International Journal of Electrical and Computer Engineering (IJECE)
Vol. 4, No. 5, October 2014, pp. 817~830
ISSN: 2088-8708

Fingerprint Direct-Access StrategyUsing Local-Star-Structure-based Discriminator Features: A Comparison Study

G Indrawan*, S Akbar*, B Sitohang*
* Data & Software Engineering Research Division,
School of Electrical Engineering and Informatics, Institut Teknologi Bandung

Article Info

Article history:

Received Aug 18, 2014
Revised Sep 20, 2014
Accepted Oct 1, 2014

Keyword:

Direct-access
ErrorRate
Fingerprint
Local-star
Penetration Rate

ABSTRACT

This paper describes a comparison study of the proposed fingerprint direct-access strategy using local-star-topology-based discriminator features, including internal comparison among different concerned configurations, and external comparison to the other strategies. Through careful minutiae-based feature extraction, hashing-based indexing-retrieval mechanism, variable-threshold-on-score-ratio-based candidate-list reduction technique, and hill-climbing learning process, thisstrategy was considered promising, as confirmed by the experiment results. For particular aspect of external accuracy comparison, thiss trategy outperformed the others over three publicdata sets, i.e. up to Penetration Rate (PR) 5%, it consistently gave lower Error Rate (ER). By taking sample at PR 5%, this strategy produced ER 4%, 10%, and 1% on FVC2000 DB2A, FVC2000 DB3A, and FVC2002 DB1A, respectively. Another perspective if accuracy performance was based on area under curve of graph ER and PR, this strategy neither is the best nor the worst strategy on FVC2000 DB2A and FVC2000 DB3A, while on FVC2002 DB1A it outperfomed the others and even it gave impressive results for index created by three impressions per finger (with or without N_T) by ideal step down curve where PR equal to 1% can always be maintained for smaller ER.

Corresponding Author:

G. Indrawan,
Data & Software Engineering Research Division,
School of Electrical Engineering and Informatics, Institut Teknologi Bandung,
Jl. Ganesha 10, Bandung 40243, Indonesia.
Email: gindrawan@live.com

1. INTRODUCTION

Recent performance comparison in area of fingerprint direct-access strategy [1], leaving further question on how far its accuracy, efficiency, and scalability performance can be improved. In general, direct-access strategy itself means any searching strategy to outputa candidate list (CL) without performing 1-to-1 matching between a query and candidates in the database. ACLfrom a query will havea list of certain Error Rate (ER) atcertain Penetration Rate (PR). The ER is the average percentage of searched queries that are not found, andthe PR isthe portion of the database to be searched on the average. The accuracy performance isthen measured bythegraph of the trade-off between ER and PR that shows atcertain ERhow low PR can be achieved. The efficiency performance is considering as strategy's search speed and memory usage.

Based on above question and motivation to answer it through new fingerprint direct-access strategy, initial work has been conducted by authors [2]. As far as authors' knowledge, the proposed strategy by this work fill in non-existing exploration area in direct-access strategy based on 1-to-1 matching using local-star-structure that was introduced first by Ratha et al. [3]. This proposed strategy will then be compared to the other state-of-the-art strategiesin this paper. Several proposed fingerprint direct-access strategies havebeen roughly classified by [4], i.e. 1) using global features such as global ridge-line frequency [5]; 2) local features such aslocal ridge-line frequency, local ridge-line orientations, andlocal features from the orientation image ([5], [6], [7], [8], [9], [10]); 3) minutiae featuressuch asgeometric features from triplets of minutiae points and perform searching through hashing strategy ([11], [12], [13], [14]);4) other features such asFingerCode, ridge curvature,and SIFT features ([14], [15], [16]), and matching scores ([17], [18]).

The proposed strategy closes to minutiae-based approach above but instead of using tripletsof minutiae, it uses local-star-structure of minutiae. This paper reports experiments on three publicly available benchmarks and its results prove that the proposed strategy was consideredas a promising strategy since it compares favorably with the other state-of-the-art strategies.

Based on previous mentioned perfomace indicator, the rest of this paper is organized as follows. Section 2 illustratively describes the proposed strategy that was initiated by Indrawan et al. [2] (its scalability performance was already given on it). Section 3 describes experiments on public data sets, internally compare itamong different concerned configurations, andexternally compare it with nine published strategies. It describes data sets that was used (Section 3.1), internal accuracy comparison among different configurationsand external accuracy comparison to the other strategies (Section 3.2), internal speed comparison among different configurations (Section 3.3), and internal memory usage comparison among different configurations (Section 3.4) related to the hashing system used by the proposed indexing and retrieval process. Section 3.3 and 3.4 cannot do external comparison because of different hardware platform with the other strategies. Finally, Section 4 draws the conclusion and Section 5 suggests immediate improvement for the proposed strategy through the future work.

2. RESEARCH METHOD

The proposed strategy was constructed by minutiae-based feature extraction process, hashing-based indexing-retrieval mechanism, variable-threshold-on-score-ratio-based candidate-list reduction technique, and hill-climbing learning process.Its mathematical perspective was elaborated in detail at [2]. Figure 1 ilustrates simplified process of the proposed strategy.

1. Feature extraction process was based on minutiae detection algorithm [19], [20], where minutia type (ending or bifurcation), minutia cartesian-coordinate, and minutia absolute-orientation (minutia angle to the horizontal line) are recorded. Based on this feature extractionresult, the algorithmidentifies certain number of minutiae edgesthrough identification of local-star structure belong to each minutia-reference (sample structurein Figure 1: dashed-line encircledminutia-reference m_1 with several dashed-lines connecting to minutia-neighbours). Each edge defines a line whose geometric features areextracted: its length from its minutia-reference to its minutia-neighbour, its minutia-referencerelative-orientation (angle difference between minutia-referenceabsolute-orientation to the edgeabsolute-orientation), and its minutia-neighbour relative-orientation (angle difference between minutia-neighbour absolute-orientation to the edge absolute-orientation).Hipotetically, the similarity between two fingerprints is defined by the number of corresponding edges that can be found under retrieval process on the next step.
2. Instead of explicitly comparing the similarity between the query fingerprint and all the candidate fingerprintsin the database (which would be very time consuming), the authors use a geometrichashing technique [21] for indexing process: a hash table is built by quantizing certain number of the edges above and for eachquantized edge, a list of pointers (ID) to the fingerprints in the database containing that specificedgeis maintained. When a new fingerprint is inserted in the database, its edges are extracted, and the hash table is updated by adding the fingerprint IDs and their corresponding fingerprint edges, to the hash values associated to the hash keys. A hash value was constructed by the*list* to accommodate *collision* that certainly happened. *Collision*will be happened when the same hash key was generated from different edges that could come from same or different fingerprint ID. Good design of hash function for the hash key would minimize *collision*. For the proposed strategy, the hashkey was based on 32-bit integer value constructed by previously described three discriminator attributes, i.e. the edge length (16-bit Least Significant Bit / LSB), minutia-reference relative-orientation (23^{th} - 16^{th} bit), andminutia-neighbour relative-orientation (8-bit Most Significant Bit / MSB).
3. At retrieval time, the edges of the query fingerprint are computed and, for each edge, the list of fingerprint IDs in which that edgeis present is retrieved. Intuitively, if the same fingerprint ID is hit by more edges in the query, then it is more likely that the corresponding fingerprint is the searched one. But through experiment, an edge comparison relatively still not reduces the search space significantly. So the proposed strategy uses comparison of connected-two-edges (samplein Figure 1: connected edges with minutia m_1, m_2, and m_3 by fingerprint ID X) and connected-three-edges (samplein Figure 1: connected edges with minutia m_1, m_2, m_3, and m_4 by fingerprint ID X). Above three edges which are connected, the computational cost is exponentially more expensive (exponentially longer time execution). Amulti-stage

similarity score computation is applied to obtain afinal ranking, which is used for visiting the database in a convenient order. It consists of relatively-fast-less-accurate similarity score computation at pre-filter stage (s_R), and relatively-slow-more-accurate similarity score computation at matcher stage (s_M). At pre-filter stage, an initial rankingwas descendingly ordered by pre-filter score, s_R.

$$s_R = s_{R_2} + w_{R_3} \cdot s_{R_3} \quad (1)$$

wheres_{R_2} is the number of connected-two-edges belong to certain fingerprint ID, s_{R_3} is the number of connected-three-edges belong to certain fingerprint ID, and w_{R_3} is the weighting value which is determined through the learning process on training data (Section 3.1). At matcher stage, a final ranking which is an update from an initial ranking was descendingly ordered by final score, s.

$$s = s_R + w_M \cdot s_M \quad (2)$$

where $s_M$is the maximum value of the addition of several weighted parameters derived from the longest constructed edges (may be disconnected) belong to certain fingerprint ID and its counterpart, the longest constructed edges (may be disconnected) belong to fingerprint ID X. Those parameters including number of edgepairs, number of same-type-minutiapairs, edge length difference accumulation of edgepairs, minutia-reference'srelative-orientation difference accumulation of edgepairs, and minutia-neighbour'srelative-orientation difference accumulation of edgepairs [2]. w_M is the weighting value which is determined through the learning process on training data (Section 3.1). Figure 1 shows simple retrieval process with w_{R_3}= 1 and w_M = 1, which are not actuall values used by the proposed strategy.

4. Candidate-list reduction mechanism was applied on retrieval process above. It outputs certain number of final candidates through combined techniques of variable threshold on score ratio [22] at pre-filter stage, and fixed-length truncation of candidate-list at matcher stage.
5. Finally, hill-climbing learning process [23] was used on training data to obtain optimal values for parameter-set of algorithm. These values for parameter-set were used on several testing data (Table 1).

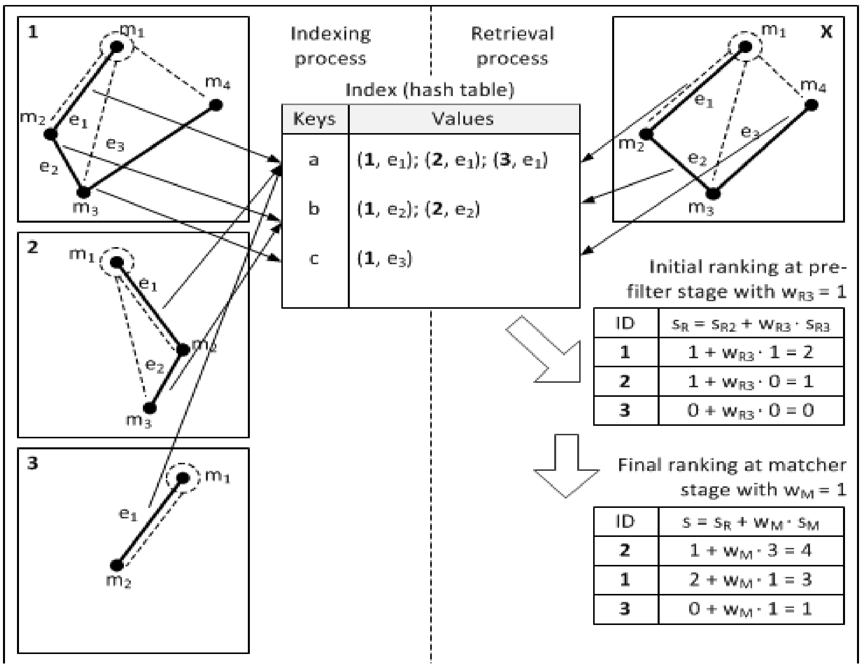

Figure 1. During the indexing phase, the extracted features from each fingerprint in the database are used to generate the *hash-keys* (a, b, c). Each keymaintains the list of fingerprint IDs (*1*, *2*, *3*) and their corresponding edges (e_1, e_2, e_3). During the retrieval phase, edges of the query fingerprint *X* are computed and the list of fingerprint IDs in which that edge is present is retrieved. Multi stage similarity score computation then outputs candidate list withits final ranking which is an update from theinitial ranking.

3. RESULTS AND ANALYSIS

This section describes experiments carried out to evaluate accuracy and other aspects of the proposed fingerprint direct-access strategy and to compare it with the other state-of-the-art in this field. The algorithm is a C# implementation, running on an Intel Core 2 Quad CPU @ 2.40 GHz. Experiments using same performance evaluation and parameters calibration in [2].

3.1. Data Sets

Table 1 shows public data sets[1] as testing data for most of the published fingerprint direct-access strategies [1]. It also shows public data set used as training data for the proposed strategy [2]. Moreover, more description of their acquisition specification [24], [25] was shown by Table 2. Image samples of those public data sets were shown by Figure 2.

Table 1. Public data sets considered for evaluation for direct-access strategies.

No.	Data Set	Data Set Status	Sensor Type	Pixels (w x h)	Image Number (w x d)	Resolution (dpi)	Published Strategies
1	FVC2000 DB1A[24]	Training Data	Low-Cost Capacitive Sensor	300 x 300	800 x 8	500	Indrawanet al. (2014) [2]
2	FVC2000 DB2A[24]	Testing Data	Low-Cost Capacitive Sensor	256 x 364	800 x 8	500	Capelli (2011) [1] Liang et al. (2006) [13] Jiang et al. (2006) [5] De Boer at al. (2001) [14]
3	FVC2000 DB3A[24]	Testing Data	Optical Sensor	448 x 478	800 x 8	500	Capelli (2011) [1] Jiang et al. (2006) [5]
4	FVC2002 DB1A[25]	Testing Data	Optical Sensor	388 x 374	800 x 8	500	Capelli (2011) [1] Shuai et al. (2008) [16] Liang et al. (2007) [26]

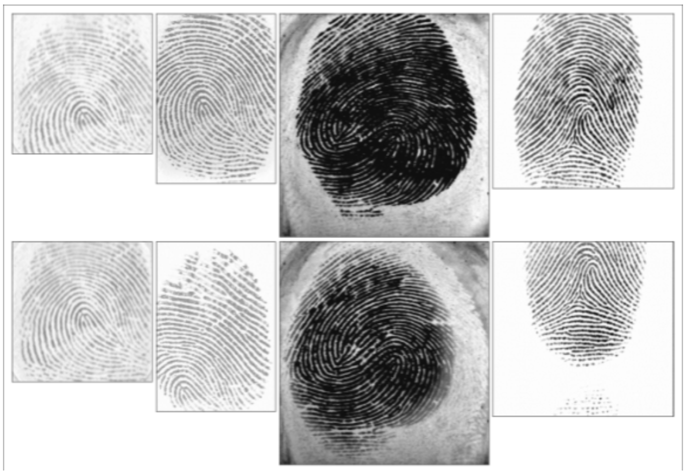

Figure 2. From left to right: sample fingerprints from FVC2000 DB1A, FVC2000 DB2A, FVC2000 DB3A, and FVC2002 DB1. First impressionswere used for index creation (top row), and the others (bottom row) were used for queries. Noted, images are not in their actual size but they are in their scale difference.

[1] Beside FVC (Fingerprint Verification Competition) public data sets, several published strategies have been also evaluated on commercial public data sets, i.e. NIST DB4, NIST DB4 (Natural), and NIST DB14

Table 2. Public data sets acquisition specification.

No.	Aspect	FVC2000 DB1A & DB2A [24]	FVC2000 DB3A [24]	FVC2002 DB1A [25]
1	Volunteer	Students (20 to 30 yearold; about 50% male).	19 volunteers (5 to 73 yearold; 55% male; 1/3 of them were over 55 yearold; 1/3 of themwere under 18 yearold; 1/6 of them were under 7 yearold).	30 students (20 year old on the average).
2	Session	Fore and middle finger of both the hands (four fingers) of each volunteer were acquired in two sessionsby interleaving the acquisition of the different fingers (e.g., first sample of left fore, first sample of right fore, first sample of left middle, first sample of right middle, second sample of the left fore, and so on).	Two images of six fingers (thumb, fore, and middle finger of both the hands) of each volunteer were acquired in four sessions without interleaving. No more than two sessions a day. The time between the first and last sessions was at least 3 days and as long as 3 months, depending upon volunteer.	Foreand middle finger of both the hands (four fingers) of each volunteer were acquired in three sessions by interleaving the acquisition of the different fingers. The time between each session was at least two weeks. At each session, four images were acquired of each of the four fingers of each volunteer. At the 2nd session, theywere requested to exaggerate finger displacement (image 1 and 2) and rotation not to exceed 35° (image 3 and 4). At the 3rd session, fingers were alternatively dried (image 1 and 2) and moistened (image 3 and 4).
3	Sensor Platen& Image Quality	The sensor platens were not cleaned systematically. The images were taken from untrained people and no efforts were made to control image quality.	The sensor platen was cleaned systematically between each acquisition. At one session, each volunteer's fingers were cleaned with rubbing alcohol and dried.	The sensor platen was not cleaned systematically. The images were taken from untrained people and no efforts were made to control image quality.
4	Core & Delta	Fingerprint cores and deltas are not guaranteed exist since no attention on checking the correct finger centering.	Fingerprint core was exist but care was taken to avoid complete overlap between consecutive images taken at a session.	
5	Rotation & Non-Null Overlapping Area	Maximum rotation is about 15 and non-null overlapping area between any two impressions of the same finger.	Maximum rotation is about 15° and non-null overlapping area between any two impressions of the same finger.	Maximum rotation is about 35°and non-null overlapping area between any two impressions of the same finger.

3.2. Accuracy Comparison

Figure 3 - 5 show the trade-off between PR and ER for several fingerprint direct access strategies on three FVC data sets.

1. All of the results, except them with an asterisk mark (Table 3), were obtained by using first fingerprint impression from each finger for index creation, and the remaining seven for queries (In this case, the proposed strategy results were represented by black dashed-lines). The results with an asterisk mark wereobtained by using first three fingerprint impressions from each finger for index creation and the remaining five for queries (In this case, the proposed strategy results were represented by grey dashed-lines). Shuai et al. [16] and Liang et al. [26] arethe exceptions. As shown by Table 3, they used random impression or random three impressions from each finger for index creation. This will have a consequence on unpredictable high quality of created index since first impression or first three impressions, as it represents or they represent higher quality image(s) than others from the same finger, was or were not selected for sure. Because of that consequence, random selection not so appropriate for head to head comparison. However, they are still worthy shown to enrich comparison study in this paper.
2. All of the results, except them with a plus mark (Table 3), were obtained by using 800 fingerprints of testing set "A" [24] [25] from 100 fingers. The results with a plus mark were obtained by using additional 80 fingerprints of training set "B" [24] [25], resulting in 880 fingerprints from 110 fingers. In another word, the results with a plus mark were obtained by using additional 10 impressionsfor index creation and 70 impressionsfor queries. Both of the results cannot be united for head to head comparison. However, the results with a plus mark are still worthy shown to enrich comparison studyin this paper.
3. All of the results, except them with a hash mark (Table 3), were obtained without selecting Top 10% Scores ($N_T = N/10$), where $N$is number of impressions used by index creation. The results with a hash mark were obtained by selecting Top 10% Scores.
4. Based on those previous points, the proposed strategy provides three kinds of results, each represented by previous mentioned black- and grey- dashed-lines. These three kinds of results each comes from three configurations that were concerned by this research for the internal comparison, and the external comparison with other existing strategies. These three configurations are for search mode: 1) uptopre-filter stage (dashed-lines with square mark); 2) up to matcher stage with N_T truncation of CL (dashed-

lines with triangle mark); and 3) up to matcher stage without N_T truncation of CL (dashed-lines with round mark). For the next discussion, three black dashed-lines with square, triangle, and round mark on the graphs, each will be referred as Pre-filter, Matcher, and Matcher2. Whilst, three grey dashed-lines with square, triangle, and round mark on the graphs, each will be referred as Pre-filter*, Matcher*, and Matcher2*. Noted, no specific square, triangle, and round legends for the proposed strategy for the sake of simplicity of the graphs. As Liang et al. [13] [26] stated that the fingerprint identification can be divided into fingerprint indexing and fingerprint verification, Matcher, Matcher2, Matcher*, and Matcher2* of the proposed strategy could be considered as sort of fingerprint verification.

5. This point refers to the internal comparison of the proposed strategy.Graph results confirmed that Pre-filter/Pre-filter* was less accurate than Matcher/Matcher*, whilst Matcher/Matcher* was less accurate than Matcher2/Matcher2*. By design, Pre-filter/Pre-filter* refer to the search powered by relatively-fast-less-accurate similarity score computation. They gave the search output in the form of initial CL with its initial candidates rank. At the other side, Matcher/Matcher*or Matcher2/Matcher2* refer to search powered by relatively-slow-more-accurate similarity score computation. They refine initial CL to become final CL with its final rank, which was more accurate on rangking order. Matcher/Matcher* loss its accuracy in certain degree compare to Matcher2/Matcher2* because of its final CL truncation by the second step of CL reduction [2] which is simply truncating length of CL to N_T if length of CL longer than N_T. Noted on Figure 4, Pre-filter result is out of graph view. Furthermore based on graph results, the proposed strategy need to improve its algorithm to gives relatively small different result between Matcher and Matcher2, or between Matcher* and Matcher2*, specifically for result on FVC2000 DB2 (Figure 3) and FVC2000 DB3 (Figure 4).
6. This point refers to external comparison of the proposed strategy (Matcherand Matcher*) to the other strategies utilizing N_T (strategies with hash mark at Table 3). On Figure 3 and Figure 4, up to PR equal to 4%, Matcher gave lower ER rather than Capelli [1]. Larger than PR equal to 4%, Capelli [1] has faster rate of decrease of its ER to PR, so it gave lower ER rather than Matcher. On Figure5, at all PR on the graph, Matcher gave lower ER rather than Capelli [1], and Matcher* gave lower ER rather than Capelli* [1]. Moreover, Matcher* gave step down curve which is ideal curve where it can maintain its PR equal to 1% for smaller ER. It means for every query, search result of top one of CL (1% of number of fingerprint at index) always give correct candidate.

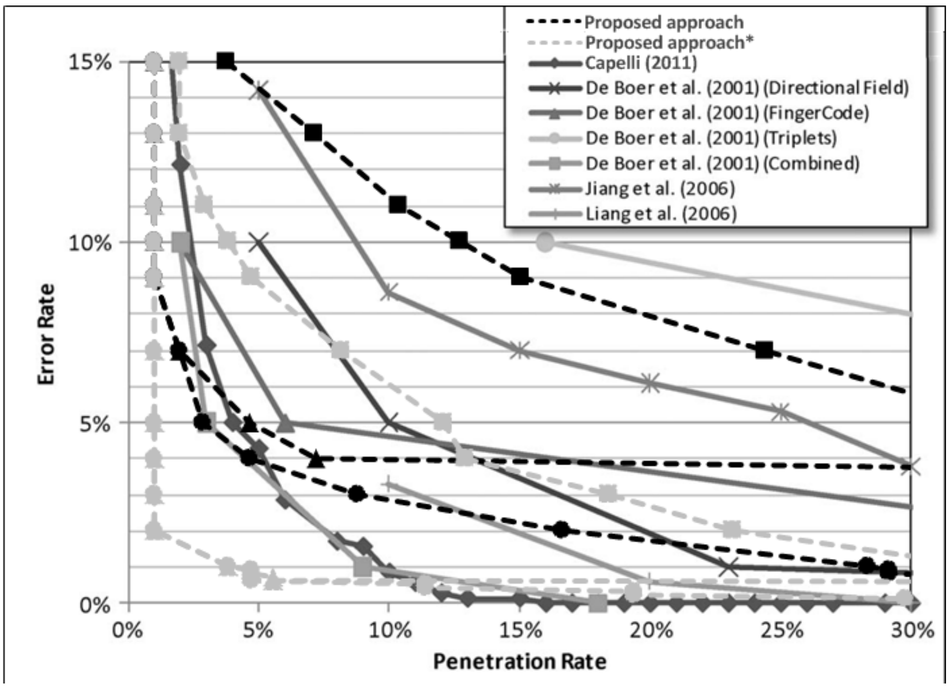

Figure 3. Accuracy of proposed direct-access strategy on FVC2000 DB2.

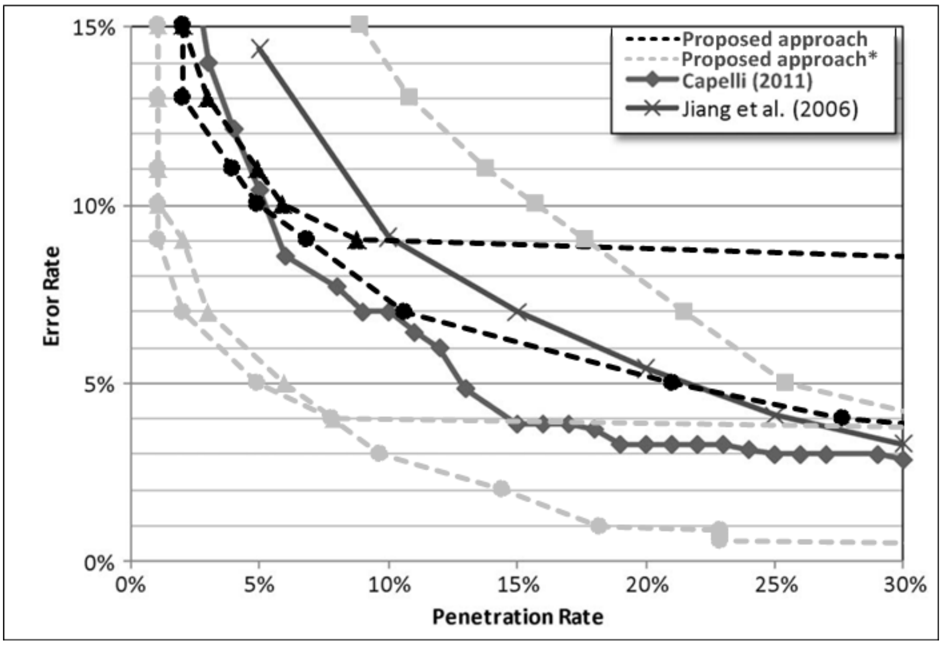

Figure 4. Accuracy of proposed direct-access strategy on FVC2000 DB3.

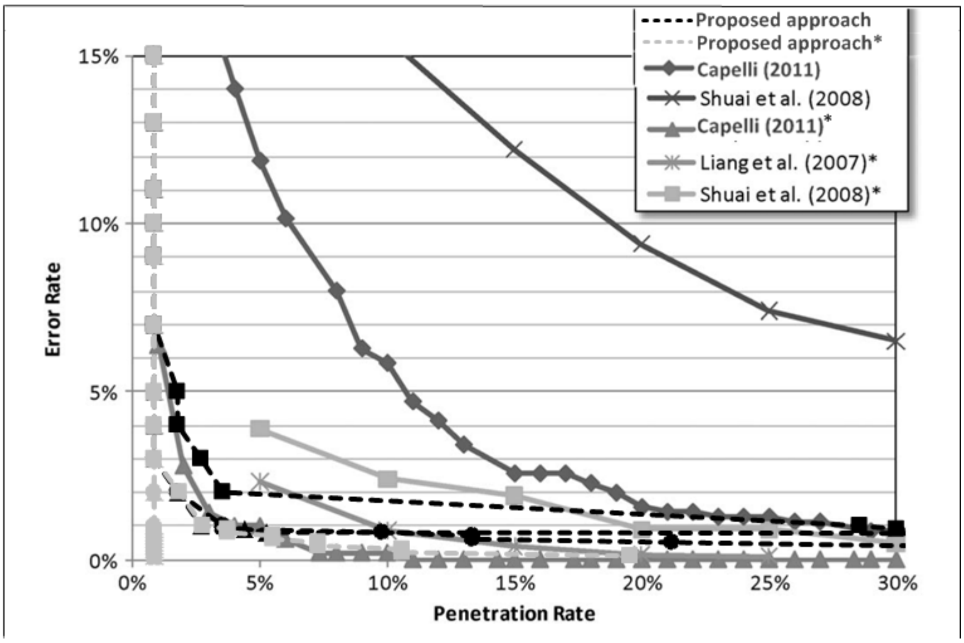

Figure 5. Accuracy of proposed direct-access strategy on FVC2002 DB1.

Table 3. The external accuracy comparisonof direct-access strategies.

No.	Data set	Proposed approach	Published strategies	Index from each finger
1	FVC2000	Matcher2	De Boer at al. (2001)+[14]	First impression
	DB2A[24]	Matcher2	Jiang et al. (2006) [5]	First impression
		Matcher2	Liang et al. (2006)+[13]	First impression
		Matcher#	Capelli (2011)#[1]	First impression
2	FVC2000	Matcher2	Jiang et al. (2006) [5]	First impression
	DB3A[24]	Matcher#	Capelli (2011)#[1]	First impression
3	FVC2002	Matcher2	Shuai et al. (2008) [16]	Random impression
	DB1A[25]	Matcher#	Capelli (2011)#[1]	First impression
		Matcher2*	Liang et al. (2007)* +[26]	Random three impressions
		Matcher2*	Shuai et al. (2008)*[16]	Random three impressions
		Matcher* #	Capelli (2011)* #[1]	First three impressions

+ Results have been obtained by using additional 80 fingerprints of training set “B” [24] [25], resulting in 880 fingerprints from 110 fingers.
* Results have been obtained by using three fingerprint impressions from each finger for index creation, instead of one.
# Results have been obtained by selecting Top 10% Scores ($N_T = N/10$).

7. This point refers to external comparison of the proposed strategy (Matcher2 and Matcher2*) to the otherstrategies not utilizing N_T (strategies without hash mark at Table 3). On Figure3, up to PR equal to 5%, Matcher2 gave lower ER rather than all other strategies. Higher than PR equal to 5%, De Boer et al. [14] (Combined) gave lower ER rather than Matcher2. Higher than PR equal to 13%, Liang et al. [13] gave lower ER rather than Matcher2. Higher than PR equal to 22%, De Boer et al. [14] (Directional Field) gave lower ER rather than Matcher2. Unfortunately, Matcher2 cannot be compared head to head to De Boer et al. [14] and Liang et al. [13] because of different number of data set used (see Table 3). Moreover, De Boer et al. obtained those results by manually correcting the core point in 13% of the fingerprints and by discarding 1% of the fingerprints because no core point could be found. Matcher2 have been performed neither such manual adjustments nor rejections. Furthermore, Matcher2 can only be compared head to head to Jiang et al. [5] start at PR equal to 5% whereMatcher2 gave lower ER rather than Jiang et al. [5]. On Figure 4, up to PR equal to 22%, Matcher2 gave lower ER rather than Jiang et al. [5]. Higher than PR equal to 22%, Jiang et al. [5] gave lower ER rather than Matcher2. On Figure 5, at all PR on the graph, Matcher2 and Matcher2* gave lower ER rather than all other strategies, even though no head to head comparison existbecause of different number of data set and different impression(s) selection mechanism for index creation (see Table 3).
8. Accuracy performance comparison results related to the used data set characteristics. Several points about these characteristics that could explain those results (point 5, 6, and 7): 1) Jiang et al. [5] stated fingerprints of FVC2000 DB2A have a higher image quality than those of FVC2000 DB3A. At a lower PR, successful retrieval needs closer similarity between the query and the candidates, which is more sensitive to the image quality. Therefore, FVC2000 DB2A has better retrieval performance than FVC2000 DB3A at lower PR. However, FVC2000 DB2A has a worse retrieval performance than FVC2000 DB3A at higher PR because of partial fingerprints whose core point is near the image edge or out of the image. FVC2000 DB2A has more such partial fingerprints than FVC2000 DB3A, which fails to be retrieved even at high PR; 2) Capelli [1] have been manually analyzed all queries that could not be found at PR equal to 30% and counted them per data set based on the most likely error cause (no errors at that PRon FVC2000 DB2): a. core not present (FVC2000 DB3A = 1, FVC2000 DB2A = 3), b. small overlapping region (FVC2000 DB3A =1, FVC2000 DB2A =2), c. large rotation (FVC2000 DB3A =2, FVC2000 DB2A =1), d. low image quality (FVC2000 DB3A =10 , FVC2000 DB2A = 0), and e. skin distorsion (FVC2000 DB3A = 10 , FVC2000 DB2A = 0); 3) About fingerprint image qualitystated by Jiang et al. [5], this paper confirmed it by measurement using NIST Fingerprint Image Quality (NFIQ) algorithm [20]. As shown by Figure 6, NFIQ outputs the image quality value (where 1 is the highest quality and 5 is the lowest quality) for 800 images per data set. Percentage of images with quality 1 and 2 (two highest image quality value) are about 37%, 88%, 26%, and 93% for FVC2000 DB1A (training data set used by the proposed strategy), FVC2000 DB2A, FVC2000 DB3A, and FVC2002 DB1A, respectively. Unlike Jiang et al. [5] and Capelli [1], the proposed strategy has nothing to do with core point and large rotation (not depend on them), so its retrieval performance on testing data sets is in accordance to their image quality results by NFIQ, i.e. best result on FVC2002 DB1A, follow by FVC2000 DB2A, and then FVC2000 DB3A. Noted, the proposed strategy arbitrarily chose FVC2000 DB1A as training data whose percentage of images with quality 1 and 2 is relatively quite low.

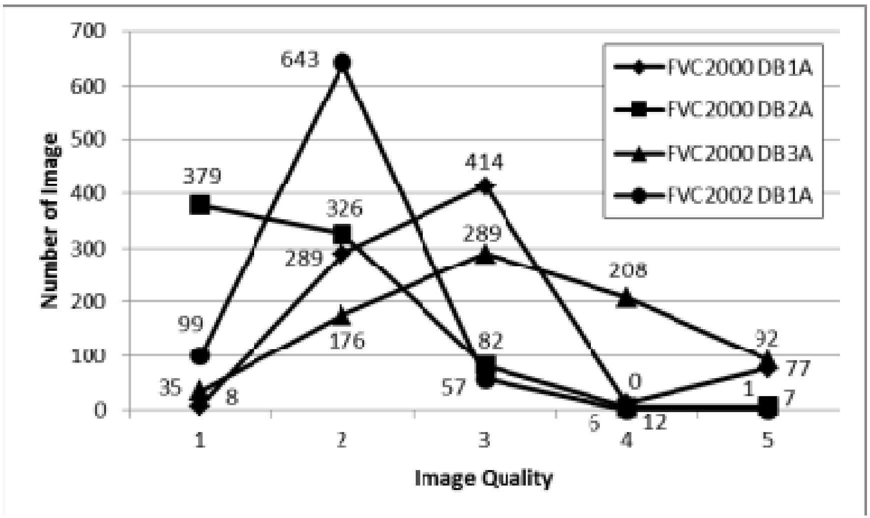

Figure 6. Histogram of fingerprint image quality based on NFIQ algorithm

3.3. Speed Comparison

Previous accuracy performance comes along with its speed performance. In general, it is not possible to make a systematic speed comparison like accuracy comparison since speed result is obtained on different hardware platforms.

1. Moreover, based on [1], unfortunately average search times are not reported for other published strategies on three data sets (Table 1), except byCapelli[1] that reported average search time about 70ms, 134ms, and 71ms, each for FVC2000 DB2, FVC2000 DB3, and FVC2002 DB1, running on an Intel Core 2 Quad CPU @ 2.66 GHz. Although it cannot be directly compared, however, that result can be used as a reference to motivate improvement of the proposed strategy. As shown by Matcher at Figure7, average search time conducted by the proposed strategy is about 247ms, 756ms, and 278ms for FVC2000 DB2, FVC2000 DB3, and FVC2002 DB1, respectively.
2. To complete internal comparison perspective, Figure7 also shows average search time comparison between Pre-filter and Pre-filter*; Matcher and Matcher* amongst three different data sets. Same template data (result of fingerprint feature extraction process) was used to provide relatively same information as an input to the proposed strategy with different configuration (search option) during experiment. On FVC2000 DB2 and FVC2002 DB1, average search time comparison gave relatively flat result trend. It means the proposed strategy relatively well applied for these data sets characteristic since triple growth number of impressions on searched-index did not make average search time triple longer, in comparison between Pre-filter and Pre-filter* (Pre-filter* on FVC2000 DB2 and FVC2002 DB1 gave additional average search time about 1.6% and -2.7%, respectively), or between Matcher and Matcher* (Matcher* on FVC2000 DB2 and FVC2002 DB1 gave additional average search time about 44.5% and 16.9%, respectively). On FVC2000 DB3, average search time comparison gave relatively sub-linear result trend, and also, triple growth number of impressions on searched-index did not make average search time triple longer, in comparison between Pre-filter and Pre-filter* (Pre-filter* on FVC2000 DB3 gave additional average search time about 32.5%), or between Matcher and Matcher* (Matcher* on FVC2000 DB3 gave additional searching time about 126.7%).

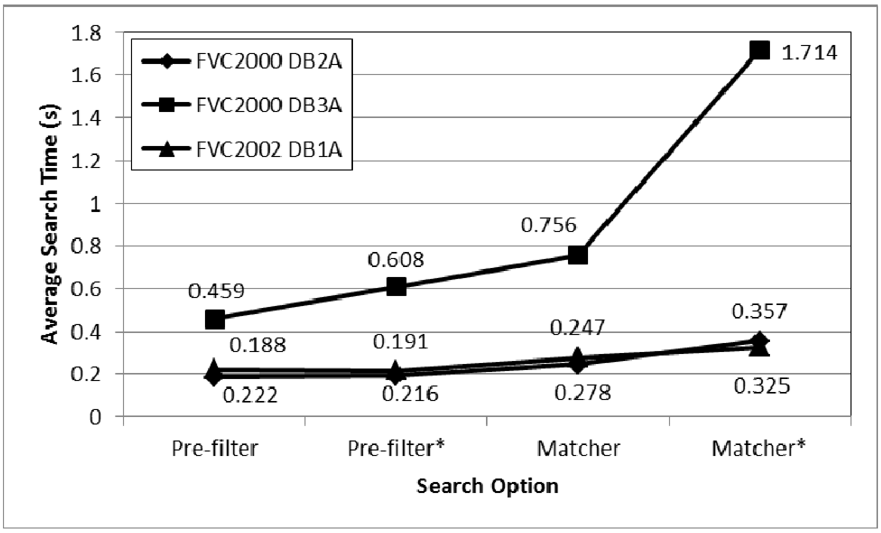

Figure 7. Average search speed comparison of the proposed strategy based on number of impression from same finger that was used for index creation.

3.4. Memory Usage Comparison

As one of the efficiency performance consideration beside speed, it is also not possible to make a systematic memory usage comparison like accuracy comparison because of different hardware platforms.

1. This section only describes internal comparison, amongst three different data sets, between indexingprocess (Section 2 Point 2) using fingerprint's first impression from each finger for index creation (referred as the Indexing), and using fingerprint's first three impressions from each finger for index creation (referred as the Indexing*). This internal comparison supposes to give more perspective of the proposed strategy.
2. Comparison uses two different configurations above (Indexing and Indexing*) because they are in accordance to the previous experimets. The Indexing process precedes the Pre-filter, Matcher, and Matcher2, whilst the Indexing* process precedes the Pre-filter*, Matcher*, and Matcher2*.
3. This indexingprocess (Indexing and Indexing*) required certain amount of memory during system running for hash-table-based searched index. Memories were occupied by hash-keys and its related retrieved objects (elements) that construct that hash-table. A hash-key holds its elements in a *list* if there is *collision* (a hash-key has more than oneelement). A hash-key was allocated by total 32-bit data, constructed by 16-bit (LSB) data of edge length, 8 bit (bit 23^{th} - 16^{th}) data of minutia-reference relative-orientation, and 8-bit (MSB) data of minutia-neighbour relative-orientation. An element was allocated by total 104-bit data, constructed by32-bit data of candidate fingerprint ID in database, 16-bit data of edge length, 8 bit data of minutia-reference relative-orientation, 8-bit data of minutia-neighbour relative-orientation, 8 bit data of minutia-neighbourtype (ending, bifurcation, or else), 16-bit data of minutia-reference number, and 16-bit data of minutia-neighbour number.
4. For the next discussion, memory usage comparison will only refers to the number of hash-keys and their related number of elements for analysis simplicity.
5. Figure8 shows increasing number of hash-key generated by the Indexing* to the base number of hash-keys generated by the Indexing. The Indexing* gave increasing number of hash-keys about 16%, 19%, and 12% for FVC2000 DB2, FVC2000 DB3, and FVC2002 DB1, respectively.
6. Related to Figure 8, Figure 9 shows hash-keys distribution based on number of its elements. Upper-row graphs (generated by the Indexing) and lower-row graphs (generated by the Indexing*) show distribution percentage of hash-key groups based on number of its elements. These hash-key groups were classified into group 1; 2-10; 21-30; 31-40; 41-50; and 51-*x*. Last group contains *x* (number with an asterix mark) that differs from each data set and represents the biggest number of elements that could fill in the list, related to certain hash-key. Noted, consecutive number of elements on last group (from 51 to *x*) was not guaranteed exists if compared with the other groups.
7. On the Indexing, distributions of hash-key group 1 are about 10%, 16%, and 10% for FVC2000 DB2, FVC2000 DB3, and FVC2002 DB1, respectively. On the Indexing*, distributions of hash-key group 1 relatively not changed much, i.e. about 9%, 15%, and 8% for FVC2000 DB2, FVC2000 DB3, and

FVC2002 DB1, respectively. This indicates scalability on data growth of hash-table. In general, this hash-key group need to obtain bigger distribution percentage for search speed performance improvement since in retrieval stage, this hash-key group represents pure time complexity $O(1)$.

8. Considering recent processor computation which is tremendous fast, it assumes hash-key group 2-10, 21-30, 31-40, and 41-50 support retrieval stage with relatively same time complexity $O(1)$ of hash-key group 1. Based on this assumption, on the Indexing, distribution of hash-key group 1-50 are about 82%, 75%, and 83% for FVC2000 DB2, FVC2000 DB3, and FVC2002 DB1, respectively. On the Indexing*, distribution of hash-key group 1-50, are about 58%, 59%, and 58% for FVC2000 DB2, FVC2000 DB3, and FVC2002 DB1, respectively. On data growth of hash-table, this indicates enlargement of distribution hash-key group 51-x that need to be suppressed on future work improvement.

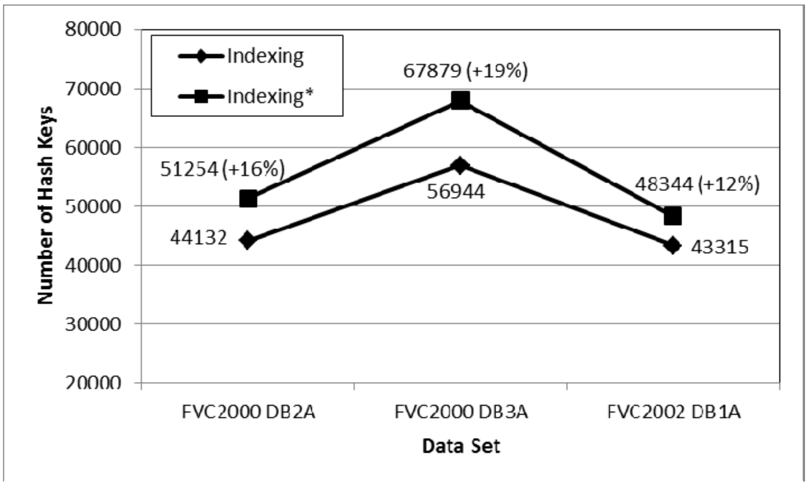

Figure 8. Increasing number of hash-keys through fingerprint impressions addition to the hashing-based searched-index.

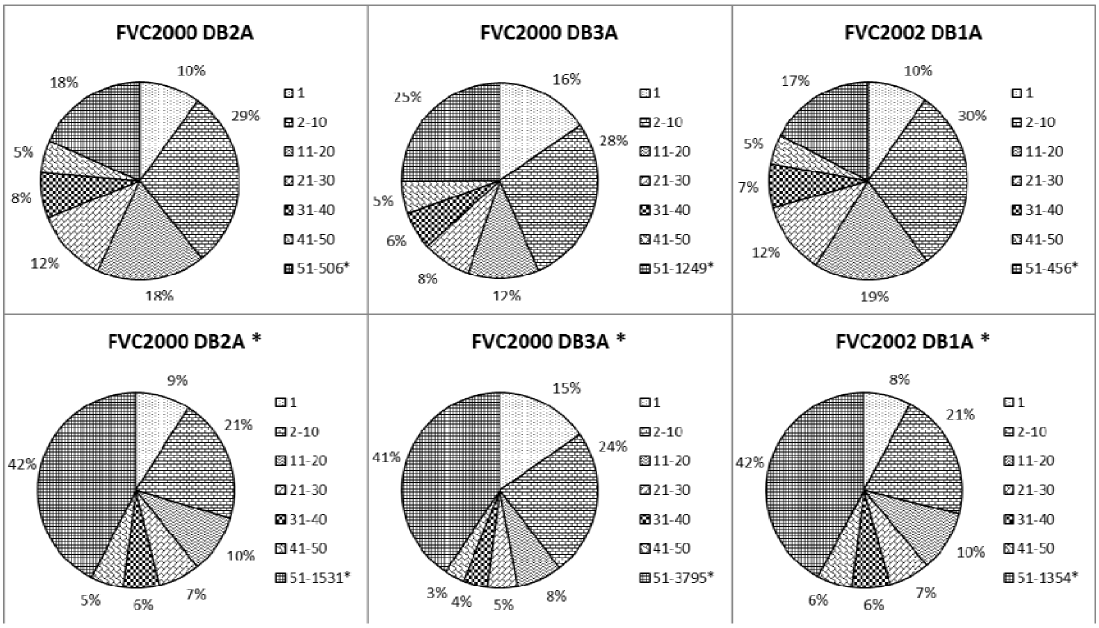

Figure 9. Hash-key distribution based on number of its element on hashing-based searched-index generated on three public fingerprint data sets. Upper row: the Indexing graphs; lower row: the Indexing* graphs.

4. CONCLUSION

Internal comparison of the proposed fingerprint direct-access strategy among different concerned configurations, and its external comparison to the other strategies have already been conducted. For external comparison, the experiment result has confirmed that the proposed strategy as a promising strategy since for particular aspect, it out performs the other strategiesover three publicly available data sets. Up to certain penetration rate (PR) equal to 5%, the proposed strategy consistently gives lower error rate (ER). By taking sample at PR equal to 5%, the proposed strategy produced ER equal to 4%, 10%, and 1% on FVC2000 DB2A, FVC2000 DB3A, and FVC2002 DB1A, respectively. At the other side, larger than PR equal to 5%, the proposed strategy needs improvement to obtain a faster rate of decrease of its ER to PR, specifically for FVC2000 DB2 and FVC2000 DB3, since several compared strategies have already achieved it. Another perspective if accuracy performance was measured based on area under curve from the graph of ER and PR, the proposed strategy neither is the best strategy nor the worst strategyon FVC2000 DB2A and FVC2000 DB3A, whileon FVC2002 DB1A it outperfomed the other strategies and even it gave impressive results for index created by three impressions per finger (withor without utilizing N_T) by producing step down curve which is ideal curve where PR equal to 1% can always be maintained for smaller ER.

5. FUTURE WORK

Several immediate improvements for the proposed strategy [2] related to obtain a faster rate of decrease of its average ER to its average PR are:

1. At indexing process, rather than generating hash-key from single edge, it is hypothetically more efficient generating hash-key directly from discriminator attributes, i.e. two edges which are connected, so no more similarity score computation needed related to these discriminator attributes.
2. At retrieval process, in certain way, merge pre-filter stage and matcher stage to reduce score computation time without losing significant accuracy of search.
3. To recent strategies, incorporate unused extracted ridge-orientation information to improve similarity score computation. Hypothetically, ridge-orientation information which is line-based discriminator attribute represents fingerprint's local area better than minutia information which is point-based discriminator attribute. Using this way will enrich hash-key information such that it can generate unique hash-key to enlarge group hash-key 1 that represents pure time complexity $O(1)$, or group hash-key 1-50 that represents time complexity close to time complexity $O(1)$ (Section 3.4). For that, hash-key can be designed with 64-bit data allocation (more memory needed as a consequence) where first 32-bit has already been implemented in this strategy, and second 32-bit can be allocated into four 8-bit segments for ridge-orientation difference information, as an example of geometric-transformation-invariant information. As shown by Figure 10, for an example, difference value between ridge orientation 1 and ridge orientation 3 of minutia-reference (Δro_{R13}) would fill in segment 1, Δro_{R24} would fill in segment 2, difference value between ridge orientation 1 and ridge orientation 3 of minutia-neighbour (Δro_{N13}) would fill in segment 3, and Δro_{N24} would fill in segment 4.

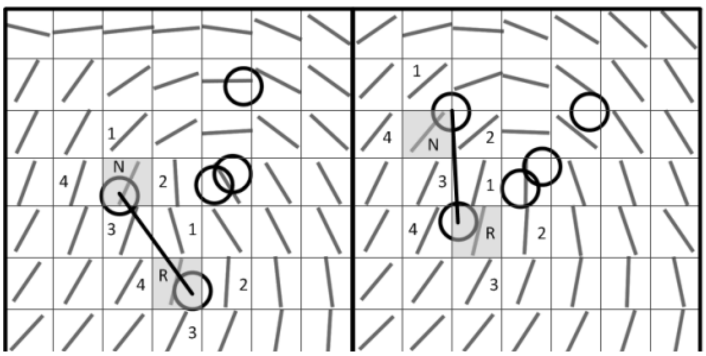

Figure 10. Ridge-orientation difference enriches hash-key information. Left: a query; right: a candidate.

REFERENCES

[1] R Cappelli. "Fast and Accurate Fingerprint Indexing based on Ridge Orientation and Frequency". *IEEE Trans. Sys., Man, Cybern. B, Cybern.* 2011; 41(6): 1511–1521.

[2] G Indrawan, B Sitohang, S Akbar, and AS Nugroho. "A Multi-Threaded Fingerprint Direct-Access Strategy Using Local-Star-Structure-based Discriminator Features". *TELKOMNIKA Indones. J. Electr. Eng.* 2014; 12(5): 4079–4090.

[3] NK Ratha, VD Pandit, RM Bolle, and V Vaish. "Robust Fingerprint Authentication Using Local Structural Similarity". in *Proc. Work. Appl. Comput. Vis.* 2000: 29–34.

[4] R Cappelli, M Ferrara, and D Maltoni. "Fingerprint Indexing Based on Minutia Cylinder-Code". *IEEE Trans. Pattern Anal. Mach. Intell.* 2011; 33(5): 1051–1057.

[5] X Jiang, M Liu, and AC Kot. "Fingerprint Retrieval for Identification". *IEEE Trans. Inf. Forensics Secur.* 2006; 1(4): 532–542.

[6] R Cappelli, D Maio, and D Maltoni. "A Multi-Classifier Approach to Fingerprint Classification". *Pattern Anal. Appl.* 2002; 5(2): 136–144.

[7] R Cappelli, A Lumini, D Maio, and D Maltoni. "Fingerprint Classification by Directional Image Partitioning". *IEEE Trans. Pattern Anal. Mach. Intell.* 1999; 21(5): 402–421.

[8] SO Lee, YG Kim, and GT Park. "A Feature Map Consisting of Orientation and Inter-Ridge Spacing for Fingerprint Retrieval". in *Proc. 5th Int. Conf. AVBPA*. 2005.

[9] J Li, WY Yau, and H Wang. "Fingerprint Indexing Based on Symmetrical Measurement". in *Proc. 18th ICPR*. 2006; 1(1): 1038–1041.

[10] M Liu, X Jiang, and AC Kot. "Fingerprint Retrieval by Complex Filter Responses". in *Proc. 18th ICPR*. 2006; 1(1).

[11] RS Germain, A Califano, and S Colville. "Fingerprint Matching Using Transformation Parameter Clustering". *IEEE Comput. Sci. Eng.* 1997; 4(4): 42–49.

[12] B Bhanu and X Tan. "Fingerprint Indexing Based on Novel Features of Minutiae Triplets". *IEEE Trans. Pattern Anal. Mach. Intell.* 2003; 25(5): 616–622.

[13] X Liang, T Asano, and A Bishnu. "Distorted Fingerprint Indexing Using Minutia Detail and Delaunay Triangle". in *Proc. 3rd ISVD*. 2006: 217–223.

[14] J De Boer, AM Bazen, and SH Gerez. "Indexing Fingerprint Databases based on Multiple Features". in *Proc. Workshop Circuits, Syst. Signal Process. (ProRISC)*. 2001: 300–306.

[15] S Biswas, NK Ratha, G Aggarwal, and J Connell. "Exploring Ridge Curvature for Fingerprint Indexing". in *Proc. IEEE 2nd Int. Conf. BTAS*. 2008: 1–6.

[16] X Shuai, C Zhang, and P Hao. "Fingerprint Indexing Based on Composite Set of Reduced SIFT Features". in *Proc. 19th ICPR*. 2008: 1–4.

[17] A Gyaourova and A Ross. "A Novel Coding Scheme for Indexing Fingerprint Patterns". in *Proc. 7th Int'l Workshop S+SSPR*. 2008: 765–774.

[18] T Maeda, M Matsushita, and K Sasakawa. "Identification Algorithm Using A Matching Score Matrix". *IEICE Trans. Inf. Syst. - Spec. Issue Biometric Pers. Authentication*. 2001; E84-D(7): 819–824.

[19] R Vazan. "Source AFIS | Fingerprint Recognition Library for .NET and Experimentally for Java". 2009. [Online]. Available: http://sourceforge.net/projects/sourceafis/. [Accessed: 25-Oct-2013].

[20] CI Watson, MD Garris, E Tabassi, CL Wilson, RM McCabe, S Janet, and K Ko. "User's Guide to NIST Biometric Image Software". Gaithersburg. 2004.

[21] A Silberschatz, HF Korth, and S Sudarshan. *Database System Concepts*. 6th ed. New York: McGraw-Hill. 2011.

[22] R Cappelli, M Ferrara, and D Maio. "Candidate List Reduction Based on The Analysis of Fingerprint Indexing Scores". *IEEE Trans. Inf. Forensics Secur.* 2011; 6(3) 1160–1164.

[23] SJ Russell and P Norvig. *Artificial Intelligence: A Modern Approach*. 2nd ed. New Jersey: Upper Saddle River. 2003: 111–114.

[24] D Maio, D Maltoni, R Cappelli, JL Wayman, and AK Jain. "FVC2000: Fingerprint Verification Competition". *IEEE Trans. Pattern Anal. Mach. Intell.* 2002; 24(3): 402–412.

[25] D Maio, D Maltoni, R Cappelli, JL Wayman, and AK Jain. "FVC2002: Second Fingerprint Verification Competition". in *Proc. 16th Int'l Conf. Pattern Recognition*. 2002; 3: 811–814.

[26] X Liang, A Bishnu, and T Asano. "A Robust Fingerprint Indexing Scheme Using Minutia Neighborhood Structure and Low-Order Delaunay Triangles". *IEEE Trans. Inf. Forensics Secur.* 2007; 2(4): 721–733.

BIOGRAPHIES OF AUTHORS

G Indrawan, Doctoral candidate at Data and Software Engineering Research Group, School of Electrical Engineering and Informatics, Institut Teknologi Bandung, Indonesia. He obtained his undergraduate degree in Electrical Engineeringat Institut Teknologi Bandung, his magister degree in Electrical Engineering at Universitas Indonesia. His field of interest is biometric systems and multimedia data retrieval. He can be reached at gindrawan@live.com.

Dr. Techn. S Akbar, Head of Lab. of Data and Software Engineering Research Group, School of Electrical Engineering and Informatics, Institut Teknologi Bandung, Indonesia. He obtained his undergraduate and magister degree in Informatics at Institut Teknologi Bandung, his doctoral degree in 3D Model Similarity Retrieval at University of Linz, Austria, and his post docs at NTNU, Norway. His field of interest is big data, open data, and multimedia database. He can be reached at saiful@informatika.org.

Prof. Dr. Ir. B. Sitohang, Head of Data and Software Engineering Research Group, School of Electrical Engineering and Informatics, Institut Teknologi Bandung, Indonesia. He obtained his undergraduate degree in Electrical Engineering at Institut Teknologi Bandung, his magister and doctoral degree in System des Gestion de Base des Donneessur Micro Ordinateur at Universite de Sciences et Techniques et Languedoc-Montpellier II, Montpellier, France. His field of interest is database systems. He can be reached at benhard@stei.itb.ac.id.