



HAL
open science

Fuzzy logic and optical correlation-based face recognition method for patient monitoring application in home video surveillance

Marwa El Bouz, Ayman Alfalou, C. Brosseau

► **To cite this version:**

Marwa El Bouz, Ayman Alfalou, C. Brosseau. Fuzzy logic and optical correlation-based face recognition method for patient monitoring application in home video surveillance. *Optical Engineering*, 2011, 50 (6), pp.067003. 10.1117/1.3582861 . hal-00607797

HAL Id: hal-00607797

<https://hal.science/hal-00607797v1>

Submitted on 11 Jul 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Optical Engineering

SPIDigitalLibrary.org/oe

Fuzzy logic and optical correlation-based face recognition method for patient monitoring application in home video surveillance

Marwa Elbouz
Alfalou Ayman
Christian Brosseau

Fuzzy logic and optical correlation-based face recognition method for patient monitoring application in home video surveillance

Marwa Elbouz
Alfalou Ayman

ISEN Brest (Institut Supérieur, de l'Electronique et du Numérique)
Département Optoélectronique
L@bISEN
20 rue Cuirassé Bretagne, CS 42807
Brest Cedex 2, Finistere 29228 France
E-mail: ayman.al-falou@isen.fr

Christian Brosseau

Université Européenne de Bretagne
Université de Brest
Lab-STICC
CS 93837, 6 avenue Le Gorgeu
Brest Cedex 3, 29238 France

Abstract. Home automation is being implemented into more and more domiciles of the elderly and disabled in order to maintain their independence and safety. For that purpose, we propose and validate a surveillance video system, which detects various posture-based events. One of the novel points of this system is to use adapted Vander-Lugt correlator (VLC) and joint-transfer correlator (JTC) techniques to make decisions on the identity of a patient and his three-dimensional (3-D) positions in order to overcome the problem of crowd environment. We propose a fuzzy logic technique to get decisions on the subject's behavior. Our system is focused on the goals of accuracy, convenience, and cost, which in addition does not require any devices attached to the subject. The system permits one to study and model subject responses to behavioral change intervention because several levels of alarm can be incorporated according different situations considered. Our algorithm performs a fast 3-D recovery of the subject's head position by locating eyes within the face image and involves a model-based prediction and optical correlation techniques to guide the tracking procedure. The object detection is based on (hue, saturation, value) color space. The system also involves an adapted fuzzy logic control algorithm to make a decision based on information given to the system. Furthermore, the principles described here are applicable to a very wide range of situations and robust enough to be implementable in ongoing experiments. © 2011 Society of Photo-Optical Instrumentation Engineers (SPIE). [DOI: 10.1117/1.3582861]

Subject terms: Vander-Lugt correlators; joint transform correlators; segmented correlation filters; surveillance video systems; fuzzy logic; hue, saturation, and value color space.

Paper 100964R received Nov. 22, 2010; revised manuscript received Mar. 16, 2011; accepted for publication Apr. 6, 2011; published online Jun. 1, 2011.

1 Introduction and Scope of Research

Modern science and technology continue to augment the average age and the quality of life. Because of the aging population of the United States and European countries, measures must be taken in order to provide health care for the elderly. The U.S. Census Bureau has projected that, by 2010, 13% of the population will be 65 or older.¹ The bureau has also projected that, by 2030, there will be nine million Americans older than 85.¹ This spells good news for everyone, especially seniors. Innovations continue to deliver better medicine, nutrition, and mental health care for seniors. Ultimately, this provides the modern senior with the opportunity to lead a robust and independent lifestyle. For a number of senior citizens, independence entails living on their own. More seniors continue to do a solid job of caring for themselves as well as their property. However, independent senior living poses inherent risks. Among several concerns, one of the most important is falling. Falls in the elderly are a problem for two main reasons: they are more likely to happen than falls in younger patients, and they are more likely to result in serious injury (e.g., two common fractures as a result of falls in the elderly are wrist fractures and fractures of the femur). Thirty percent of those over 65 fall annually. Half are repeat fallers.

Over half of those in nursing homes and hospitals will fall each year. Recurrent falls are defined as those occurring at least three times a year. Comorbidity is a serious problem both in terms of contributing to the cause of the fall and the outcome. This is one reason why mortality three months after a fall is so high. Falls can be devastating to the affected individual but are also expensive to manage. Especially when associated with fracture of the proximal femur, they carry a high morbidity and mortality. Even lesser falls lead to loss of self-confidence and reduced quality of life. This can also have significant economic consequences because of the cost of inpatient care but also loss of independence and the cost of residential care. The prevention of falls poses a challenge to caregivers and health care, in general. The main issue is to reduce the number of falls that result in serious injury and ensure effective treatment and rehabilitation for those who have fallen.

The overwhelming majority of seniors wish to continue to live in their own homes for as long as possible. However, many homes are not well designed to meet our changing needs as we age. Recent advances in the Internet protocol (IP)-based video surveillance offer new opportunities for security tools in computerized (smart) homes. A smart home or building is usually a new one that is equipped with special structured wiring to enable occupants to remotely control or program an array of automated home electronic devices by

entering a single command. While first device applications are being explored (e.g., automatic or semiautomatic control of lighting, doors, and windows; heating, ventilation, and air conditioning; and security and surveillance systems), there is growing awareness of the wealth of issues that need to be addressed to arrive at elderly monitoring applications in a home-surveillance scenario (e.g., a video system of network cameras to detect various posture-based events). These events include normal daily life activities and unusual events (e.g., fall detection). This calls for simple methods in which the development of assistive domotics makes it possible for the elderly and disabled to live by themselves and meet their potential.

The reader may wish to consult two reports,^{2,3} containing many references, for general discussions of such video systems. Here, we will briefly review various aspects of the problem that relate to our purpose. Chan et al.² proposed a general overview of video surveillance systems of the elderly and disabled in order to maintain their independence at home. Other examples of automated methods of assistive domotics are discussed by Nugent et al.³ Markovian models were introduced recently⁴ that can be used to detect occurrences of mobile or static states of a subject in a given video sequence. The analysis of such models was carried out by different types of sensors (e.g., motion and infrared sensors). However, the approach developed in Ref. 4 cannot identify the subject. In addition, the models described in Ref. 4, though adequate to describe the spatial positions of a specific subject, cannot analyze scenarios in which several persons are present in the same location. Another type of model was suggested by Zouba and coworkers.⁵ This model is based on the merging of information issued from different sensors. Thanks to these sensors, various posture-based events (using a reference database) can be recognized and motion of a subject can be detected. However, several basic issues that were not considered in that work (e.g., a change of clothes by the subject, and the presence of several persons in the scene), introduce limitations in the application of this model.

The present work investigates a system that is based on the use of a video sequence and a specific method of merging of information issued from a set of sensors. By this we mean that the complicated problem arising when many persons are present in the 3-D scene considered is circumvented by the recognition-and-tracking algorithm of a given subject following an identification step. What distinguishes our work from that of previous investigators is mainly that our chosen technique, optical correlation, allows us to make (high-level) decisions concerning behavioral change of the tracked subject.

We now briefly outline the experimental issues and our proposed strategy. In this paper, we report on the principle of a potentially promising approach for a machine-vision-based position-sensing system for detecting various posture-based events that we have conceived and implemented. Our intent in this paper is only to validate the principle of our approach using a PC (CPU Intel Dual Core E5300, 2.6 GHz, 2Go RAM). Having made the foregoing statement, however, we must point out to the reader the fact that only six images per second were considered in this study is a limitation of our system. The prospect is now raised that the numerical implementation of this algorithm on a GPU or FPGA graphics card might decrease significantly the computation time. There are

many critical issues that we want to consider: rapid detection and reliability of an anomaly, the possibility of having several levels of alarm for a given situation and nonintrusivity of the method (i.e., no need for a small wireless pendant transceiver to be worn around the neck or wrist). In addition, when falling happens, the alert goes out and a video sequence showing the scene can be sent to a medical response team. Image compression of the video sequence is mandatory to reduce its size. To preserve privacy, encryption can be also realized in order that only authorized persons can visualize the scene. Our system has focused on the goals of accuracy, convenience, and cost, which in addition does not require any devices attached to the subject. To deal with the above-mentioned issues, our system involves digital video cameras positioned in a room and processing the video stream using a specialized algorithm running on a PC. The system uses commercial off-the-shelf webcams, which are widely available and cost effective. We combine several optical correlation techniques (i.e., the VLC^{6,7} and JTC.⁷⁻⁹ More specifically, the VLC is basically used for recognition and identification of a person entering into a room, whereas the JTC is selected to follow that person in the scene. This advantage of our technique permits one to circumvent the problem of confusion, which may be observed in the subject's behavior. Our formulation of the problem is also advantageous when several persons are in the same scene, because correlation methods are adequate to track and follow a given subject in a complex environment. Face recognition is easily achieved with a camera; however, it can be affected by several potential problems, such as ambient illumination conditions and facial expression.¹⁰ We start by reducing the face area in the target scene to get a small-size image with very little background noise. This is done by segmenting the scene using a subject's skin detection method based on the hue, saturation, and value (HSV) color space. Then, this image is introduced into our correlator using an optimized, segmented composite filter^{11,12} specially designed for our application [i.e., a correlation filter combining several images of several possible references representing the image to be recognize (rotation, scale, etc.)]. Our algorithm will deal with multiple information (subject presence, identity, position in the room, local time) originating from different sensors; these data are called low-level decisions. Next, the data are analyzed and merged together using an adapted version of fuzzy logic¹³⁻¹⁵ in order to produce a trigger signal whenever the head posture shows an anomaly; this operation is called a high-level decision. Note that to validate the principle of our approach, we considered only three alarm levels: low, medium, and high. Our system keeps track of head positions, recognizing certain postures and detecting trigger events. When these triggers are detected, the system delivers alerts or security alarm notification signals. The authors emphasize that, to the best of their knowledge, the implementation of such a machine-vision-based position-sensing system for monitoring of the elderly as an application in home surveillance has never been previously documented.

The remainder of this paper is divided into three sections. In Sec. 2, an overview of the proposed algorithm to identify and track a subject moving in a given environment is detailed. Next, we present the theoretical basis of 3-D information recovery from image sequences obtained by a single camera. The succeeding section explains how the tracking is performed based on prediction, and how based on the

analysis, modeling, and data fusion, the system delivers alerts according several options (phone call to a relative or call for a medical response team by sending compressed and encrypted image sequences together with the subject's medical file). After explaining the theory behind our method, we present and discuss, in Sec. 5, an experimental study done under realistic conditions. Finally, our conclusions are summarized in Sec. 6.

2 Overview of the Algorithm

The task of the block diagram in Fig. 1 is to provide an overview of the system structure and processing procedure described in this research. The first step consists of detecting the subject at the entrance of a room. One effective way to realize this operation is by a combination of infrared sensors leading to captured video, typically a sequence of ten images. The primary function of the video or image sequence is to identify the subject using a VLC fabricated with an optimized, segmented composite filter.^{7,12} Once the subject is identified, the tracking procedure is implemented using a fringe-adjusted JTC (FA-JTC),⁹ which has been adapted for the application under study. With this machine-vision-based approach, the subject's head is detected. On the one hand, this restricts the area to be studied. On the other hand, the noise effects are minimized and the system performance is enhanced. Then, the energy of the head's image is normalized, even if this step is not mandatory. Indeed, we add references (with several ambient illumination conditions) in the references database used to design our correlation filter. Next, the subject's head is matched against an off-line database pertinent to surveillance using correlation techniques. If the subject's head is not a good match, then the system can either add the input image to the database or consider it an anomaly, depending on the intensity of the correlation peak. Identification and tracking procedures are considered at the same time. Once the subject's identification is established, the subject's identity is sent to the tracking system. If the identification step is negative, then an alert is also sent to the controller. However, because of the possible change in viewpoint, 2-D information is not enough for distinguishing 3-D postures (i.e., different head postures can have the same projection on the image plane when viewed from different perspectives and vice versa). In order to discriminate ambiguous projections, it is necessary to obtain depth information. There are various approaches to addressing this problem. Stereovision is the most natural way of extracting depth information.^{16,17} The principle of stereovision is to have several cameras focused on the same object. Depth information is then recovered based on the disparity between corresponding points in images taken by these cameras. Next, the detection of an anomaly in the subject's motion (e.g., fall) is realized by comparing the 3-D trajectory of the subject's head with that predicted by different scenarios. With this approach, the system can acquire video image recordings of the head movements, which are then merged. Because posture-based events such as a fall may look different in the video from different perspectives, the system must be able to yield the right decision regardless of where the camera is placed. To achieve this goal, we adhere to a procedure based on fuzzy logic principles.

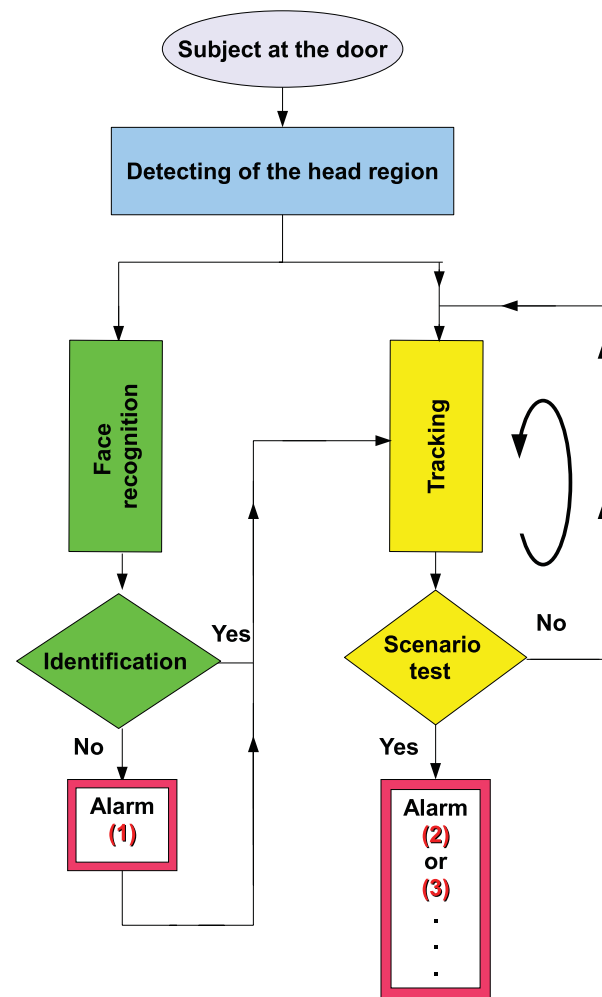


Fig. 1 Overall block diagram of machine-vision system described in this research with object detection, segmentation, and video compression.

3 Optical Correlation Techniques for Detection of Various Posture-Based Events

Our machine-vision sensing system designed for pattern recognition is based on optical correlation techniques. The authors have previously addressed some of the above issues in recent publications.⁷ Throughout the current effort, we have a threefold procedure to solve the practical issues related to the design and the implementation of a system capable of reliably communicating with the cameras, and executing the above-described algorithm. Image analysis of the subject's head is the crucial first step of the algorithm. Such analysis includes the size reduction of the images to increase the velocity processing for the comparison of the input image and the correlation filters. In this research effort, we used the HSV representation to detect subject's skin^{18,19} regions in a given image. Eye detection is also used to discriminate the head from the rest of the subject's body. Eye detection and tracking remain challenging due to the individuality of eyes, occlusion, variability in scale, and light conditions. The second step in our analysis consists of identifying the head of the subject entering the room. The final stage in this effort is the 3-D motion tracking of the subject in a given environment (i.e., typically, the subject's home).

Optical correlation techniques represent a powerful tool for face tracking and identification. Optical implementation of these techniques has been presented in the literature.⁷ However, even if these techniques are interesting, they are generally too complicated to achieve with an all-optical system, especially for reconfigurable systems.^{7-12,20-25} Hybrid methods (optical/digital) have been developed to alleviate problems associated with optical implementation, such as the need to use several spatial light modulators and the alignment problem of the optical elements. The potential of the optical correlation methods has been recognized because they possess a strong discriminating ability^{7-12,20-25} and robustness with respect to noise. In the present analysis, we introduce a new composite filter optimized for correlation purposes to deal with eye detection and tracking. While this specific filter is adapted for an all-numerical approach, which becomes an effective method for parallel processing using GPU or FPGA graphics card, it is clear that our main goal is to improve the performances of the correlator. Therefore, our focus is on composite filters that maximize discrimination (i.e., produce sharp correlation peaks so that we can easily localize the target in the scene, maximize noise tolerance, discriminate the desired class of images from all others, and easily implementable on correlators).¹¹ In addition, special attention is paid to a phase-only filter (POF). In general, this kind of filter leads to excellent performances while keeping its spectral size reasonable,²¹ for a numerical implementation.

3.1 Head Detection Based on the Hue, Saturation, and Value Color Space and the Method of Moments

Color can have its disadvantages though, some of which are the large memory requirements to store color information and additional processing overhead for conversion from one color scheme to another. These disadvantages can be additional burdens on a real-time system. Size reduction of the images, which are the input of the correlators, is done by limiting the video analysis to the head area of the subject. We first detect skin areas because skin has very specific features—whatever the type or color provided—that a relevant form of image representation is chosen. Then, eye recognition is performed inside the skin areas. Images can be represented in various ways. Skin can be detected in color images by using the HSV color space.^{18,19} In the HSV color space, hue represents the color information, which is represented by a circular wheel and varies between 0 and 360°, as displayed in Fig. 2. Saturation gives the degree of purity of hue. Intensity I or the value V gives the brightness level. Both saturation and value range from 0 to 100%. The higher S is, the more relevant the H information, and low S values indicate gray-scale pixels. A maximum value for V indicates that the color is at its brightest, and zero indicates the color black. The HSV color space is of interest because the hue values are independent of intensity. This has the advantage of intensity variations due to illumination being ignored. However, it also has the disadvantage of being insensitive to differences between objects and their shadows when the shadow is of the same color as the object. For our purpose, utilization of the HSV domain for color segmentation is well adapted. In practice, the insensitivity problem is overcome by adjusting the thresholds to be applied to the hue and saturation components according to the skin subject's skin color.

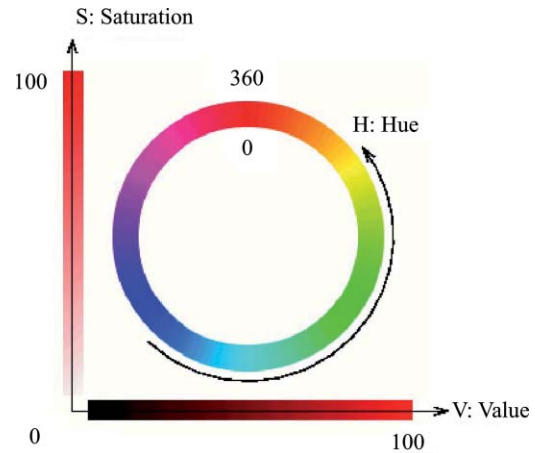


Fig. 2 HSV color space.

Figure 3 shows an example for detection of the head area in a typical image. Start with Fig. 3(a), which is obtained from a color webcam. Using our algorithm, the image can be analyzed in terms of hue and saturation, respectively shown in Figs. 3(b) and 3(c). The value component is not used here because it depends closely on the image brightness. Fig. 3(d) results from the threshold of the hue and saturation components. These two images will undergo several operations (i.e., smoothing and erosion) to reduce their noise and eliminate the isolated or small groups of pixels. Then, these two images are merged together to get the output image [Fig. 3(d)]. A restricted area centered on the head is then chosen [Fig. 3(e)]. In some situations, some complication may happen (e.g., if the subject is bare chested). To deal with these particular situations, the HSV representation of the image is associated with the RGB representation, where three channels are used to depict red, blue, and green color information, respectively, for eye detection. The idea is as follows: once the skin areas are detected in an image, we are looking for the eyes inside these areas. To reduce the negative effect of unwanted reflections from the eyes and make the system more reliable and robust, only the darker parts of the face are detected in the red component of the image. Next, the method of geometric moments is used to locate the position of the eyes within the face image with

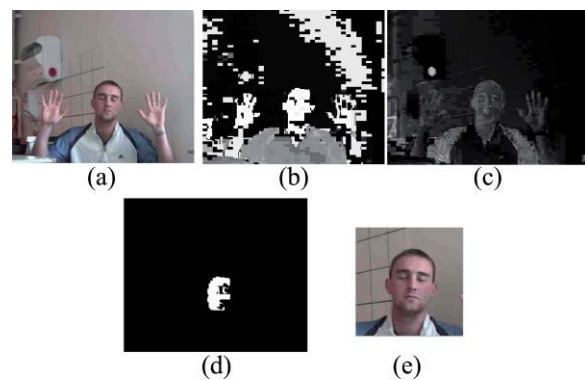


Fig. 3 Example of head detection based on the (a) target image, (b) H component thresholding, (c) S component thresholding, and (d) head-detection results. (e) This figure shows the result from running the proposed method.

a certain degree of accuracy. Many recognition techniques based on the method of moments were developed in past decades.^{26–28} Moments are used to find the center of an area of white pixels (white noise). To adapt this method to our application, we used the negative image. Thus, eyes become an area of white pixels. In our approach, we consider that the eye region is the darker part of the red component of the face. We first invert this red component so that eyes are described by a cloud of white pixels. It is of note that several defects may persist (e.g., a strand of hair). To overcome this problem, a mask is applied to the image for selection of the areas that are of interest for our application. Once this mask operation has been completed, computation of the vertical geometric moment allows us to discriminate between the right and left parts of the face. To refine the eye localization, the operation is done also for obtaining the horizontal moment. Consequently, this method locates the position of the eye within the vertical axis passing through it and between the eye and the eyebrow. With this procedure, a target image containing the information on the subject’s head is obtained. Once this first step is established, we introduce this image in a VLC filter by employing an optimized, segmented composite filter.

3.2 Vander-Lugt Correlator with Optimized, Segmented Composite Filter

The second step of our procedure is to solve a standard face-recognition issue. The proposed method is based on a VLC correlator with a new correlation filter. Utilization of the VLC for face recognition has been fairly popular.^{6–12,20–23} This method appears very effective despite possible distortions that may exist between the target face and the reference images due to vertical and horizontal rotations. Hence, we must have a large number of references leading a large number of filters. The segmented composite filter, which is used successfully in our laboratory, allows us to overcome these problems by merging several references together. This filter allows us to obtain a better optimization of the space-bandwidth product available than that of conventional composite filters.^{7,12} Indeed, with a composite filter, the phenomenon of local saturation in the Fourier plane is much more critical than for a segmented filter, due to the fact that the manufacture of the composite filter is based on the local addition of spectral information arising from different references. In this work, an optimization technique is proposed for fully exploiting the capabilities of VLC filter for our specific application.

The schematic diagram illustrating the optimized, segmented composite filter is shown in Fig. 4. Basically, it consists of segmenting the Fourier plane of the filter into several zones and allotting a reference image to each zone according to a chosen criterion. In our published work on multicorrelation filters,^{7,12} we showed that it is possible to design a segmented composite filter, merging at least 12 references. Specifically, we used a purely energetic criterion that does not take into account the resemblance between the different references. Validation of such optimization of this criterion is now illustrated. We begin by performing the Fourier transform (FT) of each image reference. Then, each spectrum is multiplied by its own pass-band function P^k , which is cal-

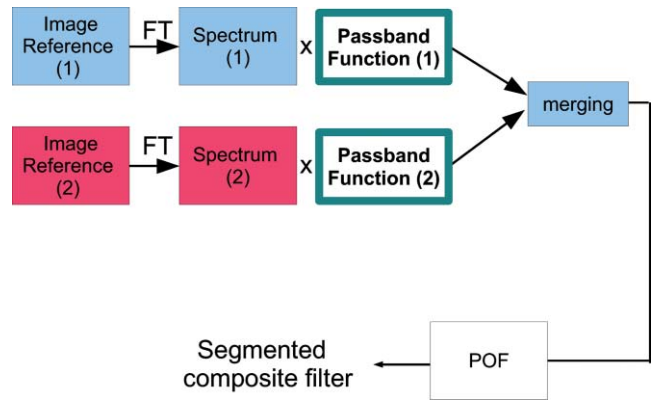


Fig. 4 Synoptic diagram of the segmented composite filter for multi-correlation (Ref. 7).

culated as

$$\begin{cases} P_{i,j}^1 = P_{i,j}^2 = 1 & V \in [-s', +s] \\ P_{i,j}^1 = 1 \text{ and } P_{i,j}^2 = 0 & \text{if } V > +s \\ P_{i,j}^1 = 0 \text{ and } P_{i,j}^2 = 1 & V < -s' \end{cases},$$

$$V = \frac{E_{i,j}^1}{\sum_i \sum_j E_{i,j}^1} - \frac{E_{i,j}^2}{\sum_i \sum_j E_{i,j}^2}, \quad (1)$$

where (i,j) denotes the coordinates of a given pixel, N is the size of the image, $E_{i,j}^k$ is the spectral energy of an image k at the location (i,j) , s and s' are two given thresholds (here, s and s' were set to 4 and 0, respectively) and $\sum_i \sum_j E_{i,j}^k$ is the spectral energy of the reference image k . If more reference images are needed, then this comparison between the reference “1” and other references should be repeated as many times as there are references. Reference 1 is chosen as the image containing the most probable position of the subject’s face. In the current study, we have chosen the reference presenting the subject’s face with horizontal and vertical rotations equal to zero.

In order to enhance the performance of the correlation decision, many algorithms have been proposed in the literature.^{23–25} In this paper, we only consider an approach based on the iterative reconstructed algorithm of the target image using its spectral phase information. It is well established that an image (scene) can be reconstructed using only the information of its phase spectrum.^{29–31} Hereafter, the relationship between the quality of the reconstructed image and the decision performance of the correlator is investigated.

The principle of the proposed method is shown in Fig. 5. First, a FT of the target image is performed. Only, the spectral phase is used later. These phases are used in the reconstruction algorithm (see, e.g., for details Ref. 32) in order to obtain the reconstructed image placed at the input plane of the correlator. Only three or four iterations of the reconstruction algorithm³² are required to reconstruct the image. This reconstructed image is placed at the input of the correlator. In Fig. 6(a), we show the correlation plane when the original image (I_0) is placed at the input of the correlator. This plane will be used as the reference plane to check the performance of the correlation decision when the reconstructed image is displayed at the input plane. The second input of the correlator is a reconstructed image from phase

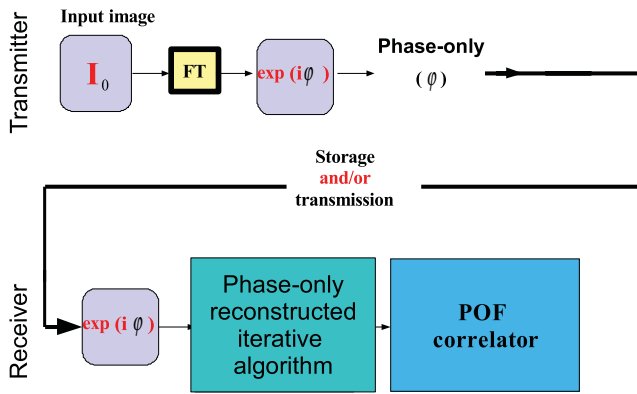


Fig. 5 Synoptic diagram of a reconstructed image as input of the segmented filter POF (Ref. 29).

information (taking four iterations). The corresponding correlation plane using the same filter is shown in Fig. 6(b). It is also interesting to note that we can see the robustness of our algorithm, which can improve the correlation between a reconstructed image and a reference image. In fact, the noise in the correlation plane has been reduced and the peak-to-correlation energy (PCE), defined as the energy of the peak correlation normalized to the total energy of the correlation plane

$$\text{PCE} = \frac{\sum_{i,j}^M E_{\text{peak}}(i, j)}{\sum_{i,j}^N E_{\text{correlation_plane}}(i, j)}, \quad (2)$$

where M denotes the size of the peak correlation spot and N stands for the size of a correlation plane, was at the same time increased by a factor of 3.8. The good results obtained are attributed to the fact that this reconstruction method, depending only on the target image's phase, may be assimilated to a contour detection-based method. In addition, it allows us to reduce noise since only several iterations are required. An important property of this method is its ability to handle shape, scale, and rotation changes. On basis of the above discussion, we suggest that the 3-D correlation technique that

is developed in Ref. 7 can be employed in order to make the system more reliable and robust.

3.3 Joint-Transfer Correlator for Subject's Tracking

Assuming that identification of the subject is positive from the image analysis, a JTC is used to guide the tracking procedure.⁷⁻¹⁰ Here the procedure is described for a single camera. However, because the main purpose here is to record the parameters related to a moving subject, we need to introduce a 3-D tracking of the subject. Hence, the depth information extracted using 3-D reconstruction must be accurate and several cameras are necessary. This case will be discussed in Sec. 4.

In Ref. 8, the principle of a new type of correlator for pattern recognition called JTC was introduced by Weaver and Goodman. In their pioneering paper, the authors demonstrated the sufficient and necessary conditions to optically achieve the convolution between two images displayed in the input plane of the JTC correlator: a target image (image to be recognized) and a reference image (image coming from a given data base). Both target and reference images are separated by a given distance. The synoptic diagram of the JTC is presented Fig. 7. A first beam coming from a laser illuminates the input plane $I(x,y) = s(x,y) + r(x-d, y-d)$, which contains the scene, with $s(x,y)$ as the target image to be recognized, and the image reference is $r(x-d, y-d)$, where d represents the distance separating the target from the reference images. The joint spectrum, obtained by Fourier transforming $I(x,y)$, is recorded and yields an expression of $|t(u,v)|^2$, which takes the form

$$\begin{aligned} |t(u, v)|^2 &= |S(u, v)|^2 + |R(u, v)|^2 \\ &+ \{|S(u, v)| \exp[\varphi_s(u, v)]\} \{|R(u, v)| \\ &\times \exp[-\varphi_r(u, v) + j(ud + vd)]\} \\ &+ \{|S(u, v)| \exp[-\varphi_s(u, v)]\} \{|R(u, v)| \\ &\times \exp[\varphi_r(u, v) - j(ud + vd)]\}. \end{aligned} \quad (3)$$

To obtain the correlation plane using the classical joint transform correlator (CL-JTC), we need to perform an inverse Fourier transform (FT^{-1}) of the recorded intensity [Eq. (3)] to get the three-peak correlation plane: a first one representing the zeroth order resulting from the FT^{-1} of the first term

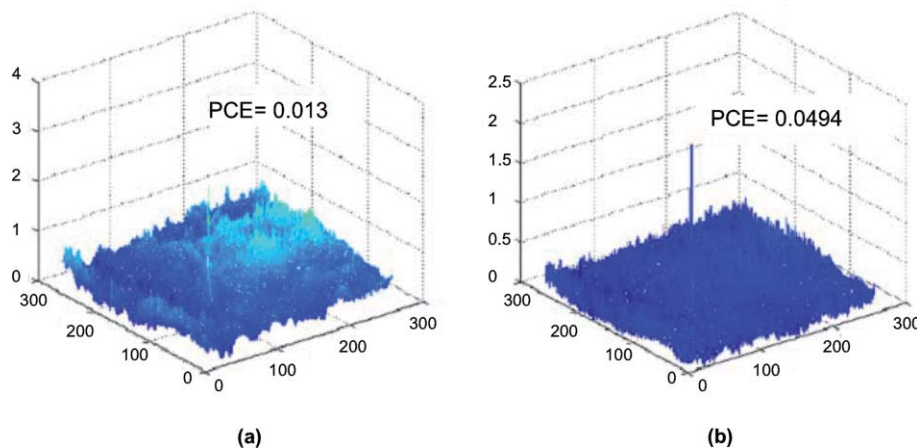


Fig. 6 3-D representation of normalized correlation planes: (a) correlation between the original image with our correlation filter and (b) correlation between the reconstructed image with the same filter.

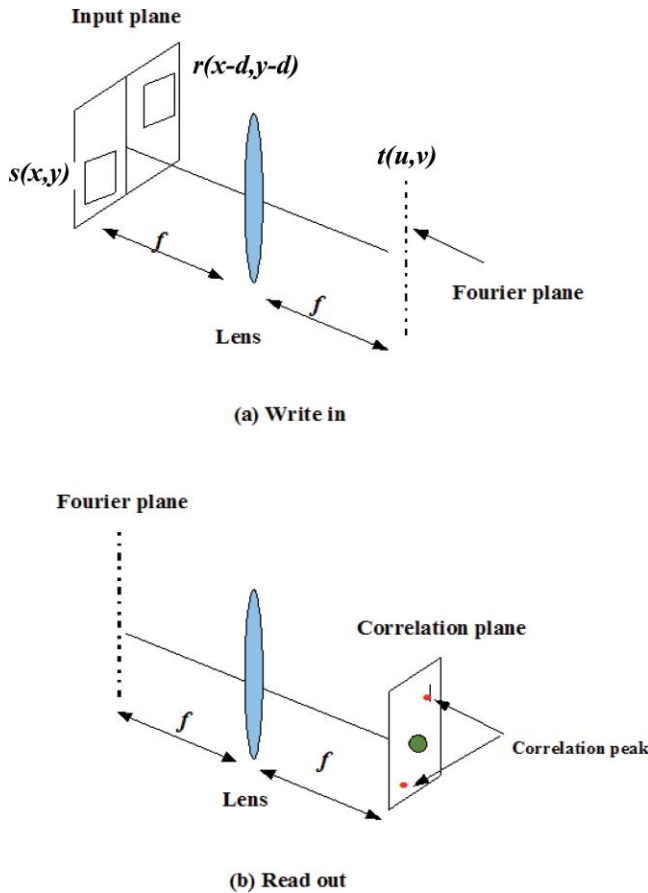


Fig. 7 Principle of the JTC (Ref. 7).

of the left-hand side of Eq. (3), and two other peaks corresponding to the FT^{-1} of the second and third terms of the left-hand side of Eq. (3). Studying these two correlation peaks determine the degree of resemblance between the target and the reference faces. However, this standard correlator has two major drawbacks rendering its use not always reliable: it presents a very large and very intense zero-order peak compared to the other peaks, and it is characterized by two low-intensity and large correlation peaks. To improve this situation, various optimization schemes were proposed in the literature. In this research effort, we will focus on the FA-JTC suggested by Alam and Karim⁹ (i.e., the FA-JTC that permits one to render the JTC more robust to noise). This procedure was adopted here because it allows us to eliminate the effect of the input image background noise (the image background noise that can exist behind the subject), and it is user friendly. Basically, it consists of multiplying the intensity of the JTC joint spectrum [Eq. (3)] with the fringe-adjusted filter

$$H(u, v) = \frac{G(u, v)}{N(u, v) + R(u, v)}, \quad (4)$$

where $G(u,v)$ denotes a function to obtain an optical gain of >1 , and $N(u,v)$ is a function used to reduce the noise effect in the correlation peak and/or to delete the band limit of the signal. In the case considered here, $N(u,v)$ denotes the image background.

The procedure for tracking the subject is realized by initializing the JTC with a reference image R_0 (in Fig. 8, left panel of a_1), i.e., the subject head obtained, as described in

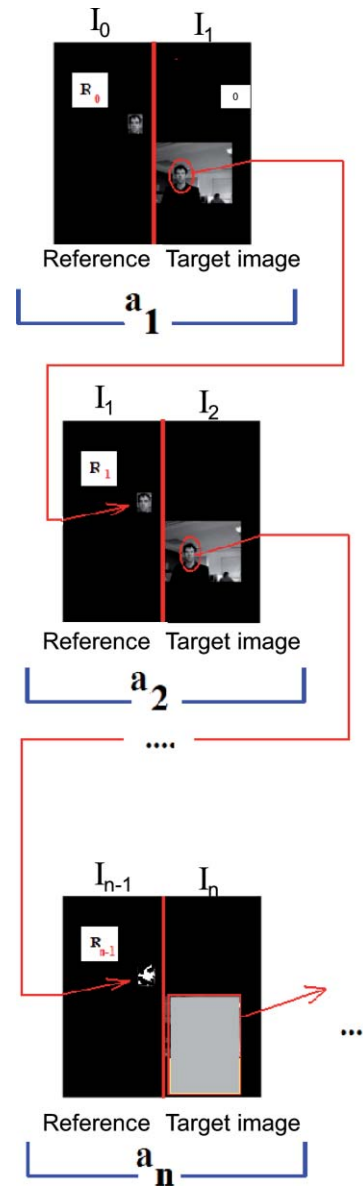


Fig. 8 Principle of the tracking system.

Sec. 3.1, at the entrance of the room. The target image (i.e., the first image of a video sequence recorded as the subject crossed the door of the room) is introduced in the corresponding right panel of Fig. 8. Because the position related to the reference image is known exactly, that of the subject's head in image I_1 can be determined by using the JTC. The latter will serve as reference image for the next iteration (panel a_2 in Fig. 8), and so on. With this kind of procedure, the head's motion and trajectory can be extracted. We have investigated this procedure with several people in the room, and it has been found to be robust.

One feature needs additional explanation. To measure the head's position in 3-D space, the depth information extracted using 3-D low-cost (low-resolution) webcams must be accurate. However, noise in the video-capture process is inevitable due to the problems of camera imaging (e.g., distortion). Therefore, the images to be analyzed contain inaccuracies that make establishing the correspondences in adjacent frames less reliable. To make the system reliable and robust,

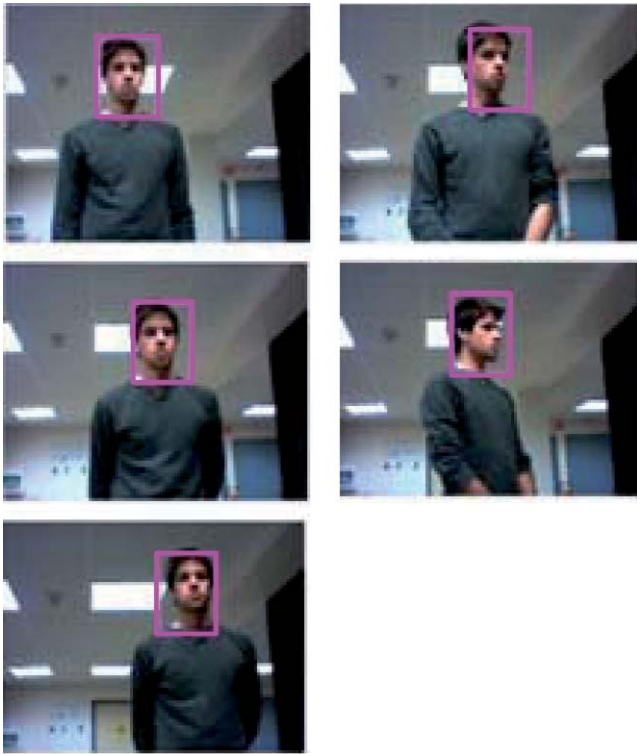


Fig. 9 Example of reference database used in our study.

we introduce a database to save the possible head's motion pattern information (Fig. 9). We also use this database when the synchronization is lost. With this kind of standard prediction and correction procedure, the noise effects are minimized and the system performance is enhanced.

4 Hardware and Software Architectures

Figure 10 shows the platform used to carry out the study for validating the principle of our approach. The measurement surface, which is the region over which the camera views overlap, is $2 \times 1.2 \text{ m}^2$. It is equipped with four static webcams and a chair. In Fig. 10(a), the stars denote the positions of the webcams, which record the parameters related to the head's motion pattern. With this approach, the depth information can be retrieved using a standard stereovision method. Each camera gives us several decisions (e.g., presence of the subject in the room, identity of the subject, information on the head's motion position, and depth of the subject in the room). This set of information, called a low-level decision, is collected and stored in different files (one file per camera). These are then merged for the purpose of making a decision concerning behavioral change of the subject; this decision is called a high-level decision.

Fuzzy logic has been applied to exposure control in the camera systems to take advantage of linguistic logic described control systems.¹³⁻¹⁵ In many image-processing applications, expert knowledge must be used for applications such as pattern recognition and scene analysis. Many difficulties in image processing arise because the data, tasks, and results are uncertain. This uncertainty is not always due to the randomness but to the inherent ambiguity and vagueness of image data (e.g., grayness ambiguity, geometrical fuzziness, and vague knowledge of image features). A related work is discussed by Kang et al.,³³ where an iris-recognition method for information security is based on fuzzy logic control. It is now well established that choices of rule sets and membership functions significantly affect achieving the performance goals.

Components for the proposed fuzzy logic system are illustrated in Fig. 11. The experimentally relevant parameters we are considering are: (i) head height, H_{head} , in the

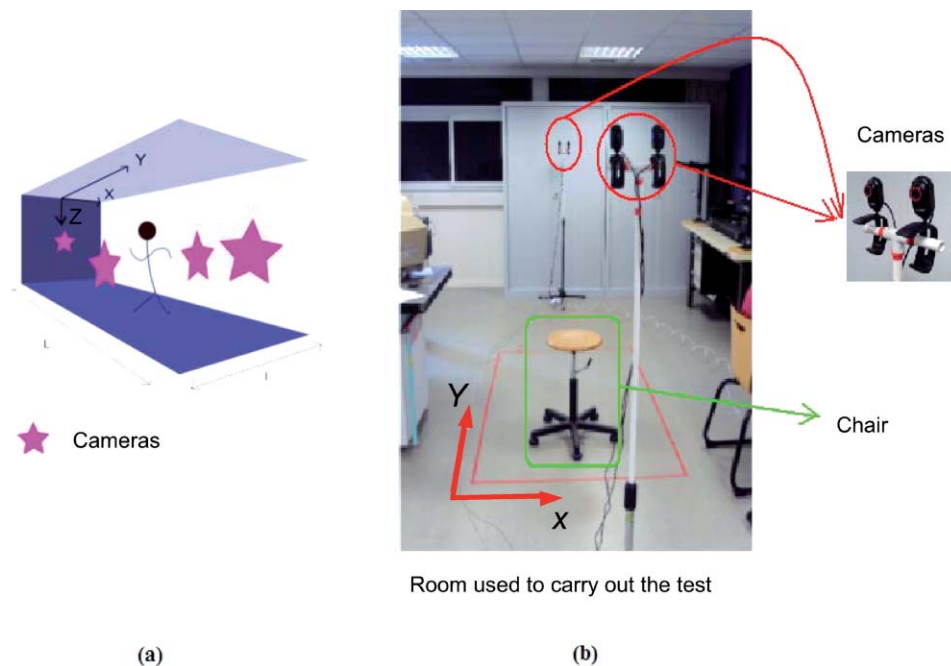


Fig. 10 Experimental platform for the study of various posture-based events, including normal daily life activities and unusual events (e.g., fall detection).

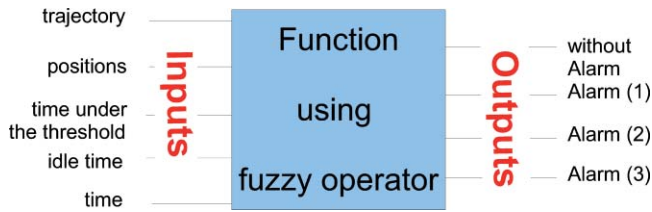


Fig. 11 Fuzzy logic controller optimizes the trigger of the alarms through fuzzy rules and membership function.

scene. This parameter is of paramount importance because a fall is characterized by an abrupt change of H_{head} . (ii) The threshold time, T_{thres} , corresponds to the time interval for which the head’s position is less than a given threshold (e.g., $S = \text{subject height}/2$). Several scenarios should be considered to avoid false alarms (e.g., distinguish the two events: “the subject bends down to pick up an object” and “the subject is feeling faint”). For this purpose, we assumed that if $T_{thres} \geq 7$ s, then the level of alarm increases and the alert goes out when it is equal to 10 s. (iii) The position, P_0 (e.g., chair). (iv) In surveillance, activity can be low most of the time, but once it does happen, it can be quite important. In addition, for several situations, T_{thres} should be significantly larger than 7 s (e.g., when the subject is drowsy in his or her chair). For these situations, it is the inactivity time, T_{inac} , that will raise the alarm. For the purposes of the experiments to follow, we assumed that this time is equal to 1000 s; and (v) the time, T .

Four levels of alarm were considered (Fig. 11): (i) no risk (i.e., the subject is moving inside the room and his or her head is above the position threshold); (ii) low risk (i.e., a security alarm is raised); (iii) moderate risk (i.e., a second security alarm is raised and a compressed and encrypted^{34,35} image of the scene can be sent to a medical response team); and (iv) high risk: a third security alarm is raised and a compressed and encrypted image of the scene is sent to the medical emergency teams and to a medical response team. To make a reliable decision, each camera is making a decision based on the five input variables considered in Fig. 11. The fuzzy value for input and output state variables is treated as a continuous function, which is called a membership function, and fuzzyAND and fuzzyOR operators. Among the many possible fuzzy membership functions, the sigmoid function

$$f(x) = A / (1 + \exp[-B(x - C)]) \tag{5}$$

was used. In Eq. (5), the prefactor A defines the maximum output gain, B is the slope, C is the offset that represents the midpoint at the transition of the sigmoid function, and x denotes the state input variables (H_{head} , T_{inac} , T_{thres}). This membership function is illustrated in Fig. 12, taking $A = 10$, $B = 1$, and $C = 5$. The membership function represents the level of alarm for each input variable. This membership function depends on the specific situation that is analyzed.

As an explicit example, Fig. 12 shows the result of a typical membership function corresponding to the situation where the head’s position is less than a given threshold (subject height/2). This situation may correspond to a variety of events (e.g., “the subject bends down to pick up an object” or “the subject is feeling faint”). For these two events, the head’s position is under the authorized threshold. To discriminate between the two events for which an alert should

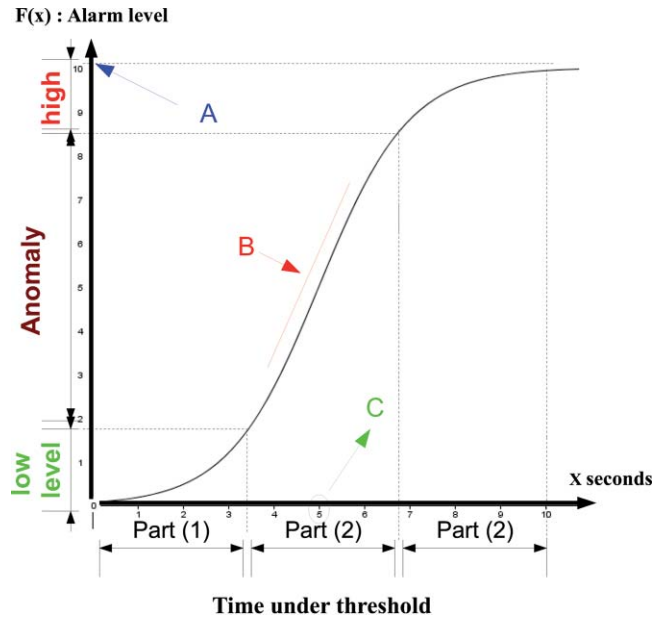


Fig. 12 Membership function.

be sent only for the latter, T_{thres} (horizontal axis in Fig. 12) is divided into three regions. Then, a level of alarm is associated to each region (vertical axis in Fig. 12). Hence, the level of alarm (expressed in percent) is seen to change much more abruptly in the second region than in the first one. The sigmoid membership functions for T_{thres} , T_{inac} , and H_{head} are similar. For T_{thres} , T_{inac} , the horizontal axis is expressed in seconds, whereas for H_{head} , the horizontal axis is expressed in centimeters.

The membership function used for time T should also consider two specific situations, i.e., the subject is present in the room during the night or may be inactive for a long time during the day. For the remainder of our discussion, we focus only on the latter situation. The membership function for position P_0 is a surface. A Gaussian function was chosen. This conical form-shaped surface is localized at the chair’s position in the scene (Fig. 13). Henceforth, if the subject sits at this point, his or her head’s position can be under the threshold and the level of alarm remains low.

5 Experimental Results

A variety of experiments were done with video recordings of several trials of subject’s body movement as input to the machine-vision algorithm. As we have noted before, the software was run on a CPU: Intel Dual Core E5300, 2.6 GHz, 2 Go RAM computer. The video files were captured by Logitech webcams, which offered video with a resolution of 352×288 pixels in size.

5.1 Fall Detection

As a practical example restriction, we consider a fall detection scenario corresponding to a rapid decrease of the head’s position below the authorized threshold. For this purpose, the position and velocity of the subject were recorded. These data allow us to decide if the subject is static (if he or she sits on a chair) or is moving (with a given velocity). The trajectory,

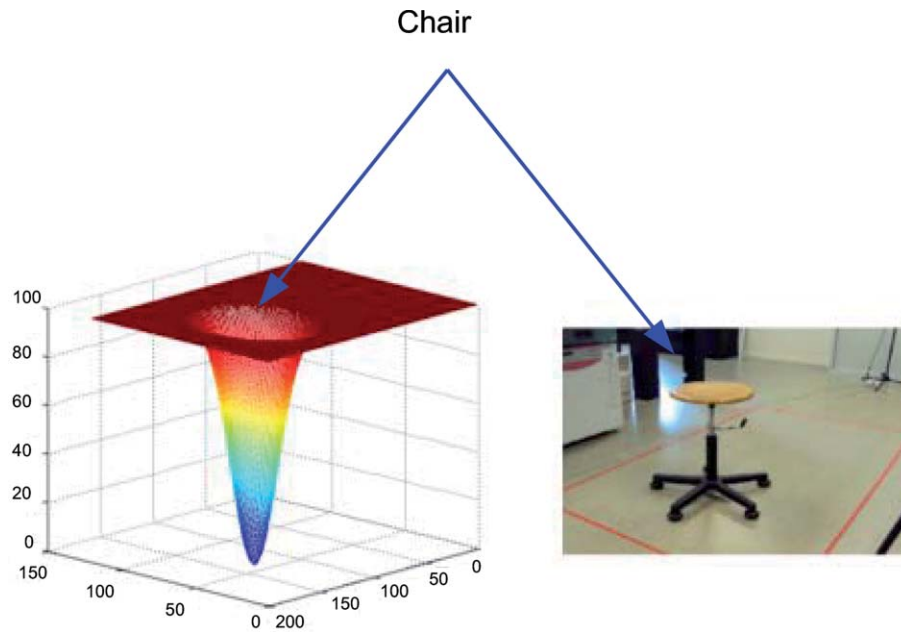


Fig. 13 Illustration of the membership function for position P_0 .

velocity, and inactivity time are the fundamental parameters of the membership function described above using fuzzy operators in order to raise a fall alarm. This procedure (limited here to six images per second due to the use of a PC) allows us to detect situations for which the head's position of the subject is falling down rapidly. However, when the subject is falling down very rapidly, the false alarms were identical. This is mainly due to the needed processing time (identification, tracking, and fuzzy operation).

To get reliable information of fall detection, it is mandatory to process in real time all information captured from the four tracking webcams. The used software is not appropriate to satisfy this constraint. In addition, a saturation problem may affect the processor. To overcome this problem, only six images per second were considered. Hence, our method cannot be relevant for detection of a rapid fall. Moreover, it remains important to improve the image quality by using high-resolution cameras and to increase the number of sensors. Powerful and rapid processors are required to overcome the software limitations and focus on the hardware part of this method. For that purpose, we are planning to use a GPU instead of a CPU.^{36,37}

A user interface applying the above-mentioned method is given in Fig. 14. Head motion is captured by four webcams and recorded from either the front, back, or side in order to extract the T_{thres} , T_{inac} , and H_{head} . Hence, four video sequences of the scene can be visualized in this interface. The identification of the subject and his or her head's position are represented by pink frames in Fig. 14. Stereovision permits one to characterize the depth information related to the subject (represented by the dot in the left panel of Fig. 14). The fuzzy alarm level (in%) is indicated at the right bottom of Fig. 14.

5.2 Experimental Validation

Many situations simulating a fall event were tested using our system. An illustrative example is shown in Video(1).³⁸ This

video clip has a duration of 2 min, in which four scenarios are considered: the subject is moving inside the scene with his or her head's position above a given threshold value (no alarm); the subject falls (flag an alert); the subject bends down slowly; and the subject sits on a chair (the security alarm is raised after a given inactivity time). For this first scenario, the results shown here suggest that the level of alarm remains very low if the head's position is above the threshold. As the subject bends down slowly (e.g., for picking up an object), the level of alarm increases gently. If the subject does not stand up beyond the 7 s, then the level of alarm exhibits a rapid rise and the alert goes out if the level of alarm reaches 100%. However, the level of alarm decreases to 0% if the subject stands up. Likewise if the subject is falling, his or her head's position is suddenly below the threshold. Consequently, the level of alarm increases abruptly and the alert goes out. The last



Fig. 14 User interface of the fuzzy logic and optical correlation-based face-recognition method developed using Matlab software and a CPU: Intel Dual Core E5300, 2.6 GHz, 2Go RAM computer. (Color online only.)

scenario shown in the video sequence concerns the situation when the subject sits on a chair. For this case, the level of alarm is calculated according the inactivity time thanks to the membership function shown in Fig. 13. On increasing T_{inac} , the level of alarm increases significantly until the alert goes out.

6 Summary

Home automation is being implemented into more and more homes of the elderly and disabled in order to maintain their independence and safety. Fundamental to the preference for home-based care over institutional care is the expectation that family caregivers will be available in the home to support patients who would otherwise be in an institution. In this paper, we have proposed a novel method for fuzzy logic and optical correlation-based face recognition for elderly monitoring in home surveillance. Our machine-vision system has focused on the goals of accuracy, convenience, and cost, which, in addition, does not require any devices attached to the subject. This is particularly important when considering the difficulty that hemiplegic subjects have in putting on attached sensors and the inherent inaccuracy introduced in their motor limitations. The system involves the use of ordinary off-the-shelf web cameras for image acquisition, and the software was implemented on a PC. The system is able to recover the 3-D position information from the camera images. The objective of the proposed research was segmentation based on motion and tracking motion during surveillance, subject (head) identification based on some descriptor on shape, and color for matching between frames. The object detection is based on HSV color space. In addition, we need a method to detect scene changes. The identification and tracking system relies on the use of a new segmented composite filter allowing us to get robust face recognition. To get a quick identification, we had to reduce the size of the target image (head area) to display at the input plane of the correlator. Considering this, quantized saturation values are used to segment the hue. Pixels with high saturation values (which give good color information) are considered similar to each other and grouped together. For the purposes of the tracking procedure, we used FA-JTC, which is adapted to eliminate the effect of the background and hence increase the efficiency of proposed fall-detection scheme. In addition, our system is cost effective because it uses low-cost webcams. Each webcam gives us information about several physical parameters, such as the head's position related to a given reference, the threshold time corresponding to the time where the head's position is less than a given threshold, and the head's velocity. Our example clearly demonstrates the feasibility of our approach to obtain a reliable system based on optics correlation methods and fuzzy logic for elderly monitoring in home-surveillance applications. The different experiments (realized with three different persons with the same scenarios shown in the video of Ref. 38, Chapter 5) demonstrate the good performances of our approach. For the four scenarios considered in this study, the level of alarm is close to 89%. The technique presented does appear to possess good performances for identifying and tracking a subject as he or she enters the room and to analyze various posture-based events, including normal daily life activities and unusual events (e.g., fall detection).

This is first time that an algorithm, combining two types of optical correlation architectures (i.e., VLC, JTC) with a merging of information based on a fuzzy logic approach is proposed for patient-monitoring applications in home video surveillance. The proposed method is novel because it is able to overcome the problem of variation of the target object (patient) with time, for which using a given base of references images may pose serious problems. To solve this problem, we replace the reference images not used in our database by new images taken more recently. Our database is composed of images taken under different conditions (e.g., positions, lighting conditions). All these aspects are included in our baseline algorithm. To deal with the problem of nonuniform illumination, the technique recently developed in Ref. 39 is included in our algorithm. To show the good performance of our algorithm several baseline algorithms were studied.^{40,41} However, these approaches do not allow us to identify the patient and have a limited performance when complex situations (e.g., several persons are presented in the scene, the patient is partially hidden) are considered. Baseline algorithms for identification based on the use of the independent component analyses (ICAs) model⁴² were also studied by us. However, these approaches necessitate much computing time and require selecting the target face present in the given scene before the identification step.^{43,44} Overall, it is found that these approaches have a lower performance than those based on correlation.⁴⁵

The main objective of this work was to present the proof of principle of a robust on-intrusive eye detection and tracking (i.e., eye localization in the image and for video images tracking of the detecting eyes from frame to frame). In the light of the preceding results, it appears to consider the limitations of the system: (i) Identification should be realized thanks to a VLC correlator with a dynamical reference image base in order to deal with the subject's position and light illumination changes and (ii) research in eye detection and tracking for face recognition poses many challenges that were not considered in the analysis (e.g., when the subject is wearing glasses). We consider that our algorithm remains efficient if >75% of the subject's face can be viewed. Otherwise, the algorithm is instable with respect to illumination and viewing angle, rotation, etc. To deal with these issues, we proposed to adapt a new correlation filter using ICA as a preprocessing stage for face recognition⁴⁶ because the ICA method is robust against the change of image versus rotation, illumination, structured noise, and scale for a database of independent components. As an outlook, we mention that we are currently trying to optimize, on the one hand, the tracking process by using a JTC and specific filters to suppress background noise and, on the other hand, the subject's localization within the scene. The only remaining systematic error in the head's position is due partly to purely numerical issues and partly to a physical feature of the specific stereovision method used. We believe that with higher computational resources, these limitations can be overcome and we could increase the efficiency of proposed fall-detection scheme. For a few cases of rapid motion of the subject's head, problems may arise due to the limitation of the CPU to deal with six images per second. A simple modification of the model (i.e., numerical implementation of this algorithm on a GPU or FPGA graphics card) will remove this shortcoming.

We close with a word of caution. This research is a first step toward building a compression and detection

machine-vision system for specific surveillance application (i.e., monitoring of the elderly). The biggest concern expressed by potential users of smart home technology is “fear of lack of human responders or the possible replacement of human caregivers by technology,”¹ but home automation should be seen as something that augments, but does not replace, human care.

Acknowledgments

Work in L@BISEN was supported by Groupe-ISEN. Work in Lab-STICC was supported in part by the CNRS. The technical assistance of A. Coum and M. P. Le Coz is gratefully acknowledged.

References

- P. Cheek, L. Nikpour, and HD Nowlin “Aging well with smart technology,” *Nurs Adm Q.* **29**, 329–38 (2005).
- M. Chan, D. Estève, C. Escriba, and E. Campo, “A review of smart homes: present state and future challenges,” *Comput. Methods Programs Biomed.* **91**, 55–81 (2008).
- C. D. Nugent, D. D. Finlay, P. Fiorini, Y. Tsumaki, and E. Prassler, “Editorial home automation as a means of independent living,” *IEEE Trans. Autom. Sci. Eng.* **5**, 1–9 (2008).
- X. H. B. Le, “Reconnaissance des comportements d’une personne âgée vivant seule dans un habitat intelligent pour la santé,” Ph.D. thesis, Université Joseph Fourier-Grenoble 1 (2008).
- N. Zouba, F. Bremond, M. Thonnat, A. Anfosso, E. Pascual, P. Mallea, V. Mailland, and O. Guerin, “A computer system to monitor older adults at home: preliminary results,” *Gerontechnology* **8**, 129–139 (2009).
- A. B. VanderLugt, “Signal detection by complex spatial filtering,” *IEEE Trans. Info. Theory* **IT-10**, 139–145 (1964).
- A. Alfalou and C. Brosseau, “Understanding correlation techniques for face recognition: from basics to applications,” in *Face Recognition*, M. Oravec, Ed., INTECH, India (2010).
- C. S. Weaver and J. W. Goodman, “A technique for optically convolving two functions,” *Appl. Opt.* **5**, 1248–1249 (1966).
- M. S. Alam and M. A. Karim, “Fringe-adjusted joint transform correlation,” *Appl. Opt.* **32**, 4344–4350 (1993).
- E. Watanabe and K. Kodate, “Implementation of a high-speed face recognition system that uses an optical parallel correlator,” *Appl. Opt.* **44**, 666–676 (2005).
- B. V. K. V. Kumar, “Tutorial survey of composite filter designs for optical correlators,” *Appl. Opt.* **31**, 4773–4801 (1992).
- A. Alfalou, M. Elbouz, and H. Hamam, “Segmented phase-only filter binarized with a new error diffusion approach,” *J. Opt. A: Pure Appl. Opt.* **7**, 183–191 (2005).
- L. A. Zadeh, “Fuzzy sets,” *Inf. Control* **8**, 338–353 (1965).
- G. J. Klir and B. J. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Application*, Prentice Hall of India (1997).
- L. A. Klein, “Sensor and data fusion concepts and applications,” *Opt. Eng.* **TT35**, Bellingham (1999).
- P. F. Luo and J. Wu, “Easy calibration technique for stereo vision using a circle grid,” *Opt. Eng.* **47**, 033607 (2008).
- J. Zhu, Y. Li, and S. Ye, “Design and calibration of a single-camera-based stereo vision sensor,” *Opt. Eng.* **45**, 083001 (2006).
- J. Terrillon and S. Akamatsu, “Comparative performance of different chrominance spaces for color segmentation and detection of human faces in complex scene images,” in *Proc. of IEEE 3rd Int. Conf. on Face and Gesture Recognition*, pp. 54–61 (2000).
- I. Kramberger, “Real-time skin feature identification in a time-sequential video stream,” *Opt. Eng.* **44**, 047201 (2005).
- B. V. K. V. Kumar, “Partial information filters,” *Digital Signal Process.* **4**, 147–153 (1994).
- J. L. Horner and P. D. Gianino, “Phase-only matched filtering,” *Appl. Opt.* **23**, 812–816 (1984).
- J. L. Horner, “Metrics for assessing pattern-recognition performance,” *Appl. Opt.* **31**, 165–166 (1992).
- J. Ding, J. Itoh, and T. Yatagai, “Design of optimal phase-only filters by direct iterative search,” *Opt. Commun.* **118**, 90–101 (1995).
- A. A. S. Awwal, M. A. Karim, and S. R. Jahan, “Improved correlation discrimination using an amplitude-modulated phase-only filter,” *Appl. Opt.* **29**, 233–236 (1990).
- F. M. Dickey, B. V. K. V. Kumar, L. A. Romero, and J. M. Connelly, “Complex ternary matched filters yielding high signal-to-noise ratios,” *Opt. Eng.* **29**, 994–1001 (1990).
- K. Grauman, M. Betke, J. Gips, and G. R. Bradski, “Communication via Eye Blinks-detection and duration analysis in real time,” in *Proc. of IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, pp. 1/1010–1017 (2001).
- M. Betke, B. Mullally, and J. Magee, “Active Detection of Eye Scleras in Real Time,” in *Proc. of IEEE CVPR Workshop on Human Modeling, Analysis and Synthesis*, Hilton Head Island, SC, pp. 8–13 (June 2000).
- J. J. Magee, M. Betke, J. Gips, M. R. Scott, and B. N. Waber, “A human-computer interface using symmetry between eyes to detect gaze direction,” *IEEE Trans. Syst., Man, Cybernet. Part B* **38**, 1–14 (2008).
- R. W. Gerchberg and W. O. Saxton, “A practical algorithm for the determination of phase from image and diffraction plane pictures,” *Optik* **35**, 237–246 (1972).
- A. V. Oppenheim and J. S. Lim, “The importance of phase in signals,” *Proc. IEEE* **69**, 529–541 (1981).
- T. Quatieri Jr. and A. V. Oppenheim, “Iterative techniques for minimum phase signal reconstruction from phase or magnitude,” *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-29**, 1187–1193 (1981).
- A. Alfalou, M. Elbouz, A. Mansour, and G. Keryer, “New spectral image compression method based on an optimal phase coding and the RMS duration principle,” *J. Opt.* **12**, 115403 (2010).
- B. J. Kang, K. R. Park, J. H. Yoo, and K. Moon, “Fuzzy difference-of-Gaussian-based iris recognition method for noisy iris images,” *Opt. Eng.* **49**, 067001 (2010).
- A. Alfalou and A. Mansour, “Double random phase encryption scheme to multiplex and simultaneous encode multiple images,” *Appl. Opt.* **48**, 5933–5947 (2009).
- A. Alfalou and C. Brosseau, “Optical image compression and encryption methods,” *Adv. Opt. Photon.* **1**, 589–636 (2009).
- Y. Ouerhani, M. Jridi, and A. Alfalou, “Fast face recognition approach using a graphical processing unit GPU,” in *Proc. of IEEE 1st Int. Conf. on Imaging Systems and Techniques*, (2010).
- S. H. Park, D. R. Shires, and B. J. Henz, “Coprocessor computing with FPGA and GPU,” in *Proc. of IEEE, 3rd DoD HPCMP Users Group Conf.*, Seattle, pp. 366–370 (2008).
- See supplementary material at <http://www.youtube.com/watch?v=3nNFTHOsm0E>
- V. H. Diaz-Ramirez and V. Kober, “Target recognition under nonuniform illumination conditions,” *Appl. Opt.* **48**, 1408–1418 (2009).
- P. J. Phillips, S. Sarkar, I. Robledo, P. Grother, and K. W. Bowyer, “The Gait Identification Challenge Problem: Data Sets and Baseline Algorithm” in *Proc. IEEE Int. Conf. on Pattern Recognition*, 385–388 (2002).
- P. J. Phillips, S. Sarkar, I. Robledo, P. Grother, and K. W. Bowyer, “Baseline Results for the Challenge Problem of Human ID Using Gait Analysis,” *5th Int. Conf. on Automatic Face and Gesture Recognition* (2002).
- A. Alfalou, M. Farhat, and A. Mansour, “Independent Component Analysis Based Approach to Biometric Recognition,” *Proc. of the 3rd Information and Communication Technologies: From Theory to Applications Conf.*, 1–4 (2008).
- X. Chen, P. J. Flynn, and K. W. Bowyer, “PCA-based face recognition in infrared imagery: baseline and comparative studies,” *Proc. of the IEEE Analysis and Modeling of Faces and Gestures Int. Workshop*, 127–134 (2003).
- Y. Jian, D. Zhang, and Y. Jing-Yu, “Constructing PCA Baseline Algorithms to Reevaluate ICA-Based Face-Recognition Performance,” *IEEE Trans. Systems, Man, and Cybernetics, Part B: Cybernetics* **37**, 1015–1021 (2007).
- A. Alsamman and M. S. Alam, “Comparative study of face recognition techniques that use joint transform correlation and principal component analysis,” *Appl. Opt.* **44**, 688–692 (2005).
- A. Alfalou and C. Brosseau, “Robust and discriminating method for face recognition based on correlation technique and independent component analysis model,” *Opt. Lett.* **36**, 645–647 (2011).



Marwa Elbouz received her PhD in telecommunications and signal processing from ENSTB-France and the University of Rennes 1 in 2001. Then, she held a two-year post-doctoral position at ISEN-Brest, during which she designed and realized a sign-language recognition system. She is now an associate professor of telecommunications and signal processing at ISEN-Brest, and a member of L@BISEN. Her research interests deal with optical engineering, optical information processing, signal and image processing, and optoelectronics. She has published 13 refereed journal articles or conference papers on these topics.



Alfalou Ayman received his PhD in telecommunications and signal processing from ENSTB-France and the University of Rennes 1 in 1999. He held a one-year postdoctoral position at ENSTB during which he designed and realized a compact and high-rate optical correlator. Since 2000, he has been a professor of telecommunications and signal processing at ISEN-Brest. At ISEN, he founded the L@bISEN. His research interests deal with optical engineering, optical information

processing, signal and image processing, telecommunications, and optoelectronics. He supervises several PhD, MSc, and engineering school students. He has published over 82 refereed journal articles or conference papers on these topics. He organized and coorganized many conferences and special sessions. He is a member of SPIE, OSA, an elected member of IoP, and IEEE senior member.



Christian Brosseau received his PhD in physics from Fourier University, Grenoble, France, in 1989. After holding a postdoctoral fellow position at Harvard University, he returned to Fourier University to become a research associate. He became an associate professor, professor, and then a distinguished professor at the University of Brest, France, in 1994, 1997, and 2010, respectively. He led the wave-matter interaction modeling and simulation group in the Physics

Department and supervises PhD students and postdoctoral research associates. He has published over 140 refereed journal articles on a wide variety of theoretical and experimental topics and has presented over 90 papers at conferences. He is an OSA Fellow, a Fellow of the Electromagnetics Academy, and a Fellow of the Institute of Physics. He has also written the book, *Fundamentals of Polarized Light: A Statistical Approach* (Wiley, 1998), and edited the book, *Prospects in Filled Polymers Engineering: Mesosstructure, Elasticity Network, and Macroscopic Properties* (Research SignPost, 2008). His research interests include polarization and coherence optics, image processing, all aspects of computational electromagnetics, and electromagnetic wave propagation in complex media.