card or an ID card for security purposes without need for further encoding. Finally, the VanderLugt correlator is better suited for spatial filtering to increase signal to noise ratio (SNR). The critical positioning tolerance problem of the VanderLugt correlator can be greatly relaxed by using in-situ recording materials, such as thermoplastic plates, as were used in this experiment. In this case, once the system is aligned,

3

new holographic filters can be generated with no fear of misalignment.

In the global correlation experiment, fingerprint images are generated from the NIST fingerprint database [3]. In the real-time correlation experiment, images are generated by a live-scan fingerprint scanner (Identicator Technology, Gatekey Plus ver. 4.1) [1]. An electrically addressable liquid crystal SLM (Kopin, LVGA kit, 14 mm diagonal)[1] is used as an input device. The SLM is mounted on a rotational stage to facilitate precise rotational tolerance measurements.

Holographic filters are recorded on a thermoplastic plate (Newport Corp. HC-300)[1] that allows fast non-chemical processing, high diffraction efficiency and high sensitivity. Although the recording process cannot be achieved in real-time (close to 1 minute), the time-consuming comparison of an input with many other images in a large database can be done very fast, once a hologram is made.

A 10 mW HeNe laser with a ND 2 filter was used as a light source, and so only 0.1 mW was used to see correlation output, due to the high light efficiency of the system.

The system is also equipped with real-time in-situ monitoring of an input image, its Fourier transform, and the correlation output. These monitoring parts, combined with a frame grabber and other analytic tools, permit real-time quantitative analyses and accurate characterization of every stage of the system operation.

The correlator system is capable of shift-invariant pattern recognition over a broad range of input positions and has high SNR due to accurate alignment using an interferometer and a microscope.

Figure 2 shows a histogram of peak correlations for gray (a) and binary inputs (b). For each of 20 randomly chosen fingerprints, an individual holographic filter was fabricated and tested against the 200 fingerprints in the NIST database stored in the control computer. Therefore each plot involves 4,000 correlations. Each peak correlation value was obtained by taking the maximum value in the correlation plane. In case of gray inputs shown in (a), all 20 autocorrelations peak at the maximum value (152). Cross-correlations distribute in a Gaussian shape with a full width half-minimum (FWHM) of around 15, and the maximum at 60.

For binary inputs shown in (b), all autocorrelations peak at the maximum value, as in gray inputs. However, in this case, cross-correlations are significantly reduced to zero except for the few cases which were found to be from the correct fingerprints of the other rolling.

For both gray and binary inputs, the autocorrelations are well separated from the cross-correlations to permit perfect 100% recognition for correct fingerprints (without considering distortions).

The exact mechanism for the significant increase in SNR for binary inputs is not completely understood. However, several previous works [17, 18] support the experimental results. Such a high SNR of binary inputs can be efficiently used to make a composite filter to permit tolerance against distortion.
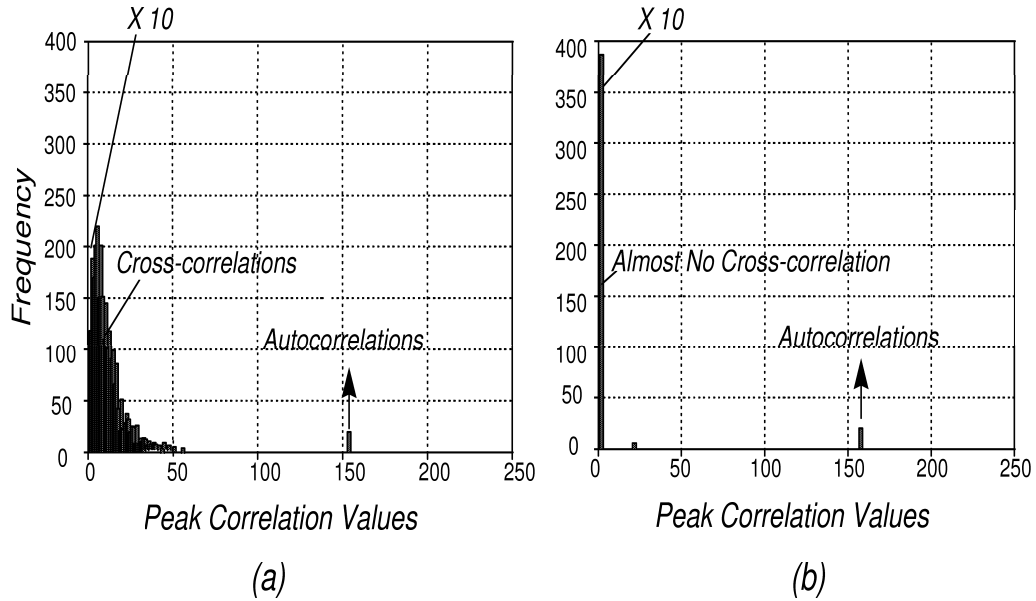

Figure 2: Histograms of peak correlations for gray (a) and binary (b) fingerprint images.

3 Fingerprint Image Characteristics

In this section we present data on fingerprint ridge pitch and frequency and the effect of image sampling frequency on image quality. The standard sampling frequency for fingerprint data for images of inked fingerprints is 500 pixels/inch or 19.7 pixels/mm, approximately 20 pixels/mm. Live scan system designed for law enforcement applications use this sampling rate but live scan systems designed for verification applications are using lower sampling rates down to approximately 5 pixels/mm.

The constraint on image quality that ridge frequency values imposed is important for both minutia matching methods and for Fourier transform methods. For minutia matching methods, the ridge structure of the fingerprint must be sampled with sufficient frequency to allow the ridge and valley structure to be accurately detected. This is observed to be approximately two points for each ridge and two point for each valley as expected form basic Nyquist sampling theory. In the FT case, the frequency of sampling is important because it effects the sensitivity of the correlation to plastic distortion. Near the center of the fingerprint ridge and valley position do not vary much with pressure but at the edges fingerprint ridges may be displaced by a full ridge width effectively interchanging the ridge and valley positions. Smaller ridge pitch values for an equally elastic finger will increase this effect. Typical effects of elastic distortion are shown in figure 14. In previous optical fingerprint correlation studies [11, 12, 13, 14, 15, 16], decreasing image size and sampling frequency has decreased sensitivity to rotational alignment and plastic distortion.
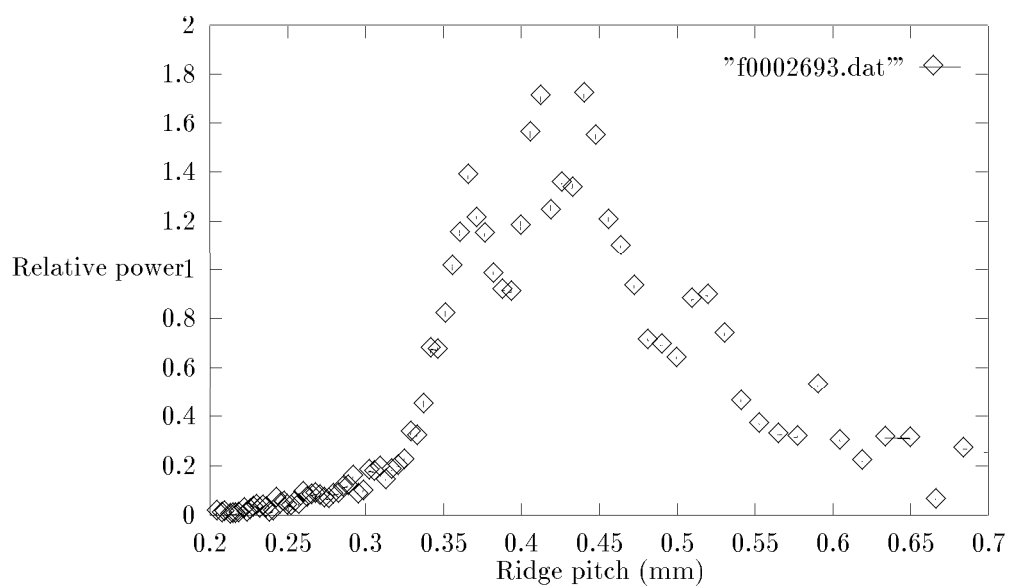
5

Figure 3: Relative power of the FT as a function of ridge pitch for fingerprint with small pitch
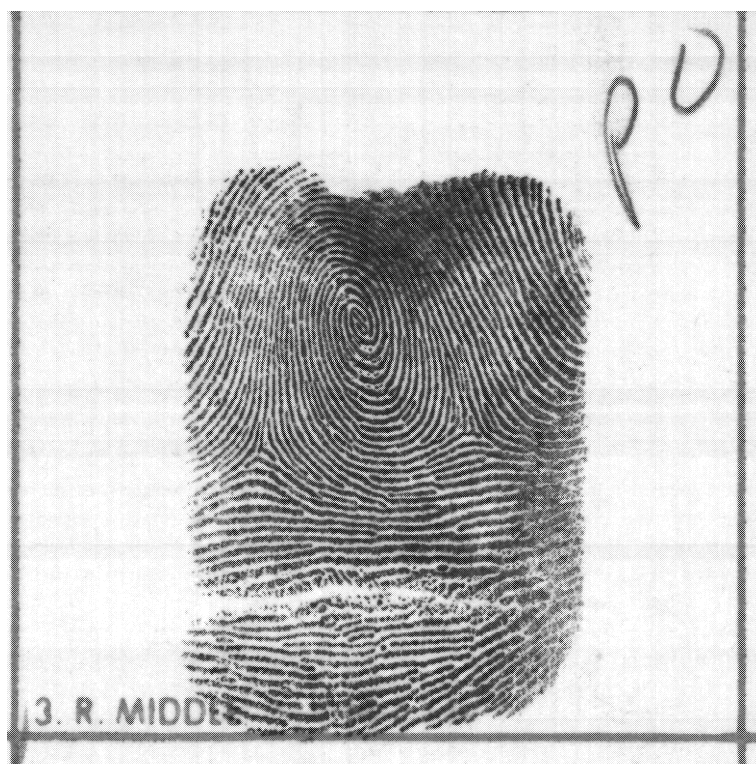

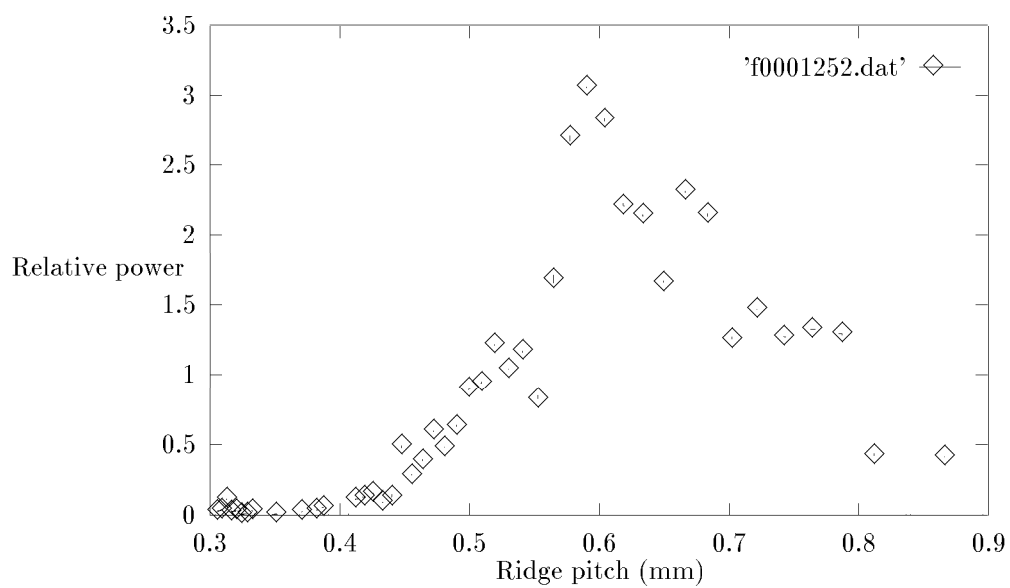Figure 4: Example fingerprint with narrow, 0.4mm, ridge pitch.

Figure 5: Relative power of the FT as a function of ridge pitch for fingerprint with large pitch.
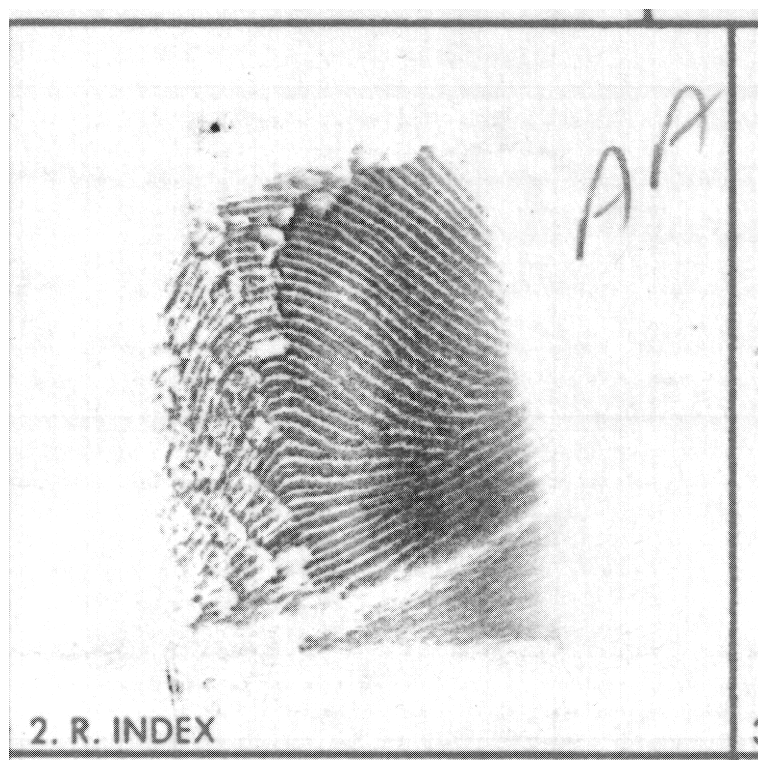

Figure 6: Example fingerprint with wide, 0.6mm, ridge pitch.

3.1   Ridge Pitch Variation

In a collection of fingerprints, two kinds of variation of ridge spacing are of interest for matcher evaluation. First we have variations in ridge pitch within individual fingerprints, and second we have variations in ridge pitch across samples of fingerprints, such as the volume 1 disk 1 of NIST database 9.

The variations discussed here were measured by performing FTs for each fingerprint. The power spectrum of the FT of the fingerprint was then sampled over angles from 0 to π at 257 different radial distances and a histogram of relative power vs spatial frequency was produced for each fingerprint. The average values of these histograms over each class for male and female subjects was also produced. The class sample sizes reflect the natural class occurrence rates; sample sizes for arches and tented arches are about 1/19 the size of those for loops and whorls.

The variations in ridge spacing for two individual fingerprints are shown in figures 3 and 5. Both fingerprints have sharp peaks in their power spectrum in the typical ridge pitch range between 0.4mm and 0.6mm and have minimum ridge spacing of about 0.2mm and maximum ridge spacing of about 0.8mm. The peak power of the two prints are near the limits for peak power observed in the special database 9 sample. The fingerprint with the smaller ridge pitch has a power distribution skewed toward smaller ridge pitch values, see figure 3, and the fingerprint with the larger ridge pitch has a power distribution skewed toward larger ridge pitch, see figure 5. The fingerprints measured to obtain the two distributions are shown in figures 4 and 6 respectively. Both images have a scale of 19.7 pixels/mm and the ridge pitch shown in figure 4 is, as expected from the distributions, 2/3 of that shown in figure 6. Figure 6 also demonstrates that using FT power to measure ridge pitch is robust enough to work well on a fingerprint with a poor quality image. Examination of the fingerprint images illustrated that the larger ridges are near the crease at the bottom of the images and the smaller ridges are near the finger tip. This is true for all of the fingerprints tested.

As the number of fingerprints used in the calculation of the power spectra vs ridge pitch is increased, the distributions become smooth and approach skewed Gaussian form. This is illustrated in figures 7 for males with fingerprints classified as whorls and by 8 for females with fingerprints classified as whorls. In these distributions, the range of ridge pitches in each distribution, 0.2mm to 1.0mm, is larger than the variations in maximum ridge pitch between males and females, and larger than the variation in maximum ridge pitch between classes. The distributions for other classes of fingerprints have similar shapes and ranges. These measurements show that ridge pitch variations in individual fingerprints and in classes of fingerprints are always larger than variations in most likely ridge pitch across groups of fingerprints or between sexes.
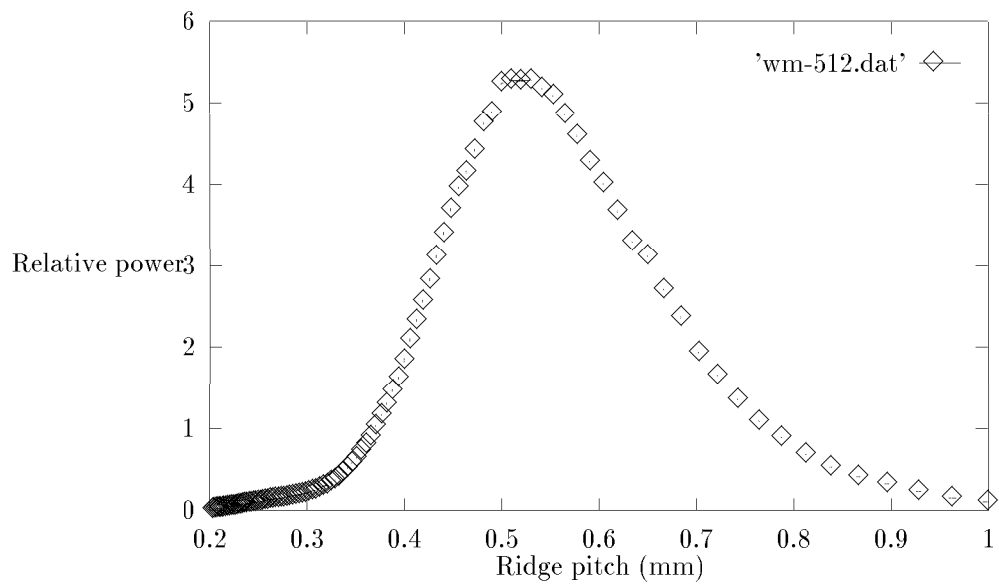
Figure 7: Relative power of the FT as a function of ridge pitch for males for fingerprint which are classed whorls.
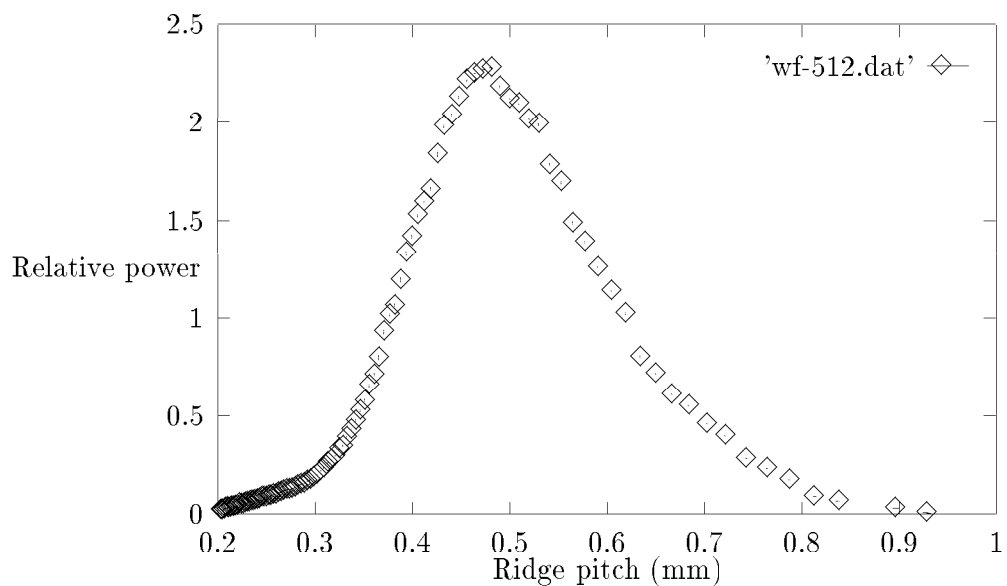

Figure 8: Relative power of the FT as a function of ridge pitch for females for fingerprint which are classed whorls.

Figure 9: Example fingerprint sampled at 20 pixels/mm.

3.2 Image Quality

The effects of sampling frequency on image quality are shown in figures 9, 10, and 11 for a single fingerprint sampled at 20 pixels/mm, 10 pixels/mm, and 5 pixels/mm. Based on the FT power spectra presented in the previous section we would expect some part of the fingerprint, usually above the core of the fingerprint, to be just adequately sample at the 20 pixel/mm sampling rate but that most of the fingerprint would be adequately sample. In figure 9 the region of narrow ridge spacing between 1 and 3 o'clock above and to the right of the core are just adequately sampled.

When the sampling rate is reduced to 10 pixels/mm, as in figure 10, some of the minutiae in this region are difficult to detect because of blurring. Most of other sections of the fingerprint has clearly defined ridge structure. From the class FT power distribution, such as figure 7, we would expect that only a few per cent of the ridges would be obscured by this 2-to-1 down-sampling and this is the case.

When the sampling rate is further reduced to 5 pixels/mm, as in figure 11, we see a large reduction in image quality. At 5 pixels/mm we would expect, from figure 7, that about 40% of the fingerprint would be sampled with a resolution less than the expected Nyquist limit. In figure 11 most of the minutia locations are blurred, and ridge locations in the area above and to the right of the core are lost at this level of resolution.

From the combined effect of FT ridge pitch analysis and from the examples given above we conclude that on most fingerprints, the effect of sampling at 10 pixels/mm will be small although some part of many fingerprints will be undersampled. Sampling at 5 pixels/mm will make minutiae detection either uncertain or impossible, and will make detection of ridges

difficult in a significant part of most fingerprints.

Figure 10: Example fingerprint sampled at 10 pixels/mm.

Figure 11: Example fingerprint sampled at 5 pixels/mm

4   Combined Optical and Neural System

Direct global correlation of fingerprints for matching has a significant failure rate caused by the elasticity of fingerprints. Two rollings of the same print can vary significantly, as seen by computing their Fourier transforms, because of the stretching variations which occur when rolling a fingerprint. Figure 12 shows the correlation of two rollings of the same print that have been rotation-and-translation-aligned based on the ridge structure around the core. It is clearly seen that the fingerprints correlate (indicated by the dark gray pixels) around the core, but away from the core the patterns have different amounts of elastic distortion.

Since the elastic distortion problem is local, a method of local correlation can be used to lower the average distortion in small subregions of the fingerprint.

4.1   Optical Features

A solution to elastic distortion that occurs in different rollings of the same fingerprint is partitioning the images into tiles and comparing the data within each of the tiles using FT based methods. For this work, each image was partitioned into 4 by 4 tiles twice so that each tile contained one sixteenth of the total image area. One partition had the core located in the center of the image, as defined by the fingerprint core, and the second partition has core shifted away from the center so that the new center is located at the corner of one if the first set of partitions. This double partitioning allowed for overlap of data (specifically data on the edge of the tiles). Since the neural network is allowed to prune any data that is not needed, excess overlap in the features can be removed during network training. Figures
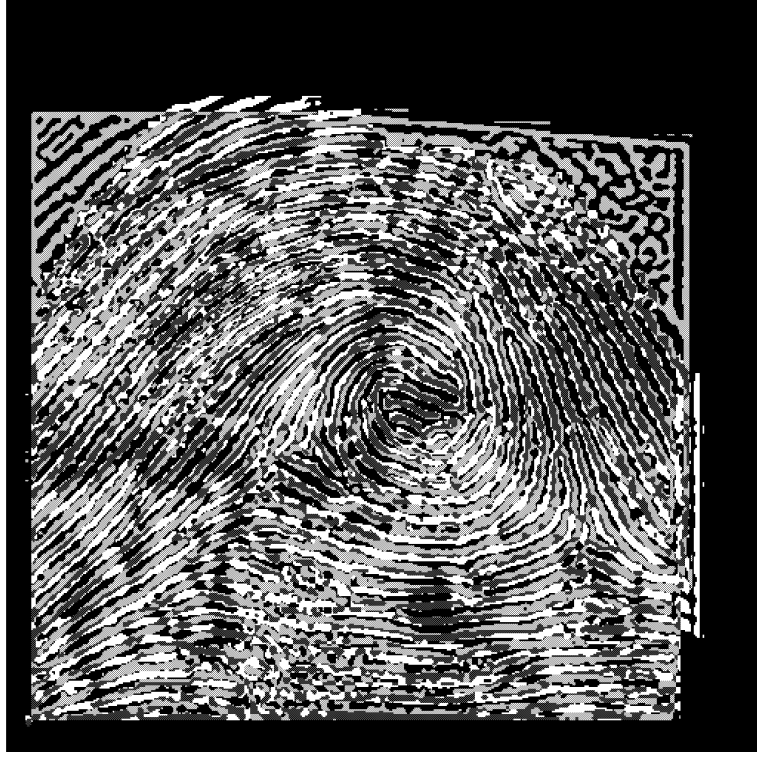
12

Figure 12: Correlation of two rollings of the same print. (Dark gray indicates correlated ridges, white and light gray indicate uncorrelated ridges.)

13a-b show the core location for each 4x4 partition.

In NIST database 9 two rollings of each fingerprint are present; these fingerprint sets labeled file prints (f(n)) and search print (s(n)) . After partitioning, each f(n) and s(m) pair are compared by correlated the corresponding tiles (32 tiles) for each print and extracting features from the correlations as inputs to the neural network. The features used are the central correlation peak height, correlation peak area and correlation peak width. Figure 14 shows two print tiles from a matched pair and the corresponding correlation output. The correlation peak data is extracted by taking a cross-section (perpendicular to the fingerprint ridge direction) of the peak at the maximum correlation value.

The correlation is computed in the Fourier domain by taking the Fourier Transform of the partitions and computing the inverse Fourier Transform of their product, using the complex conjugate of the first (eq. 1).

$$f(n) \circ s(m) = \mathcal{F}^{-1}[\mathcal{F}^*[f(n)] \times \mathcal{F}[s(m)]] \qquad (1)$$

Each f(n) and s(m) vector has 32 values for each peak feature (i.e. height, area and width) (n = 1,2, ... 900 and m = 1,2, ... 900):

$$peak\_features[f(n)_1 \circ s(m)_1]$$
$$peak\_features[f(n)_2 \circ s(m)_2]$$
$$\vdots$$

13

1	2	5	6
3	4	7	8
9	10	13	14
11	12	15	16

17	18	21	22
19	20 ○	23	24
25	26	29	30
27	28	31	32

○ = core location

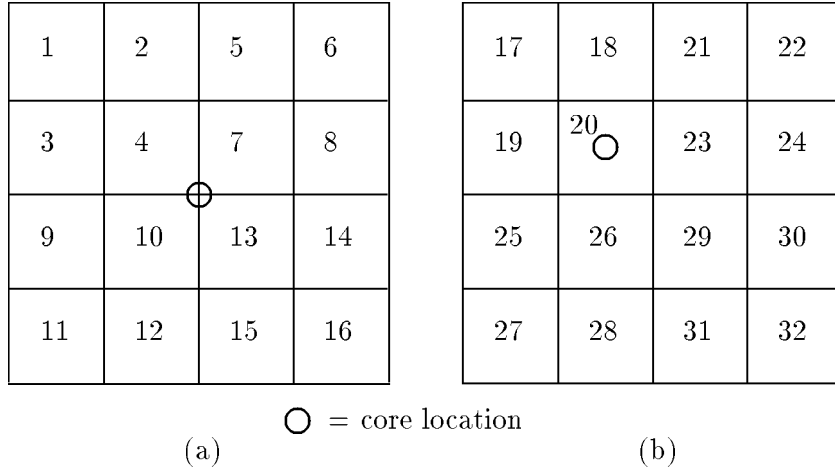(a)   (b)

Figure 13: Shows image partitioning and the corresponding feature number.

$$peak\_features[f(n)_{31} \circ s(m)_{31}]$$
$$peak\_features[f(n)_{32} \circ s(m)_{32}]$$

Automated feature detection procedures were applied to NIST Special Database 9 Vol 1, where disk2 was used as training data and disk1 was used as testing data.

For this partitioning technique to be effective, the images need to be rotationally and translationally aligned about the cores of the two fingerprints being compared. This alignment was accomplished over a large set of data using an automated technique. There are three steps in the automated alignment, filter/binarize image, detect core location, and determine alignment.

Filtering, binarization, and core detection are done using methods previously developed and discussed in detail in [6]. The only addition is that the binarized fingerprint is median filtered using a 3 by 3 window to help smooth noise in the ridge data and improve correlation performance.

The alignment step uses 128 by 128 segments that are centered about the core of the fingerprints being aligned. The correlation of the segments is computed while rotating the second segment over a range of angles. The angle which produces the largest correlation is used for rotation alignment. Since two prints can have significant angular displacements the alignment is actually done in two stages. Stage one uses an angular step size of 1 degree over a range of ±15 degrees and stage two a step size of 0.2 degrees over a range of ±1 degree from the angle determined in the first stage.

Since the correlation computed by equation 1 is translation independent, translation alignment is accomplished by using the peak correlation location from the second stage of the angular alignment. The amount that the peak correlation is off center of the 128 by 128 segment determines how much the second print needs to be shifted to achieve translational alignment with the first.

The feature extraction procedure results in a total of 96 features for each pair of fingerprints compared. In SD-9, each fingerprint has one print in the test set that matches and several thousand which do not match. Only those prints which do not match but are of the same class are included in the training set. The previously developed neural network
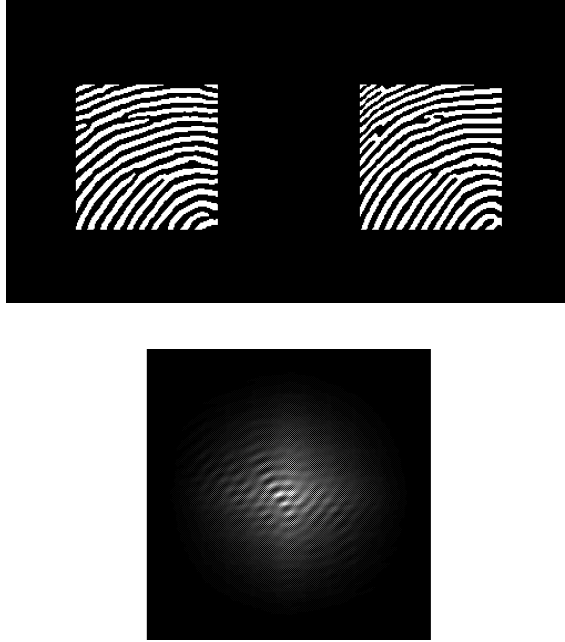
14

Figure 14: Original matched pair and corresponding correlation.

classifier [6] is used for this screening process.

The 96 features were used in three different neural networks. The first network only used the maximum correlation values as the features (32 features). Because of partitioning the prints, a main source of error resulted from non-matching prints having a maximum peak value in the ranges of matching prints. The difference was that matching prints had very narrow, well defined peaks and non-matching prints had broader flat peaks. This led to the next neural network in which correlation peak area was added as an input feature (64 features). Significant improvements were obtained in matching error rates (shown later in the paper). The final network tested used all 96 features (correlation peak maximum, area, and width) and showed a smaller improvement over the 64 feature network.

4.2   Neural Network Matching

The matching networks discussed in this section were trained using a dynamically modified scaled conjugate gradient method presented in [2]. In [2], we demonstrated that performance equal to or better than Probabilistic Neural Networks (PNN) [19] can be achieved with a single three-layer Multi-Layer Perceptron (MLP) by making fundamental changes in the network optimization strategy. These changes are: 1) Neuron activation functions are used which reduce the probability of singular Jacobians; 2) Successive regularization is used to constrain volume of the weight space being minimized; 3) Boltzmann pruning is used [20] to constrain the dimension of the weight space; and 4) Prior class probabilities are used to normalize all error calculations so that statistically significant samples of rare but important classes can be included without distortion of the error surface. All four of these changes are made in the inner loop of a conjugate gradient optimization iteration [21] and are intended to simplify the training dynamics of the optimization.

15

In this work we found that the effect of the sinusoidal activation, 1 above, was not useful, but that pruning, 3 above, and regularization, 2 above, were essential to good generalization. Since the distribution of match and do not match classes was highly unequal, the effect of prior weights, 4 above, was also very important. The optimal regularization factor for all runs was found to be 0.001 and the optimum pruning temperature was found to be 0.005.

5 Hybrid Results

Five different experiments were performed. In each experiment all fingerprints from disk one of volume one of SD9 were used as a test sample. All fingerprints from disk two of volume one of SD9 were used for network training. A total of 258,444 tests were performed in each experimental test sequence. The experiments were designed to test the effect of different methods of FT peak feature extraction and the effect of image resolution on accuracy. The feature extraction experiments used the three methods described above to obtain features from the local FT data. All images were sampled at 20 pixels/mm. In the first of these experiments the correlation peak height was used as the feature. In the second experiment the area under the correlation peak was used as an additional feature. In the third set of experiments, the width of the correlation peak was added to the feature set. For each feature set the K-L (Karhunen-Loeve) transformation was used, as in [7], to reduce the dimensionality of the feature set. Before K-L transformation, these three experiments had feature vector lengths of 32, 64 and 96-elements. In the second set of experiments, the 96-element feature vectors including correlation peak height, peak area, and peak width were extracted for sets of images which were sampled at 20 pixels/mm, 10 pixels/mm, and 5 pixels/mm. The first set of images in the resolution experiments was the same as the set of images from the third experiment.

Each of the local features sets discussed above was separated into testing and training samples both by class and as a global (all class) set. The training sets were used to calculate global and class-by-class covariance matrices and eigenvectors and to calculate K-L transform [22] features for all of the testing and training sets. The effect of the K-L transform was to reduce the feature set sizes from 32 to 13, from 64 to 58 and from 96 to 58 or 59.

When the eigenvectors of the K-L transform were examined for the peak-based 32 element feature set, the primary source of variation was found to be in 12 zones near the center of the two feature grids. The first eigenvector of each of the transforms was approximately 40 times larger than the 13th eigenvector, indicating that only about 13 statistically independent features were computed from the training sets. No large difference in K-L transform characteristics were seen between global and class-by-class data sets.

When the eigenvectors of the K-L transform were examined for the combined peak-and-area-based 64 element feature set, most elements, 58 of the 64, made a significant contribution to the sample variance. Increasing the feature set width to 96 elements by adding the correlation peak width did not increase the number of useful eigenvalues. The transformed feature vectors were still 58 elements long. We can therefore conclude that peak width and peak area are highly correlated.

The K-L transformed features were used to train neural networks for both global and class-by-class matching for each of the five data sets. The networks were trained using regularization to bound weight size and pruning to restrict the number of active weights in the network to a size. Network size, pruning, and regularization were adjusted empirically

to provide reasonable generalization. The criterion used to test generalization accuracy was the comparison of the test and training matching errors.

5.1   Correlation Peak Features

The basic network size was a 13-24-2 network with 386 weights including bias weights. 24 hidden units were needed to provide adequate coverage of the various combinations of interconnections during pruning. A sigmoidal activation function was used for the hidden nodes. With this network size and these training parameters, a typical functioning network has approximately 150 weights and has a accuracy of 62%-71%.

The results of this process are given in the table below.

Class	train	test	wts. pruned	Test set size
All	70.2	65.2	285	258444
Arch	71.8	64.9	229	1681
Left Loop	72.1	62.6	240	73984
Right Loop	72.5	71.1	209	68121
Tented Arch	75.4	67.3	247	1089
Whorl	72.0	68.5	275	113569

Table 1: Results of training and testing for global and class-by-class neural network matching using 13 K-L features. Features were taken from the correlation peak in each subregion. Images contained 512 pixels on each axis. Combining all networks with pattern classification yields 85.28% accuracy. All networks had a 13-24-2 structure with 386 weights.

These results can be significantly improved by using PCASYS or some other classification method to test only prints of the same class for matching. Assuming the PCASYS accuracy of 99% correct classification at 20% rejects and a natural distribution of classes would allow the results given above to be improved to 84.3% matching accuracy. If a perfect classifier were available, then the combined accuracy would be 90.3%. This model assumes that each print is classified or rejected by PCASYS. The rejected prints are matched with the All network given in the top line of the table. All other prints are matched by the network selected by its PCASYS class. All prints misclassified by PCASYS are assumed to be mismatched.

The process of calculating the results shown in table 1 involved training runs in which both the regularization and pruning were systematically varied to determine the correct network size and the appropriate dynamics for training. As discussed in [20], network size is an indication of the amount of information that can be transferred between the training sample and the network without learning random noise patterns. In table 1, all of the final networks had a potential weight space size of 386 weights. Larger networks were found to have poorer testing error than networks of this size. The pruning temperature was varied to produce similar testing and training errors for each class and for the global class. As the table shows, this produced weight reductions of from 209 to 285 weights, leaving 101 to 187 nonzero weights. The small network size and large pruning ratio for acceptable generalization with training set of up to 100,000 samples shows that the noise in the features used in the training is at a level where larger network sizes are not useful, because all of the information needed for generalization is learned by these small networks.

17

All of the pruning experiments require that some small amount of regularization be used to constrain the volume of weight space [2]. This allows the discriminant surfaces to remain in the part of the training space which is heavily populated by the data. All of these runs were done in the 13-feature K-L space, but numerous test pruning and regularization runs were made in the original 32-feature space. Similar effective weight spaces were found in the full 32-feature space, about 150 weights. The 13-feature data set was selected for additional testing to save on computation time during training.

5.2   Correlation Peak and Peak Area Features

Class	train	test	wts. pruned	total wts.	Test set size
All	78.3	76.0	1144	1442	258444
Arch	78.2	75.0	587	746	1681
Left Loop	79.1	84.6	608	698	73984
Right Loop	80.6	80.0	576	722	68121
Tented Arch	86.8	74.1	402	722	1089
Whorl	79.3	80.7	589	722	113569

Table 2: Results of testing and training for global and class-by-class neural network matching using 26–28 K-L features. All class networks had 24 hidden nodes while the all network had 48 hidden nodes. Features used were based on peak correlation and area under the correlation peak. Images contained 512 pixels on each axis. Combining all networks with pattern classification yields 89.95% accuracy. Typical class networks had a 58-24-2 structure.

The basic network size was a 58-24-2 network with 722 weights, including bias weights, for class networks and 59-48-2 for the all class network. A sigmoidal activation function was used for the hidden nodes. With this network size, from table 2 we see that a typical functioning network has approximately 100-300 weights and has an accuracy of 74%-84% for class networks. The global (All) class network had 298 weights and an accuracy of 76%.

This experiment differs from the peak feature experiment in that the combined feature set is only reduced from 64 features to 58 features. For some classes, such as left loops, this results in greatly improved accuracy, from 62.2% to 84.6%, and a reduced number of weights, from 146 to 90. For other classes the network does not reduce in size or improve in accuracy. For tented arches the accuracy decreases from 75.4% to 74.1% and the number of active weights increases from 139 to 320. Since the training data has a natural class distribution, the data indicates that the classes with relatively small sample sizes, arch and tented arch did not improve with more complex features but the classes with larger training set sizes, loops and whorls, improved an average of 10%.

5.3   Correlation Peak, Peak Area, and Peak Width Features

The basic network size was 57,58,59-48-2 network with 1442, 1466, 1490 weights, including bias weights, for the various networks. A sigmoidal activation function was used for the hidden nodes. With this network size, from table 3 we see that a typical functioning network

has approximately 250-530 weights and has a accuracy of 74%-84% for class networks. The global (All) class network had 298 weights and an accuracy of 76%.

This experiment differs from the two previous experiments in that the required network size has 48 hidden nodes for all of the networks and pruning on these networks with the training set sizes used is substantially less effective than it was with peak and area features. The less effective pruning of the network doubles the number of weights from 100-300 to 250-530. This shows that, even after feature correlations are removed by the K-L transform, various complex feature combinations are available that are detected in the network training. As in the previous experiment, the classes with relatively small sample sizes, arch and tented arch, did not improve with more complex features, but the classes with larger training set sizes, loops and whorls, improved an average of 10% over simple peak features.

5.4   Scan Resolution of 10 Pixels/mm

In this experiment the image scanning resolution was reduced from 20 pixels/mm to 10 pixels/mm. The basic network size was 57,58-48-2 network with 1442, 1466 weights, including bias weights, for the various networks. The network size is similar to the 20 pixel/mm experiment. A sigmoidal activation function was used for the hidden nodes. With this network size, from table 4 we see that a typical functioning network has approximately 85-190 weights and has an accuracy of 71%-81% for class networks. The global (All) class network had 302 weights and an accuracy of 76%.

The two main effects of lower image resolution are to increase pruning effectiveness and to decrease accuracy. The required network size is still about 1466 weights but the number of weights that are useful has decreased by about a factor of two and the class networks have been trained to a corresponding lower accuracy. The data needed to generate more complex weights set is missing in the lower resolution data.

5.5   Scan Resolution of 5 Pixels/mm

In this experiment the image scanning resolution was reduced from 10 pixels/mm to 5 pixels/mm. The basic network size was 62,60-48-2 network with 1562, 1514 weights, including bias weights, for the various networks. A sigmoidal activation function was used for the hidden nodes. With this network size, from table 5 we see that with these training parameter a typical functioning network has approximately 88-247 weights and has an accuracy of 74%-81% for class networks. The global (All) class network had 418 weights and an accuracy of 76%.

Further reduction in image resolution from 10 pixels/mm to 5 pixels/mm has had two effects. The K-L transform yields 60-62 features instead of 58-59 features which results in a small increase in initial network size. This larger network is then pruned to about the same size as in the 10 pixel/mm case and the final accuracy of the matching is reduced by less than 1%.

19

Class	train	test	wts. pruned	total wts.	Test set size
All	78.3	76.0	1144	1442	258444
Arch	80.1	76.0	1135	1490	1681
Left Loop	82.2	84.9	1276	1466	73984
Right Loop	82.6	79.9	1251	1466	68121
Tented Arch	96.0	74.5	963	1490	1089
Whorl	84.8	82.7	1235	1490	113569

Table 3: Results of testing and training for global and class-by-class neural network matching using 58 K-L features. All networks had 48 hidden nodes. Features used were based on peak correlation, the width of the correlation peak, and area under the correlation peak. Images contained 512 pixels on each axis. Combining all networks with pattern classification yields 90.9% accuracy.

Class	train	test	wts. pruned	total wts.	Test set size
All	78.3	76.0	1144	1442	258444
Arch	76.2	71.3	1276	1466	1681
Left Loop	77.6	81.9	1358	1442	73984
Right Loop	76.9	76.0	1343	1442	68121
Tented Arch	80.1	72.0	1276	1466	1089
Whorl	77.5	79.4	1381	1466	113569

Table 4: Results of testing and training for global and class-by-class neural network matching using 58 K-L features. All networks had 48 hidden nodes. Features used were based on peak correlation, the width of the correlation peak, and area under the correlation peak. Images contained 256 pixels on each axis. Combining all networks with pattern classification yields 89.3% accuracy.

Class	train	test	wts. pruned	total wts.	Test set size
All	78.3	76.0	1144	1562	258444
Arch	79.0	77.9	1315	1562	1681
Left Loop	79.0	81.7	1384	1514	73984
Right Loop	76.6	77.5	1430	1514	68121
Tented Arch	84.7	76.6	1360	1514	1089
Whorl	77.6	73.8	1433	1514	113569

Table 5: Results of testing and training for global and class-by-class neural network matching using 58 K-L features. All networks had 48 hidden nodes. Features used were based on peak correlation, the width of the correlation peak, and area under the correlation peak. Images contained 128 pixels on each axis. Combining all networks with pattern classification yields 88.66% accuracy.

Feature Type	Sample Resolution	Accuracy
Peak	20 pixels/mm	85.28%
Peak+Area	20 pixels/mm	89.95%
Peak+Area+Width	20 pixels/mm	90.9%
Peak+Area+Width	10 pixels/mm	89.3%
Peak+Area+Width	5 pixels/mm	88.66%

Table 6: Accuracy of match for different features and sample resolutions

5.6 Summary

In table 6 the global matching accuracy of all five matching experiments are compared. The largest improvement in accuracy occurs when the peak area is added to the peak height; this improves matching accuracy from 85% to 89%. Adding the peak width provides another 1% increase in accuracy but the peak width and area are sufficiently correlated that major improvements are not possible.

Decreasing the image resolution decreases matching accuracy. The reduction from 20 pixels/mm to 10 pixels/mm reduces matching accuracy from 90.9% to 89.3%. The most surprising result is that, by using FT-based features, the 5 pixel/mm image can still be used to train an 88.66% accurate matcher which is only 0.64% lower than the 10 pixel/mm case. This clearly shows that the FT-based features can be used for matching on images which have too low quality to provide clear minutia.

6 Conclusions

We have compared optical and combined optical-neural network methods for rolled inked fingerprint image matching. For static inked images, direct global optical correlation of inked images made at different times has very low reliability, although cross correlations and auto correlations of the original inked images are good. This difficulty can be accounted for by the plastic deformation of the fingerprint during rolling.

Combining zonal optical features with neural networks for classification and matching can yield reliable matching with an accuracy of 90.9%. This result was achieved using a neural classification network described elsewhere [4, 5, 6] and three components of the local FT correlation to drive class-by-class matching networks. The information content analysis of the features, both from the dimension of the K-L transform features and the generalization error analysis, show that the information transfer from the training data to the classification network is as high as the noise level of the features will allow for each K-L transform, feature set, and image resolution. The method used to achieve this optimal training is discussed in [2]. In principle, direct combination of multiple real-time images in a holographic matched filter can allow for greater stored information content in the matching process. This will be the subject of further study.

References

[1] A. VanderLugt. Signal detection by complex spatial filtering. *IEEE Trans. Inform.*

Theory, IT-10:139–145, 1964.

[2] C. L. Wilson, J. L. Blue, and O. M. Omidvar. Training Dynamics and Neural Network Performance. *Neural Networks*, 10(5):907–923, 1997.

[3] C. I. Watson. Mated Fingerprint Card Pairs. Technical Report Special Database 9, **MFCP**, National Institute of Standards and Technology, February 1993.

[4] C. L. Wilson, G. T. Candela, P. J. Grother, C. I. Watson, and R. A. Wilkinson. Massively Parallel Neural Network Fingerprint Classification System. Technical Report NISTIR 4880, National Institute of Standards and Technology, July 1992.

[5] C. L. Wilson, G. T. Candela, and C. I. Watson. Neural-network fingerprint classification. *Journal of Artificial Neural Networks*, 1(2):203–228, 1994.

[6] G. T. Candela, P. J. Grother, C. I. Watson, R. A. Wilkinson, and C. L. Wilson. PCASYS - A Pattern-level Classification Automation System for Fingerprints. Technical Report NISTIR 5647, National Institute of Standards and Technology, 1995.

[7] C. L. Wilson, P. J. Grother, and C. S. Barnes. Binary Decision Clustering for Neural Network Based Optical Character Recognition. *Pattern Recognition*, 29(3):425–437, 1996.

[8] C.S. Weaver and J.W. Goodman. Technique for optically convolving two functions. *Appl. Opt.*, 5:1248–1249, 1966.

[9] F.T.S. Yu and X.J. Lu. A real-time programmable joint transform correlator. *Opt. Commun.*, 52:10–16, 1984.

[10] J.L. Horner. Optical processing or security and anticounterfieting. *IEEE LEOS=9296 Proceedings, Boston, 18-21 November*, 1:228–229, 1996.

[11] Y. Petillot, L. Guibert, and J.-L. de Bougrenet de la Tocnaye. Fingerprint recognition using a partially rotation invariant composite filter in a FLC joint transform correlator. *Opt. Comm.*, 126:213–219, 1996.

[12] J. Podolfo, H. Rajenbach, and J-P Huignard. Performance of a photorefractive joint transform correlator for fingerprint identification. *Opt. Eng.*, 34:1166–1171, 1995.

[13] B. Javidi and J. Wang. Position-invariant two-dimensional image correlation using a one-dimensional space integrating optical processor: application to security verification. *Opt. Eng.*, 35:2479–2486, 1966.

[14] T.J. Grycewicz and B. Javidi. Experimental comparison of bianry joint transform correlators used for fingerprint identification. *Opt. Eng.*, 35:2519–2525, 1996.

[15] F.T. Gamble, L.M. Frye, and D.R. Grieser. Real-time fingerprint verification system. *Appl. Opt.*, 31:652–655, 1992.

[16] K.H. Fielding, J.L. Horner, and C.K. Makekau. Optical fingerprint identification by binary joint transform correlation. *Ope. Eng.*, 30:1958–1961, 1991.

[17] J.L. Horner and H.O. Bartelt. Two-bit correlation. *Appl. Opt.*, 24:2889–2893, 1985.

[18] D. Psaltis, E.G. Paek, and S.S. Venkatesh. Optical image correlation with a binary spatial light modulator. *Opt. Eng.*, 23:698–704, 1984.

[19] D. F. Specht. Probabilistic neural networks. *Neural Networks*, 3(1):109–118, 1990.

[20] O. M. Omidvar and C. L. Wilson. Information Content in Neural Net Optimization. *Journal of Connection Science*, 6:91–103, 1993.

[21] J. L. Blue and P. J. Grother. Training Feed Forward Networks Using Conjugate Gradients. In *Conference on Character Recognition and Digitizer Technologies*, volume 1661, pages 179–190, San Jose, February 1992. SPIE.

[22] P. J. Grother. Cross Validation Comparison of NIST OCR Databases. In D. P. D'Amato, editor, *Conference on Character Recognition Technologies*, volume 1906, pages 296–307, San Jose, 1993. SPIE.