Adaptive Biometrics: a comparative analysis of applications to face recognition

Rens van der Heijden
r.w.vanderheijden@student.utwente.nl
University of Twente

ABSTRACT
This paper reviews the topic of adaptive biometrics, describing two core segments seperately. These segments, selection and updating, are those that select useful templates and improve the stored template of the user, respectively. By using adaptive biometrics, we tap into the vast amount of unlabeled data collected during the operation of the system. This method has shown significant results in improving the accuracy of biometric systems.

Keywords
Adaptive Biometrics, Semi-Supervised Learning, Face Recognition

1. INTRODUCTION
Adaptive biometrics, the field of biometric systems that can adapt to changes of the underlying biometric, is a relatively new field with a large overlap with the machine learning field, in particular with what is called semi-supervised learning. The field of adaptive biometrics shows promising results for improving the recognition accuracy for current biometric systems by improving itself using data collected during the recognition phase [7, 4].

A typical biometric system consists of two phases: enrollment and identification. At the enrollment stage, a template is generated to represent a taken biometric sample. At the identification stage, the system is presented with a new biometric sample (and in the case of authentication, as opposed to recognition, an identity is also provided) and a comparison between the template and the sample is made. That is, the system attempts to identify the given sample. However, all these new samples that are identified by the system are not used to improve the performance.

Biometrics can be considered a practical example of a classification problem. A classifier is used to identify (classify, in machine learning terms) the new sample to the correct class (that is, match it to the correct template). In the case of authentication, this classification is binary (either the sample belongs to a selected class or not).

Adaptive biometrics tries to tap into the vast amount of unlabeled data collected during each identification stage in the lifetime of the biometric system. This is performed in a third phase, which consists of two steps; determining the suitable new samples and merging these into the original template. The process is illustrated in figure 1. Note that recent research points out that this phase can also be performed online, though it is usually considered as a batch process [4].

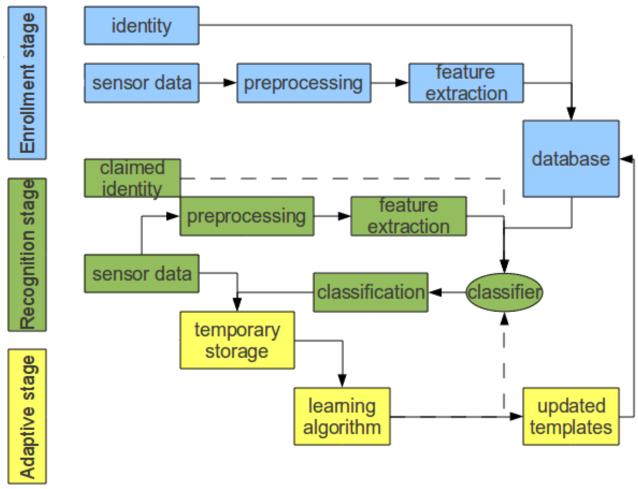

Figure 1: The adaptive phase in the whole picture of biometrics. The dashed lines are not always applicable; eg. claimed identity is only used in verification and a new classifier may be produced depending on the learning algorithm and classifier, but it is not required.

Adaptive biometrics is not to be confused with adaptive methods to perform preprocessing. Many such algorithms exists, for example [10], which uses an adaptive method to normalize lighting condtions. Though preprocessing is a very important topic in biometrics, this paper separates this task from the feature extraction phase, for the reason that preprocessing may benifit any feature extraction algorithm. Also, preprocessing algorithms are generally part of the field of computer vision and image processing, which are used in biometrics but represent a very different approach to the problem. This paper also discusses feature extraction algorithms because they influence the representation of templates, which in turn means they determine possible methods for the template update step.

To perform the first step, a form of machine learning called Semi-Supervised Learning (S-SL) is used learn from the unlabeled data. Many methods to perform S-SL exist, but the general idea is to select those unlabeled samples from which the system can learn the most, but can still classify correctly. Survey papers are available that explain the methods in detail [14, 9]. The simplest methods are self-training and co-training, which use the original classifier(s) to classify new data and subsequently update the templates with the most confident classifications they make. A notable alternative method is using graphs to describe relations between images, rather than explicitly defining a template.

In the second step, the selected samples to be processed into the template to create a new template. The simplest example of this is to just add the sample to the database of templates and relate it to the same user as the old template. However, this method is very crude and expensive in terms of storage. For example, in some cases, such as traditional Principal Component Analysis (PCA), the original algorithm for computing templates requires all the samples to be stored [3], thus effectively storing all images in addition to all samples. In state-of-the-art systems, the method to improve the template is highly dependent on the biometric and thus the algorithms used for the identification process. This method essentially entails creating a new, better template out of the previous and new samples, often called a super-template.

This paper will use biometric face recognition as an example application of adaptive biometrics. There are several reasons for this. First, face recognition is an interesting use case because the performance can vary greatly with the environment the biometric sensor is placed in and can thus benefit more from a performance improvement. Second, face recognition is a popular topic, so there are many papers available that apply their adaptive biometrics algorithm to a face recognition problem. Third, face recognition is regarded as a computationally expensive problem, which makes it an challenging topic for on-line methods.

In this paper, face recognition is used as an application scenario for adaptive biometrics to evaluate the various available methods. The various methods for performing S-SL will be descussed in section 2, followed by a brief introduction of face recognition in section 3, as well as a discussion of issues regarding the template update step. In section 4, the learning methods are compared in a qualitative analysis and in section 5, a short overview of other important topics is given. The paper will be concluded in section 6.

2. SEMI-SUPERVISED LEARNING

In this section, the most important semi-supervised learning (S-SL) methods will be explained.

2.1 Self-training

Self-training is arguably the simplest form of semi-supervised learning. It is applied as follows: a classifier, initially trained with labeled data, classifies each sample in the set of unlabeled data with a confidence level. The confidence level is the measure normally used for classification. Instead of discarding the data, as would normally happen after classification of biometric samples, the samples are stored for later use. When sufficient data is collected, those samples with the highest confidence are 'added' to the labeled data set. Note that in this context, adding to the labeled data set implies the use of a template update mechanism. The classifier is then trained using the new labeled data set. This process is repeated until some condition is met; this can be a fixed amount of iterations, until all new samples are processed or until the highest confidence is below some predefined threshold.

This method is simple but poses a significant challenge on a system designer when applied in a scenario where the self-training process also has to identify new classes, as is the case in some recognition scenarios. An example of this could be an unsupervised surveillance scenario. The problem lies in the requirement for a confidence measure in order to add templates to the system. It has been suggested that a new class can be created for samples confidence level below a set threshold. However, a significant risk is involved when this solution is used, especially in the early stages of operation, because a bad initial template for a user of the system could cause this user to be represented in the system twice. A possible solution could be to create a new class for the sample with the lowest confidence level and use it as a new class. This avoids the problem to some extent, since it is unlikely for the sample that is farest from the templates actually belongs to one of these templates.

2.2 Co-training

Co-training or co-updating is a useful method when a multi-modal biometric system needs to be improved. This method is based on the idea that two (or more) classifiers based on widely different methods can compensate the errors in each other. Since this is an assumption on which the concept of multi-modal biometrics is based, so applying co-training to such a system does not add uncertainty. The process of basic co-training is very similar to that of self-training, except there are now two classifiers. These classifiers, A and B (for example, a biometric system with fingerprint and iris identification), are first trained individually with the labeled data, just like in self-training. However, in the adaptive phase, after both classifiers have labeled the unlabeled data, they agree upon the samples to be added to the data set. There are a few possibilities for how to do this.

First off, it can be an iterative process where A 'learns from' B by adding the samples corresponding to those that B identified to the labeled data set of A, which is then retrained, relabels the unlabeled data set and 'teaches' B in the same manner, until some condition is satisfied. On the other hand, it is possible that some combination of the scores provided by A and B is used to simulate a single classifier, similar to methods used in multi-modal biometrics.

Both these options can be extended to include multiple classifiers. For multiple classifiers, the second option has gained some attention, in a scheme called tri-learning, which is generalized by [14] to the democratic co-learning algorithm. This scheme is based around an extension of the concept of ensemble classification, which performs classification by a majority vote of several classifiers.

In the past, one major challenge of co-training is the requirement that the underlying classifiers be conditionally independent. Recent advances have illustrated that this condition can be replaced with weaker requirements, paving the way for learning algorithms that use other types of multi-biometrics, in particular the multi-algorithmic (use different classifiers on the same sensor data) and multi-presentation (take several samples rather than a single one as data) approaches. Advantages of these approaches over multi-modal biometrics include reduced cost, as only one sensor is needed, and more possible applications, eg. in situations where relying on user interaction is not possible.

2.3 Graph-based methods

Graph-based methods build a graph of new samples based on various measures like time difference, sample similarity and other characteristics that can be found in the image. Similarity is defined by these various measures and represented with a (possibly weighted) edge between samples. A significant advantage of these methods is that they can easily be extended to include additional information gathered by the sensor, such the time between two samples, the time of day, overall brightness, or colors (from clothing), which can be used to link samples even though recognition based on only the face is impossible. An additional advantage is that established graph algorithms can be applied to face recognition, for which significant amounts of research into complexity and optimization have often already been done. Many of these graph-based methods exist; examples include linking the samples with edges and classify by traversing edges until the first labeled image is encountered [1] and an approach based on the mincut algorithm [6].

3. TEMPLATE UPDATING

This section explains why face recognition is used as a basis for comparison and then summarizes the basics of popular face recognition methods. In addition, possible template update methods that are specific to the recognition method are discussed. Finally, some generic methods are discussed.

3.1 Face recognition as a case study

The aim of this paper is to show that adaptive biometrics can deliver a significant improvement of the performance of biometric systems. Face recognition is a good case study for this goal, because it is a relatively difficult problem, caused by a relatively large variety between two samples in the same class (intra-class variation), when compared to the variety of two samples that belong to different classes (inter-class variation). Causes of this problem include lighting differences, aging and pose variations. Many pre-processing algorithms have been devoloped to counteract these problems, but some of them can also negatively affect recognition rates.

Face recognition itself is considered a good biometrics because it requires very little interaction from the user when compared to other systems like fingerprint or iris. It is also relatively simple to implement and has attracted much attention over the last fifty years, both in corporate and academic communities [13]. Because face recognition is considered to be a challenging topic, there is much data available on the performance of these systems, which this paper can use to quantitatively support the evaluation.

Face recognition is also challenging for various implementation reasons. The most important ones are; the considerable size of samples and thus templates, the limited quality of affordable sensors such as cameras and the previously noted problem caused by a significant amount of noise between samples of the same class. Template size is an especially important challenge for adaptive biometric systems, as some of these continuously add new samples to the template of a user, causing rapid growth in memory requirement. Of course, simple solutions such as deleting old templates are possible, but this would reduce the improvement that adaptive biometrics attempts to achieve.

A disadvantage of selecting face recognition is that it is more difficult to test the co-training method with another established biometric, such as iris, fingerprint or hand geometry, because this removes the advantage of having low user interaction, thereby excluding surveillance scenarios. However, in recent years, face recognition using moving images has begun to develop [13]. This may be a candidate for the second classifier to be used in a co-training context, but the lack of independence between the variables that the classifiers rely on may harm rather than improve the performance of this system [7].

3.2 Face recognition methods

In this section, the basics of the face recognition algorithms will be explained, followed by an explanation of how template updating can be performed by the template representation. First, Principle Component Analysis (PCA), a popular feature extraction algorithm throughout the image recognition field, is discussed, as well as its application in face recognition. The direct application of PCA is also called the eigenface method. Second, the basics of face recognition using gabor wavelets are illustrated as an example of a non-holistic algorithm.

3.2.1 Principle Component Analysis (PCA)

PCA is a statistical method that finds the eigenvectors of a matrix of input data. In face recognition, if each face image is considered as a column vector, appending all these vectors produces such a matrix. Then, after performing PCA, one will obtain a very large amount of eigenvectors, which are commonly called eigenfaces.

After applying PCA, each image can be considered a linear combination of the produced eigenfaces. Then a subset of these eigenfaces (namely the most discriminative ones, that is, those with the greatest corresponding eigenvalues) can be used to represent any image as a linear combination of these vectors. The resulting values of such a mapping into the eigenspace (the space that has a basis consisting of the selected vectors) is considered to be the representation of the image. After mapping all labeled data into this space, one can identify a new sample by also mapping it into this space and determining the nearest template by some measure (eg. Euclidean distance).

Eigenfaces have an important disadvantage when template updating needs to be performed, either by later enrollment sessions or by adaptive biometric methods. This is because to compute the eigenspace to which all samples are mapped, the original samples of all the labeled data is required, since the eigenspace should depend on all the labeled samples in order to create correct templates and to correctly extract features from new samples. In a normal, static biometric system, where templates cannot be added, this is not a problem, since the eigenspace is never recomputed, but in an adaptive biometric system, the images as well as the templates need to be stored on the system. This method is called *batch training*, as opposed to *eigenspace updating*, a more recent and computationally efficient method for computing the eigenspace. Since the eigenspace updating method relies on an iterative process that does not require the original images, but only the previous eigenspace, it avoids the above problem [3]. Thus, techniques for eigenspace updating are essential to adaptive biometrics.

3.2.2 AdaBoosted Gabor Wavelets

Gabor wavelets are functions that are used to recognize shapes in a two-dimensional image. The most important parameters of this function are its frequency and the orientation, which determine the kind of shape that is detected and its orientation, respectively. Objects recognized on the face could include eyes, lips and other facial features. A significant advantage of recognizing shapes is that Gabor wavelets are quite resistant to noise from lighting variations. However, the amount of possible gabor wavelets on an image is enormous; for each pixel on an N $\times$ M image, there can be several Gabor wavelets with different orientations and frequencies, each of which makes up a single feature.

Because it is not feasible to record thousands of features, algorithms have been devised to select distinctive features. However, just like when PCA is used, doing this brings with it a significant computational cost. An example of such an algorithm is AdaBoost, which has been successfully used in face recognition scenarios [11]. AdaBoost works by creating simple, weak classifiers, which can be created quickly and the training set classify quickly. Each of these classifiers is usually trained from one feature and classifies all images as impostor or genuine. The error rate of each classifier is then computed and used to select the best classifier. This classifier is stored for later use and the weights given to the samples in the training set are updated. This is done for T iterations, after which a good classifier can be built by using the weak classifiers identified during the algorithm. Alternatively, the features from which the classifiers were built can be used as a template, providing a low-dimensional feature vector for some other classification algorithm [11].

While it is considered to be a good way to perform face recognition, updating of AdaBoosted templates is a challenging problem. This is because either a large amount of storage is required not only to store all the images but also to store all the classifiers that are produced. Alternatively, these classifiers can all be recomputed every time an update is necessary. Apart from this storage issue, it is clear that increasing the amount of images increases the amount of times AdaBoost is run, as well as the amount of images that need to be classified. Typically, assuming an image of size 30×40 and gabor wavelets using 8 different orientations and 5 different frequencies, the amount of features is 48000 per image. Thus, for a single iteration of AdaBoost, 48000 classifiers are built to create one template. Clearly, even for a small training set, the number of classifiers is very high. Since adaptive biometrics relies on a growing 'training set' (or rather, labeled data set), this is a disadvantage of the gabor wavelets method.

3.3 Generic update methods

Besides the template specific approaches above, it would be desirable to have generic methods that can readily be applied without having to bother with the new implementation issues that the these approaches introduce. The most obvious such method is simply adding all the templates to the database of images. This method is currently quite popular, because of its ease in terms of implementation and convenience when evaluating the performance of a semi-supervised learning method. However, it is clearly not usable on the long term, since in such situations, required storage will only increase. In addition, one problem that adaptive biometrics tries to solve, that of degradation of biometric traits, such as facial features caused by aging, is not solved by this template update method, as the old templates remain in the system. Rather than having to deal with a deletion protocol, it is desirable to let the template change as the biometric traits of the user change, in order to retain the match between user and template. Another problem with the simple addition protocol is that the amount of comparisons for identification tasks, as opposed to verification tasks, goes up quickly, since identification is performed by finding the most similar template in the database.

There have also been proposals to use algorithms that seek to select the best template by collecting a set of samples and then picking out the best (most representative) resulting templates to store in the database [12]. This scheme can be extended to the above situation by collecting templates up to a certain threshold t and then selecting the best $k < t$ templates. However, it remains that newer templates may be less representative when regarded in the context of old samples. Though [12] reports that their augmenting update, which could be considered a self-training fingerprint system, improves accuracy, their test set was gathered over the course of four months (at most five samples per day), the research community should be careful before concluding that this method solves the described problem.

4. QUALITATIVE COMPARISON

In this section, available work related to face recognition using adaptive biometrics will be explained per S-SL method and compared with eachother.

4.1 Self-training

Self-training is an easy method to implement, but it offers limited improvement compared to co-training. It has been suggested that self-training can be used in an on-line manner; that is, improve the accuracy of the system as it runs, as opposed to the conventional method of an off-line adaptive phase [5]. This would essentially eliminate the extra adaptive phase and append a conditional template update to the end of the recognition phase. An additional advantage is that no stopping criteria is required, as the new sample is either used or not used. However, additional work is required to ensure the safe operation of these systems, as they may be vulnerable to additional attacks that attempt to exploit the continuous tempate updates. Samples of so-called wolves from Doddington's zoo (users that are likely to be incorrectly identified) have shown to be able to be introduced into another user's gallery quite commonly [5].

When self-training is performed offline, it may be helpful to speed up learning by adding more than one template at once, as well as different stopping criteria [8]. This is similar to co-training, but has the advantage that only one biometric system is used, which simplifies the process and increases its suitability for situations where interaction is limited, such as surveillance.

4.2 Co-training

Co-training has been shown to provide a significant matching-performance improvement. For example, in [7], is used with two classifiers; one using a fingerprint and one using a frontal face image. This work uses a simple agreement scheme, where both classifiers individually classify the unlabeled data and the union of the samples with highest confidence is added to the labeled data. The improvement in both classifiers, as well as their combination using the system as a multi-modal system with product fusion is shown. Note that here, cotraining is mostly applied to improve the face recognition system, rather than to improve the overall recognition accuracy; the fingerprint method is initially almost as accurate as the product fusion of both methods. Also note, however, that the recognition accuarcy of the fingerprint system is still improved by the face recognition system as it is co-trained. This example experiment shows that even the individual systems can benifit from co-training, showing thus that in an ensemble scenario co-training may also be helpful.

In [2], the authors propose a theoretical framework to predict co-training behavior, given the assumptions that the biometrics are independent and that the amount of different observations is bounded by some fixed number for each biometric. While the first is a reasonable assumption in most cases, the second is somewhat vague. However, given this number, the amount of new samples, the amount of templates for the slave biometric (the biometric that is 'learning') at the first step and the false reject rate (FRR) of the current master biometric (the biometric that is 'teaching'), a recursive relation is given for the size of the gallery at the ith step. Since the paper uses the generic update method of simply adding the image to the gallery, this can be further optimized by fusing similar images into one template. In addition, a relation between the size of the gallery of a class and the false reject rate of the biometric is proposed, using the average number of examples that produce a score above the threshold of the system as a measure m_i. For the experiments, this measure is used to select the initial templates, using the reasoning that initial templates are produced during an enrollment session, which is a highly supervised environment. Assuming that the initial template is good, it is reasonable to expect this template to have a high value for the measure m_i. However, it should be noted that though the initial template is indeed a 'good' image, this does not mean that it is also the best template. For example, consider enrollment is performed during daytime, while the user only passes the sensor in the morning and the evening (eg. to enter a secure building). The experiments are executed using the standard implementations of the PCA (for faces) and "String" (minutiae based fingerprints) algorithms. Results show that prediction for 'easy' classes, represented by high value of m_i, is quite accurate, while slightly optimistic for the average and hard classes. It can also be observed that the face classifier provides most of the new templates for the fingerprint classifier in the first ten co-training iterations. Though not noted by the paper, this observation could be used to speed up co-training.

4.3 Graph-Based

Several graph-based approaches have been proposed. These are difficult to compare, because they usually focus on a specific scenario and are connected to additional information (eg. timestamps and color similarity of the entire, unmodified image). However, there are significant advantages to using graphs: these algorithms focus on the underlying structure of the data and are usually well-defined in terms of computational complexity. Graph theory is a widely studied area in computer science, which means existing, well-established algorithms can be re-used.

For example, in [1] a method is proposed that correlates different images with three types of edges; time, color and face biometric edges. These edges represent short (a few seconds), mid (a day) and long (many years) term similarity, focussing on short time intervals, similar clothing and similar faces respectively. This paper focuses on the uncontrolled recognition (ie. surveillance) scenario. The dataset thus consists of people in any orientation. The paper applies simple image analysis techniques, focussing instead on the graph representation of the data. Classification is performed by a Gaussian field and harmonic function, which is explained using a gradient walk over the graph. Starting at the unlabeled node that is to be labeled, the algorithm walks to the neighbor k of the current node j that has the highest score for the predicted label y of j. This continues until a labeled example is encountered. An additional heuristic is also proposed to demonstrate that the scheme can be expanded.

A big disadvantage of this approach is that it is difficult to apply in different scenarios. For example, an authentication scenario would not have the available data required for the short and mid term edges, thus almost only relying on the biometric edges. While this scenario is obviously more controlled, it means that the system no longer benifits much from using the graph representation. In addition, a resourceful attacker may be able to exploit the time and color edges in his attempts to authenticate as another user. However, for surveillance scenarios, where users are commonly not facing the camera, the graph representation still allows the system to perform its task.

A more generic approach is [6], which creates a graph for each class that connects the available labeled data to two special nodes, representing genuine and impostor respectively. The unlabeled data is connected using a similarity matrix, which is also used in the previous algorithm, and an edge defining function called *mincut-3NN*. The total weight of the nodes whose removal disconnects the genuines and imposters node is then minimized. The graph is then cut and all the nodes still connected to the genuines node are added to the template for the class in question.

This approach uses a known technique from the field of computer vision, which in turn uses the proved max-flow min-cut theorem. However, building up a new graph for each user is a fairly expensive way to perform the update process. Though the nodes may be held in memory for the entire process, the algorithm still requires different connections for each user, and thus a completely new graph each time. In addition, the algorithm states that all samples linked to the genuine node be labeled as genuine; however, it is possible that there exists another class to which one of these samples is more similar. Thus, either a high threshold is needed, or a risk of misclassification is taken.

5. FUTURE WORK

Now that the research community has developed adaptive methods and implemented them in various applications, there are some additional issues that arise. In this section, some of these issues and the initial work to their solutions are described.

5.1 Quantitative Comparisons

In evaluating traditional biometrics, the key issue has been how to properly estimate the error of a system. Many protocols and databases have been developed and successfully applied, such as FERET and its follow-up, FRVT. However, to compare adaptive systems, it is necessary to take into account performance over time. This allows us to observe the effect of difficult samples on the performance of the system and the curve of its performance increase with respect to the amount of images, and thus how quickly the system learns.

In [4], a new protocol for testing based on specific classes (ie. users) rather than the whole set of people, considering a verification scenario (though an identification scenario could also be represented). By placing some additional constraints on the produced genuine scores; that they follow some chosen distribution and that they are continuous in time. These constraints allow the measurement of the performance at any given point in time between measurements. Experiments performed without adaptive methods have shown that already, the performance of some systems varies over time. Unfortunately, this work does not show an analysis of any adaptive systems.

5.2 Scalability

In addition to finding the measures to compare systems, it also becomes necessary to include performance constraints, especially when considering on-line applications. This aspect of biometrics has received relatively limited attention in the FRVT evaluations, which simply constrains the amount of computational power and time available. However, note that adaptive methods such as co-training and self-training increase the amount of templates stored in the system as they run longer, at least in their simplest implementation. While the amount of memory required for these methods may be limited for a simple camera, note that each additional template increases the required amount of computations at each identification and update iteration.

As pointed out in various sections of the paper, scalablility remains an important and challenging issue in adaptive biometrics. Ironically, semi-supervised learning methods are expected to perform well in situations where large amounts of unlabeled data is available, but such scenarios have not often been tested [14]. In addition, the scalability issue is a key obstacle for developing on-line methods of semi-supervised learning. An on-line method is a key element for a biometric system used for surveillance. The other key obstacle for such an on-line method is the risk of an attacker exploiting the adaptive behavior of the system [4].

5.3 Open-Set Identification

This paper has focused mainly on closed-set identification; identification of previously enrolled individuals. Open-set identification, however, focuses on being able to identify new users. Intuitively, adaptive biometric systems should be quite good at open-set identification, especially the off-line methods. Consider a surveillance scenario, such as the one proposed in [1], where a camera is placed at a random location and monitors passing users. In the work, a graph-based method is used that could be considered on-line, specifically designed for this type of scenario. In a traditional system, the only way to create a new class is by selecting some second threshold (the first being the one used for recognition), below which a new class is added to the system with the image in question as template. This method runs the risk of choosing a bad template for the new user.

However, in an off-line adaptive system, a series of images can remain after processing a batch of new images. These images have an insufficiently high recognition score. The system can then select images from this remaining set and create a new template for the image with the lowest score. While this may be a bad template, subsequent updates with the other images in the remaining set can improve this template, unlike the situation for a traditional system.

6. CONCLUSTION

In this paper, two core elements of adaptive biometrics have been subjected to a literature study. First, commonly used adaptive methods and their variants were discussed, followed by a discussion of template update methods, using face recognition as the main implementation scenario. Face recognition was established as a good use case for adaptive biometric systems, in particular in the case of surveillance scenarios. Subsequently, the adaptive methods were compared using various face recognition experiments performed in related work. These experiments were critically reviewed, pointing out areas for improvement. It was shown that there is considerable difference in computational requirements between the methods, while graph-based methods were identified as a promising compromise that does not rely on repeated batch-processing of perceived images, unlike the self- and co-training systems. Finally, upcoming topics in the adaptive biometrics field were described, noting the lack of a good evaluation method and scalability experiments.

7. REFERENCES

[1] M.-F. Balcan, A. Blum, P. P. Choi, J. Lafferty, B. Pantano, M. R. Rwebangira, and X. Zhu. Person identification in webcam images: An application of semi-supervised learning. In *Proceedings of the* 22^{nd} *ICML Workshop on Learning with Partially Classified Training Data*, 2005.

[2] L. Didaci, G. Marcialis, and F. Roli. A theoretical and experimental analysis of template co-update in biometric verification systems. In *Joint IAPR International Workshop on Structural and Syntactical*

Pattern Recognition and Statistical Techniques in Pattern Recognition S+SSPR08, pages 775–784. Springer, December 2008.

[3] H. Murakami and B. V. K. V. Kumar. Efficient calculation of primary images from a set of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-4(5):511–515, September 1982.

[4] N. Poh, R. Wong, J. Kittler, and F. Roli. Challenges and research directions for adaptive biometric recognition systems. In *International Conference on Biometrics 2009*, pages 753–764, 2009.

[5] A. Rattani, G. Marcialis, and F. Roli. An experimental analysis of the relationship between biometric template update and the doddingtonâĂŹs zoo in face verification. In *Proceedings of IEEE/IAPR International Conference on Image Analysis and Processing ICIAP 2009*, pages 429–437, 2009.

[6] A. Rattani, G. L. Marcialis, and F. Roli. Biometric template update using the graph mincut algorithm: A case study in face verification. In *Biometrics Symposium BSYM'08*, pages 23–28, 2008.

[7] F. Roli, L. Didaci, and G. Marcialis. Adaptive biometric systems that can improve with use. In *Advances in Biometrics: Sensors, Systems and Algorithms*, pages 447–471. Springer, 2008.

[8] F. Roli and G. Marcialis. Semi-supervised pca-based face recognition using self-training. In *Joint IAPR International Workshop on Structural and Syntactical Pattern Recognition and Statistical Techniques in Pattern Recognition S+SSPR06*, pages 560–568, August 2006.

[9] M. Seeger. Learning with unlabeled data. Technical report, University of Edinburgh, 2002.

[10] H. Sellahewa and S. Jassim. Image-quality-based adaptive face recognition. *IEEE Transactions on Instrumentation and Measurement*, 59:805–813, April 2010.

[11] L. Shen and L. Bai. Adaboost gabor feature selection for classification. In *Proceedings of image and vision computing*, page 77âĂŞ83, 2004.

[12] U. Uludag, A. Ross, and A. Jain. Biometric template selection and update: A case study in fingerprints. *Pattern Recognition*, 37:1533–1542, 2004.

[13] W. ZHAO, R. CHELLAPPA, P. J. PHILLIPS, and A. ROSENFELD. Face recognition: A literature survey. *ACM Computing Surveys*, 35(4):399–458, December 2003.

[14] X. Zhu. Semi-supervised learning survey. Technical report, University of Wisconsin - Madison, 2008.