AIMS AND SCOPE: Applied Artificial Intelligence addresses concerns in applied research and applications of artificial intelligence (AI). The journal acts as a medium for exchanging ideas and thoughts about impacts of AI research. Papers should highlight advances in uses of expert systems for solving tasks in management, industry, engineering, administration, and education; evaluations of existing AI systems and tools, emphasizing comparative studies and user experiences; and/or economic, social, and cultural impacts of AI. Information on key applications, highlighting methods, time-schedules, labor, and other relevant material is welcome.

Abstracted and/or Indexed in: Engineering Information, Inc. and by INSPEC.

© Printed on acid-free paper, effective with Volume 7, Number 1, 1991.
Special Issue
Artificial Intelligence: Future, Impacts, Challenges
Part 3

CONTENTS

iii  SPECIAL ISSUE ON ARTIFICIAL INTELLIGENCE: FUTURE, IMPACTS, CHALLENGES (PART 3)  □  Robert Trappl

1   SOME SEMIOTIC REFLECTIONS ON THE FUTURE OF ARTIFICIAL INTELLIGENCE  □  Julian Hilton

15  MODELING OF DEEDS IN ARTIFICIAL INTELLIGENCE SYSTEMS
    □  Dimitri A. Pospelov

29   THE HIDDEN TREASURE  □  Ernst Buchberger

39  A DEEPER UNITY: SOME FEYERABENDIAN THEMES IN NEUROCOMPUTATIONAL FORM  □  Paul M. Churchland

59  HOW CONNECTIONISM CAN CHANGE AI AND THE WAY WE THINK ABOUT OURSELVES  □  Georg Dorffner

87  THE FUTURE MERGING OF SCIENCE, ART, AND PSYCHOLOGY  □  Marvin Minsky

109  APPLIED ARTIFICIAL INTELLIGENCE CALENDAR
EYES DETECTION FOR FACE RECOGNITION

LUIGI STRINGA
Istituto per la Ricerca Scientifica e Tecnologica, 38100
Trento, Italy

A correlation-based approach to automatic face recognition requires adequate normalization techniques. If the positioning of the face in the image is accurate, the need for shifting to obtain the best matching between the unknown subject and a template is drastically reduced, with considerable advantages in computing costs. In this paper, a novel technique is presented based on a very efficient eyes localization algorithm. The technique has been implemented as part of the “electronic librarian” of MAIA, the experimental platform of the integrated AI project under development at IRST. Preliminary experimental results on a set of 220 facial images of 55 people disclose excellent recognition rates and processing speed.

INTRODUCTION

There is a growing interest in face-processing problems (Young and Ellis, 1989). The recognition of human faces is in fact a specific instance of 3D object recognition—possibly the most important visual task—and provides a most interesting example of how a 3D structure can be learned from a small set of 2D perspective views. Moreover, among several practical reasons for developing automatic systems capable of recognizing human faces, faces provide a natural and reliable means for identifying a person.

The first examples of computer-aided techniques for face recognition date back to the early 1970s and were based on the computation of a set of geometrical features from the picture of a face (Goldstein et al., 1971, 1972; Harmon, 1973). More recently the topic has undergone a revival (Samal and Iyengar, 1992), and different applications have been developed based on various techniques, such as template matching (Baron, 1981; Yuille, 1991), isodensity maps (Nakamura et al., 1991; Sakaguchi et al., 1989), or feature extraction by neural and Hopfield-type networks (Abdi, 1988; Cottrell and Fleming, 1990; O'Toole and Abdi, 1989). At present it is still rather difficult to assess the state of the art. However, a first significant evaluation is reported in (Brunelli and Poggio, 1991), where a comparison of different techniques is performed on a common database—the best results were obtained with a template matching type technique.

Following a correlation-based approach, excellent results have also been obtained with a procedure recently developed for the “electronic librarian” of MAIA, the experimental platform of the integrated AI project under development at IRST (Poggio and Stringa, 1992; Stringa, 1991a). The procedure is based on the analysis of filtered edges and grey-level distributions to allow a comparison of the directional

The author’s e-mail address is stringa@irst.it
derivatives of the entire image (Stringa, 1991d). On a set of 220 frontal facial images of 55 people, a recognition rate of 100% was obtained, at a processing speed of about 1.25 sec per face on an HP 350 workstation. A second set of experiments, using binary derivatives, disclosed excellent improvements in computing time: with a two-layer S_Net [see Stringa (1990)] the processing speed was reduced to less than 0.05 sec per face (Stringa, 1991e).

Such performance provides evidence for the validity of the approach. Moreover, the procedure proved very efficient with respect to the task of rejecting “unknown” faces, i.e., faces of subjects that are not included in the database. Apart from high recognition rates, low processing costs, and good flexibility under variable conditions, this is another important feature for a real (i.e., industrially applicable) face recognition system.

It must also be stressed, however, that such performance depends on the use of very effective normalization, registering, and rectification techniques. This is in fact a general requirement for any correlation-based approach to face recognition (and more generally to 3D object recognition), particularly when the image to be recognized is freshly captured with a video camera rather than scanned from a standardized photograph. In general it is rather natural to expect the user to look straight into the camera, for even in human interaction people tend to turn their heads so as to look at each other in the eyes. However, a certain flexibility must be tolerated concerning such variable factors as the distance and position of the user’s face from the camera. Hence, some adjustment and normalization is necessary before the system can proceed to the recognition step by comparing the input image with the available set of prototypes.

In our procedure, the normalization of the image to be recognized is obtained by first locating the position of the eyes and then rotating the image so as to align them horizontally. As a result, the need for shifting to obtain the best matching between the unknown face and a template is drastically reduced, with considerable advantages in computing time. In particular, the eyes localization algorithm developed for this purpose (Stringa, 1991c) proves very sensible, allowing very precise positioning of both pupils for each facial image included in the database.

The purpose of this paper is to illustrate this algorithm in detail. To emphasize better its crucial role for correlation-based facial recognition tasks, a brief outline of the system developed at IRST is first given, along with the experimental scenario that led to its formulation. The eyes localization algorithm is then fully described. The final sections report on the current experimental results obtained and offer some general remarks on the algorithm’s performance.

OUTLINE OF THE SYSTEM

General Background: The MAIA Electronic Librarian

As already mentioned, the reported work is part of a more general AI project (labeled MAIA, acronym for “Modello Avanzato di Intelligenza Artificiale”) pres-
ently under development at IRST. Schematically, the goal of the project is to develop an integrated experimental platform whose main "tentacles" include mobile robots capable of navigating in the corridors of IRST, an automatic "concierge" answering visitors' questions about the Institute, and an electronic "librarian" capable of managing book loans and returns and/or locating volumes requested by the user (or indicating in which office they may be found).

In this context, a system for automatic face recognition is required specifically with respect to the librarian's first task, i.e. managing loans and returns (a similar system will later be implemented in the automatic concierge). The electronic librarian must in fact be capable of identifying any user that might wish to borrow or to return a book so as to ensure that only registered personnel can have access to the IRST library. And for this purpose, the user is simply expected to stand in front of the system and look into a video camera. (In fact, our project is to use both face and speaker recognition techniques, so as to further improve the system's reliability. In the following, however, the focus will be exclusively on the vision component.)

The experimental scenario is therefore very unconstrained. No particular effort is required to ensure perfectly frontal images, and the distance of the subject from the camera as well as the location of his/her face in the image are only approximately

---

**FIGURE 1.** Functional diagram of the MAIA system and some of its tasks. Shaded blocks (connected by black arrows) indicate the contextual background of the application described in the paper.
fixed. This means that the system must be highly tolerant against variations of the head size and orientation. Moreover, background and illumination are not assumed to be constant: artificial light is used to illuminate the user’s face from the front, but the experimental environment is also exposed to sun light through numerous windows.

**Face Detection and Eyes Localization**

As is clear from the above, a most important feature of the system is that it must be capable of recognizing *dynamic* facial images, i.e. “live” images acquired by the librarian through a video camera. This is a general requirement of the MAIA project and a compelling prerequisite for most industrial applications.

To detect the user’s face from the background, the system makes use of a motion detection algorithm originally introduced in (Stringa, 1991b) and refined in (Messelodi, 1991). This is based on the general fact that a basic stimulus in the analysis of a dynamic scene lies in detecting “differences” between successive images; the algorithm proceeds by comparing pairs of sequential images captured by the camera and segments from the background those objects that determine a significant variation in the images’ matrices. Despite its simplicity, it performs well, allowing detection of faces in almost real time (about 3 images/sec) and showing a remarkable independence from background and illumination conditions.

The image of the face, detected from the background, is then adjusted and normalized before the system can proceed to the recognition algorithm. As we mentioned, in fact, this follows a template matching strategy and is formalized as a distance-based comparison between the directional derivatives of the input image

![Figure 2](image)

**FIGURE 2.** Face detection: the algorithm detects the "differences" between successive input images (left) to extract the edges of the face (right).
(the face to be recognized) and those memorized in a data-base of templates or
prototypes covering each subject known to the system. It is therefore important—as
in any correlation-based approach—that the face be accurately positioned in the
image. Otherwise the need for shifting required to obtain the best matching between
the subject and a template could increase considerably, with obvious disadvantages
in terms of computing costs.

The solution adopted in the present application is based on a technique that
localizes the position of the eyes and then align them horizontally. Eyes are in fact
a most prominent facial feature and can be detected with fair accuracy. Their
localization is then used to register, rectify, and normalize the image with respect
to the distance of the pupils, yielding a “standard” matrix that can easily be compared
with the templates.

Various approaches can be used for the purpose of locating the position of the
eyes. For instance, in Baron (1981) a procedure is used whereby a certain number
of eye templates are correlated against suitable subimages of the input image; a
correlation value greater than a fixed threshold is then taken to indicate that an eye
is successfully located. Our approach is different. It does not proceed by eye template
matching. Rather, an algorithm is used based on the exploitation of (a priori)
anthropometric information combined with the analysis of suitable grey-level
distributions, allowing direct localization of both eyes.

On the one hand, there exists a sort of “grammar” of facial structures that
provides some very basic a priori information used in the recognition of faces. Every
human face presents a reasonable symmetry, and the knowledge of the relative
position of the main facial features (nose between eyes and over mouth, etc.) proves
very useful to discriminate among various hypotheses. These guidelines can be
derived from anthropometric data corresponding to an average face and refined
through the analysis of real faces. Some typical examples [based on studies on face
animations and reported in Brunelli (1990)] are:

• the eyes are located halfway between the top of the head and the bottom of the
  chin;
• the eyes are about one eye width apart;
• the bottom of the nose is halfway between the eyebrows and the chin;
• the mouth is typically located one third of the way from the bottom of the nose
to the bottom of the chin.

On the other hand, the algorithm exploits the discriminating power of the
distribution of the image’s edges (specifically leading edges, i.e. transitions from
dark to bright) upon adequate filtering. The primary motivation is that edge densities
convey most of the information required to identify those facial zones that are
characterized by abrupt changes in image brightness [see, e.g., (Kanade, 1973)]. In
particular, eyes are typically the most structured area of a face, and their location in
the image is therefore characterized by a high number of edges. These determine a prominent peak in the edges’ vertical projection that can be indicative of the rough localization of the eyes.

The approach is schematically illustrated in Figure 3, where $R_{xy}$ and $R_{zx}$ are the vertical and horizontal resolution (number of lines and columns in the image) respectively.

The localization of the eyes proceeds from the preliminary approximate localization of the eyes-connecting line, centered on the maximum ($Z$) of the filtered vertical histogram, and of the face’s main traits, specifically the face’s side limits ($X_s$ and $X_d$) and the nose axis ($X_n$). These are not searched for in the entire image but only on those areas that correspond to appropriate ‘expectation zones’. On this basis, the search areas for the two eyes can be estimated with reasonable accuracy. Their exact localization is then obtained by computing the horizontal and vertical coordinates of a pixel belonging to the corresponding pupils.

**EYES LOCALIZATION: DETAILED DESCRIPTION**

**Preliminaries**

A detailed description of the eyes localization algorithm is given in the following paragraphs. For convenience, we first introduce some general notions that will be used throughout.

![Figure 3](image.png)

**FIGURE 3.** The eyes localization algorithm exploits the discriminating power of filtered edges and grey-level distributions.
Let $I(x,y)$ be the input image, digitized at 256 levels of grey into a matrix of size $R_{sx}$ (wide) by $R_{sy}$ (high) pixels. The binary matrix $E(x,y)$ describing the horizontal leading edges of $I(x,y)$ is computed using a thresholded directional derivative. This is defined as

$$E(x,y) = \begin{cases} 1 & \text{if} \ |I(x,y) - I(x-1,y)| > I_n/C \\ 0 & \text{otherwise} \end{cases}$$

(1)

where $I_n$ is the average value of $I(x,y)$ and $C$ is a constant parameter. In general, the exact value of $C$ will depend on the resolution of $I(x,y)$.

FIGURE 5. Once the eyes have been exactly localized, the image is parametrically registered, rectified, and normalized with respect to the distance of the pupils to produce a “standard” matrix.
The projection of the horizontal leading edges along the vertical axis defines the vertical histogram \( H(y) \). This is computed from \( E(x, y) \) by taking the value of \( H(y) \) at any point \( y \) on the vertical axis as the sum of all the leading edges of the corresponding horizontal line:

\[
H(y) = \sum_x E(x, y)
\]  

(2)

Finally, given a function \( f(x) \), the following filtered functions are defined:

\[
f_j(x) = \frac{1}{2j + 1} \sum_{-j}^j f(x - j)
\]

(3)

\[
f_{mn}(x) = f_s(x) - f_m(x)
\]

(4)

Here \( j, m, \) and \( n \) are constant parameters whose specific values depend on what \( f \) is meant to extract: \( f_j(x) \) is the result of filtering \( f(x) \) on \( 2j + 1 \) samples, and \( f_{mn}(x) \) is the result of purging \( f(x) \) with a pass-band filter based on \( f_s(x) \) and \( f_m(x) \).

**Rough Localization of the Line Connecting the Eyes’ Pupils**

Using the above definitions, the first step of the algorithm is the rough localization of the line connecting the eyes’ pupils, which allows the construction of an approximate model of the face using anthropometric standards.

The technique used for this purpose proceeds from the idea that the analysis of the vertical projection of the horizontal edges identifies the location of significant, highly structured features. The higher the peak, the more structured the feature. Eyes are the most structured part of a person’s face. Hence, their location determines a most prominent peak in the grey-level projection. Moreover, they are expected to be located somewhat below the head top, as the anthropometric guidelines reported in (Brunelli, 1990) suggest. The head top, \( Y_s \), can be computed directly from the edge of the face given by the face detection algorithm, while the upperbound of the search area for the eyes is an anthropometric parameter \( \Delta \) relative to the size of the face.

As a first rough approximation, the eye-connecting line can therefore be identified with the \( y \)-coordinate of the highest projection peak relative to this expectation zone. More exactly, this is done by searching for the \( Z \)th horizontal line of the matrix, where \( Z \) is the ordinate of the maximum of the filtered vertical histogram in the search area:

\[
H_{mn}(Z) = \max_{y} H_{mn}(y)
\]

(5)
Note that $Z$ is defined relative to the filtered histogram $H_{mn}(y)$. The analysis of $H(y)$ could already be very useful in determining the position of the eye-connecting line. However, some care is required to minimize misleading interferences such as high frequency noise or extensive areas with a high number of constant or quasi-constant brightness transitions. Accordingly, a band-pass filter is applied to purge the histogram from such disturbances, and the filtered histogram $H_{mn}(y)$ is used instead of $H(y)$ (for suitable values of $m$ and $n$).

In case there exists some other value of $H_{mn}(y)$ that is significantly high, i.e. if there exists some point $Z'$ such that

$$H_{mn}(Z') > K \cdot H(Z) \quad (K < 1 \text{ constant})$$

the eye-connecting line is identified with the value $Z$ or $Z'$ which is most plausible, relative to its distance from the head top $Y_n$, on the basis of the expected position of the eyes in a “standard face.”

**Rough Localization of the Face’s Side Limits**

This and the next step are aimed at characterizing the relative “expectation zones” of the two eyes: the search area for the left eye will be restricted to a region between the left limit of the face and the nose, while the search area for the right eye will be restricted to a region between the nose and the right limit.

It is clear that for this purpose it is not necessary to perform the search of the side limits on the entire image. Since the eyes’ expectation zones will be centered on the approximate eye-connecting line (calculated with the method indicated above), what is needed is the position of the side limits of the face relative to a region

![Figure 6](image)

**FIGURE 6.** As a first approximation, the eye-connecting line is defined as the ordinate ($Z$) of the maximum of the filtered vertical histogram.
centered on this line. In some cases these may not coincide with the “true” limits: the outermost extremities of a face could easily be located near the head top or, perhaps, far below, at the mouth level (as in people with pronounced jaws). However, for the present purposes such facial features need not be considered, and the algorithm performance can be improved by restricting the search area to the indicated region.

Let this region be defined by the interval \( Z - O_0 : Z + U_0 \), where \( O_0 \) and \( U_0 \) are fixed parameters (defining the search area Over and Under the eyes). Relative to this region, the rough localization of the side limits is determined with reference to the face detection algorithm described in Face Detection and Eyes Localization. The detection algorithm extracts the edges of the face (see Fig. 2), and the abscissae of the leftmost and rightmost edge points in the interval \( Z - O_0 : Z + U_0 \), which we denote by \( X_L \) and \( X_R \), can be taken to define the abscissae of the left and right side of the face, respectively.

It is worth observing that this approach does not require any constraint on the experimental set-up. An alternative method, based on the assumption that the experiments be performed on a light background, has also been investigated, though it has not been used for the MAIA librarian. This alternative approach moves from the remark that, on a light background, the side limits of the face are typically localized in those areas in the image that are characterized by abrupt changes in brightness, corresponding to the transition from background to object. To locate them approximately, the following operations can therefore be performed. First, the image’s horizontal average density \( A(x) \) is computed relative to the search region. This is obtained by taking the value of \( A(x) \) at any point \( x \) on the horizontal axis as the normalized sum of all the values of the corresponding vertical line in the interval \( Z - O_0 : Z + U_0 \):

\[
A(x) = \frac{1}{U_0 + O_0 + 1} \sum_{z=0}^{Z+U_0} I(x,y)
\]

Second, \( A(x) \) is filtered on \( 2p + 1 \) samples (\( p \) a fixed parameter) to produce the filtered density \( A_p(x) \) as defined in Eq. (3). A glimpse at the example in Fig. 7 will show the typical behavior of this function; on a light background, the face limits are expected to coincide with the leftmost and rightmost significant brightness changes in the image, and these correspond to the lowest values of the filtered average density \( A_p(x) \).

On this basis, the horizontal coordinates of the side limits can easily be determined. Assuming the face to be roughly centered in the image, we can locate the search area for these points within a certain interval centered on \( Rsx/2 \). Let \( \Delta_L \) and \( \Delta_R \) be the left and right extrema of this interval respectively. The left side of the face can then be defined as the abscissa \( X_L \) of the minimum of \( A_p(x) \) in the left portion of the interval:
FIGURE 7. Approximate localization of the left ($X_\ell$) and right ($X_d$) face sides.

\[ A_r(X_\ell) = \min_{\Delta_1} A_r(x) \]  

while the right side can be defined as the abscissa $X_d$ of the minimum of $A_r(x)$ in the right portion:

\[ A_r(X_d) = \min_{\Delta_r} A_r(x) \]

As we mentioned, this alternative procedure has not been implemented in the MAIA librarian due to the limiting assumption on the light background. However, our experiments have shown that when this assumption is satisfied, the procedure yields essentially the same results as that based directly on the face detection algorithm.

**Rough Localization of the Nose’s Axis**

This step is similar to the previous one. With reference to our figures, note that whereas the face limits are usually darker in the image, the nose is normally lighter than the left and right regions of the facial image and determines a prominent peak in $A_r(x)$. Accordingly, our approach is to base the search for the nose’s axis on the maximum value of $A_r(x)$.

Considering that the nose is expected to be located somewhere halfway between the face sides, the search area for its axis can safely be restricted to a central region comprised between the two vertical lines $X_\ell$ and $X_d$ calculated as above. Moreover, since the input image is assumed to provide a frontal view of the face (albeit not a
perfectly frontal one), it is not necessary to consider the entire interval $X_s : X_d$. The search area can be further restricted to a region localized at a certain distance $D_n$ from $X_s$ and $X_d$. Based on standard anthropomorphic guidelines, and considering that the eyes are usually one eye width apart (Brunelli, 1990), this distance can roughly be assessed at one fourth of the above-mentioned interval:

$$D_n = \frac{X_d - X_s}{4} \quad (10)$$

Using Eq. (10), the nose vertical axis $X_n$ is then defined as

$$A_p(X_n) = \max_{x_n + D_n} A_p(x) \quad (11)$$

i.e. as the abscissa of the maximum of the filtered density $A_p(x)$ in the interval $X_s + D_n : X_d - D_n$.

Detection of the Pupil’s Coordinates

This is the final step. Using the approximate location of the eye-connecting line, of the face sides, and of the nose axis, the expectation zones of the two eyes can be estimated with reasonable accuracy. Their exact localization is then obtained by computing the horizontal and vertical coordinates of a pixel belonging to the corresponding pupils.

![Approximate location of the nose's axis ($X_n$).](image)
**Left Pupil**

As shown in Fig. 9, the left eye is expected to be located within a region centered on the approximate eye-connecting line and comprised between the left limit of the face and the nose axis. More precisely, the search area is restricted to the rectangular region defined by the intervals \( Z - L_1 : Z + L_2 \) (high) and \( X_n - L_3 : X_n - L_4 \) (wide), where each \( L_i \) is a suitable parameter.

Relative to this area, the search of the pupil is based on the analysis of the horizontal grey-level distribution. For each line \( y \), the relative horizontal density \( G'_r(x) \) is calculated and a band-pass filter is applied to eliminate misleading interferences and high frequency noise. This is obtained as in Eq. (4):

\[
G^y_{rs}(x) = G'^r_r(x) - G'^r_s(x)
\]  

(12)

where the values of \( r \) and \( s \) are based on the a priori knowledge of the relative position of sclera, cornea, and pupil in the eyeball. The second derivative \( g''(x) \) is then calculated as

\[
g''(x) = \frac{d^2G^y_{rs}(x)}{dx^2}
\]  

(13)

The second derivative allows those areas where the brightness changes are most rapid to be detected and is therefore a most efficient solution to determine the exact location of the eye-pupils.

A typical plot of \( g''(x) \) is schematized in Fig. 10 (compare also the examples in Figs. 4 and 9).

**FIGURE 9.** The rectangular regions located between the face limits and the nose axis indicate the expectation zones of the two eyes.
FIGURE 10. A typical plot of the second derivative used to localize the eye pupils: note the peak in the middle, corresponding to the pupil’s position.

Note that the "symmetry" axis in the distribution of the curve is clearly marked by a peak, corresponding to the eye’s pupil (recall that artificial light is used to illuminate the user’s face from the front, thereby producing a typical white spot in the middle of the eye). The two adjacent peaks, of less intensity, clearly indicate the discontinuity represented by the cornea, while the vertex of the main peak indicates the center of the pupil. The precise localization of the pupil is therefore obtained by taking as a vertical coordinate the value of $Y_1$ such that

$$\max g^y(x) > \max g^r(x) \quad \text{for all } y \neq Y_1$$

(14)

and as a horizontal coordinate the abscissa $X_1$ of $\max g^y(x)$

*Right Pupil*

The procedure is similar. The right eye is expected to be located more or less symmetrically, in a region comprised between the nose axis and the right limit of
the face and centered on the approximate eye-connecting line. The search area for the right pupil is localized in a rectangle defined by the intervals \( Z - R_1 : Z + R_2 \) (high) and \( D_s - R_3 : D_s - R_4 \) where each \( R_i \) is a suitable parameter and

\[
D_s = 2X_n - X_1
\]  

(15)

The exact coordinates of the pupil, \( Y_p \) and \( X_p \), are then computed with the same method described for the left pupil.

**EXPERIMENTAL RESULTS**

The performance of the eyes localization algorithm has been tested on a database of 220 frontal images of faces of 55 people working at IRST (34 males and 21 females, 4 images per person). Subjects of various age were considered (from 23 to 37 years). Some males had beards, though none of the people wore glasses.

The algorithm has been implemented on an HP 350 workstation. Each input image was digitized at 256 levels of grey into a matrix of size 524 \( \times \) 342 pixels \((R_{xy} = 524, R_{xx} = 342)\) acquired with a b/w CCD camera and detected from the background with the method described in Face Detection and Eyes Localization. Some examples are given in Fig. 11, while the distance and orientation distributions of the pupils in the data base are summarized in Fig. 12.

Several steps of the algorithm, such as the computation of the binary matrix of the horizontal leading edges or the band-pass filter applied to purge their vertical histogram, involve parametric functions. The parameters used in the experiment, based mainly on heuristic evaluations and previous knowledge, are reported in Table I.

![Figure 11: The database comprises several images (prototypes) for each subject known to the system. Input facial images are assigned to the class of the closest prototype, relative to a fixed threshold, or rejected as 'unknown'.](image-url)
TABLE 1. Parametric Values Used in the Experiment

<table>
<thead>
<tr>
<th>Parameter description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Threshold for the binary matrix ( E(x, y) ) of the horizontal leading edges</td>
<td>( C = 16 )</td>
</tr>
<tr>
<td>Search area for the eye-connecting line [see equation (5)]</td>
<td>( \Delta = \frac{R_s y}{6} )</td>
</tr>
<tr>
<td>Parameters of the filtered vertical histogram ( H_{mv}(y) ) [see equation (5)]</td>
<td>( m = 18; \ n = 5 )</td>
</tr>
<tr>
<td>Threshold on ( H_{mv}(y) ) [see equation (5)]</td>
<td>( S_{yo} = 4 )</td>
</tr>
<tr>
<td>Evaluation of ( Z' ) [see equation (6)]</td>
<td>( K = .85 )</td>
</tr>
<tr>
<td>Expectation zones of the face’s side extremities [see equation (7)]</td>
<td>( O_0 = 32; \ U_0 = O_0/2 )</td>
</tr>
<tr>
<td>Search area for the face’s left side [see equation (8)]</td>
<td>( \Delta t = R_s x/8 )</td>
</tr>
<tr>
<td>Search area for the left face’s right side [see equation (9)]</td>
<td>( \Delta r = .875R_s x )</td>
</tr>
<tr>
<td>Filtered horizontal density ( G^T_{tv}(x) ) for the localization of the pupils [see equation (12)]</td>
<td>( r = 6; \ s = 1 )</td>
</tr>
<tr>
<td>Search area for the left pupil</td>
<td>( L_1 = L_2 = U_0/4; \ L_4 = L_3/4 )</td>
</tr>
<tr>
<td>Search area for the right pupil</td>
<td>( L_3 = \max [(X_d - X_n, X_n - X_d)]; \ L_4 = L_3/4 )</td>
</tr>
<tr>
<td></td>
<td>( R_1 = R_2 = U_0/2 )</td>
</tr>
<tr>
<td></td>
<td>( R_3 = R_4 = 7 \cdot \max (X_d - X_n, X_n - X_d) )</td>
</tr>
</tbody>
</table>
With this choice of parameters, the algorithm correctly identified the location of the eyes by determining the coordinates of a pixel belonging to the left pupil and a pixel belonging to the right pupil for each facial image included in the data base. The processing time was less than 0.2 sec per face.

The results of the experiment can thus be summarized as

<table>
<thead>
<tr>
<th>Correct eye localization:</th>
<th>100%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Processing time per face:</td>
<td>&lt; 0.2 sec</td>
</tr>
</tbody>
</table>

**FINAL REMARKS**

These results provide good evidence of the efficiency of the approach. Indeed, although it should be emphasized that the reported experiments are to be considered preliminary, the size of the database and the structure of the testing set are already indicative of the system’s performance. The algorithm is simple, fast, and reliable. Further improvements are to be expected if the inherent pragmatism of the approach—a successful trade-off between the analysis of grey-level distributions and a priori anthropometric information—is fully exploited.

The efficiency of the algorithm for use in a correlation-based face recognition system has also been successfully verified. In the system implemented for “electronic librarian” of MAIA, the eyes localization algorithm is applied to register, rectify, and normalize the image parametrically with respect to the distance of the pupils, producing a “standard” $44 \times 27$ matrix at 16 levels of grey that can easily be compared with the templates in the Testing Set (220 frontal images of faces of 55 people working at IRST, 34 males and 21 females, 4 images per person). As already mentioned in the introduction, the experimental results obtained so far indicate a recognition rate of 100% at excellent processing speed (about 1.25 sec per face on an HP 350 workstation and less than 0.05 sec on a two-layer S_Net using binary derivatives). Considering that the experimental set-up of the MAIA librarian is very unconstrained, the proposed solution appears to provide a simpler and more efficient ground for face recognition than other methods described in the literature.

On-going research is focused on generalizing the approach to more general problems such as the extraction of facial features (nose, shape of the eyes, eyebrows, forehead, etc.) and the evaluation of the rotation angle around the vertical axis of the face.
REFERENCES


Stringa, L. 1991b. MOD: Moving objects detection per l’estrazione da una scena di oggetti e dei loro bordi. IRST Internal Report #9101-06, gennaio.


