JOURNAL OF FORENSIC SCIENCES

J Forensic Sci, May 2012, Vol. 57, No. 3
doi: 10.1111/j.1556-4029.2011.02018.x
Available online at: onlinelibrary.wiley.com

TECHNICAL NOTE

DIGITAL & MULTIMEDIA SCIENCES

Stefano Arca,[1] *Ph.D.; Paola Campadelli,*[1] *Ph.D.; Raffaella Lanzarotti,*[1] *Ph.D.; Giuseppe Lipori,*[1] *Ph.D.; Federico Cervelli,*[2] *Ph.D.; and Aldo Mattei,*[3] *Ch.E.*

Improving Automatic Face Recognition with User Interaction*

ABSTRACT: Face recognition systems aim to recognize the identity of a person depicted in a photograph by comparing it against a gallery of prerecorded images. Current systems perform quite well in controlled scenarios, but they allow for none or little interaction in case of mistakes due to the low quality of images or to algorithmic limitations. Following the needs and suggestions of investigators, we present a guided user interface that allows to adjust from a fully automatic to a fully assisted modality of execution, according to the difficulty of the task and to amount of available information (gender, age, etc.): the user can generally rely on automatic execution and intervene only on a limited number of examples when a failure is automatically detected or when the quality of intermediate results is deemed unsatisfactory. The interface runs on top of a preexistent automatic face recognition algorithm in such a way to guarantee full control over the execution flow and to exploit the peculiarities of the underlying image processing techniques. The viability of the proposed solution is tested on a classic face identification task run on a standard publicly available database (the XM2VTS), assessing the improvement to user interaction over the automatic system performance.

KEYWORDS: forensic science, user interaction, automatic system, face recognition, end user, biometrics

The problem of face recognition can be formulated in different ways, according to the applicative context. In forensic sciences, the expression is used sometimes to refer to the practice of face comparison or face verification: in the former case, the task is to establish whether two pictures depict the same subject; in the latter formulation, one is shown a picture and should answer the question "is this an image of person X?," given some notion of how X looks like.

These are variants of the same problem that can arise during investigations when the police have a picture of a person taken in the act of committing illegal activities (e.g., from the camera embedded in an ATM station) and suspects a certain subject who is already known to the police (e.g., he or she is recorded in the database of mug shots). Several studies prove that this is an extremely difficult task to accomplish with a high degree of confidence even for human observers (1), and surely no available computer vision technique is accurate enough to provide evidence which would be accepted as proof in a legal trial. Nonetheless, automatic processing can be of great support to investigators as explained later.

In this work, we refer to the following scenario of face recognition:

> given the gallery G, a set of photographs of subjects already enrolled (with known ID),
> given a new image depicting a test subject with unknown ID, taken from a set T of probes,
> the goal is to reorder the pictures in the gallery so that the known subjects are found in descending order of similarity with respect to the probe subject,

where the similarity is measured by a function that takes as input the description of two faces, and gives as output a value in the interval [0, 1] that represents their degree of correspondence, ranging from orthogonality to identity of the descriptions. The description (or characterization) of a face is formulated in terms of the chosen facial features (anthropometric distances, shape, textures, particular signs, etc.), and their automatic and accurate extraction is arguably the most difficult step in the design of a face recognition system.

The aforedescribed scenario is referred to as *face identification* or *face matching*. Notice that *face identification* is ideally more general than *face verification* because, in principle, one could estimate an optimal threshold for the similarity measure such that two subjects are judged as being the same person if and only if their similarity exceeds that value. In practice, this approach gives poor results as it is very difficult to estimate such a threshold, as face similarity depends very much on the employed facial features, and the detail of the resulting description is in general not sufficient to state that two subjects are the *same* person with high confidence; one can only conclude two subjects are *similar* to a certain degree in the sense of the extracted facial features.

If all the subjects in T have been already enrolled in G (*closed world* assumption), it is possible to quantitatively evaluate the performance of a face matching system (during testing, the IDs of the probe subjects are obviously kept unknown to the face

---

[1] Dipartimento di Scienze dell'Informazione, Università degli Studi di Milano, Via Comelico, 39/41, 20135 Milano, Italy.

[2] Dipartimento di Ingegneria Industriale e dell'Informazione, Università degli Studi di Trieste, Via Valerio, 10, 34100 Trieste, Italy.

[3] Raggruppamento Carabinieri Investigazioni Scientifiche, Viale di Tor di Quinto, 151, 00191 Rome, Italy.

*Partial support by the PASCAL2 Network of Excellence under EC Grant No. 216886. This publication only reflects the authors' views.

Received 22 Sept. 2010; and in revised 20 Jan. 2011; accepted 12 Mar. 2011.

© 2011 American Academy of Forensic Sciences

identification system, while they are known to the experimenter for evaluation purposes).

Let us assume for simplicity that each subject is represented exactly once in the gallery; let us define the rank of a probe image as the position in the ordered gallery at which is found the correct match; then, we can compute the *Cumulative Match Characteristic* (CMC) (2) as:

$$P_I(r) = \frac{|C(r)|}{|G|}$$

with $C(r) = \{t \in T \mid \text{rank}(t) \leq r\}$ where the subscript I of P in the original notation stands for "identification," $C(r)$ is the set of probes recognized at rank r or better, and $|\ |$ is the notation for set cardinality (size); the CMC measure is hence a nondecreasing function of r and of course $P_I(|G|) = 1$. Although the system performance is usually summarized with the value $P_I(1)$, that is, the percentage of probe images correctly identified at rank 1, the trend (the steepness) of $P_I(r)$ is also relevant to describe the global behavior of the system; in particular, an interesting value is the smallest rank r^* for which $P_I(r^*) = 1$. Notice that the CMC value is dependent on the gallery size in two ways: not only $|G|$ appears under the fraction in the definition of CMC, it also defines the domain of CMC, that is, $r \in \{1,\ldots,|G|\}$. Consequently, to compare the results obtained on two different experiments of face identification, one should consider the rank relatively to the gallery size: for instance, given G_1 and G_2 such that $|G_1| \neq |G_2|$ one cannot directly compare the CMC for some r, but should rather consider two values r_1 and r_2 such that $r_1/|G_1| = r_2/|G_2|$. This will be better clarified by an example in the next paragraph.

It is important to stress the fact that the applicability of a face recognition system is not limited to the ideal case; it can be useful also in realistic situations, like in normal investigation activities, when no assumption can be made about the identity of the person committing an offense, nor about his previous enrollment. In fact, if the similarity measure used to order the gallery is accurate enough, investigators can run the system over the suspect image against the entire gallery at their disposal, and search for a one-to-one correspondence by visual inspection only for the top of the ordered gallery. The maximum number of images they will have to go through, to be reasonably confident that the subject is either present or absent from the gallery, depends both on the cardinality of the gallery and on the accuracy of the system as measured on previous experiments in terms of CMC. For instance, if face recognition reaches accuracy of 100% at rank 10 ($r^* = 10$) with respect to a gallery of 1000 subjects ($r^*/|G| = 1\%$), when applying the system in an *open world* situation (as opposed to the *closed world* assumption discussed earlier) to a gallery of 20,000 pictures, one is expected to inspect at least 200 images before discarding the photo as depicting someone probably not present in the gallery.

In this paper, we present an extension to our previous work on face recognition by adding on top of the system an interactive graphical user interface (GUI) that allows the user to get a better control over a face matching test. The interaction aims at improving the system performance in three respects: (i) the user can filter out some pictures from the gallery to narrow the search domain by specifying some attributes of the probe subject (gender, age, eye color, hair color, race, presence or absence of spectacles, facial hair, etc.); (ii) the user can validate or correct the output of feature extraction, resulting in a more reliable similarity measure; and (iii) the user can inspect as many subjects of the ordered gallery as he or she wishes.

The next section introduces the gallery and probe images used in this work for testing the proposed system. Then, we present an overview of the original automatic method and its extension with the interactive capabilities. Section "Results" details the performance improvement owing to user interaction with respect to fully automatic face recognition. Finally, we discuss the merits and weaknesses of the approach, and draw some conclusions on its validity.

Database

To test the system performance, we refer to the XM2VTS database http://www.ee.surrey.ac.uk/Research/VSSP/xm2vtsdb/ (accessed January 17, 2011), which is a publicly available data set that is considered a baseline for performance evaluation and has been employed in several publications on the topic (see [3] for a face verification competition using the XM2VTS database). In particular, it consists of 1180 frontal color images of 295 subjects taken with neutral expression; the pictures have been acquired in four sessions that span a period of over 4 months to favor some changes in appearance, beard-cut, clothing, and makeup. Although the illumination conditions are controlled, the data set still presents certain variability with respect to race, the presence or absence of spectacles, facial hair, etc. (see Fig. 1 for a sample of XM2VTS images). In our experiments, the gallery consists of the images taken from the first session, while the probe set is composed of the remaining 885 images taken from the last three sessions (three probes per subject). No image enhancement was applied to those images.

Proposed Method for Interactive Face Recognition

In the last two decades, the problem of face recognition has been widely investigated (4) bringing to the commercialization of several products, most of which have participated to the *Face Recognition Vendor Tests* (FRVT) (5,6). These contests have proven that:

From 2002 to 2006, the error rate is dropped by an order of magnitude.

Existing systems work very well on controlled, high-resolution images, that is, frontal face, no occlusion, uniform background, and homogeneous illumination.

In these conditions, algorithms performed as well as or better than humans.

FIG. 1—*A sample from the XM2VTS database (originally in color).*

> The recognition performance decreases significantly on uncontrolled or low-resolution images.

Although the FRVT outcomes are very interesting, it is understood that the tests were conducted in a relatively *ideal* scenario, and it is not known how the results would extend to real-world situations. In fact, one would expect user interaction to be most useful especially in the case of images acquired by low-quality and low-cost C/C TV, and in situations where the subject is noncollaborative, and hence, his pose and expression are not controlled. To our knowledge, most commercial systems work almost like *black boxes*, lacking in control over the different phases of the recognition process. What is more, the proprietary nature of the user license they release makes it impossible to intervene on the code to adjust it on specific needs. On the contrary, it would be desirable to present the user with the following choice: no interaction when the image quality is high, requests for correction or validation when the image is of low quality.

Another useful feature regards the results of the recognition process: often a limited number of matching subjects is returned (if not only the best match), and their scores of reliability are unspecified or given with no further specification of their interpretation. The investigators would appreciate to have a list of matches (choosing the list length), each one associated with a similarity score which is made meaningful to them. In this way, the human operator is in full control of the system, but at the same time, he or she is assisted in the recognition process being guided and presented with a limited number of options, as it happens in other similar forensic databases, for example, for Automated Fingerprint Identification Systems (AFIS).

Following these considerations, we extended a previously developed automatic system for face recognition by adding an interface through which the user can (optionally) drive the process step-by-step.

Face Recognition Algorithm

In this section, we give a very brief overview of the employed face recognition method which has been developed by some of the authors. As the technique has been published in different journals and conference proceedings (7–10), we do not intend here to go into the finest details, we will just recall its functional requirements and constraints: the technique takes in input images of faces with nearly frontal pose (with a tolerance to rotation of about 20° around all three axes), in which the face resolution is such that the interocular distance is equal or >75 pixels, while no particular assumption is made on the illumination of the scene (might be controlled or uncontrolled, natural or artificial), on the facial expression (neutral pose is preferred, however to some extent expressions can be modeled too [8]), on the presence of partially occluding objects (spectacles, beard, and hair), and on the level of background clutter. On the contrary, we are interested in giving a short description of the facial feature extraction, on which it is based the design of the GUI.

There is a general agreement in the face processing literature that feature extraction is a crucial step in the design and operation of a technique (4). In our system, a face is described by a constellation of 27 points which are called *fiducial points* and shown on the left of Fig. 2. Given a fiducial point, its characterization is obtained by convolving the neighborhood of the point with a bank of 40 Gabor filters (5 scales × 8 orientations, see the right part of Fig. 2) as described in (7). This operation produces for each fiducial point a vector of 40 real coefficients called *jet*, which represents the texture around that point at different scales and orientations, thus capturing both the finest details (moles, rims of spectacles, etc.) and the general appearance (lightness and contrast of facial features, eyebrow thickness, etc.) of that limited image region. Hence, the entire face is characterized by 27 jets, and the similarity between two faces is defined in terms of the similarity of corresponding jets as measured by the internal product of the underlying vectors (see [7] for further details).

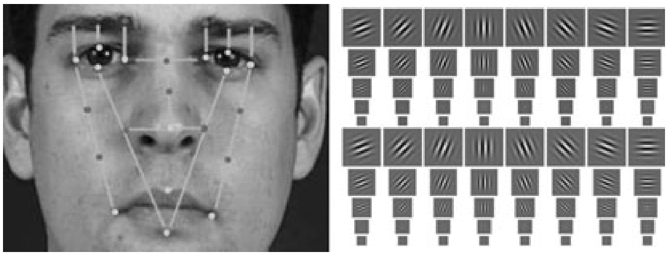


FIG. 2—*Left: A face is described by 27 fiducial points: 13 are directly extracted from the image (light gray dots), 14 are inferred from the former ones (dark gray dots). Right: the bank of 40 Gabor filters; the bank is applied to the neighborhood of each fiducial point to produce a jet vector.*

We remark that the information involved in the face description is *local* and consists of the *texture* around the facial fiducial points, as represented by the Gabor filter responses. One could argue that, given such positions, it would be interesting to characterize faces using anthropometry (point-to-point distances in the image plane). From our experience and according to dedicated studies (11), this characterization has limited value. In fact, distance measurements have little consistency because of various reasons: face expressions cause high deformations; variability of camera-to-subject positioning affects distances in the image projection; the error made by automatic localization of landmarks is in the same range of variability of interpersonal differences.

We also point out that the local nature of the description presents important advantages. In fact, the extraction of fiducial points lying around different facial components (eyes, nose, mouth, eyebrows, etc.) is practically independent, so local errors do not propagate, and their impact on the global quality of the characterization is limited. This property of locality is further exploited in our system by modifying the definition of the similarity function: the internal product is not computed over all jets, but only on the 14 fiducial points that manifest the highest similarity. This technique gives better results than considering the average of all similarities (12), as it allows to discard those points that introduce ambiguity in the subject identification; there could be several causes of this ambiguity such as the wrong positioning of a fiducial point, its partial occlusion, a significant feature deformation owing to facial expression, or even a distortion owing to acquisition.

We stress the fact that the fiducial point extraction is fully automatic in the noninteractive version of our system, and it involves the application of many modules. At first, the input image is scanned with the aim of locating the face bounding box using the Viola and Jones algorithm (13); this is the only subtask accomplished using a technique that has not been expressively devised by the authors. Subsequently, the face area is scanned looking for the eyes and mouth centers, with the purpose of normalizing the face with respect to scale and in-plane rotations, aligning all faces to a common reference; we observe that the accomplishment of this task is crucial as its failure makes it impossible to carry on any further processing. Then comes the extraction of the actual fiducial points. The process proceeds in a top-down fashion: first the eye, nose, and mouth subimages are derived on the basis of the localized centers, following some simple geometrical considerations. Then, the eyes and mouth contours are analyzed to extract the corresponding fiducial points (the light gray points in Fig. 2). Regarding the nose, symmetry and geometrical plausibility are used for localizing its tip. Finally, to enrich the face description, further 14 fiducial points (the dark gray points in Fig. 2) can be inferred on the basis of the position of the extracted points. Refer to (7–10) for a detailed description of the entire system.

Interactive Functionalities

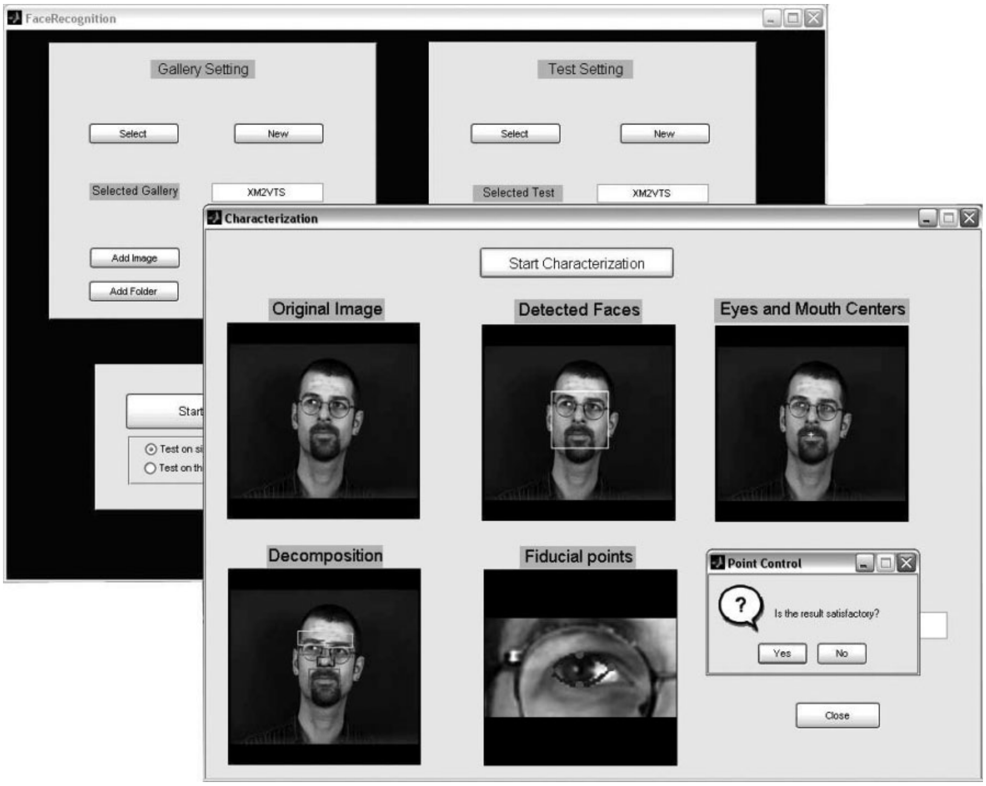
The system interface has been designed together with the investigators, so its features are driven by their needs. It offers several possibilities of interactions, regarding:

- *Gallery size*: the archive dimension can be significantly reduced on the basis of the information immediately available to the investigators, gender, age, race, hair color, eye color, etc.
- *User validation/correction*: the user can ask to stop the process at all intermediate steps, to correct the partial results if he or she wishes so.
- *Presentation of results*: the user can choose the number of subjects to receive as output, each one associated with a score of reliability.

However, if desired, the system can still be set to run in a fully automatic way as explained in the previous section.

The first point involves the definition of the gallery in which to search for the unknown subject and moves from the evident fact that the difficulty of the face identification task is proportional to the gallery size. Nowadays, the police have a gallery of subjects with size in the order of hundreds of thousands, and this number increases continuously. It is thus desirable to exploit all the information available to investigators to reduce as much as possible the gallery cardinality. The interface allows the investigators to label the subjects with several tags that can be exploited as search criteria: gender, age, particular signs on the face, eye color, hair color, the presence or absence of spectacles, and facial hair. The insertion of all or part of this information allows to produce a more specific query to the database, significantly reducing the search domain, and thus increasing the system success rate. Obviously the user should take care in not over-trusting some features that can be easily counterfeited, such as eye color and hair color.

The second point regards the core of the interaction. As described in the previous subsection, the face identification method is composed of several modules working in cascade; therefore, a failure at a certain point along the execution chain can provoke the failure of the whole recognition task. To treat also the cases in which the automatic execution experiences a failure, the interface provides two modes of execution, *recovery* and *assisted*, which correspond to two levels of intervention. In case of *recovery* execution, the system runs quietly by itself and it requires human intervention only when one of the modules experiences a hard failure (the face detector does not locate any face, the eye detector localizes only one or no eyes, the mouth is not detected, etc.) or else when the fiducial point extraction misses a few points (the eyebrows points, the nose tip, etc.); this latter event does not interrupt the execution, as the number of extracted points is still sufficient to compute the similarity; however, their reduction could jeopardize the correct recognition so the human intervention is suggested. The *assisted* case is meant to treat qualitative errors too: the system pauses at all partial outputs regarding face detection, face normalization, fiducial points extraction, and the interface asks the user to perceptively evaluate them to decide whether to manually correct those that he or she considers incorrect or just imprecise (see Fig. 3). In this mode, the system can work under ideal conditions.

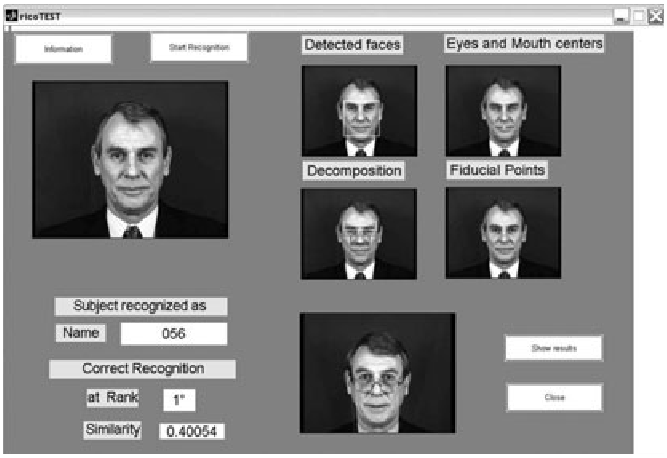


Selected Gallery: XM2VTS

Selected Test: XM2VTS

Is the result satisfactory?

FIG. 3—*A screen shot of the interactive user interface when the system is used in the assisted modality on a typical XM2VTS image.*



Subject recognized as
Name 056
Correct Recognition
at Rank 1°
Similarity 0.40054

FIG. 4—*A screen shot of a face identification experiment: on the left is shown the probe image, on the right some partial results of execution, while at the bottom is shown the most similar image in the gallery. The interface also gives the information about the ID of the subject from the gallery (056 in the example), the similarity score, and, when available (in a closed world scenario), the information about the rank of the experiment (1 in the example, as it is a case of correct identification).*

The last option of interaction concerns the visualization of results (Fig. 4). Recall from the introduction that one should not expect always correct identification, that is, $P_I(1) = 1$, even if the matching is based on fiducial points as precise as the ones annotated by the user under the assisted mode of operation. A wrong match can happen for several reasons: because of the limits of the identification similarity measure; because of differences in appearance of the subject in the probe image with respect to the corresponding gallery image regarding pose, aging, facial expression, presence of shadows or partial occlusions, etc.; even because of the presence of look-a-like person in the gallery. For these reasons, the solution worked out with the investigators consists in presenting the user with a list of matches ordered according to their similarity (according to the same logic behind the AFIS system) so that the user can investigate the list of matching subjects, as shown in Fig. 5. The length of the list can be varied by the user and should be long enough to make it probable that the subject of interest is included, as explained in the introduction. This is a very helpful instrument as it dramatically facilitates the search operations.

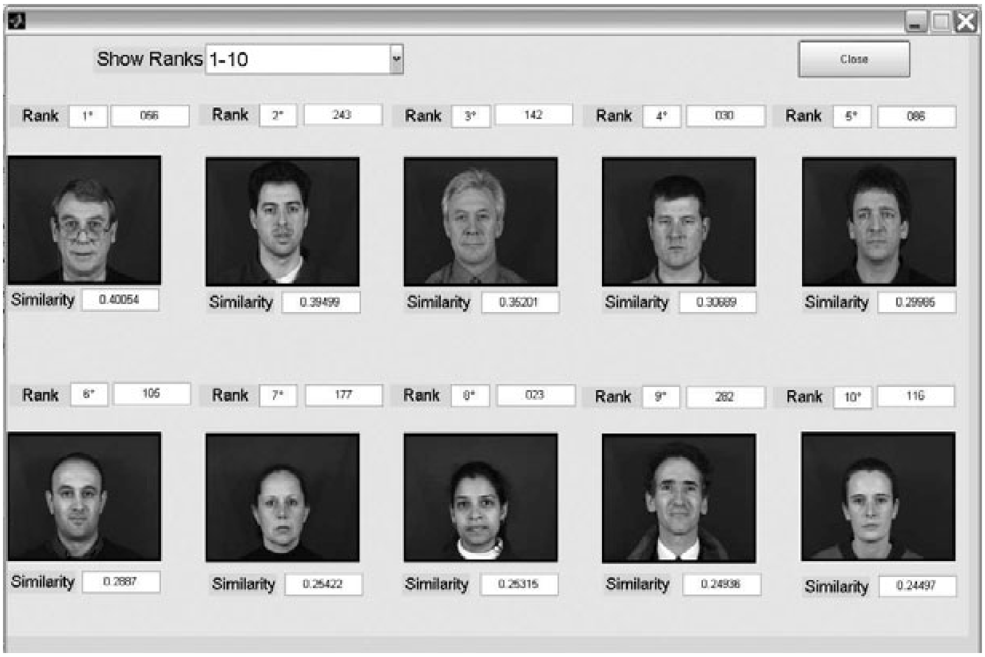
Results

To evaluate the system performance, as well as the benefit of human intervention, we carried out three recognition experiments setting the execution mode (for both the gallery and the probe sets) respectively to *automatic*, *recovery,* or *assisted*.

The *automatic* experiment is meant to draw a baseline to the system performance; by definition, this mode does not involve any human intervention, so the subjects whose enrollment in the gallery caused a hard failure are discarded from the experiment (this happened only to one of the 295 subjects); on the other hand, the two probe images on which the system experienced a hard failure are counted as recognition errors.

We observe that in the *recovery* experiment, the system suggests to intervene on three images of the gallery (one because of missed face detection and two because of missing fiducial points) and 11 of the test set (twice for missing faces, nine times for missing fiducial points); after intervention, the improvement of the performance over the fully automatic experiment is 0.7% at rank 1.

Regarding the *assisted* experiment, we asked five persons to separately check the quality of all intermediate outputs on gallery and probe images and to correct the results they judged unsatisfactory. The reason why we repeated the experiment five times is the subjectiveness of both the quality evaluation and the possible correction. With this modality, the recognition rate at rank 1 increases by another 1.3% over the recovery experiment, for a total of 2% over the fully automatic case (the standard deviation of the results obtained asking assistance to five different human operators



Show Ranks 1-10

Rank 1° 056
Similarity 0.40054

Rank 2° 243
Similarity 0.39499

Rank 3° 142
Similarity 0.35201

Rank 4° 030
Similarity 0.30689

Rank 5° 086
Similarity 0.29986

Rank 6° 105
Similarity 0.2887

Rank 7° 177
Similarity 0.25422

Rank 8° 023
Similarity 0.25315

Rank 9° 282
Similarity 0.24938

Rank 10° 116
Similarity 0.24497

FIG. 5—*This is what would pop up if the Show results button in Fig. 4 were to be pressed: the figure shows the first 10 subjects in the gallery that "resemble" the most the probe image. For every subject, it reports the gallery ID and the similarity score.*

TABLE 1—*Automatic: results obtained running the face recognition system fully automatically; recovery: results obtained correcting only the fault errors automatically raised by the system; assisted: results obtained averaging the system performance when assisted by five different users.*

Experiment	$P_I(1)$	$P_I(2)$	$P_I(3)$	$P_I(4)$	$P_I(5)$	$P_I(6)$	$P_I(7)$
Automatic, %	93.5	95.2	95.9	96.5	96.8	97.0	97.1
Recovery, %	94.2	95.8	96.3	96.8	97.1	97.2	97.3
Assisted, %	95.5	97.1	97.5	98.2	98.4	98.6	98.8

Experiment	$P_I(8)$	$P_I(9)$	$P_I(10)$	$P_I(20)$	$P_I(30)$	$P_I(50)$	$P_I(100)$
Automatic, %	97.5	97.5	97.6	98.2	98.7	99.2	99.6
Recovery, %	97.7	97.7	97.8	98.3	98.8	99.2	99.6
Assisted, %	98.8	98.8	98.9	99.4	99.4	99.7	99.7

is about 0.1%). The five operators corrected on average 43.6 images, which is about three times as much as in the recovery experiment; this means that the quality of automatic fiducial point extraction was judged as satisfactory on about 96% of images.

In Table 1, we report the obtained recognition rate at ranks 1–10, and at ranks 20, 30, 50, and 100, which correspond approximately to one-fifteenth, one-tenth, one-sixth, and one-third of the subjects' cardinality. Notice that at rank 30, that is, approximately 10% of the gallery cardinality, the recognition rate is about 99% whichever the experiment modality.

Discussion and Conclusions

Although the results reported in the previous section are obtained solely on the XM2VTS data set, we can already draw some interesting conclusions. First of all, the system performance reaches a certain degree of reliability even when executed in fully automatic mode: 93.5% of images are correctly identified at rank 1. What is more, the *recovery* and *assisted* interactive modalities of operation significantly improve the identification, by 0.7% and 2% respectively. As expected, the *assisted* modality presents the best performance increase, however, at the cost of a great deal of interaction: the user is asked for validation (and possibly correction) of every partial processing, which requires several clicks on the mouse. The *recovery* modality is probably the best compromise as it does not require the user to continuously control the system behavior, while it requires his/her intervention only on few critical cases: in particular, this happened on 1.3% of images from the XM2VTS database.

There is surely some room for improving these results; however, we emphasize here again that the system need not be perfect to be useful; in fact, it is not meant to provide legal evidence in court trials; on the contrary, it is devised as a tool supporting investigators in going through a long list of suspects as rapidly as possible. At present, investigators are considering the system for experimentation on the data sets at their disposal.

In future work, we plan to put the system to test bigger data sets, to measure how these promising results scale up with size, as well as on less controlled images which are the kind investigators are likely to come across during their activity (aside from the "mug shots" like pictures which could constitute the reference gallery). In real conditions, as long as the aforementioned functional constraints of our technique are respected, we do not expect to observe major changes in the behavior of the system. In general, it is possible to predict that the more unalike are the gallery and probe images with respect to all conditions of acquisition, the lower would be the rate of automatic recognition and the more helpful would turn to be the interactive capabilities of the system. In future works, we plan to investigate low- or mixed-quality scenarios to quantify these effects.

Moreover, we stress the fact that we did not make use of the filtering capabilities of the interface in the experiments on XM2VTS because we intended to estimate a baseline for the system performance; using those criteria, the user can narrow very much the search domain (e.g., by about 50% only specifying gender), and the system performance will surely benefit from their utilization. The interface makes it simple to use the standard filters (gender, age, eye color, hair color, race, presence or absence of spectacles, and facial hair), and it also allows to define custom ones; the only requirement to the user for enabling filtering criteria is to manually label the gallery pictures (during enrollment) with the same tags as the filters he or she intends to use on probes.

Acknowledgments

The research on which this (report) is based acknowledges the use of the Extended Multimodal Face Database and associated documentation. Further details of this software can be found in: Messer K, Matas J, Kittler J, Luettin J, Maitre G. XM2VTSbd: the extended M2VTS database. In: Akunuri S, Kullman C, editors. Proceedings of the 2nd Conference on Audio and Video-base Biometric Personal Verification (AVBPA99); 1999 Aug 20–24; Washington, DC. New York: Springer Verlag, 1999;72-7, http://www.ee.surrey.ac.uk/Research/VSSP/xm2vtsdb.

References

1. O'Toole A, Phillips P, Jiang F, Ayyad J, Penard N, Abdi H. Face recognition algorithms surpass humans matching faces over changes in illumination. IEEE Trans Pattern Analysis Mach Intell 2007;29(9):1642–6.
2. Grother P, Micheals R, Phillips P. Face recognition vendor test 2002 performance metrics. In: Kittler J, Nixon MS, editors. Proceedings of the International Conference on audio-visual based biometric person authentication; 2003 June 9–11; Guildford, U.K. Berlin, Germany: Springer-Verlag, 2003;937–45.
3. Messer K, Kittler J, Sadeghi M, Marcel S, Marcel C, Bengio S, et al. Face verification competition on the XM2VTS database. In: Kittler J, Nixon MS, editors. Proceedings of the International Conference on audio-visual based biometric person authentication; 2003 June 9–11; Guildford, UK. Berlin, Germany: Springer-Verlag, 2003;964–74.
4. Zhao W, Chellappa R, Phillips P, Rosenfeld A. Face recognition: a literature survey. ACM Comput Surv 2003;35(4):399–458.
5. Phillips P, Grother P, Micheals R, Blackburn D, Tabassi E, Bone J. Face recognition vendor test 2002: overview and summary. 2003, http://www.frvt.org (accessed January 17, 2011).
6. Phillips P, Scruggs W, O'Toole A, Flynn P, Bowyer K, Schott C, et al. FRVT 2006 and ICE 2006 large-scale results. Gaithersburg, MD: National Institute of Standards and Technology, Technical Report NISTIR 7408, 2007.
7. Arca S, Campadelli P, Lanzarotti R. A face recognition system based on automatically determined facial fiducial points. Pattern Recogn 2006;39(3):432–43.

8. Arca S, Campadelli P, Lanzarotti R, Lipori G. A face recognition system dealing with expression variant faces. In: Tang YY, Wang SP, Lorette G, Yeung DS, Yan H, editors. Proceedings of the International Conference on Pattern Recognition; 2006 Aug 20-24; Hong Kong. Los Alamitos, CA: IEEE Computer Society, 2006;1242–5.
9. Campadelli P, Lanzarotti R, Lipori G. Automatic facial feature extraction for face recognition. In: Kresimir Delac K, Grgic M, editors. Face recognition. Vienna, Austria: I-Tech Education and Publishing, 2007;31–58.
10. Campadelli P, Lanzarotti R, Lipori G. Precise eye and mouth localization. Intern J Pattern Recognit Artif Intell 2009;23(3):359–77.
11. Kleinberg K, Vanezis P, Burton A. Failure of anthropometry as a facial identification technique using high-quality photographs. J Forensic Sci 2007;52(4):779–83.
12. Wiskott L, Fellous J, Kruger N, von der Malsburg C. Face recognition by elastic bunch graph matching. In: Jain LC, Halici U, Hayashi I, Lee SB, Tsutsui S, editors. Intelligent biometric techniques in fingerprints and face recognition. Boca Raton, FL: CRC Press, 1999;355–96.
13. Viola P, Jones M. Rapid object detection using a boosted cascade of simple features. Int Conf Comput Vis Pattern Recognit 2001;1:511–8.

Additional information and reprint requests:
Federico Cervelli, Ph.D.
Università degli Studi di Trieste
Via Valerio, 10
34100 Trieste
Italy
E-mail: federico.cervelli@gmail.com