A Novel Approach to Improve Biometric Recognition Using Rank Level Fusion

Jay Bhatnagar, Ajay Kumar, Nipun Saggar
Biometrics Research Laboratory
Department of Electrical Engineering, Indian Institute of Technology Delhi
Hauz Khas, New Delhi 110 016, INDIA
jbhatnagar@ee.iitd.ac.in, ajaykr@ieee.org, nipun.saggar@gmail.com

Abstract

This paper proposes a novel approach for rank level fusion which gives improved performance gain verified by experimental results. In the absence of ranked features and instead of using the entire template, we propose using K partitions of the template. The approach proposed in the paper is useful for generating sequential ranks and survivor lists on partitions of template to boost confidence levels by incorporating information from partitions. The proposed algorithm iteratively generates ranks for each partition of the user template. Ranks from template partitions are consolidated to estimate the fusion rank for the classification. This paper investigates rank level fusion for palmprint biometric using two approaches: (1) fixed threshold and resulting survivor list, and (2) iterative thresholds and iteratively refined survivor list. The above approaches achieve similar performances as related manifestations of fusion architecture. The experimental results support the proposition of high in-template similarity of palmprint for a user and its relevance to the intra-modal fusion framework. Experimental results using proposed approach on real palmprint data from 100 users show superior performance with recognition accuracy of 99 % as compared to recognition accuracy of 95% achieved with the conventional approach.

I. Introduction

Biometric based secure access systems are becoming increasingly popular for personal authentication. Biometrics is an integral part of a person's identity such as behavioral or physiological attributes conveyed in Face, Iris code, Fingerprint, Palmprint, Gait etc. Multi modal biometrics is a significant evolution in biometric authentication systems for the additional performance gains that can be harnessed by use of fusion algorithms [1]-[2]. Combining biometric information from multiple descriptions of a user is possible at several levels: sensor level, feature level, score level, rank level, and decision level [2]. Multiple copies of the desired signal are known to provide redundancy against channel errors at the receiver [5], this in turn reflects in reduced error rate at the receiver. Likewise, employing signal diversity substantially increases the classification performance for a MIMO (multi-transmit multi-channel) system. Incorporating signal diversity from use of multiple modes provides greater performance gain than achieved by using any one of the component modes [5]. The total information measure after fusion cannot be algebraic addition of the information measures from component modes if the modes are correlated [2]. For fusion to achieve the claimed performance enhancement, fusion rules must be chosen based on the type of application, biometric traits, and level of fusion. There has been substantial work in the biometric literature reporting fusion paradigms at score level using multiple matchers or scores from different biometric traits [6]-[7]. Authors in [8] extensively discuss classifier level fusion to conclusively illustrate superior performance using sum-product rule. Rank level fusion [9] is relevant in identification systems where each classifier associates a rank with every input template (a higher rank indicating a good match). Thus, fusion entails consolidating the multiple ranks associated with an identity and determining a new rank that would aid in establishing the final decision. There are not many techniques seen in the biometrics literature which employ rank level fusion [9]. Customarily, biometric identification is equivalent to testing hypothesis using likelihoods based on Neyman - Pearson rule. Rank level approach has appeared in the literature for testing sample size requirements and in non-parametric estimates based on order statistics [10]. Sequential algorithms in decision theory were first applied for decoding (detection-estimation) over noisy channels by Elias-Wozencraft [5]-[13]. This approach was further generalized by Robert Gallager [5] introducing the concept (in his Doctoral thesis) of low density codes as long partitions of blocks that employ multiple passes for iterative estimation (detection) on a noisy received set. Typically, for identification the input user template is matched to the database of the biometric system. Matching gives a score which is then compared to nearest user-specific score from the distribution for each user, for some pre-specified threshold. Classification is possible merely using the computed matching score, which is considered as pre-classification. Specifically, likelihood or other rule based algorithms can be incorporated to design a classifier (post-classification) under a cost criterion. In the following sections we will propose the basis for and the reasons that motivate rank level fusion. The rest of this paper is organized as follows:

(i) Give motivation for proposing rank level fusion (section II)

(ii) Show relevance of game theory to fusion rule at rank level (section III)
(iii) Propose the two models for rank level fusion architecture (section IV)
(iv) Experimental results that establish usefulness of algorithms proposed in (iii) (section V).

II. Motivation – Rank level fusion

Information content (feature, score, and decision) from user biometric template has statistical nature, wherein portions of a user template show a high degree of dependence or intra-correlation. The measure of which is self correlation or self similarity between adjoining partitions of the full template. The correlation between partitions of a template implies redundant information in the template which can be used against noise or errors. Conceptually, this is similar to the use of redundancy employed in error control (detection-estimation) mechanisms [5] to combat errors (noise) in data. It is likely that some biometric traits exhibit local feature level intra correlations in a template. This is particularly seen in the time-frequency windows for speech biometric. The Hidden Markov model and regression model for speaker recognition incorporates such in-class similarity. It is also evident in the morphological features (line and surface geometry) as in the case of palmprint. Another approach to fusion using single biometric combines multiple representations of template to extract different feature level information, with a particular representation emphasizing some type of features [3]. There are other fusion techniques cited from biometric literature [2]-[4]. The following precursors may be noted with regard to the approaches proposed in this paper:

(1) An algorithm for optimal *universal* source coding /representation [11] is also usable for low complexity decoding (classification). Universality assures and is understood in context of long block lengths and varying types of data strings;
(2) *Sliding window technique* to partition a source set has been employed widely in Lempel-Ziv algorithm [5] for long block length compression, including the maximum likelihood sequential search (detection) rule employed by Viterbi decoder [5];
(3) Essentially, memory units records favorable instances for every window and rank these instances as the window slides over the length of the observed data. A final winner is selected based on best overall rank which can be computed using *sum of ranks* rule;
(4) The proposed technique employing sliding window can be perceived to be a sequential Neyman – Pearson (N-P) framework [12], using fixed/variable thresholds, along with iterative updates of ranks at every node. The final decision from this process is a sum of likelihood ranks obtained by maximizing the conditional likelihood ranks from each iteration or partition;
(5) Identification can be split into smaller steps of template windows, wherein the approach iteratively eliminates and admits users based on a rank test. In a non-iterative framework, the sum rank of users from all windows is used for classification. Rank based non-parametric tests are widely used for test statistics in the literature [2] - [10].

In order to develop a decision theoretic basis of the approach proposed in this work, we revisit some results from probability theory - union bound, Bonferroni's inequality and the Bayes rule. Consider all such $a_i \in A$ for a user *m*, as the i^{th} window (partition) of the complete template *A*. The union bound states [12],

$$\sum_{\forall a \in A_i} P(a_i) \geq P\Big(\bigcup_{\forall a_i \in A} a_i\Big) \qquad (1)$$

Furthermore for choice of *j* and *i* as neighborhood (adjoining) partitions in user template, the partitions *j* and *i* show dependence resulting in,

$$\begin{aligned}&\sum_{\forall i} P(a_i) \leq \sum_{\forall j} P(a_j) \sum_{\forall i} P(a_i / a_j) \approx \\ &\sum_{\forall j} P(a_j) \max{}_{\forall i}[P(a_i / a_j)] = \\ &\sum_{\forall j} P(a_j) P(a_{j+1} / a_j)\end{aligned} \qquad (2)$$

In (2), R.H.S. incorporates conditional information into the measures thereby improving the total log likelihood. If only, a_i, a_j are replaced by matching scores for i^{th} and j^{th} window then likelihoods can indicate the threshold sets per partition $i, j \in K$. For high in-class similarity and high inter-class separability the approximation given in (2) works well. Thus, it is possible to operate on multiple matching scores from template partitions than using single score using the full template. Increase in confidence levels can be seen using Bonferroni's result [12]. This result together with (1) and (2) provides a basis for improved confidence levels in utilizing statistical data from dependence between partitions of a user template. The Bonferroni method shows an important result that allows many comparison sets (ranked lists used in list decoding) to be made (or confidence intervals to be constructed) while still assuring an overall confidence coefficient is maintained. The Bonferroni method is valid for equal and unequal sample sizes. Formally, the Bonferroni general inequality is presented by:

$$P\Big(\bigcap_{\forall i=1}^{k} a_i\Big) \geq 1 - \sum_{\forall i=1}^{k} P(\overline{a_i}) \qquad (3)$$

where, a_i and its complement are component scores from the windows. Probability in equations (1), (2), and (3) can be replaced by respective distributions to convey the same measure in terms of the confidence coefficients. In particular, if each a_i is the event that a calculated confidence interval for a particular matching score includes the true rank of the user, and then the left-hand side of the inequality is the probability that all the confidence intervals simultaneously cover the respective true rank values. The right-hand side is one minus the sum of the probabilities of each of the intervals missing their true values. Use of t-statistics helps compute the confidence level of true ranks combined for all the partitions, as can be applied in evaluating fusion performance. Incidentally, the normalized matching scores that generate ranks in proposed algorithms follow a student t-statistics [12] (section III). In the inception of convolutional codes, authors of [13] cite that random errors occur with greater probability in windows close to a noisy data window. Authors in [10] - [11] incorporate soft likelihoods in the sequential decoding process in working with partitions of the observed (received) data vector. As will be shown in section IV, the proposed work employs a near optimal threshold which we term as semi-soft from usage of median values [12].

III. The Algorithms

In this section we formulate the algorithms for rank level fusion. Training of the classifier is studied, where by semi soft thresholds will be generated per level for all the users. Thresholds are semi soft since during training the values of thresholds defined per level per user are chosen as median values [12]. In this paper the term window will be used interchangeably with partition of a template, while stating iteration level will refer to the matching score from the respective partition of the template.

Algorithm A:

Training Phase:

Training set comprises R templates of the total T templates for any given user. Each of the R templates per user is partitioned into K partitions or windows of fixed size. For each window of the R templates we perform matching giving ${}^{R}C_2$ genuine score combinations. The median matching score from ${}^{R}C_2$ scores will be referred as $\overline{s_i^m}$ and will be used to label the i^{th} level in the training phase [12]. K such labels (levels) will be generated for each user. The $\overline{s_i^m}$ will be later useful for computing ranks from matching score obtained from a test template. The procedure is repeated for all M users. This provides a K- basis set of scores for each of the M users as shown in Figure 2. Finally, a mean value of respective features from seven templates will be used to construct a reference template for each user as shown in Figure 2. There will be M such reference templates. This completes the training phase for the proposed framework. We summarize the first algorithm which will consolidate ranks from windows using sum rule. (Refer Figure 2):

Test Phase:

(i)Accept the test template and partition it into K windows,

(ii)Compute matching scores for the K windows of input template with respect to corresponding reference template window,

(iii)Translate this matching score into ranks for this window using following transformation:

$$\alpha_m^i = \frac{s^m{}_i - \overline{s_i^m}}{\max(s_i^m - \overline{s_i^m})} \qquad (4)$$

where, s_i^m is the i^{th} level score computed by matching the i^{th} test window with i^{th} reference window. $\overline{s_i^m}$ are matching scores from training phase corresponding to the i^{th} component of K-basis set of m^{th} user. We also note that α actually lies in the interval [0, 1]. For computing ranks at the i^{th} partition for all m users' use $w_{m(i)} = 1 - \alpha_m^i$. Here, $w_{m(i)}$ are weights which lie in [0, 1] intervals and denote ranks of each user,

(iv) Store the *list* of ranks from each level in lists: $\{l_i\}, \forall i \in K$, by rejecting all users with rank < γ,

(v)Repeat steps (ii) to (iv) until all partitions (levels) have been exhausted,

(vi) The algorithm can now output a final winner using the lists $\{l_i\}$ based on the simple *sum rule* optimization given in (3). Make note that (4) gives semi-soft thresholds incorporated into the rank lists per level, thus a sum rule can be based using a variation of (3).

$$winner = \{S : \max(\sum_{\forall l_i} w_{l_i})\} \qquad (5)$$

(Experimental results on this algorithm will follow in section IV.) An optimal rule in the choice of window size can be based by employing Kullback-Leibler divergence measure, if incorporating uniform window sizes as follows: The following Information theoretic measure is also useful in choosing the partition size $\boldsymbol{i}$ that minimizes the Kullback-Leibler divergence or Information Divergence [5] given by,

$$i^m = \{k : | \max \ln[p(user = m / f_k{}^m) / p(user = m / f_i{}^m)]\} \qquad (6)$$

Fusion rule using variable window size is not in the scope of the present work.

Summary of rank level fusion architecture (Figure 1)

A test template is partitioned into K windows. K-scores are computed from matching the partitions of the test template with corresponding partitions from each of the M reference template. Each of the K-scores generated from above is compared with the corresponding score level from M user training set. The matching scores at each level are converted into corresponding ranks by the score to rank converter block. This results in a rank for each user at every level with independent ranks per user from each level. Since, ranks lie in the interval [0, 1], probabilities or ranks can be considered as a random variable which gives maximum entropy at 0.5 and we incorporate this as threshold γ, per window. The final ranks assigned to survivor (user) from the threshold block are summed in the summer block. Finally, the user with maximum sum rank is declared a winner. This algorithm comes close to techniques using Majority voting and Borda count [2].

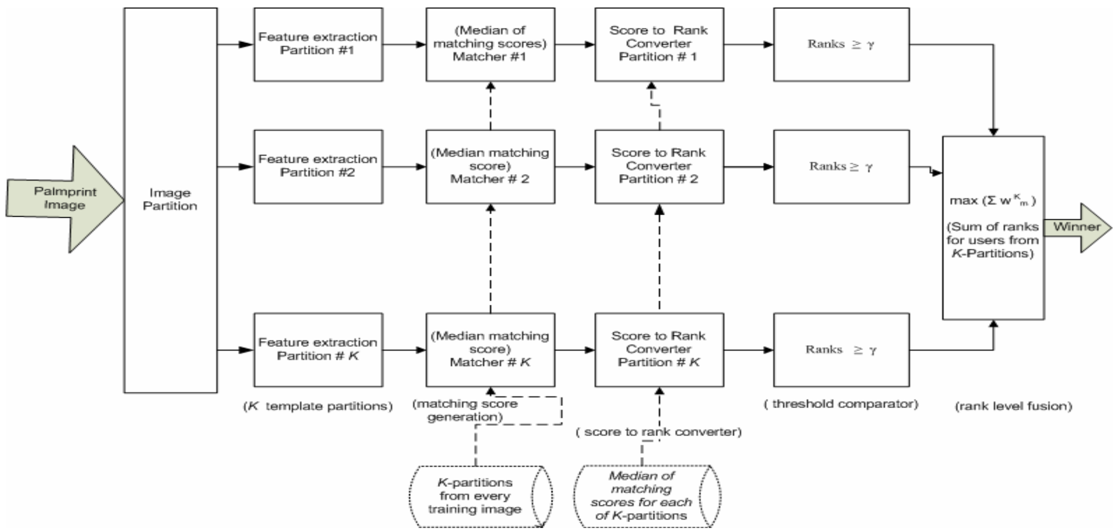

Figure 1. Architecture for rank level fusion using algorithm A.

***Algorithm B*:**
Training Phase: This is identical to algorithm A.
Test phase:
The following summarizes the second algorithm in which consolidation of ranks is achieved by successive refinements in the list (rank level data) from a level to next level. The size of threshold also increases for every successive partition. The maximization of rank in the final list at last iteration (level) provides the rule for classification.

(i) Accept the test template and partition it into *K* windows,
(ii) Compute matching scores for the *K* windows of input template with respect to corresponding reference template windows (Figure 2),
(iii) Translate matching scores into ranks for this window using transformation given in (4). Also, to compute ranks at the i^{th} level for all *m* users we use, $w_{m(i)} = 1 - \alpha_m^i$. Here, $w_{m(i)}$ are weights which lie in [0, 1] intervals and denote ranks of each user,
(iv) In case of first list, reject all users with rank below γ. In case of every subsequent list consolidate ranks of the user from the previous list and current list using:

$$\hat{w}^m{}_{i+1} = \hat{w}^m{}_i + w^m{}_{i+1} \qquad (7)$$

Where, $\hat{w}^m{}_{i+1}$ is current update of rank for m^{th} user at end of *(i+1)* th level and $w^m{}_{i+1}$ is the rank within the *(i+1)* th level. This shows that every current list is input to next level and rank consolidation is a sequential logic.

(v)Reject all users in the list, at the end of *(i+1)* th level that satisfy $\hat{w}_{i+1}^m \le i\gamma$; and retain users in the list if, $\hat{w}_{i+1}^m \ge i\gamma$ note that the threshold set at the end of level *i* will be $i\gamma$. Compute the lists at the end of each level as $\{l_i\}$
Transmit the list at end of level *i* (fusion list) to next level (*i*+1), at each pass of the iteration.
(vi) Repeat steps (ii) to (v) until $i = K$. The final list $\{l_K\}$ comprises all survivors consolidated iteratively from component lists or levels. The final winner is chosen using, $Winner = \{S : | \max(w_{l_K})\}$
Alternate order statistics than that proposed above can also be employed to achieve fusion rank.
Summary of rank level fusion architecture (Figure 2):
The first two blocks remain identical in both algorithms. As given in step (iv) of the algorithm B, the third block in Fig.3 will consolidate ranks from immediate prior level and add the rank at the present level to generate a new rank. The new rank is compared with a threshold for this level (threshold are iteratively defined) to reject all users below the threshold. Finally, the process of successive refinements (short listing of users) continues till the last level at which the highest ranked user from the list is declared winner.

IV. Experimental work

In order to estimate the recognition accuracy for rank level fusion algorithms in section III, we perform experiments on real biometric samples. Palmprint images of 100 users (10 images of each user) employed in [14] were used to extract palmprint features. The image normalization and feature extraction is same as detailed in [14]. The matching scores were generated using Euclidean distance. The classifier was trained using seven samples per user and was tested with three samples per user. Figure 3 shows the performance for varying threshold using window size of 12. The window size was varied to ascertain the performance of algorithms A and B for varying number of partitions (2, 4, 8, 12 and 24)

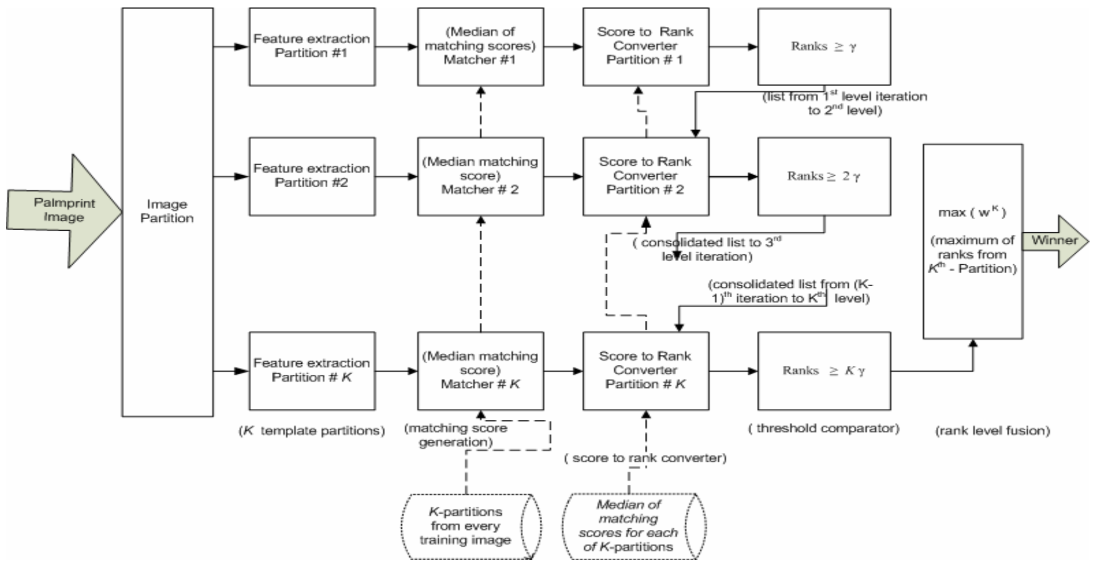

Figure 2. Architecture for rank level fusion using algorithm B

for threshold γ fixed at 0.5. As illustrated in the Figure 4, algorithms A and B achieve the same performance gains for the variations on window size. A viable explanation can be offered accounting for the similar performance gains achieved by both algorithms. The partitioned matching scores are highly dependent from a level to next level. Figure 4 substantiates that strong dependence between inter-level scores for input template gives similar recognition accuracy both for independent list (algorithm A) and sequential list approach (algorithm B). Typically, a useful observation made during experiments was that the recognition accuracy in the proposed framework of training and testing remained consistent for all users in terms of total number of errors for any user. However, algorithm B has lower complexity than algorithm A since algorithm B incorporates successively smaller size of list of survivors with each of the iterations. When employing the conventional approach of using entire template (without partitions) and incorporating user-specific thresholds, the recognition accuracy varied for users from 0.78-0.95 with an average of 0.91 for the same data base. In this paper, the recognition accuracy has been abbreviated as R.A.

The performance gains with smaller window size can be explained from results that were introduced in section I. As observation space of a biometric template is reduced (with decrease in size of partition) the confidence levels contained in partitions become essentially uniform, therefore a sequential setting of iterative costs (ranks) from a window to next window will give superior average confidence estimate. Experimental work in this section may not directly include the case of variable window size, but we conjecture that variable size will not improve performance (Figure 4).

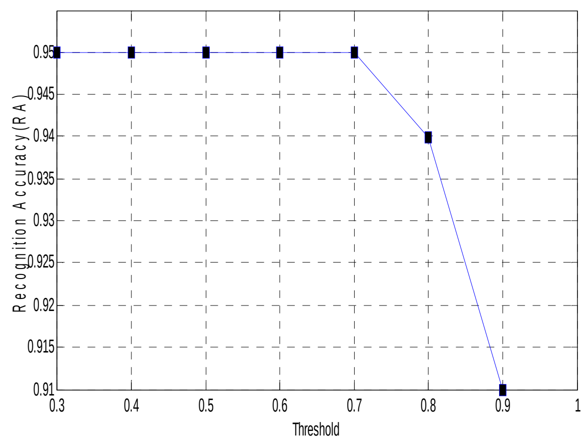

Figure 3. R.A. for variable thresholds, window 12

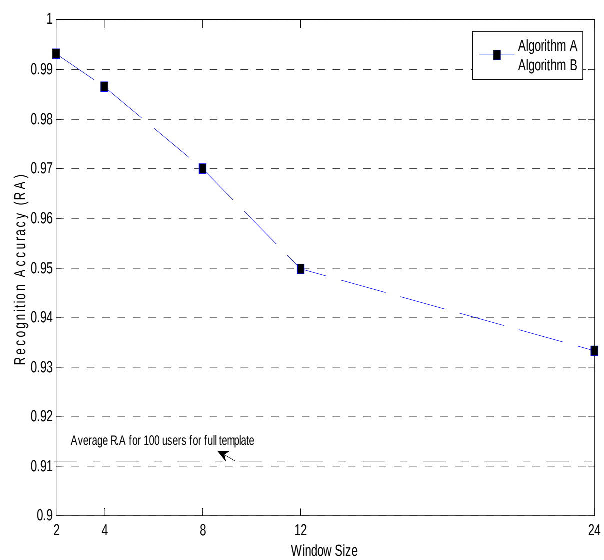

Figure 4. R.A. for variable window, threshold 0.5

V. Conclusions and future work

(a) Conclusions

This paper presented a new approach to rank level fusion which achieves a superior performance gain than conventional approach. Experimental results (section IV) using algorithms A & B on a real palmprint for 100 users confirm that recognition accuracy indeed improves to 99 % and it is consistent for all users. It is also pointed that algorithm B gives lower computational complexity by incorporating iteratively refined lists through multiple passes as compared to algorithm A. Thus a variation of Elias-Wozencraft approach successfully achieves the proposed goal of rank level fusion. Pattern classification based on the proposed algorithm is a statistical recursive algorithm which has been justified using results in statistics (section I), further substantiated by experiments (section IV). Based on criterion using second order statistics, a low cross correlation and high self correlation is required of biometric sequences. The proposed approach will not be very effective with biometric traits that show low self similarity in morphological feature variations or temporal feature variations.

(b) Future work

Our current focus is to extend the proposed work to co-operative (and live biometrics) multi modal biometrics such as speech-face, face-lips etc. to estimate the performance gains and understand computational complexity. Another insight which springs from algorithm A proposed here is variation of sum rule to achieve rank level fusion. Particularly, a method such as weighted sum fusion rule which is known to give a superior fusion classifier performance is a good rule to consider in extending the present work.

VI. Acknowledgement

This work is partially supported by the research grant from Ministry of Information and Communication Technology, Government of India, grant no. 12(54)/2006-ESD.

VII. References

[1] A K. Jain, R M. Bolle and S. Pankanti (Eds.), *BIOMETRICS*: *Personal Identification in Networked society*, Kluwer Academic Publishers, 1999

[2] A. Ross, K. Nandakumar, and A K. Jain, *Handbook of Multibiometrics*, Springer Verlag, 2006

[3] A. Kumar, and D. Zhang, "Personal authentication using multiple palmprint representation," *Pattern Recognition*, vol. 38, pp. 1695-1704, 2005

[4] A. Kumar, and D. Zhang, "Integrating shape and texture for hand verification," *International Journal of Image & Graphics*, vol. 6, no. 1, pp. 1-13, 2006.

[5] R G. Gallager, *Information Theory and Reliable Communication*, John Wiley and Sons, 1963.

[6] M. Indovina, U. Eludes, R. Snelick, A. Mink, and A K. Jain, "Multimodal Biometrics Authentication Methods: A COTS Approach," *Proc. MMUA*, pp 99-106, CA, 2003.

[7] S. Dass, K. Nandakumar, and A K. Jain, "A Principled Approach to Score Level Fusion in Multimodal Biometric Systems," *Proc. of AVBPA 2005*, pp. 1049-1058, NY, 2005.

[8] F. Roli, and J. Kittler, *Multiple Classifier Systems*, Springer–Verlag, Lecture notes in Computer Science, vol. 2364, 2002.

[9] A. K. Jain, A. Ross and S. Pankanti, "Biometrics: A Tool for Information Security," *IEEE Trans. Info. Forensics & Security, v*ol. 1, no. 2, pp. 125-143, 2006.

[10] J. Bhatnagar, and A. Kumar, "On Estimating Sample Size Requirements for Reliable Personal Identification," *Proc. of CVPRW*, pp. 36, NY, 2006.

[11] A. D. Wyner, J. Ziv, and A J. Wyner, "On the Role of Pattern Matching in Information Theory," *IEEE Trans. Info. Theory, v*ol. 44, pp. 2045-2056, 1998.

[12] V K. Rohatgi and Md. E. Saleh, *An Introduction to Probability and Statistics*, Wiley Series in Probability & Statistics, 2001.

[13]J M. Wozencraft and B. Reiffen, *Sequential Decoding*, Cambridge, MA, MIT Press and Wiley.

[14] A. Kumar, and D. Zhang, "Personal recognition using hand-shape and texture," *IEEE Trans. Image Process.*, vol. 8, pp. 2454-2461, Aug. 2006.