A Generic Template Update Framework: The Solution for Intra-Class Variations in Biometric Systems

XI CHEN¹,², MING YU¹,³, (MEMBER, IEEE), FENG YUE², BIN LI² AND YINGCHUN GUO³.

¹School of Electronic and Information Engineering, Hebei University of Technology (HEBUT), Tianjin, P. R. China
²Beijing Institute of New Technology Applications, Beijing Academy of Sciences and Technology, Beijing, P. R. China
³School of Artificial Intelligence, Hebei University of Technology (HEBUT), Tianjin, P. R. China

Corresponding author: Ming Yu (e-mail: yuming@hebut.edu.cn).

This work was supported in part by the Natural Science Foundation of Hebei Province F2019202464 and F2019202381; the National Natural Science Foundation of China 61806071; the Open Project Program of the National Laboratory of Pattern Recognition (NLPR) of China 201900043; Hebei Provincial Education Department Youth Foundation QN2019207.

ABSTRACT It is already a consensus that intra-class variations will decline the performance of biometric systems. Many studies have explained how intra-class variations affect the biometric systems, while few approaches have proposed the effective solutions of template update. Since template update is a crucial procedure in biometric recognition systems, the problem is challenging because the newly selected templates must be representative of a large amount of intra-class variations such as posture changes and lighting conditions. To this end, existing techniques proposed to perform template update by periodically selecting new representative templates. However, all of them are designed for specific biometric traits rather a generic framework, and they are either ineffective because selected templates are less representative or prone to selection errors due to the presence of outliers. In this paper, we propose a novel generic template update framework. We use pass tables to accumulate information for outlier removal and representative templates selection, and further, propose a novel criterion to select the optimal template set according to current templates and similarity threshold. To confirm the effectiveness of our method, we carry out several experiments on synthetic and real datasets from different biometric traits. The results show that the proposed method is more robust than existing ones, and at the same time, achieves a higher accuracy according to the specified performance indicator. It also demonstrated the superiority of our method to random selection for different biometric traits.

INDEX TERMS Intra-class Variations, Pass Table, Template Selection, Template Update

I. INTRODUCTION

BIOMETRICS refers to automatically recognizing a person based on one’s unique biological and behavioral traits without carrying password and token, which significantly facilitates every aspect of daily life. Systems extract features from enrolled samples as template sets. By this way, the systems based on the nearest neighbor classifier only able to correctly classify the samples with adequate similarity to a template, while wrongly reject other genuine samples due to the difference between templates and query. Thus a critical difficulty arises when biometric traits undergo intra-class variations.

There are two main reasons for the intra-class variations. The first is template aging. Template aging is defined as the increase in error rates caused by time-related changes, meaning changes in the biometric pattern, the presentation of the pattern and the sensor [1], [2]. Aging affects the human body and the acquisition sensor. Therefore, several physical and/or mental characteristics and acquired biological data will change over time. The second is non-permanent changes. For example, illumination and posture vary from image to image, temporary decorations and stains may appear on the skin on different acquisition sessions [3]–[5]. For the intra-class variations that are not included in the template set, the systems are very likely to reject genuine users falsely. Hence, biometric systems are expected to be adaptive to such variations by updating the templates constantly. Therefore, how to update them by selecting a small yet representative templates...
set and how to keep or even improve the representativity is thus a challenging problem and is crucial for biometric systems.

In practical applications, the system managers often expand the template set to force it to be sufficient representative or perform re-enrollment to deal with intra-class variations. However, for the sake of storage and speed, the number of templates per subject cannot be very large, and in many applications, to ensure the security of systems, the enrollment stage requires special supervision. Therefore, re-enrollment is a very time-consuming and troublesome process.

Many approaches have analyzed the impact of intra-class variations on biometric systems [6]–[8], while few works focus on the solution to the problem. Especially in recent years, there is almost no literature on this subject, and it seems that the research on this topic reached a bottleneck. Previous methods on the solution to intra-class variations can be roughly classified into two categories, i.e., sequential mode and batch mode. The former are usually utilized to modify the template by embedding all the useful information into a single “super template” [9], [10], in this way, the biometric traits should be well-aligned and under a consistent acquisition environment. These methods select useful information instead of the whole feature and update the template by embedding the information. Some contact biometric systems (e.g., fingerprint recognition system) adopted such a framework, while they are not suitable for contactless systems (e.g., face recognition system). The latter are commonly adopted by prototype-based systems to select new templates from input samples [11]–[13]. The purpose of this kind of framework is to update the template set by selecting new template set with adequately representativity, and most of the existing works related to template update focus on the development of effective template selection algorithms. For instance, DENT [12] employs a hierarchical clustering strategy to choose a template set that best represents the intra-class variations, and MDIST [12] selects templates that exhibit maximum similarity with the rest of the samples. Later, other clustering-based [14], [15] and similarity-based [16] techniques have been proposed to improve the performance of DENT and MDIST. Besides, the template selection methods making use of Shannon entropy [17], ant colony optimization [18], and graph min-cut algorithm [19] have also been proposed and tested. However, all of the proposed methods are designed and tested for a specific biometric trait, and no published research has proposed the template update framework for the generic purpose.

Besides, the template update is a semi-supervised online procedure with jointly using labeled and unlabelled data [20], it is inevitable to introduce some imposters in the candidates. The template selection methods which are sensitive to such noise would be prone to selection errors and subsequent false acceptance [21]. A practical solution is to adopt an updating threshold, which is more restrictive than the recognition threshold, to select highly genuine candidates. However, in this way, only the samples with high similarity to the templates will be retained, and they would not well represent the difficult and informative intra-class variations [20], [22]. Another solution is “co-updating” the template candidates of different biometric traits [23]–[25]. One classifier operating at a high confidence level can help another to identify difficult patterns. Although co-update outperforms self-update when initial templates are non-representative, the limitation is that it is only applicable to multi-modal biometric systems.

Based on the current research status, we summarize the key issues related to the generic update framework of biometric systems in a batch mode as follows.

- **How to select a representative template set?**
  - The template set is expected to have no imposters (outliers), and at the same time, contains large intra-class variations with acceptable storage space.

- **When to perform updating?**
  - The update procedure should be performed in a periodically update round to ensure adequate intra-class information has been collected, and the procedure should preserve the potential templates (candidate templates) from in-coming samples constantly between the update rounds.

- **What to do after an update round?**
  - While the typical practice of updating is replacing the old template set with a new one, the replaced template is simply discarded. However, in our opinion, these templates still contain useful information. Compared with the current template set, they are non-representative for the in-coming samples so far but may be selected as templates in the future.

In this paper, we propose a novel generic template update framework based on robust template selection. We use pass tables to accumulate information for outlier removal and representative templates selection. We further propose a novel method to select the optimal template set according to current templates and similarity threshold. The flow chart of proposed framework is depicted in Fig.1. For a genuine user, the nearest neighbor classifier based biometric system with template update procedure will select an appropriate subset of a sequence of new-coming unlabeled samples as candidate template set, and generate a new template set from the union of current template set and candidate template set with better representativity to intra-variations.

We have carried out several experiments to evaluate the performance of the proposed method. Our previous work [26] has been upgraded by a novel template selection method. At the same time, the research has been extended by designing a novel generic framework. A more in-depth discussion of candidate template set selection is also introduced. We propose an additional palmvein dataset as well as an experimental report on face, palmprint and palmvein datasets, on which the analysis of distinct results in different biometric traits are given. On synthetic and real datasets, the comparative study with the existing methods has demonstrated its robustness and effectiveness. Our main contributions can be summarized...
as follows.
(a) We propose the key issues of a generic update framework that satisfy the majority of biometric systems rather than a specific biometric trait.
(b) We introduce a potential template (candidate template) selection method by similarity and closeness threshold. The method could ensure the candidate template is a valid sample as well as is useful to update procedure.
(c) We provide a long-term collection of palmprint and palmvein datasets with multiple samples.
(d) We evaluate the proposed framework on two types of biometric traits and give the analysis of experimental results.

In the remainder of this paper, we introduce our update framework in Section II. In Section III, we provide a quantitative and qualitative experimental results of proposed framework as well as existing methods on both synthetic palmprint dataset and real face, palmprint and palmvein datasets. Finally, in Section IV, we conclude with some final remarks and drawbacks of our experiments and also give a discussion of our future work.

II. PROPOSED TEMPLATE UPDATE FRAMEWORK
The aim of proposed template update framework is to maintain or even improve the representativity of templates.

After the template set has been selected at enrollment stage, the framework accumulates information and selects the better representative samples as new templates. To illustrate how proposed framework works, a simplified update procedure on face dataset is depicted in Fig. 2. The red dotted box indicates the templates that will be replaced in the next update round and the blue dotted box indicates the samples (candidate templates) that will be selected as new templates. As shown in the right of Fig. 2, when the framework works, the templates with weak representativity are replaced by the in-coming samples with better representativity. As a consequence, the false rejection rate (FRR) decreases from the initial stage to round $t$, and similarly, from round $t$ to current update round. That is, the proposed framework will significantly reduce the false rejection rate while improve the recognition performance for face recognition systems.

A. TEMPLATE UPDATE PROCEDURE
The purpose of template update is replacing an existing weak-representative template to a new one. Between two update rounds, the framework will collect the potential templates as candidate template set from in-coming sample sequence, and then select the new template set from the union of candidate template set and current template set. In this section, we will provide the details of the information accumulation based update procedure.

1) Candidate Template Set Selection Method
The candidate template set is a collection of samples which contains potential templates for the next update round. How to select a potential template from a sequence of unlabeled in-coming samples is a crucial step in the template update. It is expected that the candidate template set would contain large intra-class variations, while at the same time introduce imposters as few as possible. The proposed candidate templates selection and information accumulation method is given in Algorithm 1.

A fixed similarity threshold $t_{sim}$ is usually adopted to determine the incoming sample is a genuine or imposter for biometric systems with nearest neighbor classifier. For the sake of system safety, a sample from genuine user can be considered as a valid sample and fed to the template update procedure. However, due to the potential false acceptance occurred at recognition stage, several researchers [20], [23] propose to use a more restricted similarity threshold to reduce the risk of false updates, i.e., templates contamination. We also adopt this strategy. Besides, we propose a robust template selection method to further reduce the adverse effects of imposters.

When a valid sample arrives, we should evaluate its usefulness for template update. We introduce another threshold called closeness threshold $t_{close}$ to determine the closeness of a sample to the current templates and candidate templates. If the match score $Score$ exceeds the closeness threshold,
an imposter and abandoned directly. However, we believe
that we cannot pass the similarity threshold would be determined as
for other update methods [9], [10], [29], the sample that
is passed to sample pass table, which will be described in the next
subsection.

2) Pass Table
For other update methods [9], [10], [29], the sample that
cannot pass the similarity threshold would be determined as
an imposter and abandoned directly. However, we believe
that this kind of samples may also contain useful information. For
example, a template set $T = \{t_1, t_2, t_3\}$, a candidate template $c$,
and incoming sample $s$ which has large intra-class variations
compared with $T$ while has moderate intra-class variations
compared with $c$, as illustrated in Fig. 3. Existing works
simply discard this type of samples. The main drawback is
the information loss, which would lead to the underestimate
of the importance of template $c$. In our work, we eliminate
the drawback by utilizing the information provided by this type
of samples. Since candidate templates may be noisy, these
samples will not be considered as new candidates. Instead,
we set up a sample pass table to record the importance of
each template and update the table when these samples arrive
(line 10 in Algorithm 1).

For the sake of storage, the system will not store a large
number of candidate templates, so we employ a pre-defined
size of candidate template set $N$. At the initial stage, the
candidate template set $T_c$ is empty. The update procedure will
select an appropriate candidate template and insert it directly
if current candidate number $N'$ satisfies

$$N' < N, \,
\quad (1)$$
or replace one with minimum column summation according to
sample pass table, which will be described in the next
subsection.

\begin{algorithm}
\textbf{Algorithm 1} Candidate templates selection and information
accumulation \\
\textbf{Require:} A template set $T = \{T_1, \ldots, T_K\}$, the similarity
threshold $t_{sim}$, the closeness threshold $t_{close}$, and a se-
quence of unlabeled samples $s_1, \ldots, s_M$ \\
\textbf{Ensure:} A candidate set $T_c$, and a sample pass table $A_s$ \\
1: $T_c = \emptyset$ \\
2: for $i = 1$ to $M$ do \Comment{for $i = 1$ to $M$}
3: \hspace{1em} if $Score(T, s_i) > t_{sim}$ then \Comment{if $Score(T, s_i) > t_{sim}$}
4: \hspace{2em} if $Score(T, s_i) > t_{close}$ then \Comment{if $Score(T, s_i) > t_{close}$}
5: \hspace{3em} Update sample pass table $A_s(s_i, T \cup T_c)$ \Comment{Update sample pass table $A_s(s_i, T \cup T_c)$}
6: \hspace{2em} else \Comment{else}
7: \hspace{3em} $T_c \leftarrow T_c \cup s_i$ \Comment{find a new candidate}
8: \hspace{2em} end if \Comment{end if}
9: \hspace{1em} end if \Comment{end if}
10: Update sample pass table $A_s(s_i, T_c)$ \Comment{Update sample pass table $A_s(s_i, T_c)$}
11: end if \Comment{end if}
12: end for \Comment{end for}
\end{algorithm}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig2.png}
\caption{A simplified update procedure on face dataset.}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig3.png}
\caption{Using unclassified samples to update the importance of candidate
templates.}
\end{figure}

Pass table $A$ is a two-dimension binary table that records
the pass-fail information with $W$ columns and $H$ rows, $W = K + N$ where $K$ and $N$ indicate the size of the template
set $T$ and candidate template set $T_c$ respectively. $H$ equals
to the number of recorded samples. We first introduce the
sample pass table $A_s$. $A_s$ records the pass-fail information
of incoming samples concerning to current templates and
a similarity threshold. If the incoming sample sequence
contains $M$ instances, $A_s$ is a $M \times (K+N)$ binary table. When
a new sample $s$ arrives, if its similarity with template $t$ is larger
than the threshold, then $A_s(s, t) = 1$. Otherwise $A_s(s, t) = 0$.
The sample pass table is at the core of template selection
as it reflects the relative importance of each template. As far
as a single template is concerned, a larger pass count, i.e., the
summation of column $t$ indicates a more important template.
An example of sample pass table with $T = \{t_1, \ldots, t_5\}$ and
$T_c = \{tc_1, tc_2\}$ is illustrated in top-left of Fig. 4. A sequence

\begin{eqnarray}
N' < N, \quad (1)
\end{eqnarray}
with five samples compares with $T$ and $T_c$. The columns of the table indicate the union of template set and candidate template set. From Section II-A1, if a sample is selected as candidate template, the sample pass table will add one column in the right of $A_s$, under the condition of Eq. 1, or replace the template which has the minimum pass count in $T_c$ by a new one. If the predefined $N = 2$, $tc_2$ which has the pass_count = 3 will be replaced. Samples that will be used to update the table $A_s$ are in the following condition.

(a) The sample $s$ is too similar to one or more templates in $T$.

If $s$ satisfies the closeness condition(line 4 in Algorithm 1), i.e.,

$$S core(T, s) > t_{close},$$

it indicates that $s$ is too similar to one or more current templates, the representativity of $s$ could be covered by $T$. Therefore, $s$ is not a suitable potential template to expand the representativity of template set.

(b) The sample $s$ is rejected by $T$ with $t_{sim}$, while accepted by $T_c$.

If the score between $s$ and $T$ exceeds $t_{sim}$, while the score between $s$ and $T_c$ is less than $t_{sim}$ as described in Line 10 of Algorithm 1, i.e.,

$$S core(T, s) > t_{sim} > S core(T_c, s),$$

it indicates that even $s$ is an unclassified sample, it may contain intra-class variations information of $T_c$, which as shown in Fig. 3.

3) Update Round

Update round is the cycle of templates re-selection. Template update is usually carried out in a batch mode. For a practical biometric system, we should take the number of incoming samples, the period of update, and the available space requirement into consideration. For proposed method, we can also determine the update round by the size of the sample pass table.

4) Generation of New Template Set

There are two kinds of methods for new template generation. The first is template fusion, in which the incoming samples are always fused to an existing template. As a sequence, the “super template” will embed all the information together, which is typically adopted in fingerprint recognition by adding and removing minutiae [9]. These methods assume that the templates and samples are well-aligned by key points. However, for contact-less biometric systems, non-linear deformations appears during the samples acquisition, the acquisition environment and posture of users are also more complicated, as shown in Fig. 5. Therefore, this assumption is not valid for contact-less systems.

The second is instance based template selection methods, in which the candidate templates are always added as new instances to the template set. The new template set is then selected from the enlarged template set to replace the old one. In this paper, we focus on the instance based template selection.

5) Post-processing

Once template update has been carried out, the templates belong to the new template set will replace those of the old template set. There are two types of templates as well as the information stored in the sample pass table. The first is the candidate templates that not be selected as templates. The second is the templates that have been replaced by new ones. They would be still useful for the future template update. According to the experimental results, for face recognition, 9.8% and 11.2% templates will be reselected as templates in the future update rounds for these two types of templates.
They are 10.3% and 12.5% for palmprint, and 10.9% and 13.1% for palmvein. If we discard them all, the FRR on the test datasets will increase about 3.0%, 2.3% and 2.2% respectively.

Thus instead of discarding them all, we suggest to retain a portion of the remaining templates for the next update round. Actually, we have analyzed the influence on recognition performance in different size of remaining templates $S_r$. The results are distinct on different test datasets. For face, a large $S_r = 0.60 \ast N$ brings the best accuracy. The palmprint and palmvein obtain the best performance at $S_r = 0.42 \ast N$ and $S_r = 0.47 \ast N$ respectively, where $N$ is the predefined candidate template size. The experimental results also show that if the $S_r$ is between $0.45 \ast N \sim 0.65 \ast N$, the recognition performance is almost the same to all the datasets, while a larger $S_r$ even declines the performance. Hence, for a generic purpose, we suggest the retained templates $T_c^*$ can be selected as the half of candidate templates set $T_c$ according to the summation of match scores with the new template set

$$T_c^* = \bigcup_{i=1}^{N^*} \arg\min_{t' \in T} \sum_{j=1}^{K} \text{Score}(t_j, t'),$$  \hspace{1cm} (4)

where $K$ is the size of template set, $N$ and $N^*$ are the size of $T_c$ and $T_c^*$ respectively, $N^* = N/2$. We select a $t'$ into $T_c^*$ if $t'$ obtains the minimum similarity scores.

**B. TEMPLATE SELECTION**

Template selection is the procedure of selecting new templates from previous templates plus several candidates. The new template set is expected to be representative of all the templates and the candidates, thus would be adaptive to intra-class variations. In the following, we first give the problem formulation of template selection, and then provide a practical solution based on joint pass table $A_j$.

1) Problem Formulation

In an instance based biometric system with the nearest neighbor classifier, the recognition performance, especially the false rejection rate, is closely related to the template set. Considering a template set $T = t_i, i = 1, 2, \ldots, K$ and a fixed similarity threshold $t_{\text{sim}}$ of a specified user, the total samples that can be classified as genuine is determined by

$$X(T) = \bigcup_{i=1}^{K} \{ s | \text{Score}(t_i, s) > t_{\text{sim}} \}. \hspace{1cm} (5)$$

To reduce the false rejection rate, $X(T)$ should be maximized. Formally, let the merged set of samples $T_M$ has $K$ templates and $N$ candidates, the problem is defined as selecting of new template set $T^*$ that satisfies

$$T^* = \arg\max_{T^* \subset T_M} X(T^*). \hspace{1cm} (6)$$

It is evident that $T^*$ in (6) is prone to update errors due to potential outliers. Besides, with the prior knowledge of high confidence of $T_M$, it is expected to maximize the number of correctly classified templates in $T_M$. We then modify (6) as

$$T^* = \arg\max_{T \subset T_M} | T_M \cap X(T) |. \hspace{1cm} (7)$$

where $| \cdot |$ denotes the number of elements in a set.

We then build a template pass table $A_j$ for $T_M$, in a similar way to that of sample pass table $A_s$. $A_j$ records the pass-fail information in the union of template and candidate template set. Remember that the sample pass table $A_s$ is used to record the pass-fail information of useful samples or potential templates. We now merge it with $A_j$ by the conjunction of their columns which is called the joint pass table $A_j$. An example is shown in the bottom-left and right of Fig. 4 respectively. It is convenient to compute $T_M \cap X(T)$ by logical OR operation on corresponding columns of $A_j$. Specifically, we select the subsets with $K$ instances of $T_M$ as $T_M$, $l = 1, 2, \ldots, L$, where $L$ is the types of the combinations of subsets. In a greedy way, we compute the permeability of $T_M$ in $T_M \cap X(T)$ by count 1s after performing logical OR operation on the corresponding columns of the $T_i$, the subset with the most 1s indicates that in $T_M \cap X(T)$, this type of combination obtains the maximum permeability.

2) Outlier Removal

Candidate set may contains a few impostors due to the compromise between false acceptance and false rejection. If some of them are selected as new templates, the biometric system will have the risk of performance degradation with noticeably increased false acceptance. Therefore, it is desirable for the template selection method to be robust against such outliers.

Afterward, the pass count in $A_j$ can also be used for outlier detection as

$$\left\{ \begin{array}{l} \text{pass count} = \sum_{j=1}^{K+N} A_j(t^*, t_j) \\ \text{pass count} \leq t_{\text{sum}} \end{array} \right. \hspace{1cm} (8)$$

If $\text{pass count}$ is smaller than a pre-defined threshold $t_{\text{sum}}$, e.g., $t_{\text{sum}} = 3$, then the corresponding sample is probably an outlier and should be discarded before template selection. We denote the set after outlier removal as $\hat{T}_M$. An example of sample pass table, template pass table and the joint pass table in Fig. 4. The candidate template $t_1$ obtains the $\text{pass count} = 2$, and smaller than the pre-defined threshold 3. Hence, $t_1$ is regarded as an outlier and is discarded from template selection. As a result, $T_M = (t_1, t_2, t_3, t_{c1}, t_{c2})$ and $\hat{T}_M = (t_2, t_3, t_{c1}, t_{c2})$.

3) Proposed Selection Method

Our solution of template selection now is straightforward: Choose $K$ templates from $\hat{T}_M$ by calculating the pass count of the templates. We test all the combinations with size $K$ and retain the one with the largest pass count. If there is a tie, the template set obtaining the smallest average similarity score will be selected as the final result. The complete approach is given in Algorithm 2.
Algorithm 2 Robust templates selection

Require: A template set $T_M = T \cup T_C$, a sample pass table $A_t$, and a pass count threshold $t_{pass}$

Ensure: A new template set $T^*$

1: Calculate the template pass table $A_t$
2: $A \leftarrow A_t \cup A_t$
3: $P(t_i) \leftarrow \sum A(., t_i)$ \{pass count of $t_i$\}
4: $\hat{T}_M \leftarrow \emptyset$
5: for $t_i \in T_M$ do
6: \hspace{1em} if $P(t_i) > t_{pass}$ then
7: \hspace{2em} $\hat{T}_M \leftarrow \hat{T}_M \cup t_i$
8: \hspace{1em} end if
9: end for
10: $T^* \leftarrow \emptyset$
11: for $T' \subset \hat{T}_M$ with $|T'| = K$ do
12: $P(T') \leftarrow \sum \bigcup_{t_i \in T'} A(., t_i)$ \{pass count of $T'$\}
13: $\text{Score}(T') \leftarrow \sum_{t_i \in T'} \text{Score}(t_i, t_i)$
14: if $P(T') > P(T^*)$ or $(P(T') = P(T^*)$ and $\text{Score}(T') < \text{Score}(T^*)$) then
15: $T^* \leftarrow T'$
16: end if
17: end for

TABLE 1. The Selection of Template

<table>
<thead>
<tr>
<th>The Subset Combination of $T_M$</th>
<th>Pass Count</th>
<th>Permeability</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_{M1} = { t_2, t_3, t_{C1} }$</td>
<td>11</td>
<td>100.0%</td>
</tr>
<tr>
<td>$T_{M2} = { t_2, t_3, t_{C2} }$</td>
<td>9</td>
<td>81.8%</td>
</tr>
<tr>
<td>$T_{M3} = { t_2, t_{C1}, t_{C2} }$</td>
<td>8</td>
<td>72.7%</td>
</tr>
<tr>
<td>$T_{M4} = { t_3, t_{C1}, t_{C2} }$</td>
<td>9</td>
<td>81.8%</td>
</tr>
</tbody>
</table>

Tab. 1 illustrates the templates selection procedure of the example in Fig. 4 with outlier removal. For this example, $\hat{T}_M = \{ t_2, t_3, t_{C1} \}$, the types of subset combination of $T_M$ is $C_2 = 4$. For these four types of combination, we compute pass count for each of them and select the one with the maximum, which is $\hat{T}_{M1} = \{ t_2, t_3, t_{C1} \}$. In this manner, the new template set is selected for next update round.

III. EXPERIMENTAL RESULTS

In this section, we first evaluate our template selection method using synthetic 2D points dataset, and then evaluate the performance of our template update framework using synthetic and real palmprint, palmvein and face datasets.

In the experiments, the update round is $Rd = 50$. If $Rd$ is small, the pass tables cannot accumulate enough information, and the size of test datasets is not big enough to support a larger $Rd$. As opposite, the times of update will be insufficient. The size of template set and candidate template set are $K = 5$ and $N = 25$, respectively. The similarity threshold of face is $t_{sim} = 0.75$, and they are 0.78 for palmprint and palmvein. All the test datasets adopt a closeness threshold $t_{close} = 1.25 * t_{sim}$, and the pass count threshold for outlier removal is $t_{pass} = 3$.

A representative template is irreplaceable and can be repeatedly matched. In a template set, some templates can be totally covered by other templates, i.e., if we remove these templates from the dataset, there is no effect on recognition performance, we call them replaceable templates. Replaceable templates cause data redundancy. We have designed an experiment to evaluate data redundancy. In more detail, for a sequence of in-coming samples, if a sample is matched and only matched to a template, the template is irreplaceable. Otherwise, the templates can be covered by others. With same parameters that we will implement in our experiments, for randomly selected template datasets, there are 19.8%, 20.1% and 19.7% templates are replaceable in our face, palmprint and palmvein datasets. Besides, although some templates are irreplaceable, they have not been matched for quite some time. They are likely unable to adapt to time-related changes, i.e., template aging. Hence, this kind of templates are useless for future recognition. A. EXPERIMENTS ON TEMPLATE SELECTION

To demonstrate the intrinsic property of our template selection method, we carried out several experiments on the 2D points dataset used in [30]. The classical template selection methods DENT [12], MDIST [12], as well as their competitors MMS [16], GMMS [16], FCM [14], are selected for comparison. Although the method in [30] is not designed for biometric template selection, and its purpose is to be more precise rather than more reasonable for our topic, we still take it into our comparison as a state-of-the-art clustering method. In the experiment, we manually set $K=2,3,4,5$, and depict the templates selected by each method, as shown in Fig. 6.

From Fig. 6, we can see that DENT tends to select the points which are far away from the others. It is therefore quite vulnerable to outliers, which has been pointed out by previous works [12], [31]. The results of MDIST, MMS, and GMMS are almost the same, with the selected points lie within the central area of the dataset. FCM and method in [30] seem to give much better results, as the cluster centers span the entire dataset, and each of them is representative of a portion of data points. However, when they assign the attribution of data, the size of clusters is a significant basis. Hence, if the number of tightly connected data is small, they are probably assigned to a cluster of massive data regardless of whether they are a reasonable cluster, e.g., the two points in the top-left. This is also the difference between the proposed generic-purpose method and specially designed methods for biometric template selection.

Besides, all the above methods do not take the operational threshold into consideration, which has a crucial influence on the performance of biometric systems. As opposite, the proposed method attaches much importance to the operational threshold. The finally selected templates are the most scattered, while at the same time, their coverage of the dataset with respect to the similarity threshold is maximized.
### B. EXPERIMENTS ON TEMPLATE UPDATE FRAMEWORK

In this subsection, we report the experimental results on synthetic and real datasets. Note that our method is a generic template update framework that is applicable to any biometric modality. In this work, we use synthetic palmprint data and real palmprint, palmvein and face data to demonstrate its effectiveness.

**FIGURE 6.** The template selection results of DENT, MDIST, MMS, GMMS, FCM, method in [30] and the proposed method. When $K = 2, 3, 4, 5$ templates are required, the results are shown in columns from left to the right, with final templates denoted by $\times$. 

<table>
<thead>
<tr>
<th>Template</th>
<th>DENT</th>
<th>MDIST</th>
<th>GMMS</th>
<th>MMS</th>
<th>FCM</th>
<th>Method in [30]</th>
<th>Proposed Method</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><img src="image1" alt="DENT" /></td>
<td><img src="image2" alt="MDIST" /></td>
<td><img src="image3" alt="GMMS" /></td>
<td><img src="image4" alt="MMS" /></td>
<td><img src="image5" alt="FCM" /></td>
<td><img src="image6" alt="Method in [30]" /></td>
<td><img src="image7" alt="Proposed Method" /></td>
</tr>
</tbody>
</table>

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/.
1) Synthetic Data of Palmprint

In order to simulate the variations in the biometric samples, we synthesized a dataset by capturing the palmprint images with more and more dot and line drawings on hand. Starting from a clean hand, in each session we drew one more dot or line on the hand, and captured around 10 ~ 20 images with different gestures. In a total, we got 571 images in 31 sessions. Some examples are shown in Fig. 7. We further add about 5% noisy data to the dataset to evaluate the robustness against outlier. The noisy data is generated by the feature-level fusion of genuine data with imposter data.

We first demonstrate the time variability in palmprint images by investigating the matching scores with different session intervals. After pre-processing the palmprint images, we perform feature extraction and matching using the competitive code method [32]. The matching scores are averaged with respect to session interval, as shown in Fig. 8. We can observe that the larger the session interval, the smaller the similarity score. As a result, the dataset shows a clear time variability and is suitable for the evaluation of template update methods.

In the experiment, the 571 images were partitioned into two sets. The 12 images with no drawings were used for initial template selection. The other images, together with noisy data, are then used by the template update framework in ascending order of drawings. Template selection was carried out using half of the images in each session, and the others were used as test samples. The performance with respect to recognition accuracy is evaluated by the false reject rate (FRR). We further evaluate the times that some outliers were selected as templates, which is the so-called creep in of errors. The introduction of outliers would cause an increased false acceptance rate (FAR), which is extremely harmful to practical biometric systems.

We tested the template update framework with DENT, MMS, k-Means, method in [30] and the proposed template selection method. Random selection of initial template set without update was adopted as a baseline, and its results were averaged over dozens of trials. GMMS and MDIST were omitted here because of the similar characteristics with MMS. K-Means replaced FCN since palmprint feature was composed of binary codes and its weighted mean required.

It can be observed from Tab. 2 that the initial template set without an update procedure is not adaptive to subsequent variations. As a result, a considerable portion of samples is falsely rejected. All the update methods are much superior to Random in terms of FRR, which demonstrates the effectiveness and necessity of the template update.

As for the update methods, MMS has the largest FRR because the templates cannot well be represented by the partition with the sum of match scores maximized. DENT, k-Means and the method in [30] have smaller FRR, at the expense of several creep in errors. As expected, the proposed method has the lowest FRR while at the same time is more robust against outliers, thus is better than both update and no-update methods.

Table 2. Results on Synthetic Palmprint Dataset

<table>
<thead>
<tr>
<th>Method</th>
<th>FRR</th>
<th>Creep in Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random(No-update)</td>
<td>61.52%</td>
<td>0.64</td>
</tr>
<tr>
<td>DENT</td>
<td>23.51%</td>
<td>15.18</td>
</tr>
<tr>
<td>MMS</td>
<td>32.58%</td>
<td>0.00</td>
</tr>
<tr>
<td>k-Means</td>
<td>25.10%</td>
<td>4.29</td>
</tr>
<tr>
<td>Method in [30]</td>
<td>27.60%</td>
<td>5.34</td>
</tr>
<tr>
<td>Proposed</td>
<td>21.28%</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Next, we investigate the indices of the selected templates with the progress of template update. The initial templates are randomly selected from samples of session 1, then the other samples are sequentially taken as the input of our template update framework. Candidates, as well as other information, is collected with the progress of sample classification. Template selection is carried out once the update condition is satisfied. As shown in Fig. 9, the average indices of selected templates increase monotonically with update round, which means the templates are gradually replaced by new coming samples. Besides, the FRR decreases from 55.932% to 15.569%. Thus we can conclude that the proposed template update framework is robust and effective for biometric data with evident time variations and outliers.

2) Real Data

In this section, we simulate the proposed method in three datasets.

- Palmprint Dataset: We have collected a palmprint dataset with images of 60 hands over 12 months. In each
day palmprint images are captured in 0 ~ 10 separate sessions, with only 1 image per session retained. Totally we have obtained about 300 ~ 600 images per hand. The long-term collection aims to evaluate the performance of the template update framework.

- **Palmvein Dataset**: This dataset is collected just like palmprint. Actually, the palmvein dataset is collected at the same time as the palmprint dataset.
- **Face Dataset**: We use the CMU Multi-PIE face database [34]. It contains 337 subjects, captured under 15 view points and 19 illumination conditions in four recording sessions.

We first test the time variability in palmprint and palmvein images by calculating the matching scores with different time intervals. Surprisingly, no obvious time variability can be observed. Since the results of different hands are quite similar, we only report those of 3 hands in Fig. 10. Except for the big wave around the right end, the similarity scores with a different time interval, from 1 day to 360 days, are almost unchanged for a specific hand. We found similar results in palmvein test. The result is inconsistent with previous research that noticeable intra-class variations can be captured with long-term biometric data acquisition [2], [28], [33]. The probable reason is that palmprint and palmvein are more robust against aging, injury, hairstyle, etc., which would cause large variations of other biometric traits, e.g., face. As a result, during the 12 months, the palmprint and prlmvein images under consideration only undergo neglectable changes. Besides, the face is considered as a biometric trait contains plenty of intra-class variations during long term data acquisition [33].

On this foundation, we can simulate the effectiveness of the method for these two types of biometric traits: one contains large or moderate intra-class variations. For each sample, we use half of the samples for template selection ($K = 5$) and the rest samples for testing. The FRRs in different datasets with respect to different similarity threshold are depicted in Fig. 11. We can see that the superiority of the proposed method is evident. In most cases, the proposed surpasses random selection by a large margin. Furthermore,
TABLE 3. Decrease of FRR on Different Datasets

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Similarity threshold</th>
<th>FRR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Face</td>
<td>0.80</td>
<td>35.5%</td>
</tr>
<tr>
<td></td>
<td>0.78</td>
<td>31.2%</td>
</tr>
<tr>
<td></td>
<td>0.76</td>
<td>24.7%</td>
</tr>
<tr>
<td></td>
<td>0.74</td>
<td>19.3%</td>
</tr>
<tr>
<td></td>
<td>0.72</td>
<td>15.1%</td>
</tr>
<tr>
<td></td>
<td>0.70</td>
<td>9.8%</td>
</tr>
<tr>
<td>Palmprint</td>
<td>0.85</td>
<td>24.3%</td>
</tr>
<tr>
<td></td>
<td>0.83</td>
<td>19.9%</td>
</tr>
<tr>
<td></td>
<td>0.81</td>
<td>12.1%</td>
</tr>
<tr>
<td></td>
<td>0.79</td>
<td>9.6%</td>
</tr>
<tr>
<td></td>
<td>0.77</td>
<td>8.2%</td>
</tr>
<tr>
<td></td>
<td>0.75</td>
<td>4.4%</td>
</tr>
<tr>
<td>Palmvein</td>
<td>0.85</td>
<td>23.9%</td>
</tr>
<tr>
<td></td>
<td>0.83</td>
<td>20.1%</td>
</tr>
<tr>
<td></td>
<td>0.81</td>
<td>14.7%</td>
</tr>
<tr>
<td></td>
<td>0.79</td>
<td>8.2%</td>
</tr>
<tr>
<td></td>
<td>0.77</td>
<td>3.1%</td>
</tr>
<tr>
<td></td>
<td>0.75</td>
<td>5.7%</td>
</tr>
</tbody>
</table>

Conversely, if the biometric traits have moderate intra-class variations like palmprint and palmvein, the randomly selected template set with a sufficient size may contain enough information, so the effect of FRR decrease is not significant. As mentioned before, more templates bring more space requirement and more pressure to compute. In this context, the proposed framework is more attractive for systems that are subject to the storage and have a stronger demand for processing speed.

IV. CONCLUSION

In this work, we have proposed a generic template update framework as the solution to intra-class variations for biometric systems. It is applicable to the systems using prototype and nearest neighbor classifier. We have carried out several experiments on synthetic palmprint and real datasets have moderate and large intra-class variations. The comparative study has demonstrated the robustness and effectiveness of the proposed method. According to the experiment results, the proposed framework is more effective to the systems have a strong demand for security, speed and lightweight storage pressure.

It is worthy to emphasize that our method is a generic framework that can be applied to various biometric traits. In the paper, we adopt three contact-less traits which have larger intra-class variations than the majority of contact based traits. That is for the purpose of a clearer experiment results. We believe that the effectiveness of the method will be better if datasets contain a longer-term collection or more intra-class variations. In the future, we will collect more samples and conduct more experiments on other datasets, such as iris and fingerprint to extensively evaluate its performance and generality. At the same time, to obtain a long-term experiment results, we are ready to apply the proposed framework to an application system.

REFERENCES


[14] A. Lumini, L. Nanni, A clustering method for automatic biometric tem-


XI CHEN received the M.S degree from School of Computer Science and Engineering, Hebei University of Technology(HEBUT), Tianjin, P. R. China in 2013. He is currently pursuing the Ph.D. degree in School of Electronic and Information Engineering, HEBUT and a research assistant in Pattern Recognition Research Center, Beijing Institute of New Technology Applications. His research interests covers image processing, pattern recognition and computer vision.

MING YU (M’15) received the Ph.D. degree in communication and information systems from the Beijing Institute of Technology. Since 1989, he has been teaching at the Hebei University of Technology and a Full Professor at the School of Computer Science and Engineering since 2000. His research interests focus on biometrics of voice and image vision information fusion, image mathematical transformation, efficient algorithm of image and video coding, computer network, and high compression image video transmission under wireless mobile network.

FENG YUE received the B.Sc. and Ph.D. degrees in computer science from the Harbin Institute of Technology (HIT), in 2003 and 2010, respectively. He is now a researcher in the Pattern Recognition Research Center, Beijing Institute of New Technology Applications. His current research interests include computer vision and pattern recognition.

BIN LI received his MSc. and Ph.D. degrees in computer science from Harbin Institute of Technology (HIT), Harbin, China in 2000 and 2006, respectively. From 2006 to 2008, he worked in the School of Computer Science and Technology, HIT as a lecturer. From 2008 to now, he has been working in Key Laboratory of Pattern Recognition of Beijing Academy of Science and Technology as a professor, and in Beijing Institute of New Technology Applications as a director. So far, he has published over 30 papers and two books. His research interests include signal processing, pattern recognition, and biometrics, etc.
YINGCHUN GUO received the Ph.D. degree from the School of Information, Tianjin University, Tianjin, China, in 2006. She is currently an Associate Professor with the School of Computer Science and Engineering, Hebei University of Technology, Tianjin. Her research interests are image saliency and its application, image processing, artificial intelligence, and image compression.

***