Face-recognition intelligent interactive system

Ana Carina Palumbo
Graduate School of Informatics
University of Amsterdam, The Netherlands.

Abstract

The aim of this report is to provide insight in the principles for the design of a face recognition intelligent interactive wearable system. Relevant developments and applications in this field are analyzed, in order to detect efficient methods, techniques and guidelines for the design of the system. The selected techniques are combined to achieve the design of a face-recognition intelligent interactive system on a wearable device, capable of recognizing people's faces on a real time context. Apart from recognizing individuals, the system allows the users to manage, edit and interact with information about friends, acquaintances, familiar strangers and completely strangers, enhancing social interactions in urban, academic, professional and leisure settings.

1. Introduction and motivation

In social interactions, flows of personal and identity information occurs. Data about people is needed to interact and influences involvement, engagement, and expression. Sharing names, ages, interests, views, opinions, tastes, and any other kind of personal information, enhances the ability to create a sense of connectedness with others. This connection allows that individuals gain a sense of belonging cultivating comfort, safety and inclusion (Goodman & Paulos, 2004).

People develop working models of their relationships that function as cognitive maps that help them navigate the social world (Baldwin, 1992). As they interact with other people based on this schemes, each time they hold a conversation, share an activity or have any other kind of interaction with people, they retrieve information from cognitive maps. These complex mental models involve a wide variety and diversity of information, that ranges from the smell, voice and face to even the interests and opinions of people.

Models of relationships come into play each time we recognize a friend, acquainted or somebody who seems familiar, or when a stranger awakens our curiosity. In any of these cases, providing additional social and personal information to the mentioned maps, could enhance social interactions with people (Kostakos et al., 2005). Individuals' behavior in any social setting differs and is dependent upon with whom or what they are interacting. How to interact appropriately in a given situation is drawn from the environment, individual's expectations, cultural and social norms of behavior.

One way to access information about people within visual range in a real-time context, without having to approach them, is through their faces. A wearable face-recognition system that retrieves social networks profiles and any other kind of online information about a person, would represent a new way of navigating social spaces, either in professional or leisure environments. In the same way that Quick Response Barcodes are used to get instant access to digital information about a company, product or event, face recognition in a real time context can be used to obtain information about a person that is not analogically available or that is not part of the mental model of the user.

A tool that makes this possible is proposed in this report: a face recognition system that turn faces into "links" that connect face-to-face interactions with virtual information, combining spatial and transpatial social networks, and transgressing the lines and protocols between public and private spaces (Kostakos & Venkatanathan, 2010). The technology designed on this report would act as a bridge between spatial and transpatial networks, creating an individuals' platform for social interaction and engagement (Kostakos & Venkatanathan, 2010).

Considering the work of Drascic & Milgram (1996), it can be stated that the proposed system is an Augmented Reality (AR) tool, in which views of the real world (faces) are combined in some proportion with views of a virtual environment (virtual information about people). The system enhances the user's ability to read situations and to understand the social context that influences behavior. Some of the key dimensions of social intelligence (Kostakos et al., 2005) are particularly supported by this system. Presence, which is the external sense of self that others perceive (e.g. confidence, self respect and self worth), and empathy, which is the ability to create a sense of connectedness with others.

When being used by people, the face recognition intelligent interactive system, would support various modes of interaction according to the degree of the user's knowledge about the person in question, meeting different needs respectively: (a) Friends and acquaintances: When interacting with friends and acquaintances, the system would provide contact and personal information about the person, avoiding the process of manually entering contact data. Besides, as Singletary & Starner (2001) state, in casual social interaction, it is easy to forget the names and identities of those we meet, so it would be useful as well to avoid the need to be reintroduced to somebody, for instance a business contact. Augmenting one's senses by getting alerted when a friend or a friend of a friend is nearby would be useful functions for this particular case. (b) Familiar strangers: Familiar strangers form a border zone between people we know and the completely unknown strangers we encounter once and never see again. For this cases, the face-recognition system would allow impromptu social gathering (Goodman & Paulos, 2004), fetching information to/from nearby users, learning more about people in proximity, or finding social partners (Jung et al., 2006). (c) Strangers: When unacquainted people are around, users may wish to identify a potential, appropriate, or inappropriate conversational party, or a potential topic for conversation (Jung et al., 2006). In this sense, the system would encourage acceptable social contact with strangers, making up for the lack of mobile devices to explore and play our subtle connections to them (Goodman & Paulos, 2004).

Related work

The face recognition system proposed in this report makes use of face detection, face tracking and face recognition. These techniques are applications of computer vision (a) and biometrics (b). To introduce different works on this area, general definitions about the mentioned fields of study are provided below.

a. Computer vision can be defined as the enterprise of automating and integrating different processes and representations used for vision perceptions. It involves techniques such as image processing, statistical pattern classification, geometric modeling and cognitive processing (Ballard & Brown, 1982). Visual perception is the relation of visual input to previously existing models of the world. As there is a representational gap between the image and the models that describes the image information, computer vision systems determine ranges of representations to connect the input with an output (Ballard & Brown, 1982).

Machine vision can be used for detection, identification, and tracking extending human visual abilities by delegating those tasks to the system that are hard or impossible to perform (Crowley, 2006). Nowadays, it is used in a wide variety of real-world applications, which include, among other, optical character recognition (OCR), match move, surveillance, fingerprint recognition, photo-based breakthroughs, face detection and face recognition (Szeliski, 2010).

Face recognition can be defined as the identification of individuals from images of their faces using a facial database with people's identities (Cevikalp et al., 2005). The technique, widely studied within the past decades, can be applied to a wide variety of problems, including criminal identification, security systems, image and film processing, and human computer interaction (Turk & Pentland, 1991).

b. Biometrics is the science of establishing the identity of an individual based on the physical, chemical or behavioral attributes of the person (Jain et al., 2008). However, the term biometric authentication is perhaps more appropriate than biometrics since the latter has been historically used in the field of statistics to refer to the analysis of biological (particularly medical) data (Wayman et al., 2010). Although biometrics emerged from its extensive use in law enforcement to identify criminals, it is being increasingly used today to establish person recognition in a large number of civilian applications. Human physiological and behavioral characteristics can be measured with biometric techniques as long as they are universal (each person should have the characteristic) distinctive (any two persons should be sufficiently different in terms of the characteristic), permanent (it should be sufficiently invariant over a period of time), and collectible (it can be measured quantitatively). Performance (achievable recognition accuracy and speed, resources required, and factors that affect the accuracy and speed), acceptability (the extent to which people are willing to accept the technology), and circumvention (how easily the system can be fooled using fraudulent methods) are issues that should be considered as well (Jain et al., 2004).

According to Jain et al. (2004) biometrics system may operate in one of the following operation modes: verification, that implies a comparison of a captured

Figure 1. Smart glasses: Android-App and Goggle-based interface/app(Kurze & Roselius, 2011)

> biometric with a stored template to verify if the person is who he claims to be; and identification, that consist on the comparison of the biometric information with a database in order to identify an individual. Face recognition is one of the several applications of biometrics. The term recognition can be used when it is not required to make a distinction between verification and identification (Jain et al., 2004).

Using biometrics and computer vision methods, several face recognition systems have been developed over the last years. Kurze & Roselius (2011) have recently developed "Smart Glasses" (Figure 1) to perform real time face recognition. The system consist of a wearable display, a camera, some local processing power and an uplink to a backend service in order to identify individuals. The application runs on three devices: mobile phones with Android, a conventional wearable augmented reality display, and a goggle-based wearable display. The system executes face detection and face tracking locally on the device (e.g. smartphone) and then links to the service running in the cloud to perform the actual face recognition based on the user's personal contact list. It allows users to recognize a person sending back information about this person: person's name, affiliation, profile-image and latest social network activities.

Krishna et al. (2005) designed the iCare Interaction Assistant: an assistive device for helping the individuals who are visually impaired during social interactions. In this work, they address face recognition as the initial step towards building a comprehensive social interaction assistant that will identify and interpret facial expressions, emotions and gestures. The researchers run experiments in order to select a face recognition algorithm that works despite changes in facial pose and illumination angle. The system uses face detection, face recognition, a database and it also includes a text to speech converter in order to give feedback to the users (Figure 2).

Han et al. (2010) introduce a prototype of a wearable system for observation and recognition: the WOSS. The system, which aims at supporting human vision, allows real-time face detection, recognition and identification, by using a wearable camera that captures video. They particularly deal with the problems related to the constantly and arbitrary moving of the camera. These researchers also propose the “iWearSA”, a similar system but mainly focused on the security surveillance field.

Mukherjee et al. (2008) propose a secure face recognition system for mobile-device. Their work is particularly interesting because it deals with problems associated with limited storage and processing power of the mobile device, connection instability, security and

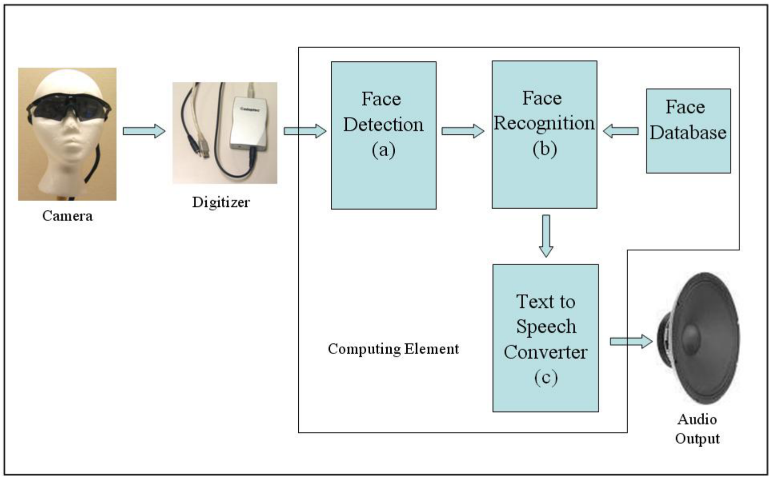

Figure 2. Diagram of the wearable face recognition system for visually impaired individuals (Krishna et al., 2005).

privacy concerns, and limited network bandwidth. As a solution to this challenges, a DCT[1]-based compression method is proposed. The system has been implemented on a commercially available phone and experimental results demonstrate its potential.

One of the latest works on this field has just been introduced by Android and Samsung: the Galaxy Nexus. This novel phone is the first commercial smartphone that runs face recognition to be unlocked (*Nuevo Android 4.0 lleva el reconocimiento facial al móvil | ELESPECTADOR.COM*, 2011).

Other interesting application is being developed by the Brazilian police: face recognition glasses to identify and root out criminals at the 2014 World Cup ("Brazilian police to use 'Robocop-style' glasses at World Cup - Telegraph", 2011). The glasses will make use of a small camera, capable of capturing 400 facial images per second. The system will be used to scan faces in crowds up to 50 meters away. If adjusted, it will also be capable of searching for a specific target as far as 12 miles away. The recognition process will involve the comparison of biometric data at 46,000 points on a face. Matches to known criminals or people wanted by the police will be notified.

The described related work, states the feasibility of real-time face recognition. Several applications on this field have been developed and tested, achieving efficient face detection and recognition on different contexts and with diverse aims. Even tough the development of a computational model for face recognition seemed difficult some years ago because of the complexity, multidimensional and meaningful visual stimuli it implies (Turk & Pentland, 1991) it is nowadays possible, even considering difficulties such as poor lightening and variances in facial expressions.

2. Methods

2.1. Interaction design

[1]Discrete Cosine Transform

2.1.1 Scenarios and use cases

Scenario and use cases are described below. The different contexts of use and goals analyzed, allow the exploration of interaction variables and hence, the modeling of the system's interaction design.

Scenarios are organized in two main categories: informal scenarios, which include leisure and urban environments, and formal scenarios, associated with business, professional or academic contexts.

Informal scenarios.

Use case 1: Afternoon in the park

Jazmin Smith (32) is visually impaired. She is meeting some work friends at the park to spend a sunny afternoon. She arrives alone at the park and turns on her recognition wearable device, which automatically synchronizes with her mobile. She accesses the menu and selects the "flagging function". She looks up for the "work friends" group of contacts and flagged the group as "searched". The system notifies her when it recognizes one or more faces of Jazmin's work friends.

Use case 2: Stranger in the metro

John William (21) is on the metro, traveling with his friends. He sees a girl he finds attractive on a seat next to him. He turns on his recognition device, making sure that her face is on the camera's field of view. On his mobile screen, he sees a list of the detected faces. He selects the face of the girl he likes. As she is not available to be recognized by everybody, John decides to send her a recognition request to get access to information about her. The girl, who has already seen him shooting her, receives a notification stating that somebody named "John William" wants to recognize her and accesses her information. She accepts the request giving him only access to her "informal" profile. After her acceptance, John reads her profile and finds out that she is called Elizabeth Jones and studies Psychology. He decides to start interacting and sends her a text message.

Formal scenarios.

Use case 1: Academic conference

Jack (55) is an important researcher on Climate Change. He attends an international conference on the field that takes place in California. He is highly interested in adding new PhD students to his research group. He uses the face-recognition wearable device to search for them. To achieve his goal, he turns on the device during a coffee break and configures it in the following way:

1. He edits his privacy settings in order to be recognizable by anyone who is interested in climate change and is enrolled as a PhD at a University. He allows access to his contact data and professional and academic profile.

2. He also activates the "Flagging" function and assign a green flag ("searched") to all the people that match the profile he would like to meet: PhD students enrolled at any University in the United States, interested, related to or working on Climate Change. As

an optional requirement, he establish that they could be members of the Climate Change Association.

During the coffee break he will be notified if somebody with the described profile is on his visual range or if somebody with this characteristics recognizes him.

Use case 2: Business cocktail reception

John Anderson (42) is the manager of an International Company that develops Software. He is invited to a business cocktail reception party where he meets Arnold Thompson (48), an important investor that is interested in establishing a partnership to develop mobile applications. After talking about the interesting project, they want to exchange contact information. They both use the recognition wearable device, but Arnold is not wearing it today. John takes a picture of Arnold and sends him a recognition request. Arnold accepts the request and gives him access to his basic profile, contact data and professional profile. They both save each other as contacts.

2.1.2 User goals and interaction processes

From the scenarios and use cases described, different user goals were identified (Figure 3). Even tough the main goal of the users of this system may be to recognize individuals, secondary needs emerge from this goal. The identified goals are the following:

- Recognizing an individual: the user may want to recognize an individual on his visual range. The system gives him two alternatives in this case. The face to recognize can be selected from a list ("Basic Recognition" function) or it can be selected from real-time video ("Exploratory mode" function).
- Being notified about the presence of an individual: the user may want to get alerted if somebody he knows or somebody he would be interested in meeting is nearby. The system can be configured to achieve this goal by using the "Flagging" function.
- Searching for people: the user may want to look up for an specific individual or a target in the global database of the system, either to access available information about them or to get alerted when they are nearby.
- Viewing or editing contacts: the user may also want to save certain people recognized or found on the database as personal contacts, and view information about them. The user may also want to assign groups or relationships to them. Different properties of contacts can be edited, such as type of relationship (close friend, friend, acquaintance, family member), group and flags (searched or avoid). In this sense it is important to mention as well that the recognizing application is connected to all the basic functions of the mobile. This means that when the user is visualizing somebody's profile he has the option to call or send a text message to this person through his mobile. The user can also add his contacts in his social networks accounts.
- Editing personal and privacy settings: when creating an account in the face recognition system[2], the user provides personal information. This is the data that other users will retrieve when recognizing him. The system provides three main categories of personal information: "Contact information", which must be manually entered by the user, "Academic and professional profile", that can be imported from LinkedIn, and "Informal profile", which

[2] Prerequisites to use the system are described below.

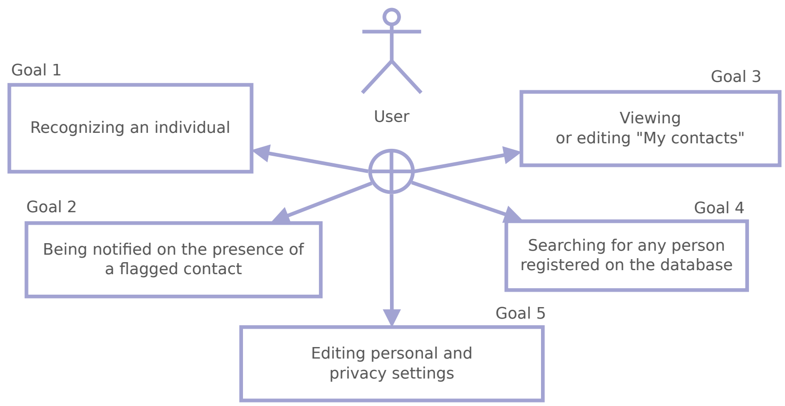

Figure 3. Interaction goals

can be also imported from Facebook. The system allows the user to edit, update or delete this information. On the other hand, the user may want to edit privacy settings to manage who can or cannot recognize him and to which categories of personal information he wants to allow access in each case.

The users interact with the face-recognition system using two devices: a wearable device, which is is camouflaged under the shape of a fashionable unisex brooch in diverse colors, textures and combinations; and their mobile phones. The mobile provides the main interface for interactions. Most of the inputs from the user and the feedbacks of the system take place on the mobile's screen. Further information about this devices can be found on section 2.2, which contains the system description.

Prerequisites . The system requires the prerequesites below to work properly:

- An installation software has to be run before the first time the system is used. This software configures the camera, installs the mobile software and establish a unique encoded Bluetooth connection between the wearable device and one specific mobile.
- The user has to create an account on the face-recognition system. For this, he has to upload at least 5 pictures of his face and fill in his personal information: contact information, that must be entered, professional and academic profile, which can be imported from LinkedIn, and informal profile, that can be imported from Facebook.
- The mobile and the wearable device must be in the range of their bluetooth connectivity to be able to perform face recognition.
- To be recognized, people must be on the wearable device camera's field of view, and no further than 5 meters from the user.
- Though the user can use the mobile to review, edit and add contacts, search for people or edit personal settings, other functions (exploratory mode, basic recognition and flags) only work if the user is carrying both the mobile phone and the wearable device.
- The system requires massive use to be successful.
- The user must have Internet connectivity on his mobile.
- Both the camera and the wearable device must have enough battery to operate.
- The user must be competent using his mobile phone and have basic digital literacy.

Interaction processes. The interaction with the system is initiated by the user in any of the following ways:

- The user presses the wearable device button[3] during 3 seconds to turn it on and then opens the recognition software on his mobile: this is used for achieve any of the identified goals.
- The user presses the wearable device button during 3 seconds to turn it on and then presses it simply to start detection: this initial step can be done to start more quickly the process to achieve goal 1. This turns on the bluetooth connectivity, synchronizes with the mobile and activates the camera, which starts sending video signal to the mobile. This also means that the mobile application[4] is opened as well. The button on the wearable device was introduced to allow faster processes of recognition. If the user needs to recognize somebody that is walking, he may not have enough time to open the application on the mobile phone and then start the face detection.

The interaction process finishes when the user closes the system's application on his mobile phone or presses the wearable device button for 3 seconds.

Based on the scenarios and use cases described, interaction diagrams were constructed to show the process that each possible user goal imply (Figures 4, 5, 6, 7 and 8). Inputs, processes, decisions, displays and outputs are modeled below in order to present an overview of the interaction design of the face-recognition wearable system.

2.1.3 Mobile graphical user interface

The core interface for the user's interaction with the face-recognition system is the mobile's GUI. By using this interface, the user is able to give inputs to, and receive feedback from the system. Figure 9 shows different screenshots of the interface. The functions that it allows, organized as menu buttons or quick accesses are explained below.

Main menu.

The first screen that appears when the application is opened, contains the main menu (Figure 9a). From this screen, the users can select to view and edit their personal contacts (people they already added - Goal 3), search for people (individuals registered in the system, but not added on their contact list - Goal 4), or edit their personal settings (privacy settings and personal information - Goal 5). Quick accesses are also provided on the header (they allow the users to achieve goals 1 and 2).

- My contacts

It shows the contact list, which includes basic information about each person, and the flagging status. The user is given the option to call, send a message or delete individuals (Figure9b). Pictures of the person are only shown if the user has recognized that person before.

[3] More details about functions of the wearable device's button are described in the "System Description" section (2.2)

[4] The mobile application software is described in the section 2.2.1.2; Figure 9 provides screenshots of the graphical user interface.

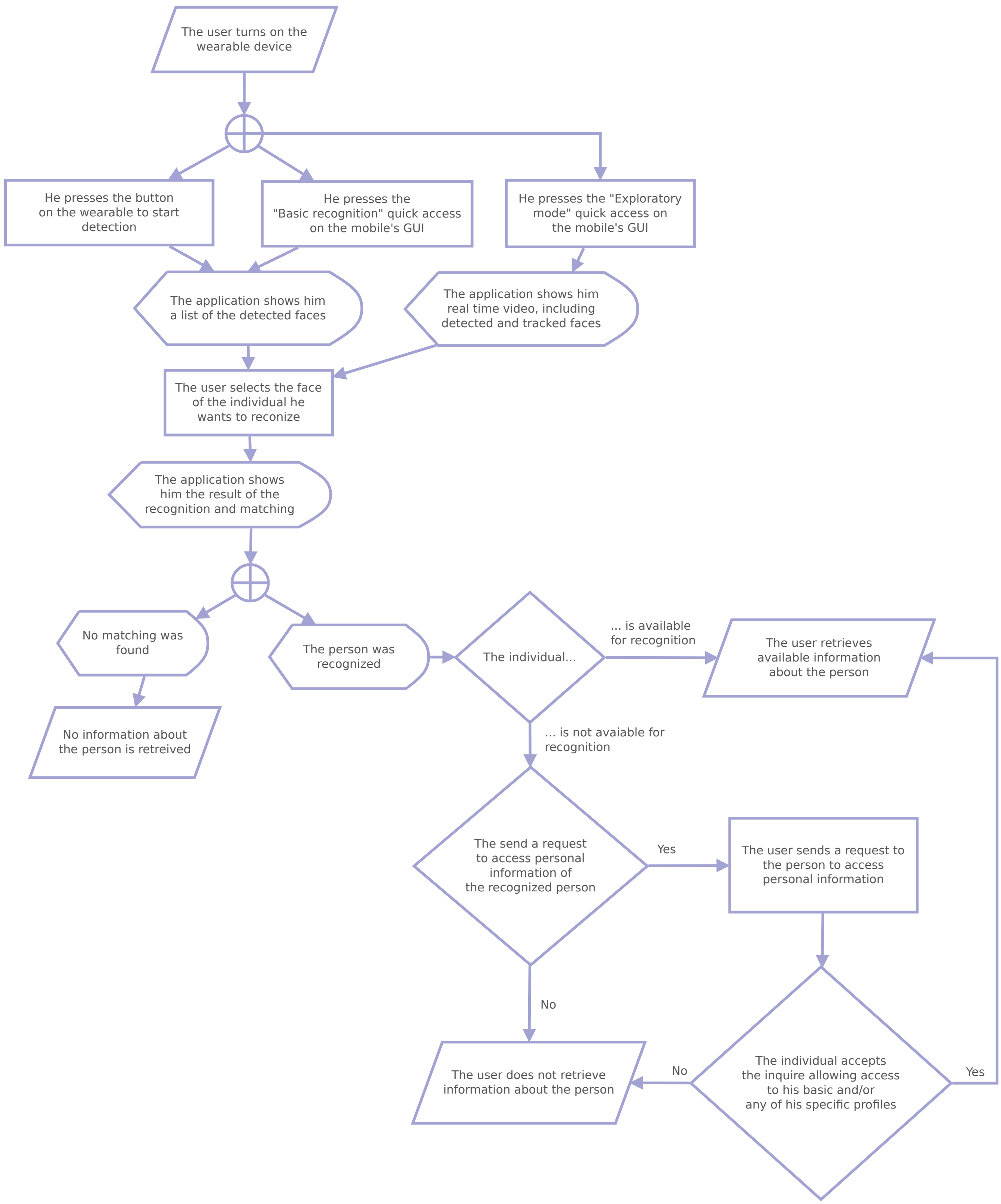

Figure 4. Interaction steps to achieve goal 1

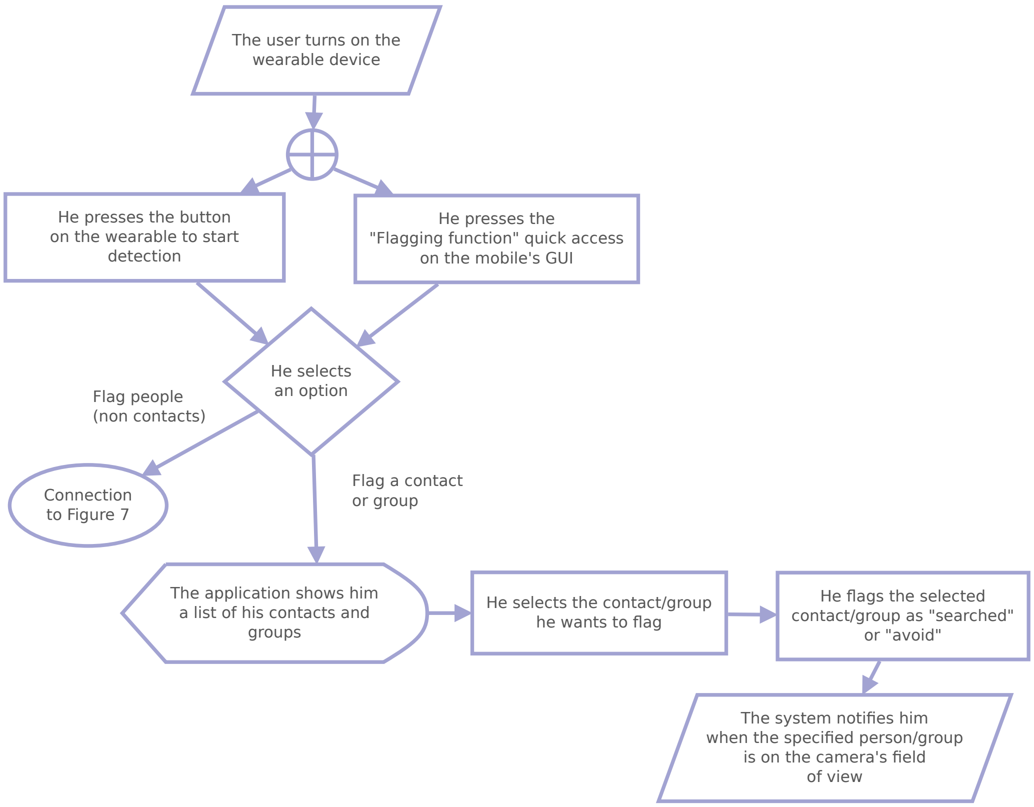

Figure 5. Interaction process for goal 2

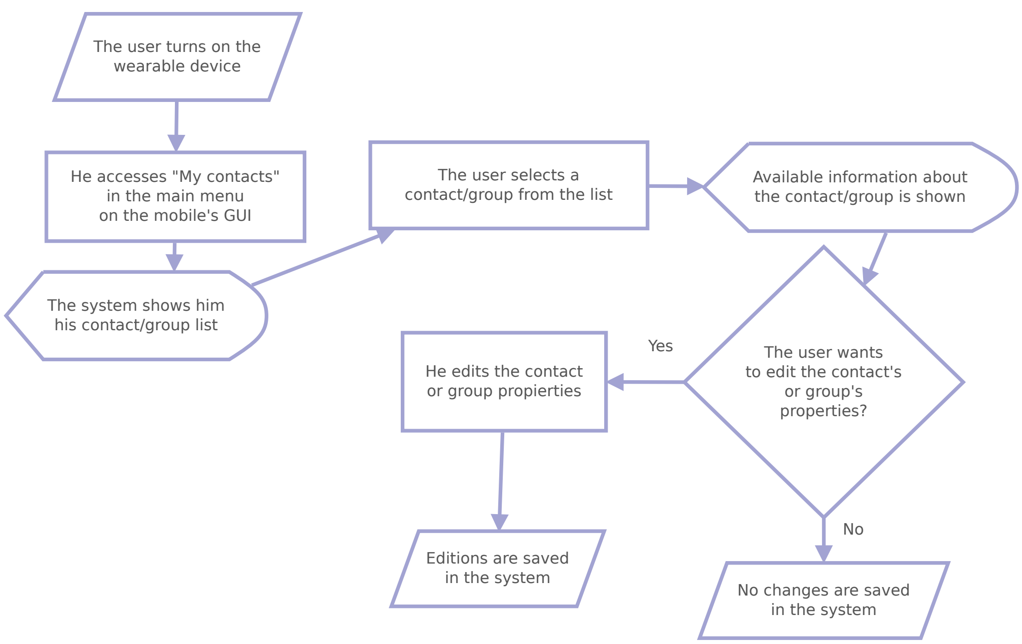

Figure 6. Interaction process to achieve goal 3

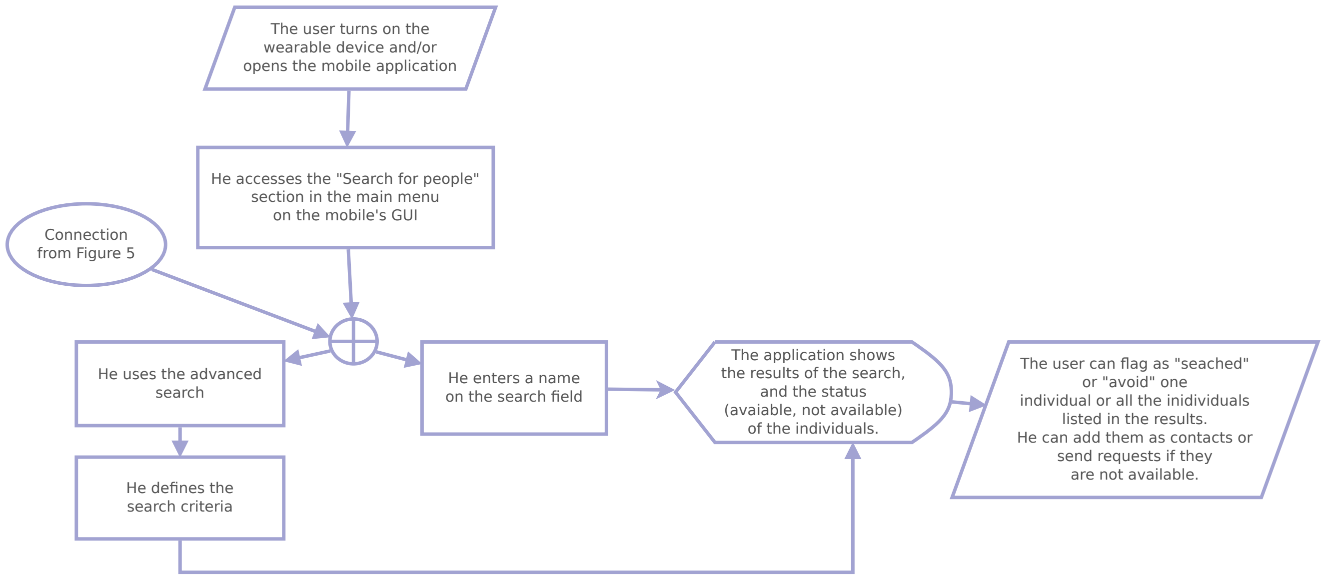

Figure 7. Interaction process to achieve goal 4

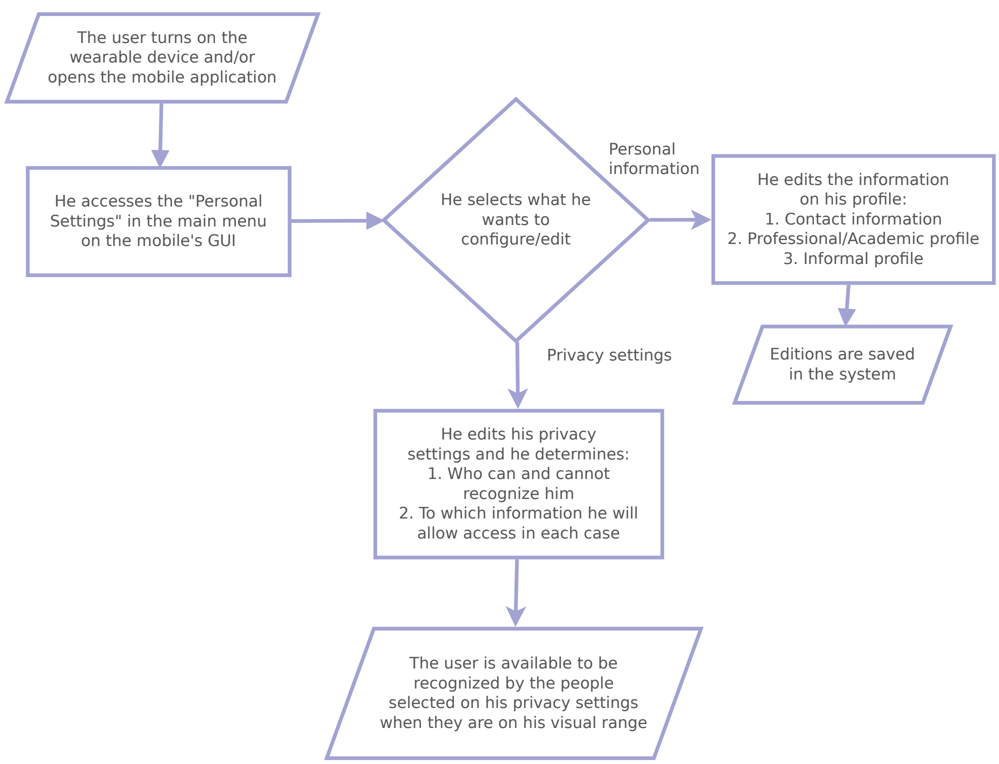

Figure 8. Interaction process to achieve goal 5

The user can retrieve more information about a contact if he selects it from the list (Figure9c). In this case the user can see the information available (depending on the privacy and profiles settings of the person in question). From this screen the user can edit some of the properties of that person: group or groups assigned, type of relationship and comments. Furthermore, it is possible to add or search the contact in social networks.

The contacts are organized in groups. The user can see his group list, and add or delete members from it. Groups have a name and contacts assigned.

- Search for people

The user can search for people that are not among his contacts by using this function. He can either enter a name on a search field or use the advanced search engine. By entering a name the user get all the results available on the global database that matches that name (an example of this can be seen on Figure 9g). When using the advanced search, he can specify the following types of information: age, type of education, work, academic and professional interests, memberships, related organizations, hobbies, activities, leisure interests, sports, ands art and entertainment.

The user can add the people from the list of results as contacts, if they are available, or send a request to add them in case they are not. He can also flag one individual or all the individuals listed.

- Personal settings

Personal information is comprised of: (1) basic and contact data: includes telephone number, e-mail, address, and birthday; (2) academic and professional profile: includes education, work, academic and professional interests, memberships, related organizations; and (3) informal profile: includes hobbies, activities, leisure interests, sports, ands art and entertainment.

When trying to recognize an individual not available for recognition, the user can send a recognition request (Figure 9h). The individual can add the person trying to recognize him as a contact or not, and can also reject or accept the request. In the latter case, this individual will have to define to which types of personal information he will allow access (Figure 9i).

Quick accesses.

The quick accesses are shown on the blue header of the interface, which includes 3 functions, each of them represented by an icon: the basic recognition, the exploratory mode and the flagging function.

- Basic recognition (list view) [5]

[5] The user can select the detected faces on the camera's field of view either by using the basic recognition mode or the exploratory mode. The user is able to change from one view to the other while doing the face detection.

In this mode, the system displays a list of the faces that are being detected at the moment and gives the user the option to select one of these faces to be recognized (Figure 9e). The user can refresh/update the screen to see new faces detected on the camera's visual range. After the recognition is made, the system displays the status of the person recognized. If this person is not available, the user can send a request. In any case the user will be shown a screen similar to Figure 9c, of course that with different information depending on the privacy settings and availability of that person.

- Exploratory mode (radar view)

The user can explore (watching real time video) the faces on the camera's visual range (Figure 9f). Face detection is performed immediately. However, to see both the status (available/not available) and information of the person, the user has to select a face.

- Flagging function

Different flags can be assigned either to contacts or other people on the global database (non contacts). The gray flag indicates that the person is "not flagged"; the red one shows the person was flagged as "avoid"; and the green one indicated that the user is searching for that person (an example of a flagged contact can be seen on Figure 9b). When the system identifies a flagged person, it notifies the user with a short sound and a message on the screen similar to the one shown in Figure 9d.

2.2 System description

2.2.1 System overview

The system consists of 3 main hardware components: a wearable device, a mobile phone, and a server. Each of these components, have a software that allows the user to achieve any of the 5 mentioned goals in the previous section.

In this section processes regarding face detection, recognition and matching are described. Processes related to contacts, personal settings and this kind of editions are not considered because they do not differ from general processes involved in social networks.

Figure 10 shows an overview of the complete system process to achieve the user goal 1 (recognizing an individual), including inputs and feedbacks from/to the user. The diagram shows all the steps involved, including the algorithms used in each case and the types of connections used to tranfer data from the wearable to the mobile phone and between the mobile phone and the server.

2.2.1.1 HARDWARE

The hardware required to make use of this system includes a wearable device, a mobile phone and a server.

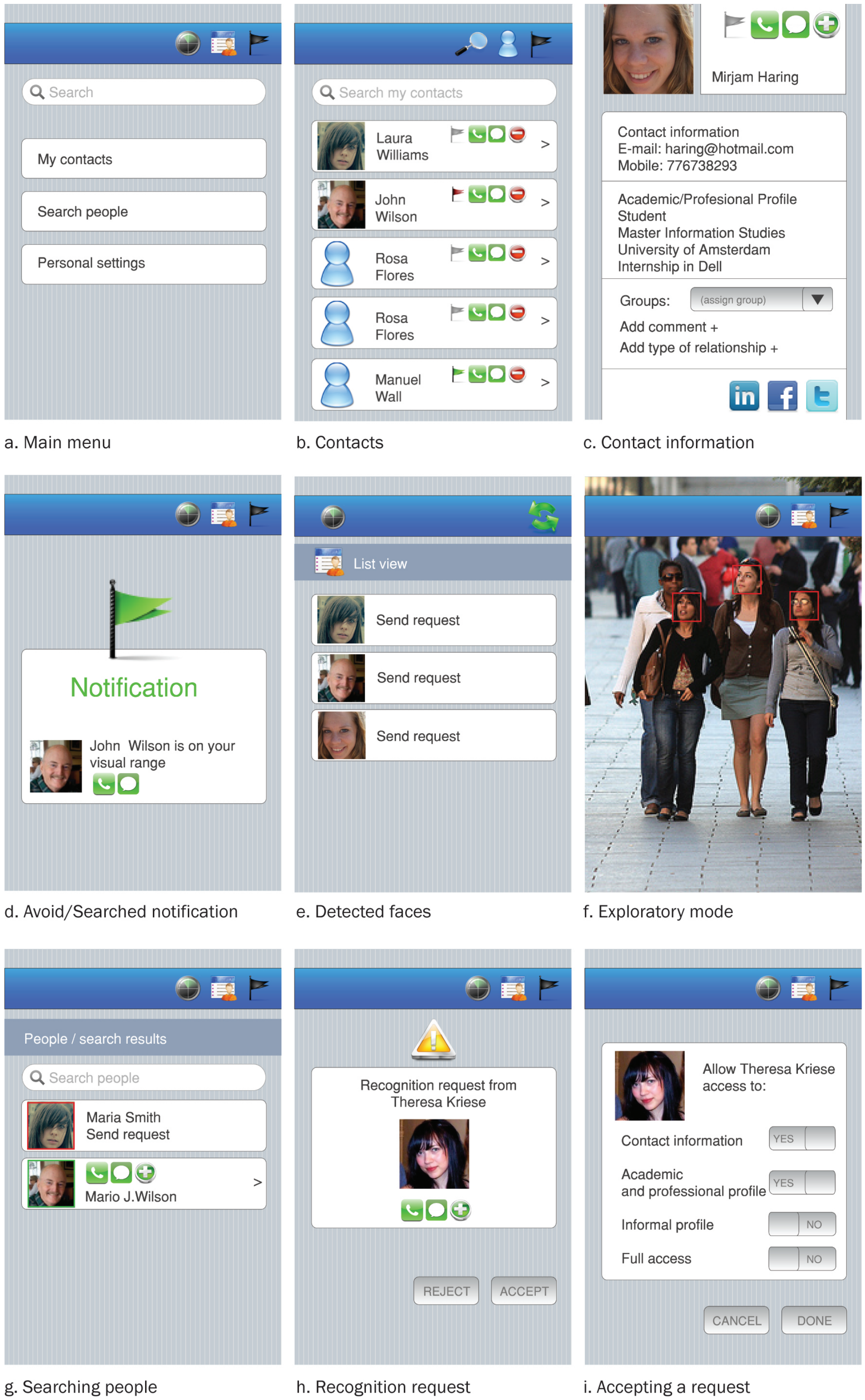

Figure 9. Screenshots of the mobile graphical user interface

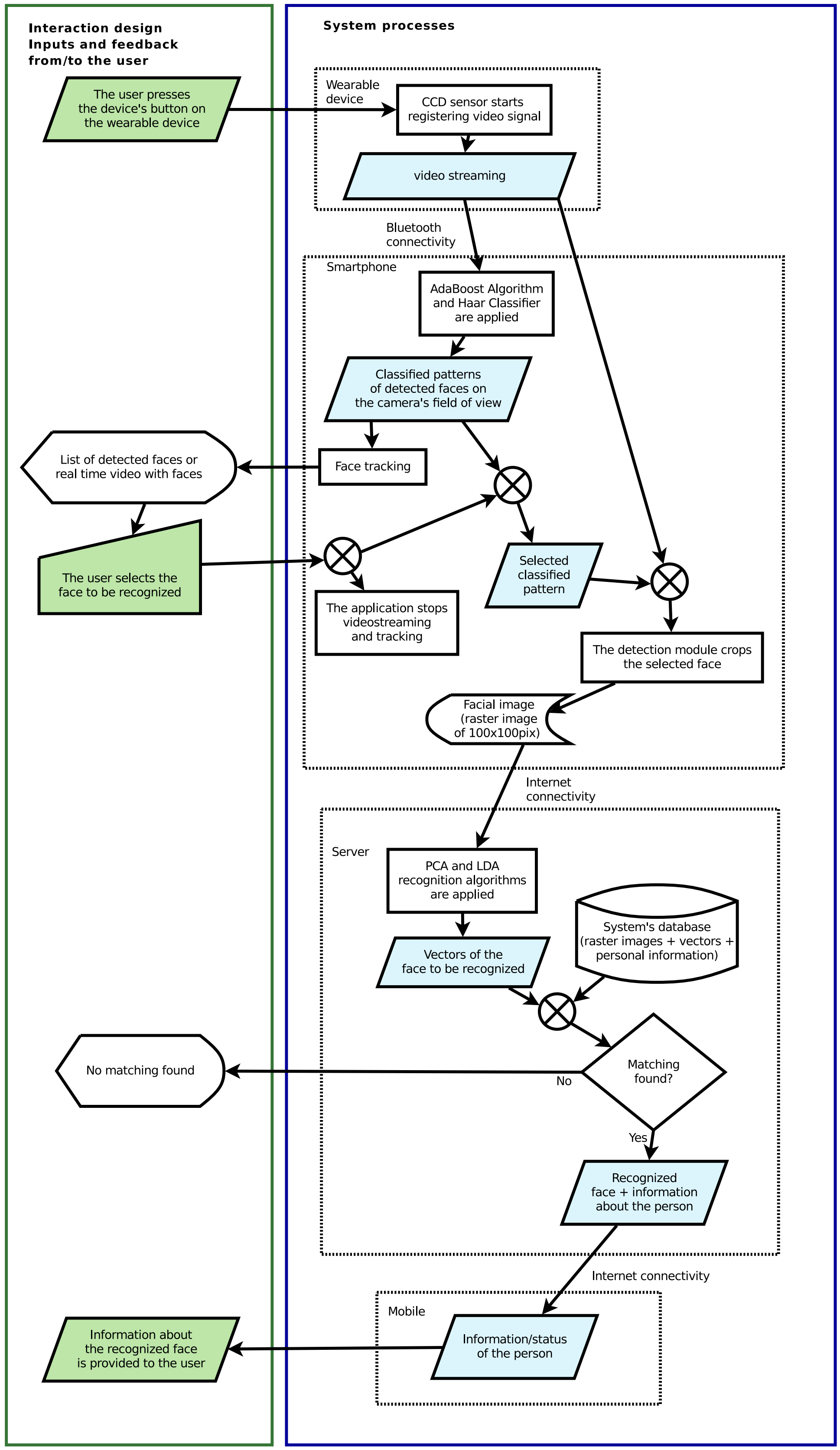

Figure 10. Overview of the system's recognition process

Wearable device

The recognition device is camouflaged under the shape of a fashionable unisex brooch in diverse colors, textures and combinations. The device comprises a tiny camera (only its lens is visible), flash memory, universal storage platform, a button, a 9V battery (Krishna et al., 2005), and an aluminum painted cover. A similar device can be seen in Figure 11 [6].

The wearable device does not have high processing power because all the processes take place in the mobile phone of the user. This is aimed at decreasing its size and cost, taking advantage of the capacities of the mobile.

The button has two functions according to the way it is pressed: (1) turn on/off: when pressing the button for more than 3 seconds, the device turns on/off its bluetooth connectivity. Once the bluetooth is on, it is able to synchronize with its matching mobile. This must be done always before using any of the recognition functions of the system. (2) Start detection: when pressing and instantly releasing the button, the the camera starts recording video which is temporary stored on the flash memory and sent to the mobile by Bluetooth connectivity. This only works if the wearable device was turned-on previously, as explained in (1).

The camera stops recording either when the user selects a face to recognize, when he closes the mobile application or when he turns off the device completely by pressing its button for more than 3 seconds.

The characteristics of the camera are similar to the ones that wearable cameras in the market have (the Looxcie camera is a similar example[7]. Susheel Kumar et al. (2011) used a basic web camera to develop face recognition. The recognition system developed by Kurze & Roselius (2011), uses either the camera of a mobile phone (running Android OS), the camera of a conventional wearable AR display (Nomad by Microvision[8]) or the goggle-based wearable camera. These two cases show that face recognition can be easily achieved using a basic camera. However, this particular system uses a camera with a broader field of view, as in several contexts of use, many faces are detected, including background and other contextual elements. Hence, to provide a good coverage an analog CCD camera, with a light sensitivity of 0.2 Lux, and a 92 degrees field of view is used (Krishna et al., 2005). The position of the camera in the wearable device is fixed, but the brooch can be moved by the user in order to achieve the desired visual range. The 1.3 megapixel camera captures 640x480 quality and color images in RGB (Han et al., 2010). Real-time video is captured in NTSC video format Krishna et al. (2005). It captures 25 fps streaming video (CSR gives the first demonstration of streaming video via a Bluetooth link, able to handle this amount of frames per second[9]).

Mobile phone

This system is expected to make use of any smart phone that fulfill the following minimum requirements:

- Bluetooth connectivity

[6] http://www.takespy.com/blog/Smiling-Face-Brooch-Versatile-Miniature-Pinhole-Camera/
[7] http://looxcie.com/specifications
[8] http://www.microvision.com/
[9] http://www.csr.com/

Key Figure

NO	ICON	Chinese	NO	ICON	Chinese
1		Camera	5		LED Light
2		Recording	6	T-FLASH	Micro SD Card Slot
3		Video Record /Take Photo	7	USB	USB Plug
4		Power	8		

Figure 11. A similar wearable device: a smiling face brooch spy camera commercially available. Source: http://www.takespy.com/blog/Smiling-Face-Brooch-Versatile-Miniature-Pinhole-Camera/

- Wifi, EDGE, 3G and/ or wimax connectivity
- 1.2 GHz processor CPU (Han et al., 2010).
- 512 MB of RAM memory (Krishna et al., 2005).

Server

A dynamic and global database, with all the users registered on the recognition system, is stored on a server. Considering the potential growing of the database, this server would have the following characteristics:

- Fujitsu PRIMERGY RX200 with Intel D2786 Chipset
- 2x Xeon E5520
- 8 GB DDR3 DIMM-RAM
- 2x500 GB SATA II HDDs, 7.200rpm
- 10.000 GB data transfer

2.2.1.2 SOFTWARE

Different software run on each component of the system. The wearable device software, the mobile application and the software in the server.

Wearable device software

This software is basically capable of starting and stopping the video capture, temporary saving the video in the flash memory and streaming the video signal to the mobile device, using the Bluetooth connection. Installers must be run before using the system for the first time, as mentioned in the later section as a prerequisite.

Mobile software

The mobile software is expected to be compatible with the most popular operating systems (Android, RIM OS, Symbian, iOS, Windows Mobile). Details about its interface, can be found on the previous section of this report (item 2.1.3). With regards to the recognition process, the mobile software performs the required face detection and tracking.

- Face tracking

The tracking procedure is based on particle filtering and a strong classifier trained with an AdaBoost algorithm, following what (Fahn et al., 2009) propose. This research group uses a particle filter with a low computational cost, by employing only a fusion of color (skin and hair) and motion cues to reliably measure the likelihood of each particle. The skin and hair color models are adapted to deal with color variation. With this technique, the face tracking rate is more than 97% in general situations and 89% when the face suffering from temporal occlusion.

- Face detection

According to Susheel Kumar et al. (2011), face detection determines the locations and size of a human being face in a digital image. Several methods can be used to detect faces: pattern-classification, controlled background, by color, by motion, or model-based approaches. The mentioned researchers state that detection by color and by motion are not accurate techniques, and that the modeled-based method gives high level of accuracy but only if it is used as a complement of other techniques. It is for this reason that pattern classification and controlled background techniques are used in this system.

With the aim of detecting faces extremely rapidly, face detection is performed in this system by combining the AdaBoost Algorithm with the Haar Classifier (Viola & Jones, 2002). As in the system proposed bySusheel Kumar et al. (2011), AdaBoost is used here to select a small number of critical visual features from a very large set of potential features (Freund et al., 1999), whereas the Haar classifier machine learning algorithm is used to approach for the visual face detection, as it combines classifiers in a "cascade" that allows background regions of the image to be quickly discarded while spending more computation on promising face-like regions (Viola & Jones, 2002).

This system captures video, allowing the selection of good angle image of the face by selecting from the frame sequence an individual calibrated image (Han et al., 2010).

Apart from detecting the faces in the captured image, the mobile software also crops the faces when the user selects the face of the person he is willing to recognize. This is aim at reducing the weight of file transfers to the server, where the face recognition process takes place. As regards resolution, Han et al. (2010) develop face recognition with images of 128 pixels wide and 128 pixels high. However, in this system facial information is saved in a 100 x 100 matrix size, which is feasible as shown by the system presented in Susheel Kumar et al. (2011). Each matrix have three layers, corresponding to red, green, and blue color in the image. Only this facial information in JPEG format is then used to carry out the recognition and matching process.

Server software

The role of this software is to allow the database administration and the running of the face recognition and matching process, as well as receiving/sending information from/to the mobiles using Internet connectivity.

- Database administration

The server stores a dynamic database that grows continuously either when a new user create an account or when a person is recognized (in the latter case an additional pattern of his/her face is added). The database includes at least 5 pictures of each person and personal information and profiles. 5 images of each individual in different position and with different emotions should be uploaded by the user when creating an account on the system.

- Face recognition and matching module

The face recognition and matching algorithms are run on the server, when the user selects one of the detected faces (either in basic recognition or in exploratory mode) to recognize. The face picture of the selected person is sent to the server through wireless Internet to perform the recognition. Details and reasons about the techniques selected to achieve the recognition process are described below.

Susheel Kumar et al. (2011) present different face-recognition techniques for real-time recognition: Neural Networks (trained on a large set of faces), Geometrical Feature Matching (vector representing main features of face), Graph Matching (employs elastic graph matching to find the closest stored graph), Eigenfaces (uses Principal component analysis -PCA- to construct eigenvectors of faces), Fisherface (PCA and Fisher's Linear Discriminant Analysis -LDA-), and Feature Extraction Techniques (Geometry-based, Template-based, and Color segmentation techniques).

Considering Susheel Kumar et al. (2011) analysis, the following methods are not adopted for this system: Neural Networks approach, because it has problems when the number of classes, i.e., individuals increases and the system has to be able to deal with this cases; Geometrical Feature Matching, because it is based on precisely measured distances between features and it is particularly useful for finding possible matches in a large database (which our system won't have in the begging); Graph Matching, because the matching process is complex and computationally expensive. On the other hand, the following methods are interesting for this system: Eigenfeatures (based on Eigenfaces), because it can deals with lighting variations; and Fisherface, because it is one of the most successful widely used methods, it takes advantage of within-class information minimizing variation within each class, yet maximizing class separation, and it can overcome problem with variations in the same images, such as lightening (Susheel Kumar et al., 2011).

Facial feature extraction is necessary for identification of an individual face. As facial features, the shape of facial parts is automatically extracted from a frontal face image. There are three main techniques for feature extraction: Geometry-based, Template-based, and Color segmentation techniques. The following techniques are discarded for this system: the geometry-based method because it requires threshold, which, given the prevailing sensitivity, may adversely affect the achieved performance; the color segmentation approach because there is no algorithm in this sense that can work in real time face recognition considering limitations such as illumination and pose (Susheel Kumar et al., 2011). The proposed face-recognition system uses the template-based approach.

Considering all this, the face recognition process in the proposed system is performed using a combination of Eigenfunctions and Fisherface methods, which means, combining both PCA and LDA techniques. The assessment provided by Krishna et al. (2005), also supports this decision. This research group states that the Hidden Markov Model (HMM) algorithm is the poorest performing algorithm. On the other hand, performances of PCA and LDA are considered useful. Though LDA is twice as fast as PCA, (Krishna et al., 2005) opted for PCA to be the face recognition algorithm on the wearable device they designed, due to its higher recognition rate.

Once PCA and LCD are applied to the picture of the face in question, the system carries out the recognition by searching a matching face in the database. The database, configured considering the work of Huang et al. (2007), is particularly designed to be capable of dealing with variation in pose, lighting, expression, background, race, ethnicity, age, gender, clothing, hairstyles, camera quality, color saturation, focus, and other parameters.

3. Discussion

According to current scientific knowledge and techniques analyzed in this report, the design of a face recognition intelligent interactive system is nowadays feasible. The provided descriptions show that it is theoretically possible to build an efficient system of this kind. The problems faced in this work, the solutions that were reached and the identified challenges for future work are discussed.

3.1 Interaction design Building the interaction design for the system, implies the identification of user needs and goals (Sharp et al., 2007). Even tough the interaction design was initially based on the main user goals, particularly related to recognize somebody on the camera's field of view, additional tools were introduced later both to improve the user experience and to meet secondary needs that emerged when designing the basic recognition functions of the system. Hence, the system was enriched to allow the users to be notified about an individual's presence, to manage information about their contacts and strangers, and to settle privacy aspects and information about themselves.

The introduction of these additional services led the complexity of the system's interaction process to a higher level. This represented a challenge in terms of organization and clarity, that was solved by analyzing the interactions involved by goals. This goal-oriented approach, provides a clear structure of each part of the system, at the same time that it allows a panoramic view of the user's interaction.

On the other hand, it is worth mentioning that this work focused on identifying, understanding, and examining user experience issues to ensure an appropriate goal-oriented interaction design. Interface development, and cognitive and affective aspects of the interaction (Sharp et al., 2007), could be deepened in future analysis. The aesthetic characteristics and the usability of the system have not been analyzed in this work. Future analysis should evaluate if the user's expectations are met, if the feedback from the interface is clear and instructive, and if the system elicits positive responses from the users, such as feeling at ease, being comfortable, and enjoying the experience of using it (Sharp et al., 2007).

3.2 Hardware and software This reports states the theoretical possibility to build and develop the required hardware and software to obtain an efficient face-recognition system. However, some of the decisions made during the design process still represent interesting areas of analysis for future work.

For instance, an alternative option could be to run real time face recognition for all the detected faces, instead of doing it only with one face selected by the user. This would have resulted on a more interesting augmented reality (AR) tool, particularly in the exploratory mode on which the user could retrieve real-time information about the faces on the camera's field of view. However, this was not implemented on this system as it could imply considerable longer time lags and higher processing requirements.

Another issue that deserves discussion is the use of an external wearable device, instead of taking advantage of the mobile phone camera. Whereas the wearable device allows the user to discreetly recognize people in many cases, it implies acquiring and carrying an additional device. However, this system uses an external camera device because it is ergonomically convenient: at the same time that images are captured by the camera, the users can see the detected faces on their mobile phones and interact with the application. Including a wearable device avoids the users to hold uncomfortable positions when trying to reach the face to be recognized and at the same time reading and interacting with the application on the mobile screen.

Exploring detection and recognition with less resolution is another challenge for future work. Solving resolution and processing difficulties of video streaming would allow recognition in larger distances. The limit imposed by this system (approx. 5 meters) may not be always convenient for the users, particularly when they want to recognize discreetly.

Furthermore, battery power limitation deserve future analysis and solutions. In this system video streaming and face detection are not continuous, which forces the user to initiate the recognition process in specific moments, but that could save battery power. If battery capacities were augmented, the system could provide larger amounts and different kinds of information, for example allowing the users to keep the recognition function always on.

Image quality problems due to the constantly and arbitrarily movement of the wearable camera in this system is another challenge for future work. Different techniques to select the best frame to perform the recognition and matching (Han et al., 2010) could be analyzed in eventual research as well.

As it is stated by the analyzed research, real time face recognition is feasible in spite of the difficulties that face images can have (Susheel Kumar et al., 2011; Krishna et al., 2005; Han et al., 2010). Whereas it is true that head pose, illumination problems, hair style, occlusion (i.e. glass, scarf) and aging can reduce the accuracy and performance of the system, the building of an efficient database, specially designed to achieve face recognition in "natural" variability in pose, lighting, focus, resolution, facial expression, age, gender, race, accessories, make-up, occlusions, background, and photographic quality can deal with these problems (Huang et al., 2007).

The system should be prototyped and evaluated. Multi-recognition processes taking place at the same time is another challenge to face.

3.3 Social implications To be successful, the proposed face-recognition system requires massive use. This ensures a populated database, allowing users to retrieve any information about the recognized individual (at least about availability status). This means that to be useful, the system should become popular. Marketing strategies should be designed to ensure this.

The proposed system would make people become portals with digital identity and awareness, offering new ways of communication and new possibilities for interaction

(Kostakos et al., 2005). As any pervasive technology, a system of this kind would not only extend existing social conventions, but it will also enable new ways of acting and interacting, and will stimulate a reassessment of what constitutes human action and interaction. In some cases, social actions will occur in entirely new ways, and in other cases, completely new social actions will appear. What will those new social actions be is not known yet (Kostakos et al., 2005).

On the other hand, privacy issues deserve relevant attention on a system of this kind. If people's faces, names and other personal information are stored on a database, people may be recognized whenever they walk the streets. The polemic discussion awakens different perspectives and reactions as people could not be "anonymous no more" ("Face recognition: Anonymous no more", 2011) in any case. This means that such a system could be either used for good purposes or for negative aims. Beyond the different uses and opinions, face-recognition methods may still spread and open the "floodgates" to a new, privacy-sapping world ("Face recognition: Anonymous no more", 2011). Avoiding this process would definitely be hard, so probably social adaptations and new privacy visions will take place. Besides, new challenges for designers will arise. Would it be possible to design useful and efficient augmented reality tools that reasonably protect privacy? (Friedman & Nissenbaum, 1997)

Furthermore, emerging technologies of this kind touch upon other important social issues, apart from the often-cited issue of privacy (Kostakos et al., 2005). At the same time AR is complementary to human cognitive processes and its potential appears promising and worthwhile (Drascic & Milgram, 1996), several issues must be taken into account. Human well- being, human dignity, justice, welfare, human rights, autonomy, access, accountability, freedom from bias and trust are problems that have to be faced in systems of this kind (Friedman & Nissenbaum, 1997).

3.4 Future works Develop prototypes and test the system on a real context is a pending task. Exploring the way it works on different real life situations, would probably introduce new requirements and needs.

Besides, the extension of the system's possibilities in terms of social interactions and network tools would be interesting. Adding resources similar to the ones applied on social networks, such as statistics on recognitions performed on each user, list of places on which the person was recognized by others, among other, would improve the proposed system.

Adding this face-recognition system to other augmented reality tools that are already in the market would be an interesting application. This would allow the users to augment their perceptions of their realities by including among other environmental elements, peoples identities. For instance, faces and familiarity -as defined by Goodman & Paulos (2004)- of visited places, geotags, time stamps and instant chat functions could be introduced.

Finally, special applications of this system would be interesting to provide tools for visually impaired individuals, prosopagnosic persons and people with memory disorders.

4. Conclusion

Insight in the principles for the design of a face recognition intelligent interactive wearable system were provided in this report. Relevant developments and applications in this field were analyzed in order to select the most efficient and suitable methods and techniques for the design of a usable and efficient face recognition system.

Interaction processes were modeled to allow the users to achieve the different identified goals. By interacting with their wearable devices and mobile phones, the users could recognize people, explore faces on the camera's field of view, be notified about somebody's presence, view and edit information about contacts, search for people on the database and edit their personal information and privacy settings.

Efficient potential hardware and software configurations were described. Characteristics of a possible wearable device were depicted. The device would work in combination with any smart phone accomplishing particular requirements. AdaBoost Algorithm and Haar Classifier were selected to perform face detection. Particle filtering and a strong classifier trained with an AdaBoost algorithm were chosen for face tracking, and PCA combined with LDA were selected to perform the face recognition process.

Social implications about the face-recognition system were also discussed. Privacy, well- being, human dignity, justice, welfare, human rights, autonomy, access, accountability, freedom from bias and trust are important issues concerned with a system of this kind. In this sense, several research questions remain unanswered, but it seems that technology is likely to continue influencing our lives in ways that are not known yet.

Areas for future works were analyzed as well, both to improve the proposed system and to adapt it to particular users, such as visually impaired individuals, prosopagnosic persons and people with memory disorders. Furthermore, improvements to lead the system towards a more interesting augmented reality tool were also considered.

References

Baldwin, M. (1992). Relational schemas and the processing of social information. *Psychological Bulletin*, *112*(3), 461.

Ballard, D., & Brown, C. (1982). *Computer vision*. Prentice Hall Professional Technical Reference.

Brazilian police to use 'Robocop-style' glasses at world cup - telegraph. (2011, April). *The Telegraph*. Available from `http://www.telegraph.co.uk/news/worldnews/southamerica/brazil/8446088/Brazilian-police-to-use-Robo`

Cevikalp, H., Neamtu, M., Wilkes, M., & Barkana, A. (2005). Discriminative common vectors for face recognition. *IEEE transactions on pattern analysis and machine intelligence*, 4–13.

Crowley, J. (2006). Things that see: Context-aware multi-modal interaction. *Cognitive Vision Systems*, 183–198.

Drascic, D., & Milgram, P. (1996). Perceptual issues in augmented reality. In *PROCEEDINGS-SPIE THE INTERNATIONAL SOCIETY FOR OPTICAL ENGINEERING* (p. 123–134).

Face recognition: Anonymous no more. (2011, July). *The Economist*. Available from `http://www.economist.com/node/21524829`

Fahn, C., Kuo, M., & Wang, K. (2009). Real-time face tracking and recognition based on particle filtering and AdaBoosting techniques. *Human-Computer Interaction. Novel Interaction Methods and Techniques*, 198–207.

Freund, Y., Schapire, R., & Abe, N. (1999). A short introduction to boosting. *JOURNAL-JAPANESE SOCIETY FOR ARTIFICIAL INTELLIGENCE*, *14*, 771–780.

Friedman, B., & Nissenbaum, H. (1997). Software agents and user autonomy. In *Proceedings of the first international conference on autonomous agents* (p. 466–469).

Goodman, E., & Paulos, E. (2004). The familiar stranger: anxiety, comfort, and play in public places.

Han, L., Li, Z. S., Zhang, H., & Chen, D. Y. (2010). Wearable observation supporting system for face identification based on wearable camera. , *7*, 91–95.

Huang, G., Ramesh, M., Berg, T., & Learned-Miller, E. (2007). Labeled faces in the wild: A database for studying face recognition in unconstrained environments. *University of Massachusetts, Amherst, Technical Report 07*, *49*, 1.

Jain, A., Flynn, P., & Ross, A. (2008). *Handbook of biometrics.* Springer-Verlag New York Inc.

Jain, A., Ross, A., & Prabhakar, S. (2004). An introduction to biometric recognition. *Circuits and Systems for Video Technology, IEEE Transactions on*, *14*(1), 4–20.

Jung, Y., Blom, J., & Persson, P. (2006). Scent field trial: understanding emerging social interaction. In *MobileHCI '06: Proceedings of the 8th conference on human-computer interaction with mobile devices and services* (pp. 69–76). ACM Press. Available from `http://dx.doi.org/10.1145/1152215.1152230`

Kostakos, V., Oneill, E., Little, L., & Sillence, E. (2005, September). The social implications of emerging technologies. *Interacting with Computers*, *17*, 475–483. Available from `http://dl.acm.org/citation.cfm?id=1712150`

Kostakos, V., & Venkatanathan, J. (2010, August). Making friends in life and online: Equivalence, Micro-Correlation and value in spatial and transpatial social networks. In (pp. 587–594). IEEE. Available from `http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5591361`

Krishna, S., Little, G., Black, J., & Panchanathan, S. (2005). A wearable face recognition system for individuals with visual impairments. In *Proceedings of the 7th international ACM SIGACCESS conference on computers and accessibility* (p. 106–113).

Kurze, M., & Roselius, A. (2011). Smart glasses linking real live and social network's contacts by face recognition. In *Proceedings of the 2nd augmented human international conference* (p. 31).

Mukherjee, S., Chen, Z., Gangopadhyay, A., & Russell, S. (2008). A secure face recognition system for mobile-devices without the need of decryption. In *Workshop on secure knowledge management.*

Nuevo android 4.0 lleva el reconocimiento facial al móvil / ELESPECTADOR.COM. (2011). http://elespectador.com/tecnologia/articulo-306311-nuevo-android-40-lleva-el-reconocimiento-facial-al-movil. Available from `http://elespectador.com/tecnologia/articulo-306311-nuevo-android-40-lleva-el-reconocimiento-facial`

Sharp, H., Rogers, Y., & Preece, J. (2007). Interaction design: beyond human-computer interaction. *Book*.

Singletary, B. A., & Starner, T. (2001). Symbiotic interfaces for wearable face recognition. In *HCII2001 workshop on wearable computing, new orleans, LA.*

Susheel Kumar, K., Prasad, S., Bhaskar Semwal, V., & Tripathi, R. C. (2011, July). Real time face recognition using AdaBoost improved fast PCA algorithm. *International Journal of Artificial Intelligence & Applications*, *2*, 45–58. Available from http://www.airccse.org/journal/ijaia/papers/0711ijaia05.pdf

Szeliski, R. (2010). *Computer vision: Algorithms and applications*. Springer-Verlag New York Inc.

Turk, M., & Pentland, A. (1991). Eigenfaces for recognition. *Journal of Cognitive Neurosicence*, *3*, 71–86.

Viola, P., & Jones, M. (2002). Robust real-time object detection. *International Journal of Computer Vision*, *57*(2), 137–154.

Wayman, J., Jain, A., Maltoni, D., & Maio, D. (2010). *Biometric systems: Technology, design and performance evaluation*. Springer Publishing Company, Incorporated.