

Journal of Biomedical Optics

BiomedicalOptics.SPIEDigitalLibrary.org

Blinking supervision in a working environment

Bernardo Morcego
Marc Argilés
Marc Cabrerizo
Genís Cardona
Ramon Pérez
Elisabet Pérez-Cabré
Joan Gispets

Blinking supervision in a working environment

Bernardo Morcego,^{a,*} Marc Argilés,^b Marc Cabrerizo,^a Genís Cardona,^b Ramon Pérez,^a Elisabet Pérez-Cabré,^b and Joan Gispets^b

^aUniversitat Politècnica de Catalunya, Advanced Control Systems Research Group, Rambla Sant Nebridi, 10, 08222 Terrassa, Spain

^bUniversitat Politècnica de Catalunya, Department of Optics and Optometry, School of Optics and Optometry, c/Violinista Vellsolà, 37, 08222 Terrassa, Spain

Abstract. The health of the ocular surface requires blinks of the eye to be frequent in order to provide moisture and to renew the tear film. However, blinking frequency has been shown to decrease in certain conditions such as when subjects are conducting tasks with high cognitive and visual demands. These conditions are becoming more common as people work or spend their leisure time in front of video display terminals. Supervision of blinking frequency in such environments is possible, thanks to the availability of computer-integrated cameras. Therefore, the aim of the present study is to develop an algorithm for the detection of eye blinks and to test it, in a number of videos captured, while subjects are conducting a variety of tasks in front of the computer. The sensitivity of the algorithm for blink detection was found to be of 87.54% (range 30% to 100%), with a mean false-positive rate of 0.19% (range 0% to 1.7%), depending on the illumination conditions during which the image was captured and other computer–user spatial configurations. The current automatic process is based on a partly modified pre-existing eye detection and image processing algorithms and consists of four stages that are aimed at eye detection, eye tracking, iris detection and segmentation, and iris height/width ratio assessment. © 2016 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.JBO.21.2.025005]

Keywords: automatic eye blink detection; blinking frequency; image processing algorithm; computer vision syndrome; video display terminal; dry eye; webcam.

Paper 150475R received Jul. 15, 2015; accepted for publication Jan. 6, 2016; published online Feb. 2, 2016.

1 Introduction

Users of visual display terminals (VDTs) commonly complain of visual fatigue after prolonged work in front of the computer. The term “computer vision syndrome” was coined to describe the diverse symptoms reported by computer users including eye-strain, tired eyes, irritation, a burning sensation, dry eye, redness, blurred far vision, and double vision.¹ Among these, dry eye is the most frequently reported ocular complaint of VDT users.^{2,3} Computer use has been associated with an alteration of the blinking patterns and with a larger palpebral aperture, which is influenced by screen position. The joint contribution of both factors results in a greater exposure of the ocular surface to the environment and in an increased tear film evaporation and instability, leading to dry eye symptomatology.⁴

Spontaneous eye blink rate (SEBR), which is usually measured in blinks per minute (blinks/min), has been found to be a very sensitive parameter to changes in the cognitive demands. For instance, SEBR was observed to have increased from 4.5 blinks/min while reading to 17 blinks/min at rest, with a further increment to 26 blinks/min during conversation.⁵ Similarly, several authors described a sharp decrease in SEBR when subjects perform a highly demanding task with the computer. Indeed, Skotte et al.⁶ noted a change in SEBR from 16 to 5 blinks/min when comparing passive (watching a film) to active computer tasks (this required subjects to connect a sequence of small dots on the screen). Similarly, Himebaugh et al.⁷ evaluated SEBR while participants conducted a series of low to high levels of concentration tasks on VDT (looking at a blank computer screen or watching a film and playing a

computer game or viewing a series of rapidly changing letters, respectively). They observed a comparatively reduced blinking rate during the high concentration activities, in addition to a higher level of fluctuation in SEBR values, particularly during the computer game trial.

2 Background

Given its widespread application in multiple fields of science, psychologists, psychiatrists, ophthalmologists, and neurophysiologists have studied blinking of the human eye for decades. Some authors used electro-oculography for this purpose,⁶ that is, a relatively complex technique, not easily applicable for blink monitoring in a real-life working environment. However, in more recent times, the incorporation of cheap integrated cameras in computers suggests the possibility of using image processing techniques for the evaluation of SEBR instead of using other more invasive or intrusive methods. Several efforts have subsequently been conducted to develop automatic blink detection strategies.

Won et al.⁸ described a blinking detection algorithm based on binary images. The binarization threshold is critical and depends on the illumination conditions of the image; that is, it requires a previous normalization process whereby the threshold is automatically determined.⁹ Won used two features to detect whether the eye was open or closed: first, consecutive frames were compared to determine the number of cumulated black pixels, since in the open eye conditions the presence of the iris/pupil leads to a greater number of black pixels; second, the relation between iris height and width was measured. These two factors were

*Address all correspondence to: Bernardo Morcego, E-mail: bernardo.morcego@upc.edu

combined using a support vector machine to determine the frames, and thus the time during which the eyelids were closed.

Similarly, Jiang et al.¹⁰ were able to detect the beginning and the end of an eye blink. The difference between two consecutive frames was binarized, and morphological operations were employed to determine the presence of the iris. The detection of the iris was based on dimension parameters requiring the definition of several thresholds, which needed to be optimized in advance. With optimal values for these thresholds, the authors reported true-positive rates (TPRs) of 90.3% and false-positive rates (FPRs) of 0.1%; this is an accuracy of 99.7% using their technique. A level of precision of 66% was reported by Tan and Zhang,¹¹ in their proposed method for iris detection through pattern recognition, which was subsequently improved to take into account the different configurations resulting from the actual position of the iris with respect to the maximal response zone.¹² With this method, the authors reported an accuracy of 88% in blink detection.

Finally, Mitelman et al.¹³ developed the semiautomatic eye state detection algorithm, with which the authors were able to detect the differences between open and closed eye conditions by examining the corresponding brightness and the frequency distribution of the image. This method, which requires a training process to define several thresholds, relied on brightness peaks arising from the iris and the pupil regions. Later, Bernard et al.¹⁴ implemented an accurate image processing analysis to detect the two lines that correspond to the margins of the eyelids, whereupon the distance between these two lines was monitored to identify eye blinks.

3 Blinking Supervision with Image Processing

The blink counting algorithm that was developed in this study consists of a combination of known image processing algorithms, with the addition of a new algorithm which was inspired by the work of Jiang et al.¹⁰ The present algorithm is conceptually divided into two tasks: eye segmentation and blink counting.

The first task, eye segmentation, is to carry out two key procedures: eye detection and eye tracking. This combination of procedures improves the efficiency, in terms of the actual computing time required for eye detection in each frame, since eye tracking only requires a portion of the image to operate. The second task also involves the combination of two algorithms: iris detection and the iris height/width ratio evaluation. Redundancy was introduced to improve the accuracy and to avoid false blinking detection results (false positives). Blinks are only counted when both the algorithms detect a blink within the same set of consecutive frames. A detailed description of these algorithms is given in the following sections.

3.1 Eye Detection

For eye detection, the rapid object detection algorithm developed by Viola and Jones¹⁵ is applied. This algorithm was first created to identify faces that are in an image, using a learning cascade feature detector. The implementation in MATLAB[®] (MathWorks, Inc., Natick, Massachusetts) of this algorithm can detect eyes, mouths, and noses, and it may also be trained to detect other user-defined objects (facial features). The algorithm works by locating the left eye of the subject in the first frame, whereupon the eye tracking algorithm becomes active until the eye has disappeared (see Fig. 1). At that time, the cascade

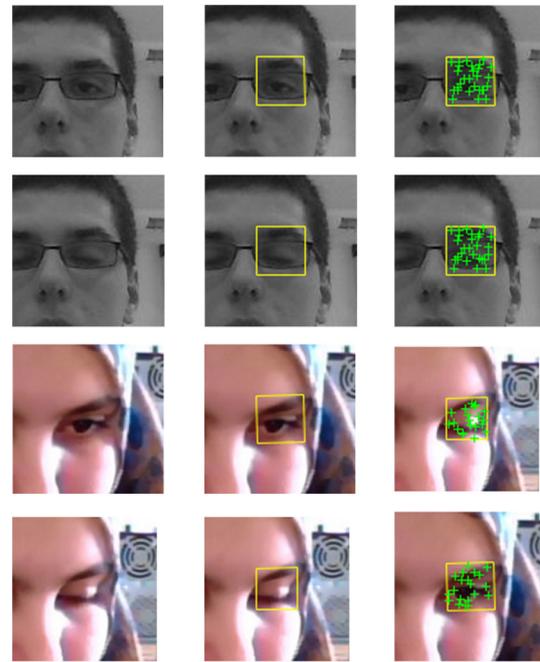


Fig. 1 Eye segmentation. Each row shows the result of the algorithms on the sample frames. From left to right: original image portion; Viola and Jones eye detection algorithm, which is highlighted with a yellow square; and Kanade–Lucas–Tomasi eye tracking algorithm, with calculated features highlighted with green crosses.

learning feature detector recommences and the process continues. It must be noted that blinking does not interfere with eye detection. In fact, even if the iris is lost during the interval of a blink, the eye is still detected.

3.2 Eye Tracking

After the eye is detected, the region where the eye is located is used as the input region for the Kanade–Lucas–Tomasi feature tracking algorithm.^{16,17} The MATLAB[®] implementation of this algorithm is very thorough, including the tracking features, and it also allows updates in regard to the size and the location of the feature search space, according to the relation between the positions of these features in consecutive frames.

In the present application of the algorithm, the configuration that is considered optimal is as follows: the number of pyramid levels where the tracking points are looked for is 4; the maximum bidirectional error, which is a parameter to help check good tracking points and eliminate uncertain ones, is 2; the maximum number of search iterations is 40; and the type of assumed transformations between frames is similarity, which means changes in scale, position, and orientation of the object are allowed.

Finally, provided that the eye is successfully tracked from one frame to the next one, the region where the eye is located is used as the input region for blink detection. Blinks are counted in the next two algorithms.

3.3 Iris Detection

The aim of this algorithm is to identify and segment the iris in each frame so that when the iris is not detected, the algorithm assumes a blink has taken place. This algorithm is inspired by the work of Jiang et al.,¹⁰ although several important

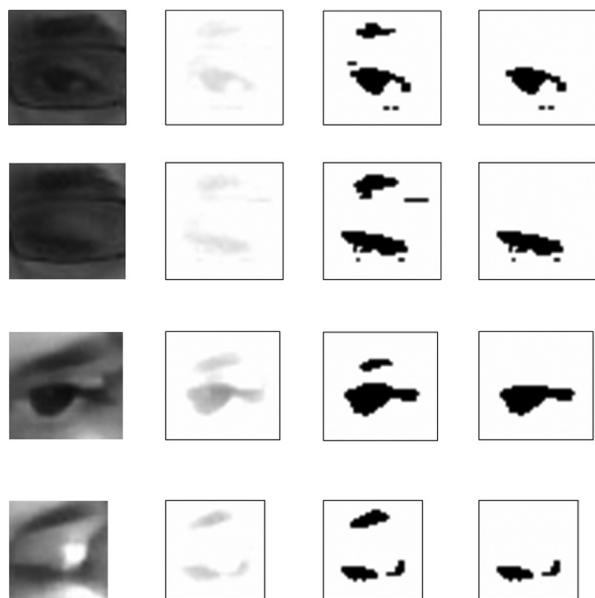


Fig. 2 Binarization process. Each row shows the result of the algorithms on the same sample frames as in Fig. 1. From left to right: original image; luminosity normalization; binarization; and eyebrow masking.

modifications were implemented. First, the luminosity of the image is normalized using a 31×31 pixel median filter, after which the Otsu¹⁸ optimal threshold binarization is applied. Finally, the eyebrow is erased with a mask and the borders are removed (to eliminate fortuitous portions of glasses, hair, and so on) (see Fig. 2).

At this point of the process, the image contains a black eye shape with fragments of the eyelids and, less frequently, of the eyebrows. In order to remove all but the iris (see Fig. 3), an opening process^{9,18} is applied with a disk structuring element of four pixels of radius. This operation keeps the iris and sometimes other round-shaped elements. Thereupon, the image is labeled, and only the largest object is kept. Provided that the iris is the only visible shape, an erosion process^{9,18} is subsequently applied, which results in the homogenization of the shape of the iris.

Finally, the Hough transform^{9,18} is used to detect circular shapes in the image and thus to segment the iris. If one circular shape is detected, a no-blink condition is registered. Conversely, when a circle is detected in a frame but lost in the following one, the algorithm registers the beginning of a blink. The sensitivity of the circular Hough transform, which is set to 0.82, determines whether the shape under consideration is circular or not.

It must be noted that this algorithm is not perfect. Indeed, in some cases, the iris is not properly segmented from the surrounding anatomical structures, such as the margins of the eyelids, thus failing to detect a circular shape, which leads to a false blink count (false positive). Therefore, as noted previously, the iris detection algorithm is combined with a second algorithm based on the iris height/width ratio evaluation in order to improve the accuracy in blink detection.

3.4 Iris Height/Width Ratio Evaluation

This algorithm computes and compares the width and height of the iris. It is based on the work of Won et al.,⁸ although only the iris is used in the present modification of the algorithm, whereas

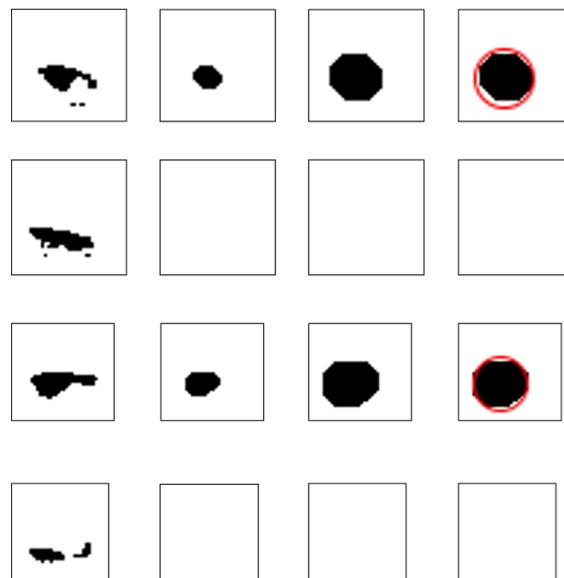


Fig. 3 Iris detection and segmentation is a step-by-step process. Each row shows the result continued from the samples of Fig. 2. From left to right: initial image; opening operation; the largest object preservation and homogenization; and the detection of circular shapes.

these authors assessed the entire eye. Iris width and height were selected since their ratio suffers significant changes during an eye blink.

The maximum horizontal width and vertical height of the iris are measured from the image obtained in the last step of the iris detection algorithm. Assume the algorithm is processing frame j . Since the image contains a round-shaped object, the first column, starting left, that contains a black pixel is the leftmost end of the iris, $l(j)$. Similarly, the rightmost, top, and bottom ends are identified, $r(j)$, $t(j)$, and $b(j)$. The height–width ratio of frame j is

$$w(j) = \frac{b(j) - t(j)}{r(j) - l(j)}.$$

An adaptive threshold, $\bar{w}(j)$, is calculated from the previous N frames as follows:

$$\bar{w}(j) = \frac{1}{N} \sum_{i=j}^{j-N} w(i),$$

where N was empirically set at 10 frames. Once the adaptive threshold is obtained, this value is compared with the current height–width ratio to determine the presence of a blink:

$$\text{blink}(j) = \begin{cases} \text{true} & \text{if } w(j) < K_{th} \bar{w}(j) \text{ and } w(j-1) > K_{th} \bar{w}(j-1) \\ \text{false} & \text{otherwise} \end{cases}.$$

It must be noted that, to ensure the correct detection, this equation contains the parameter K_{th} , which needs to be adjusted depending on the illumination conditions, the camera configuration, the distance to the subject, and other factors too. K_{th} is a value between 0 and 1, typically around 0.9.

4 Algorithm Testing (Preliminary Results)

Preliminary trials revealed that the current version of the algorithm was not fast enough to be useful for real-time video stream monitoring. Consequently, in its current state of development, it was only used with recorded videos.

The algorithm was tested on 17 one-minute videos of subjects undertaking different actions on personal computers (reading texts, playing games, browsing the web, and so on). All participants provided written informed consent after the nature of the study was explained to them.

A variety of illumination conditions, working distances, face configurations (with and without glasses), skin tones, and webcam resolutions were included in the preliminary trials to assess the performance of the algorithm in less than ideal conditions, albeit closer to real-life ones. Each video was manually revised to determine the true blink count, and then this value was compared with the value obtained by our algorithm to calculate true blink positives (TP), false blink positives (FP), true blink negatives (TN), and false blink negatives (FN). Furthermore, true-positive [TPR = $TP/(TP + FN) * 100$] and false-positive rates [FPR = $FP/(FP + TN) * 100$] were determined to compare the performance of our algorithm to that of previously described algorithms.

After reviewing the 17 one-minute videos, the total number of blinks was 269, with a range from 2 to 45 per video. The mean TPR was 87.54% and the mean FPR was 0.19%, in accordance with the published report by Jiang et al.¹⁰ The range of TPR was from 30% to 100%, and the range of FPR was from 0% to 1.7%. It must be noted that the worst values of TPR and FPR corresponded to the combination of very challenging illumination conditions, dark skin tone with features of interest that were more difficult to discriminate from background, and low camera resolutions, resulting in video captures in which it was very difficult to observe the eye of the participants. In contrast, TPR and FPR values were close to 100% and 0%, respectively, provided that illumination conditions were close to those recommended by ergonomic standards such as, for example, ISO 9241-6,¹⁹ which notes that the average room illumination should be between 320 and 600 lx, uniform, and without large differences between the surrounding environment and the workstation. In addition, a webcam resolution of at least 720 pixels was considered a requirement for quality image acquisition. Given that these minimum criteria were met, it was found that the parameter K_{th} did not need for further adjustments prior to video analysis.

For instance, good blink detection conditions are shown in Fig. 4, in which the subject blinked 19 times. Although our algorithm slightly overestimated the number of true blinks, all real blinks were detected, resulting in a TPR of 100% and a FPR of 0.52%. Conversely, Fig. 5 depicts a more challenging situation,



Fig. 4 Blink detection in optimal conditions (TPR: 100%; FPR: 0.52%).

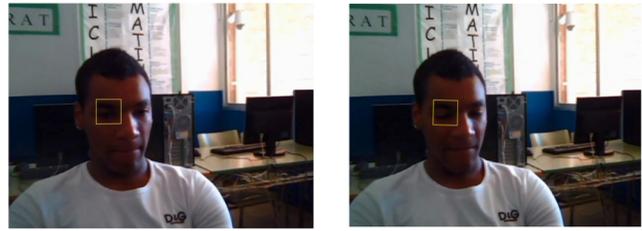


Fig. 5 Blink detection in challenging conditions (TPR: 30.43%; FPR: 0.06%).

in which the combination of darker skin tone, low webcam resolution, and unsatisfactory illumination conditions leads to an underestimation of true blinks (only 9 of 23 real blinks were successfully detected), with TPR and FPR rates of 30.43% and 0.06%, respectively.

5 Conclusion

The present research aims to develop and implement an algorithm for automatic blink detection and counting. Preliminary trials on recorded videos show good sensitivity of the algorithm to detect blinks, provided that normal illumination conditions and webcam resolutions are present. Given the relevance of blink frequency in the visual fatigue symptoms experienced by most computer users, noninvasive and nonintrusive blink monitoring strategies are a first step toward developing biofeedback mechanisms for blink re-education. The innovation of the present algorithm relies on requiring the configuration of only one parameter, K_{th} , which may be kept constant if the workplace has normal illumination conditions, and on being functional on most computer-integrated webcams, thus supporting the need for further research to advance its implementation on other ubiquitous devices, such as tablets and smart phones. Further research is being carried out to make the algorithm operational on real-time video streaming and with standard computing languages and tools. An application incorporating biofeedback for blinking re-education is currently under development.

Acknowledgments

Authors were grateful to all the staff and students of the Secondary School Josep Lladonosa of Lleida, Spain, who have participated in the video recording process to collaborate in this research study. E. Pérez and G. Cardona thank the Spanish Ministerio de Economía y Competitividad and Fondos FEDER for financial support (Project Number DPI2013-43220-R).

References

1. C. Blehm et al., "Computer vision syndrome: a review," *Surv. Ophthalmol.* **50**, 253–262 (2005).
2. J. R. Hayes et al., "Computer use, symptoms, and quality of life," *Optom. Vision Sci.* **84**, E739–E745 (2007).
3. M. Uchino et al., "Prevalence of dry eye disease among Japanese visual display terminal users," *Ophthalmology* **115**, 1982–1988 (2008).
4. K. Tsubota and K. Nakamori, "Effects of ocular surface area and blink rate on tear dynamics," *Arch. Ophthalmol.* **113**, 155–158 (1995).
5. A. R. Bentivoglio et al., "Analysis of blink rate patterns in normal subjects," *Mov. Disord.* **12**, 1028–1034 (1997).
6. J. H. Skotte et al., "Eye blink frequency during different computer tasks quantified by electrooculography," *Eur. J. Appl. Physiol.* **99**, 113–119 (2007).

7. N. L. Himebaugh et al., "Blinking and tear break-up during four visual tasks," *Optom. Vision Sci.* **86**, E106–E114 (2009).
8. O. L. Won, C. L. Eui, and R. P. Kang, "Blink detection robust to various facial poses," *J. Neurosci. Methods* **193**, 356–372 (2010).
9. R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 2nd ed., Prentice Hall, Upper Saddle River, New Jersey (2002).
10. X. Jiang et al., "Capturing and evaluating blinks from video-based eyetrackers," *Behav. Res. Methods* **45**, 656–663 (2013).
11. H. Tan and Y.-J. Zhang, "Detecting eye blink states by tracking iris and eyelids," *Pattern Recognit. Lett.* **27**, 667–675 (2006).
12. Y. Tian, T. Kanade, and J. F. Cohn, "Dual-state parametric eye tracking," in *Proc. Fourth IEEE Int. Conf. Automatic Face and Gesture Recognition*, pp. 110–115 (2000).
13. R. Mitelman et al., "A noninvasive, fast and inexpensive tool for the detection of eye open/closed state in primates," *J. Neurosci. Methods* **178**, 350–356 (2009).
14. F. Bernard et al., "Eyelid contour detection and tracking for startle research related eye-blink measurements from high-speed video records," *Comput. Methods Programs Biomed.* **112**, 22–37 (2013).
15. P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proc. 2001 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR 2001)*, pp. 511–518, IEEE, Kauai, Hawaii (2001).
16. B. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Int. Joint Conf. on Artificial Intelligence*, pp. 674–679 (1981).
17. C. Tomasi and T. Kanade, "Detection and tracking of point features," Technical Report CMU-CS-91-132, Carnegie Mellon University (1991).
18. R. Szeliski, *Computer Vision: Algorithms and Applications*, Springer, New York (2010).
19. European Standard EN, ISO 9241-6, "Ergonomic requirements for office work with visual display terminals (VDTs)—part 6: guidance on the work environment" (1999).

Bernardo Morcego is an associate professor at the Universitat Politècnica de Catalunya (UPC). He received his PhD in computer science from the UPC in 2000. He has been teaching several subjects in the area of automatic control in the schools of engineering and aeronautics in Terrassa and Barcelona. He is a member of the Research Center for Supervision, Safety, and Automatic Control of UPC. His research interests include UAV control systems and computer vision applications.

Marc Argilés graduated in optometry in 2009 and was awarded a Master of Science in vision in 2011 by the UPC. He is currently

undertaking a PhD in optical engineering at UPC, with research interests related to dry eye symptoms, computer vision syndrome, and the relationship of ocular blinking and cognition. He is actively involved in various projects linking new visual display terminals with related vision problems.

Marc Cabrerizo received his degree in industrial electronics and automatic control from the UPC, Terrassa, Spain, in 2014. He is currently working as PLC and robot programmer in a Spanish manufacturer of packaging machinery (SR INNOVA) located in Barcelona, Spain.

Genís Cardona received his degree in optometry from the UPC in 1992, and MSc and PhD degrees from the University of Manchester Institute of Science and Technology, UK, in 1994 and 1996, respectively. He is currently employed as a full-time lecturer at the Department of Optics and Optometry in the UPC. His research interests include ocular surface and tear film, contact lenses, refractive surgery, blinking, and intraocular lenses.

Ramon Pérez received his MSc degree in physics from the University of Barcelona in 1993 and his PhD in physics from UPC, Terrassa, Spain, in 2003. He currently holds a position as a lecturer at the Department of Automatic Control in the same university. His current research interests include control and supervision, particularly focused in water systems. He is a part of the Advanced Control Systems (cs2ac) research group at the UPC and of the Technological Center of Manresa (CTM).

Elisabet Pérez-Cabré received her PhD in physics from the UPC in 1998. Her academic activities at the School of Optics and Optometry in UPC involve lecturing, mainly on fundamental optics and optical instruments. Her current research interests include encryption techniques, programmable diffractive optical elements, and biomedical optics. She is a senior member of the International Society for Optical Engineering (SPIE). She is also a member of the Spanish Optical Society (SEDOPTICA) and the European Optical Society (EOS).

Joan Gispets was awarded his degree in optometry from the UPC in 1992, his MSc degree in optometry and vision science from the University of Manchester in 1993, and his PhD from UPC in 2009. He has been a faculty member at the Department of Optics and Optometry at UPC since 1995. He is currently dean of the faculty. His research interests are related to contact lenses, keratoconus, noninvasive diagnostic techniques, and myopia.