Image and Vision Computing 73 (2018) 17–27

Contents lists available at ScienceDirect

Image and Vision Computing

journal homepage: www.elsevier.com/locate/imavis

Found a good match: Should I keep searching? — Accuracy and performance in iris matching using 1-to-First search ☆

Andrey Kuehlkamp*, Kevin Bowyer

University of Notre Dame, Notre Dame, IN, United States

ARTICLE INFO

Article history:
Received 15 September 2017
Received in revised form 15 January 2018
Accepted 16 March 2018
Available online 27 March 2018

Keywords:
Biometrics
Iris recognition
Error rates
Identification
Accuracy
Search
1:First
1:N
Open-set

ABSTRACT

Iris recognition is used in many applications around the world, with enrollment sizes as large as over one billion persons in India's Aadhaar program. Large enrollment sizes can require special optimizations in order to achieve fast database searches. One such optimization that has been used in some operational scenarios is 1:First search. In this approach, instead of scanning the entire database, the search is terminated when the first sufficiently good match is found. This saves time, but ignores potentially better matches that may exist in the unexamined portion of the enrollments. At least one prominent and successful border-crossing program used this approach for nearly a decade, in order to allow users a fast "token-free" search. Our work investigates the search accuracy of 1:First and compares it to the traditional 1:N search. Several different scenarios are considered trying to emulate real environments as best as possible: a range of enrollment sizes, closed- and open-set configurations, two iris matchers, and different permutations of the galleries. Results confirm the expected accuracy degradation using 1:First search, and also allow us to identify acceptable working parameters where significant search time reduction is achieved, while maintaining accuracy similar to 1:N search.

1. Introduction

One of the most powerful biometric modes is *Iris Recognition*. It is based on images from the area of the eye surrounding the pupil, called the iris. Each iris contains a complex pattern composed of elements like crypts, freckles, filaments, furrows, pits, striations and rings. These texture details are what make the iris particularly useful for recognition [1].

Since its first demonstration by Daugman [2], iris recognition has evolved to become one of the best-known biometric characteristics. The largest biometric database in the world, the Aadhaar program in India, has already collected 1.13 billion people's irises (and fingerprints) for enrollment [3].

In 2016, Somaliland started to register voters using iris biometrics [4]. The motivation is to prevent voting fraud, after authorities found a large number of duplicate registrations, even with the use of facial and fingerprint recognition [5]. The decision was made after months of testing and preparation, aided by a feasibility study by a team of academic researchers [6].

Since 2002, countries like the United Kingdom, Canada and Singapore have used iris biometric systems to perform border-crossing checks on frequent travelers. Similarly, the United Arab Emirates (UAE) has employed an iris-based biometric system to keep track of banned travelers since 2001. The UAE system is known for performing approximately 14 billion IrisCode comparisons daily [7].

Iris recognition is being deployed in an increasing number of applications, and with larger and larger database sizes. Although iris matching can be performed in an extremely rapid manner, the need for optimization becomes stronger as the number of enrolled persons in applications becomes larger. In this sense, we analyze one search technique that is known to have been utilized in some operational scenarios, but whose performance and accuracy have not been considered in the research literature.

In iris databases, the traditional search approach for identification is called 1:N, which means the entire biometric enrollment is scanned and the best match is selected. For nearly a decade, the NEXUS border-crossing program [8] employed a variation of this search technique, called 1:First, in order to improve search speed. In 1:First search, the search of the biometric enrollment is terminated when the first biometric template that satisfies the matching threshold is found. This approach generally speeds up the search. However, the biometric template selected by this approach may not be the best match. The biometric template matched in 1:First search is more

☆ This paper has been recommended for acceptance by Vitomir Štruc.

* Corresponding author.
E-mail addresses: akuehlka@nd.edu (A. Kuehlkamp), kwb@nd.edu (K. Bowyer).

https://doi.org/10.1016/j.imavis.2018.03.003
0262-8856/

likely to not correspond to the same subject as does the biometric probe than in 1:N search.

Kuehlkamp and Bowyer [9] found a significant difference in the 1:First False Match Rate (FMR) in comparison to 1:N search, especially with larger enrollment sizes, and higher rotation tolerances. However, as pointed out by the authors, Kuehlkamp and Bowyer [9] were not complete in some aspects. The enrollment sizes used in the experiments were rather small. The evaluation was not done using a "commercial quality" iris matcher. And the experiments did not contemplate open-set scenarios. The objective of this work is to address each of these issues.

In this work, experiments are performed on two iris matchers, the second being a well-known commercial matcher. Also, we use data augmentation techniques to increase the size of our dataset, and perform experiments in enrollment sizes that can be more representative of real world applications. This also allows us to create and test open-set scenarios. In addition, we perform experiments using different permutations of the same enrollment, in order to verify how that could affect 1:First accuracy.

2. Background and related works

Iris recognition, like other biometric modalities, can be used in two types of identity management functionality [1]: verification and identification.

In **verification**, the task of the system is to verify if the identity claimed by the user is true. In this case, the biometric reference from the user is compared to a single biometric probe in the database (*one-to-one matching* [10]). In turn, when performing **identification**, the user does not claim an identity. Consequently, the system has to compare the user's biometric probe with the biometric references of potentially all the persons enrolled in the database (*one-to-many matching* [10]).

Within identification, it is possible to distinguish *closed-set* and *open-set* identification tasks. With a closed-set, the user is known to be enrolled in the database, and the system is responsible to determine his or her identity. On the other hand, when doing open-set identification, the system must, before trying to identify a user, determine if he or she is enrolled in the database [10]. This work is concerned with one-to-many matching as used in an identification system, and particularly, with exploring the difference between two possible implementations of one-to-many.

The *comparison* procedure is a core part of every biometric identification or verification system. In this procedure, the system compares the biometric probes acquired from the user against previously stored biometric references and scores the level of similarity or dissimilarity between them. According to a predetermined threshold, the system then makes a decision about the user: either it is a *match* or a *non-match*. Declaring a match means to assume that the system accepts both biometric samples as being originated by the same human source [10].

2.1. Iris comparison output

Two types of errors can be made by biometric systems: *False Match* (FM) and *False Non-Match* (FNM). A FM occurs when biometric probe and biometric reference from different individuals are incorrectly classified as a match. Conversely, FNM occurs when biometric probe and biometric reference of the same individual are not recognized as a match [1].

These errors are very similar to *Type I* (false-positive) and *Type II* (false-negative) statistical errors. However, this traditional standpoint usually does not contemplate a scenario variation: open-set *vs.* closed-set [11]. In both of these cases, there is an enrollment G of biometric references, and the comparisons made against that enrollment come from the biometric probe set P. If the identities in P are a proper subset of G ($P \subseteq G$), then the scenario is said to be *closed-set*. On the other hand, if any of the identities in P are not contained in G, that is, $P = \{P \cap G \wedge P \nsubseteq G\}$, the scenario is called *open-set*.

The different search methods, 1:N and 1:First, can produce different results for the same biometric probe and list of biometric references [9]. These situations are described in detail in [9]. Ultimately, this distinction is the source of the accuracy difference between 1:N and 1:First.

2.2. Comparison output in closed and open set scenarios

As mentioned by ISO 19795-1:2006 [11], and shown in Table 1, the conventional definition of comparison results is a little different when considering open-set and closed-set scenarios. In a closed-set scenario, we have the typical cases of TM and FNM for biometric references. An interesting peculiarity of the closed-set scenario is that TNMs cannot happen, because all the references are enrolled (Table 1[b]).

On the other hand, in an open-set scenario, all four typical cases occur, but there is a distinction to be made: false matches (Table 1[c]) can occur either as Enrolled False Matches (EFMs), like in a closed-set, or as Unenrolled False Matches (UFM), when a biometric probe of an unenrolled individual is similar enough to match one of the enrolled biometric references.

2.3. Traditional searching: 1:N

Mukherjee and Ross [12] define the problem of iris identification in terms of comparing a probe iris sample q, with enrolled iris samples $D = \{d_1, d_2, d_3, \ldots d_n\}$, in order to determine the identity y of the query sample. Each enrollment sample d_j, $j = 1, 2, \ldots, n$ is associated with an identity y_j. Consequently, the computational complexity of the process is directly linked to the number of enrolled biometric references $|D| = n$ in the enrollment.

Matching iris samples based on Daugman's approach is an operation that involves the accumulation of bitwise XOR operations between the biometric templates, and can be done quite efficiently. However, the computational complexity of the task grows linearly with the increase in enrollment size, and the complexity for tasks like de-duplication grows quadratically regarding the size of the database, as noted by [13].

Unlike other numeric or lexicographic data, biometric samples do not have any natural ordering [14]. This hinders any attempt to index biometric databases. Since there is no order for the enrollment records, the obvious approach used in automatic iris identification is to compare the probe to every enrollment record.

Other efforts have been made in the sense of improving the search performance in iris databases. In [15], the parallelization of the algorithms involved in the iris recognition process is proposed, including the template matching. However, their parallelized version still has its overall performance directly associated to the size of the database. In another attempt to address the issue, Hao et al. [16] propose an

Table 1
Possible outputs for matching against an enrollment in Closed-set and Open-set scenarios.

	Matching result			
	Closed-set		Open-set	
	True	False	True	False
Enrolled reference	TM	FNM	TM	FNM
Non-enrolled reference	EFM[a]	N/A[b]	EFM/UFM[c]	TNM

[a] *Enrolled False Match*: An enrolled non-mated biometric reference is similar enough to be considered a match.
[b] True Non-Match cannot happen, because there are no unenrolled references.
[c] *Unenrolled False Match*: An unenrolled biometric reference is similar enough to be considered a match.