



# Blink and wink detection as a control tool in multimodal interaction

Piotr Kowalczyk<sup>1</sup> · Dariusz Sawicki<sup>1</sup> 

Received: 14 January 2018 / Revised: 3 August 2018 / Accepted: 15 August 2018 /  
Published online: 25 August 2018  
© The Author(s) 2018

## Abstract

The problem of multimodal interaction is discussed. The use of blinking and winking, interpreted as “eye gestures,” is considered. The main aim of this study is to propose a simple method that allows the recognition of the state of the eye: open or closed; and to distinguish between blinking and winking. Wearable technology has been used in the introduced solution. Placing the camera close to the eye allows us to simplify the complicated image analysis. The proposed method works irrespective of the user’s location and his/her gaze direction. Further, the use of infrared radiation limits the influence of external disturbing factors such as lighting conditions and pollution. The new solution is tested in two types of experiments: with  $2 \times 5000$  pictures of open and closed eyes and with a group of 30 participants. The total correctness of the eye state recognition is 99.68% (99.94% for open eyes and 99.42% for closed ones). This result implies that the proposed solution can be effectively applied to real-world scenarios. Two applications are considered. In the first one, blinking recognition allows us to check whether safety glasses are used appropriately. In the second application, the control of mouse keys is replaced with eye gestures interpreted from a winking analysis. The introduced solution allows for effective and correct eye state recognition. The new method of control by eye gestures was accepted by the participants of this study.

**Keywords** Multimodal interaction · Blinking · Winking · Eye state recognition

## 1 Introduction

Multimodal interaction is a method of communication that plays an increasingly important role in the technical applications of everyday life. From year to year, new devices use more and more interesting methods of control and communication: touch, gestures, and body language; what else does the future hold? The use of eye gestures (as blinking/winking interpretation) is a relatively new form of communication, but with “visible” potential.

---

✉ Dariusz Sawicki  
Dariusz.Sawicki@ee.pw.edu.pl

<sup>1</sup> Warsaw University of Technology, Warsaw, Poland

Blinking is a simple physiological activity to which we do not attach much significance. However, this activity can be very useful in multimedia applications. Supporting communication with a computer for disabled or paralyzed people is a difficult but very important task. For these people, blinking/winking can be an effective way of transmitting information or controlling devices. Certain advanced applications of this method of communication have also been reported thus far. In contrast, winking can be an additional control element in complex systems, for game players, and in situations where the users have busy hands. A computer mouse is the most popular pointing device today, which allows not only the selection of an appropriate object but also the generation of an appropriate system event. To control such events, the mouse is equipped with a few (1 to 3) keys. An application that replaces the clicking of the mouse keys by eye gestures can be considered. Such an application would be universal, but the basic problem is the correct recognition of the user's winking. We can also consider a different application where the registration of blinking can be a sign of human presence. Such identification can sometimes be simpler than the analysis of a thermogram or a search for other signs of life such as heart rate or periodic local changes in the level of carbon dioxide. In both these situations, the success of the application depends on the correctness of the recognition of the closing and opening of the eyes.

## 2 Motivation and aim of the article

There are many known solutions that allow for moving the cursor without using hands. For this purpose, the identification of head movements based on the head image [23] or facial image analysis [14] is often used. It is also possible to use additional sensors, e.g. accelerometers [27], or gyro sensors [17]. However, cursor movement control is not enough to replace a computer mouse. The basic motivation for undertaking the research is the need for effective and hands-free generation of system events that are normally provided by the mouse keys. We chose to use blink/wink to solve the problem. It is the simplest action that is always available and does not interfere with the tasks in most situations. The main aim of this study is to develop a simple method that allows the recognition of the state of the eye (open/closed) and to distinguish between blinking and winking. We also consider two applications where the blink and wink action plays the decisive role. In such a situation, the eye state recognition should work rapidly, effectively, and correctly irrespective of the environment and the external factors.

## 3 Recognition of the eye state and multimodal interaction

Face and eye images as well as gaze tracking are analyzed in various applications of multimodal interactions. Face recognition and the analysis of closing and opening eyes are presented in [35]. The authors proposed a nonstandard control of computers for the disabled. Mandal et al. [26] performed a complex analysis (skin color, shape, etc.) to recognize the position and movement of the head. Jaimes and Sebe [13] presented a general review of methods used in multimodal human–computer

interactions. They also described the use of face image analysis and gaze tracking. Blinking itself can be the basis of various applications. Grauman et al. [11] proposed a method of communication for the disabled based on blinking and eyebrow movement. Driver's fatigue, drowsiness, and degradation of concentration can be detected by analyzing the blinking parameters [19, 21].

The visual analysis of a facial image is the most popular method (also the oldest one) for blinking detection. Such an image can be captured with a camera. The face shape in the first step and the eye region in the second are separated. Since then, various methods have been used for blinking detection: statistical analysis [2], pupil detection in the image [18], or image comparison to the templates [10]. Le and Liu in [22] built an open/closed classifier based on a feature vector obtained using a principal component analysis. To make the image analysis independent of the external lighting conditions, infrared (IR) radiation is sometimes applied. Kapoor and Picard [15] used IR light-emitting diodes (IR LEDs) and an IR camera for the head position (and movement) analysis.

Blinking can be detected in a simple way by using the eye tracking method. Eye tracking (oculography) is used in many fields of research and interaction applications. Singh and Singh [33] presented a review of eye tracking implementations related to multimodal interactions. A review of methods used in eye tracking is presented in [1, 29]. Eye tracking also allows the recognition of eye blinking as part of the standard task, but the use of eye tracking only for such a purpose would be ineffective.

Blinking can be also recognized with other techniques: the use of surface electromyography (sEMG) for the identification of appropriate muscle tension [5] or the use of electrooculography (EOG) [20].

On the other hand, multimodal interaction (with which our problem is related) is a wider issue. It touches the recognition of human activities during various types of interactions [7, 24, 25]. A review of various solutions with regard to many aspects and comparative analysis is not the purpose of our article. However, it is worth to refer our proposal to similar solutions. We present such a concise comparison in the scope of the considered solutions in summary of this article.

## 4 Eye physiology: Blinking and winking

Eyesight is the most complex sense of humans. The eye ensures the reception of visual impressions; it is a complex instrument, but a very delicate one. The eyelids (upper and lower) serve a very important protective function. During the normal functioning of the sense of sight, blinking ensures the appropriate wetting of the cornea of the eye through an even distribution of tears. At the same time, this mechanism causes the natural removal of fine impurities from the surface of the cornea. Under difficult ambient conditions, closing the eyelids protects against dust and other impurities. The closing of the eyelids blocks (in fact significantly reduces) the amount of light during sleep. The physiology of the eye and the blinking mechanism have been well known for a long time and have been documented [9, 16, 30, 36]. Depending on individual predispositions and external conditions, a human being blinks approximately 6–30 times/min (on average, approximately 16 times/min) [20]. Mental and emotional activities affect the frequency of blinking. Talking and

verbal engagement, as well as anger, excitement, stress, fear, and fatigue, can increase the frequency by several times. In contrast, reading or absorbing visual work (the concentration of drivers, for example), and above all, working at the computer, reduces the frequency of blinking. Environmental factors have an independent influence. Low humidity, cigarette smoke, and pollution naturally stimulate the blinking mechanism to work more frequently.

A single blink can be divided into three phases: fast eyelid closing, closure state, relatively slow (approximately two times slower than closing) opening [20]. The full cycle of a single blink takes from 100 ms to 400 ms [32]. It is often assumed that a longer time of closing the eye (approximately 1000 ms) means micro-sleep and is most often associated with fatigue. Modern research [28] indicates the relationship between the eyes' closed state (blinking) and temporary rest of the brain. The natural blink includes synchronously both the left and the right eye and is a physiological activity that does not require control, although humans can consciously extend the period between blinks. In contrast, a human being can close his/her eyelids in an intended manner—as intentional winking. Winking is the conscious closing of the eyelids: in a conscious way, we can close independently the left and the right eye and control the closing time. For a certain group of people, winking with one eye can be very difficult. The work of the eyelids is controlled by a complex muscle system. Further, the *oculi* and the *levator palpebrae superioris* muscle are associated with the upper eyelid, and the *inferior palpebral* muscle is associated with the lower eyelid [20]. Such muscles can be identified using sEMG.

## 5 Recognition of the eye state—The main idea

The expected solution requires the consideration of various aspects. On the one hand, we took into account the solutions reported in previous studies. On the other hand, we analyzed the physiological properties of the eyes and the versatility of the considered applications. Therefore, we propose the following assumptions for our solution of the blink and wink recognition:

- It should allow us to distinguish blinking from winking on the basis of the analyzed eye closure time.
- It should work correctly with any position of the head and any gaze direction.
- It should work correctly when the face is obscured (partially and almost completely).
- It should work effectively in real time, so the method should be simple and fast.

The simplest solution, known from past research, would be to use a set of cameras, implement an appropriate face recognition algorithm, separate the eye image, and finally, analyze the eye state. However, such a solution does not meet our assumptions, and under real conditions, it would be very difficult to implement and ineffective. To simplify the issue, we concentrate on the wearable technology [6]. Placing the camera locally very close to the eye allows us to limit the image only to the image of the eye (or its fragment). This solution allows us to simplify the complicated image analysis as much as possible. Further, the solution works irrespective of the user's location and his gaze direction. Additionally, we propose the use of IR radiation and IR cameras to limit the effect of external disturbing factors (lighting conditions, pollution, etc.). The application of wearable technology should meet an additional requirement: the used sensor (camera, etc.) cannot obscure the field of view.

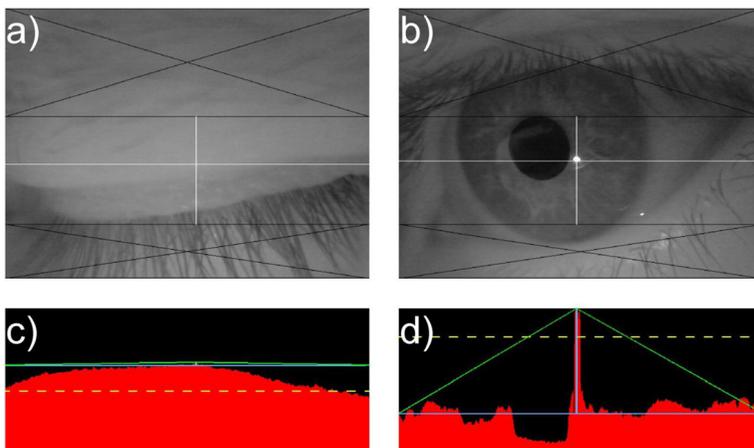
The images of closed and open eyes differ significantly. In the first case, we see the skin of closed eyelids, and in the second case, the pupil, iris, and sclera (white of the eye). Human skin has a very characteristic color [3, 31], which in combination with the white color of the eye allows us to effectively distinguish between eye states (closed/open), in both daylight and artificial light, even from a distance. Such an application of skin color recognition would be very interesting. However, a preliminary study has shown that the color analysis for the diagnosis of eye closure is very complicated.

We propose the determination of the state of the eye on the basis of the analysis of the light reflection in the eye image [8]. The light that falls on any surface (also the surface of the human body) is partially reflected [3, 8]. We can distinguish between specular (directional) or diffused (scattered) reflection depending on the type of surface. The surface of the eyelid (skin) reflects in a diffused manner, and the eyeball reflects, primarily, directionally. To carry out the analyses irrespective of the external lighting conditions, we decided to use IR radiation (instead of visible light). In addition, it prevented the occurrence of glare, which could interfere with the user's work. In our solution, a micro camera recorded the image of the eye that was illuminated by an IR LED. The IR radiation reflected from the eyeball (when the eye was open) was clearly visible (Fig. 1b). In this case, we observed a small but very bright point of reflection. The IR radiation reflected from the eyelid (if the eye was closed) did not create such a visible bright point of reflection (Fig. 1a) because the human skin had very strong diffusion properties.

Therefore, we could very easily determine the state of the eye only on the basis of the reflection analysis. If a bright point of specular reflection was observed, the eye was open; if not, the eye was closed.

## 6 Algorithm of the reflection analysis

We implemented the algorithm for the analysis of the images recorded by the micro cameras in three steps:



**Fig. 1** Captured image of the left eye: (a) closed and (b) open. The horizontal line passing through the brightest pixel is marked. Brightness distribution in this line: (c) closed eye and (d) open eye

- 1. Determination of the point of light reflection in the image.** In the first version of the method, we observed that most often, the point of reflection was represented as a group of bright, neighboring pixels. Therefore, we searched for a group of pixels by analyzing the brightness difference between neighboring pixels. We analyze images stored in 8 bits of gray scale - which allows considering 256 levels of luminance from 0 (black) to 255 (white). In the image of the eye (Fig. 1b) we selected points with the highest luminance. If there are several such neighboring points, they form the considered group. If there is only one such point, among the neighboring points we choose those that have a luminance level by 1 degree lower. These points together with the brightest point form the considered group. Then, we determined the exact point/pixel of the reflection as the center of gravity. For several bright groups, we selected the brightest one. After conducting experiments on a set of eye images (opened and closed), we simplified the method. We found it considerably easier and sufficiently effective to simply look for the brightest pixel in the eye image. Additionally, because we did not have to search the entire image, we limited the search area. The excluded areas are marked (crossed) in Fig. 1a and b. Excluded area means an area where the camera does not record the reflection. It depends, simply, on the position of the camera and the IR diode. We determined it experimentally.
- 2. Determination of the brightness profile for the image.** After the first step, we found the brightest point in the image. Through this point, a horizontal line (of one pixel width) was analyzed. For each point on this line, the brightness of the image was determined. Then, we generated a brightness profile as a graph based on the brightness level of each pixel from this line (Fig. 1c and d). The 1-byte grayscale was used; therefore, for each point of the brightness profile, a line segment of up to 255 pixels was obtained.
- 3. Determination of the eye state.** In this step, we analyzed the local differences in the brightness profile. Starting from the maximum point of the profile, we found the largest decrease in the neighborhood, taking into account the local level differences. The graph for the diffused reflection (Fig. 1c) was characterized by small differences in the luminance levels and a slow increase and decrease in the luminance around the highest point. In this case, the eye was closed. The graph of specular reflection (Fig. 1d) was characterized by large differences in the levels of luminance and a rapid increase and decrease in the luminance around the point of the highest level. In this case, the eye was open. To show the difference between “small differences” (for diffuse reflection) and “large differences” (for specular) we marked triangles with the highest level of luminance in Fig. 1c and d. The base of the triangle corresponds to the level of luminance beyond the highest decrease in the luminance value near maximum. Typical decreases in the luminance value for diffuse are 5–10 of gray levels, for specular approx. 100 or more. The differences between the decreases are so large that any threshold in this range allows distinguishing the eye condition correctly. We adopted threshold = 50. This is indicated in the figures: the dashed line means the decrease of luminance value by 50 in relation to the maximum. The brightness profile was analyzed only to distinguish between two cases (specular or diffused reflection) rather than determine the reflection properties. Therefore, the analysis was simple and fast.

The most important part of the algorithm (determination of the eye state based on the brightness profile in horizontal line) can be described in symbolic form as a pseudo code (pidgin code) as follows:

Let  $LBP(l, i)$  be the luminance level of the brightness profile in line  $l$  for pixel  $i$ . Let  $il\_max$  be the index of pixel where exist the maximum of luminance (point of reflection) in line  $l$ . So  $LBP(l, il\_max)$  be the luminance value in point of reflection. Let  $LTh$  be the threshold assumed in the algorithm. For simplicity, the maximum difference in the levels of luminance is searched for pixels in the range of  $[i - LTh, i + LTh]$ .

```
procedure OPEN_or_CLOSE ( $l, il\_max$ )
```

```
begin
```

```
   $k:=1$ ;  $Lmax:=LBP(l, il\_max)$ ;
```

```
  while ( $k<LTh$ ) do
```

```
    begin
```

```
       $diff\_L:=Lmax - LBP(l, il\_max - k)$ ;
```

```
       $diff\_R:=Lmax - LBP(l, il\_max + k)$ ;
```

```
      if ( $diff\_L > LTh$  OR  $diff\_R > LTh$ ) then return (OPEN);
```

```
       $k:=k + 1$ ;
```

```
    end
```

```
  return (CLOSE);
```

```
end
```

## 7 Analysis of signals generated from eye states

We analyzed the channels for the left and the right eye separately. In each channel, the eye closure signal was recorded and its duration ( $t_x$ ) was measured. To correctly distinguish between blinking and winking, we experimentally determined the longest time of natural blinking ( $t_0$ ) and the shortest time of intentional winking ( $t_1$ ). Taking into account the time  $t_1$ , we experimentally determined two additional times  $t_2$  and  $t_3$ , such that  $t_1 < t_2 < t_3$ . The set of gestures for the left and the right eye is shown in Fig. 2. We did not analyze cases when one eye was closed and the other was winking because doing so is difficult for many people.

Therefore, on the basis of the duration ( $t_x$ ), we propose a simple classification for recognizing eye gestures.

**Natural blinking** Natural eye blinking is characterized by the simultaneous occurrence of very short signals in both channels ( $t_x < t_0$ ), shown as gesture  $B_0$  in Fig. 2.

**Unnatural (non-standard) blinking** Very rarely, but it can so happen that  $t_0 < t_x < t_1$ , both synchronously and asynchronously. This denotes a special case caused by, for example, external factors (air blast, dust, etc.). Such a closing of the eye cannot be included as intentional winking. However, this eye gesture is the closest to natural blinking. Such a situation corresponds to the  $BL_0$  and  $BR_0$  gestures. We analyzed these gestures separately because the durations for the left and right eyes were different.

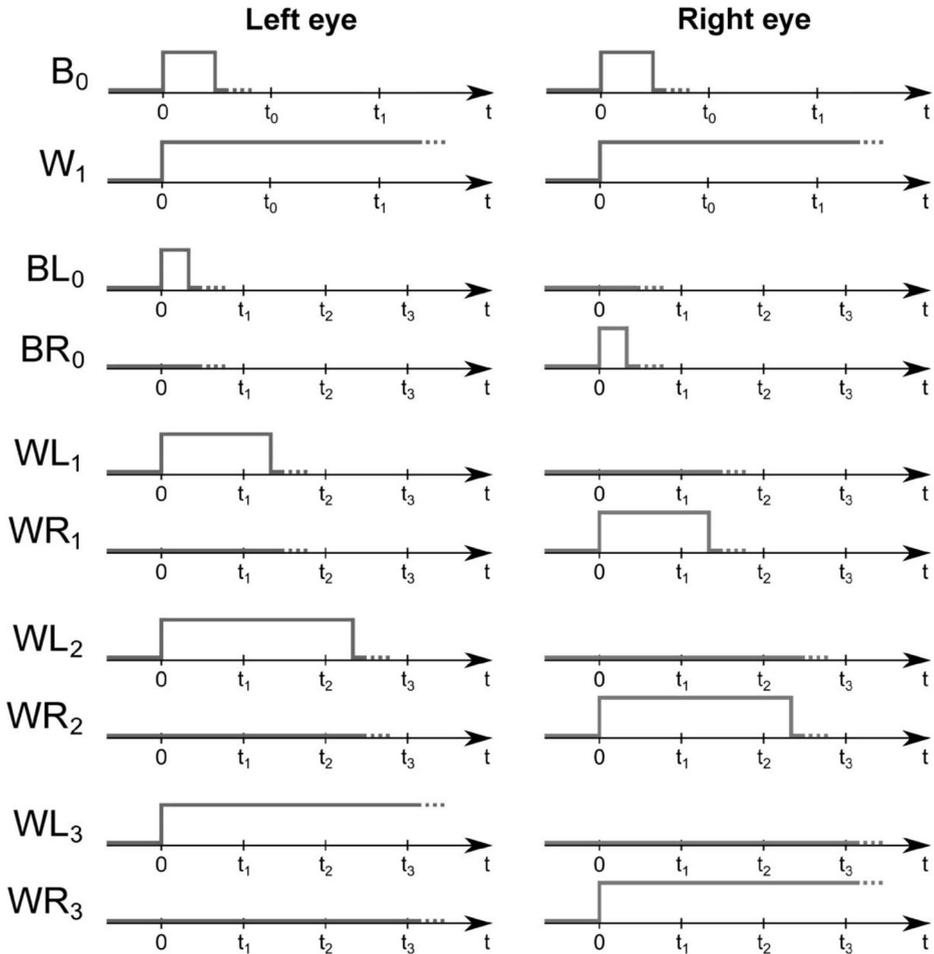


Fig. 2 Eye gestures:  $B_0$  and  $W_1$ ,  $BL_0$  and  $BR_0$ ,  $WL_1$  and  $WR_1$ ,  $WL_2$  and  $WR_2$ , and  $WL_3$  and  $WR_3$

**Simultaneous winking of both eyes** In this case, the user intentionally closes both his/her eyes simultaneously for a certain period of time ( $t_x > t_1$ ). This state is interpreted as gesture  $W_1$ .

**Shorter wink with one eye (the other is open)** If  $t_1 < t_x < t_2$ , then there is a shorter intentional wink. It corresponds to gestures  $WL_1$  and  $WR_1$ .

**Longer wink with one eye (the other is open)** If  $t_2 < t_x < t_3$ , then there is a longer intentional wink. It corresponds to gestures  $WL_2$  and  $WR_2$ . If  $t_3 < t_x$ , then the eye is closed intentionally for a relatively very long time (gestures  $WL_3$  and  $WR_3$ ); this gesture is strictly not a wink.

The values of time parameters were determined experimentally, taking also into account result of other studies concerning blinking mechanism [9, 16, 20, 32]. Parameter values are as follows:  $t_0 = 0.4$  s,  $t_1 = 0.5$  s,  $t_2 = 1$  s,  $t_3 = 3$  s. Such parameters have been proposed to all participants. Most of them mastered the using of eye gestures after a few minutes of exercises. A few participants, having difficulties in winking with one eye, asked for extending the time. After the individual correction, the participants worked correctly.

## 8 Prototype and technical solutions

We developed and manufactured the prototype. The main purpose of this was to replace the pointing device based on head movements, but the application described here is very important because it allowed us to replace mouse clicks with eye gestures. The prototype allowed us to study the first step of the problem analysis and helped in development of the algorithm for the recognition of eye states. The prototype also helped us to test the entire introduced solution.

In the prototype, we used two modules of IR LEDs and cameras, separately for the left and the right eye. Each module recorded reflections from the surface of one eye, and the signal was transmitted via a USB cable to a computer where our algorithm analyzed the image of the eye. We used an eyeglasses frame as a tool for mounting the rigid cameras and the LEDs in appropriate places before the eyes.

The most important parameter for the entire project was the quality of the recorded images. We experimentally determined that the image (from the micro camera) had a low resolution ranging from 4800 pixels ( $80 \times 60$ ) to 0.3 megapixels ( $640 \times 480$ ). This was consistent with the findings of previous research [22]. The use of a very large resolution increased the computational cost and did not improve effectiveness. Note that today there is a tendency to manufacture cameras with a considerably higher resolution. This implies that to analyze the blinking gesture, we should use a significantly reduced resolution (averaging the image) or use non-standard camera modes if possible. Additionally, in our solution for low resolutions of camera (max  $640 \times 480$ ) and close device settings, noise problem does not occur. So we do not need preprocessing of the acquired images.

The second important parameter was the frequency of registration (and analysis) of the eye image. After analyzing the time dependencies, we assumed that the eye state examination was performed 25 times/s, i.e., every 40 ms. This frequency allowed us to register at least two eye images even during the fastest closing cycle (100 ms).

The crucial condition for practical applications was to ensure work safety. The main problem in our solution was the IR radiation, which had a harmful effect on the eyes. Therefore, we had to select safe ranges of the IR LED parameters.

IR radiation has been documented to be safe for the human eye when two parameters meet the appropriate range criteria: the wavelength should be greater than 1400 nm, and the emission at the retinal level should not exceed  $100 \text{ W/m}^2$  [37]. We used IR LEDs having a wavelength of 1550 nm and appropriate rated power. Additionally, we repeatedly attempted to reduce the power relative to the nominal value. Finally, we received the smallest level of the power at which the algorithm of the reflection analysis worked correctly.

An additional problem is the speed and work efficiency of our solution. However, our method of eye state recognition is so simple that the experiments from the first trials showed the possibility of working in real time. Moreover, the analysis of the computer's processor load has shown that the blinking analysis works in the background and practically does not take up computer resources.

## 9 Applications of the recognized eye gestures

The proposed method for the time analysis of the eye states is the simplest possible and has the ability to separate different eye gestures. We distinguished among sets of gestures, which have practical applications. We proposed two different applications where the entire set of the analyzed eye states and gestures could be used.

### 9.1 Safety glasses—Check whether they are worn

Ensuring work safety under industrial conditions is one of the most important tasks of an employer. Eyes are particularly vulnerable to factors that can lead to eye damage. In many workplaces, there are potentially dangerous factors for our eyes: mechanical (filings and other machining waste), chemical and biological (corrosive or toxic substances), optical (ultraviolet radiation). Therefore, in many workplaces, the use of personal protective equipment is strictly required [4]. The eyes are protected by the use of appropriate protective goggles and safety glasses or full face protective gear. A separate problem is the awareness of threats among employees. Unfortunately, very often, despite training and restrictions, not all of the employees use personal protective equipment in all of the required situations [38]. Even safety glasses matched in the best way are always an uncomfortable addition. The reluctance to wear these glasses is heightened by the conviction that if we have been able to work without an accident for a long time, we will be able to do so in the future. Sometimes, the employee, when going to the workplace, gets glasses but does the work without them. Glasses instead of protecting his/her eyes go into his/her pocket. The U.S. Bureau of Labor Statistics (BLS) analyzed the number of workers who experienced eye injuries and noticed that nearly three out of five of them had not worn safety glasses when the injury occurred [34].

Therefore, we are faced with a non-trivial problem: how do we check whether the employee correctly uses the personal protective equipment? Are the glasses worn to protect the eyes or whether they are in the pocket or on the table? To check this, we could mount sensors onto the frames of the glasses and analyze the selected parameters from the following sample list:

- Temperature (and its changes over time) of the nearest surface—whether it corresponds to the temperature of the human body.
- Distance from the surface of the skin—does it correspond to the correct position of the glasses.
- Color of the nearest surface—does it correspond to the color of the human skin.
- Vibration analysis—whether the human heart rate can be recognized.

In the laboratory, such analyses would lead to interesting research findings. However, under industrial conditions, these solutions are completely useless. Therefore, we developed the following simple method for check whether the safety glasses are worn and whether they are worn appropriately:

- Mount a micro camera and micro LED IR on the frame of the safety glasses.
- Install a microcontroller with the software for analyzing the state of the eyes in the frame of glasses (software should implement points 6 and 7 mentioned above in this paper).
- Recognize eye gestures  $B_0$ ,  $BL_0$ , and  $BR_0$ .

- Perform this analysis every few seconds; the result will be transmitted by a simple wireless transmitter to the information system of the industrial plant.

Note that we used all of the blink states ( $B_0$ ,  $BL_0$ , and  $BR_0$ ), although it would have been sufficient to check whether the eyes were closed. However, extracting blinking (with a specific duration and frequency) provided additional information regarding whether the employee was tired or whether he/she fell asleep. In contrast, the correctness of the operation of the closing eye recognition algorithm confirmed that the micro camera recorded the image of the eye and not any blinking shape. This confirmed that the glasses protected the eyes.

### 9.2 Eye control in multimodal interaction

When we analyzed the known system commands generated with the mouse, we could assign system events to eye gestures (Table 1). The eye gestures  $B_0$ ,  $BL_0$ , and  $BR_0$  were not used in this application. Eye gesture  $W_1$  fulfilled a special role in our application. We assumed that it was used to enable/disable the detection system of winking. To practically use the gestures and the system events from Table 1, we developed software that after recognizing the eye gestures, assigned the appropriate system events. This software was installed as a driver instead of a standard mouse key driver.

Note the differences in the system events between the use of the mouse and the use of eye gestures. For many years, we have used the single click and the double click of a mouse. Performing such activities by using our hands is not a problem. However, a similar implementation for the system events with the eye winking gesture would be very difficult for the user. Therefore, we proposed replacing the single click and the double click by differentiating between the winking duration. In addition, we individually adjusted the times for each user; this provided a comfortable working environment.

Because the use of eye gestures is not a new idea, it is worth paying attention to the advantages of our solution. Applications that use eye tracking [33] or electromyography (sEMG) [5] require specialized equipment, are difficult to implement and practically used only in the laboratory. Our solution is simple and the device is light and easy to use. Applications that use image registration [35] from one or more cameras require complex and difficult algorithms for image processing. The method proposed in our solution is very simple and practically does not burden the computer resources. However, the most important

**Table 1** Used eye gestures and corresponding events

Eye gesture	Closing time ( $t_x$ )		System event
	Left eye	Right eye	
$W_1$	$t_1 < t_x$	$t_1 < t_x$	Switch on and off (of the interaction)
$WL_1$	$t_1 < t_x \leq t_2$	$t_x = 0$	Single click of the left mouse button
$WL_2$	$t_2 < t_x \leq t_3$	$t_x = 0$	Double click of the left mouse button
$WL_3$	$t_3 < t_x$	$t_x = 0$	Hold down of the left mouse button
$WR_1$	$t_x = 0$	$t_1 < t_x \leq t_2$	Single click of the right mouse button
$WR_2$	$t_x = 0$	$t_2 < t_x \leq t_3$	Single click of the central mouse button or scroll(double click of the right mouse button is not popular)
$WR_3$	$t_x = 0$	$t_3 < t_x$	Hold down of the right mouse button

advantage of our solution is mobility. The user with our glasses can control eye gestures at any position of the head and anywhere in the room. This is never possible with even a professional eye-tracker or a set of cameras.

## 10 Tests

We prepared two independent types of tests for our solution. In the first type, we used  $2 \times 5000$  pictures of the eyes (open and closed). The collection of images has been prepared by us previously, based on the registration of eye images of several people. The photos have been carefully selected and divided into two groups: eyes open and eyes closed. The set of pictures was selected such that it contained examples of different eye positions (different eye gaze positions) and of various lighting conditions (natural and artificial light, and low and high levels of light). In the second type of tests, we performed the experiments in a real environment, on a group of 30 participants.

We tested our solution on the basis of the following two categories:

- Correctness of eye state recognition (closed/open). The result of this test also showed the correctness of the safety glasses application. In this category, we performed tests based on eye pictures and tests with a group of participants.
- Efficiency of using eye gestures to control events instead of mouse keys. In this category, we performed tests with a group of participants.

### 10.1 Tests of eye state recognition

The test results of the correctness of recognizing the state of the eye on the basis of an image analysis are summarized in Table 2. The correctness of recognizing the open eye was very high. The introduced algorithm for recognition, based on the analysis of the reflection of radiation, allowed us, with almost 100% correctness, to identify open eyes (99.94%). The correctness of the recognition of closed eyes was slightly lower; it was 99.42%. The incorrect recognition was in most cases caused by a diffused reflection of the radiation in the area of the eyelashes. The total correctness of the eye recognition was 99.68%. This was a very good result, particularly as there was a low probability of two neighboring recorded images to be incorrectly identified.

The experiments performed on a group of 30 participants showed that practically, a 100% efficiency of determining the state of the eye was obtained in real time. Here, the term “practically” implies that the errors of the eye state recognition were observed only when the participant wore the glasses incorrectly or deliberately carelessly.

**Table 2** Results of the correctness of eye state recognition

	The real state of the eye: open	The real state of the eye: closed
The algorithm: the eye state is recognized as open	4997	29
The algorithm: the eye state is recognized as closed	3	4971
Correctness	99.94%	99.42%

## 10.2 Tests of using the eye gestures for mouse control

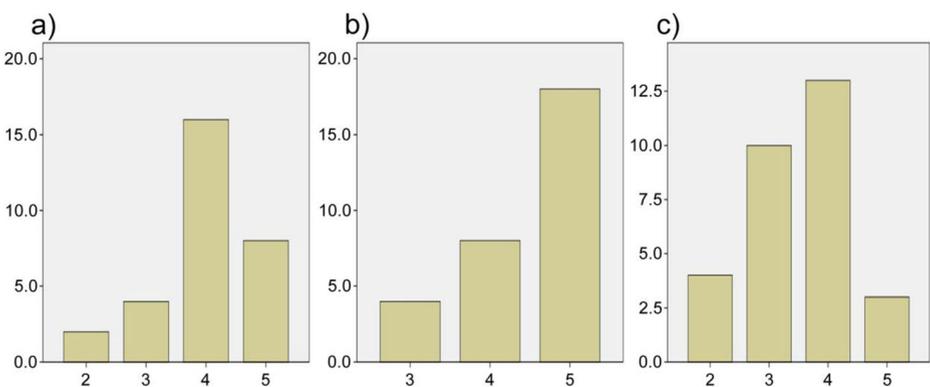
The participants (from a group of 30) had computer experience, but the proposed solution was new to all. They had to perform a set of simple tasks of moving or selecting elements on the screen. However, they had to use winking instead of mouse clicking. After the tests, the participants were asked to evaluate the new solution. In the assessment, we used methods consistent with the Standard ISO 9241–411:2012 [12]. The participants evaluated our solution using a five-step subjective scale, in three independent subjects: operation speed, accuracy, and general comfort (Fig. 3).

The high values of the evaluation parameters showed the acceptance and the correctness of our solution. In the discussion, the participants drew attention to the novelty of our solution and to the problem resulting from the habit of using a standard mouse. Our algorithm allowed for individual tuning of the time conditions for each participant; however, the participants noticed sometimes problem with double clicking replaced by a long period of closed eye. The participants also pointed out that for some people, winking at the right moment could be a very difficult task. This limitation seemed to be the most serious problem of our solution.

## 11 Summary

The main aim of this study was to develop a simple method for recognizing the eye states with the possibility of distinguishing between blinking and winking. We introduced a new solution in which we proposed a simple and effective algorithm for recognizing open and closed eyes on the basis of an eye image analysis. To simplify this analysis IR radiation was used. The concise comparison of our solution and other methods discussed in literature is presented in Table 3.

We proposed two different applications concerning multimodal interactions. In the first application, we used blinking recognition to check whether safety glasses were used appropriately. The correct use of personal protective equipment is one of the most important tasks for ensuring work safety under industrial conditions. Our solution solves this problem in a simple and effective manner.



**Fig. 3** Frequency of grades in the participants' evaluation: **(a)** Operation speed: from 1 (unacceptable) to 5 (acceptable), the average result was 4.0 ( $\delta = 0.83$ ). **(b)** Accuracy: from 1 (very inaccurate) to 5 (very accurate), the average result was 4.47 ( $\delta = 0.73$ ). **(c)** General comfort: from 1 (very uncomfortable) to 5 (very comfortable), the average result was 3.5 ( $\delta = 0.86$ )

**Table 3** The comparison of methods

The method	Ease of application; impact of external factors	Complexity of algorithm	Mobility	Additional comments
The visual analysis of facial images/head images (only cameras) [2, 10, 18, 22]	Impact of the lighting environment; impact of position	Complex and difficult image analysis	YES but in a limited space only	Universal method
The visual analysis of facial images/head images (cameras + IR devices) [15]	Independent of the lighting environment; impact of position	Complex and difficult image analysis	YES but in a limited space only	Universal method
Eye tracking [1, 29, 33]	Independent of the lighting environment; working only in strictly defined place	Depend on tracking method and device	NO	High effectiveness; very high precision
sEMG [5]; EOG [20]	Specialized and complicated device required	Depend on device	NO	Discomfort of the work (electrodes on the head)
Our solution	Simple and light device; easy in use; independent of the lighting environment	Simple algorithm; does not take up computer resources	YES (any position of the head; any place in the room)	Universal

The second application involves controlling system events with winking. Initially, the aim of this application was to replace the mouse keys by eye gestures interpreted from the winking analysis. However, we realized that our solution could be used in many fields of multimodal interactions such as situations where user hands are busy and the mouse button cannot be clicked, and situations where an additional form of control is needed: an extra control element for game players or a nonstandard control form in professional systems for the operating room.

Moreover, we built a simple prototype in which the proposed solution was used. The intended purpose of the prototype was to replace the pointing device by head movements. The application of control by eye gestures replaced mouse clicking and was used as an independent but very important module of interaction.

Our solution was tested with two independent types of tests. In the first one, we used  $2 \times 5000$  images of open and closed eyes; we conducted the second experiment on a group of 30 participants. We tested the new solution for different eye gaze positions and for various lighting conditions such as natural and artificial light, and low and high levels of light. The total correctness of eye recognition was 99.68%. This was a very good result that proved the practical feasibility of the proposed solution. Additionally, the new method of system control with eye gestures was accepted by the participants and allowed for quick and comfortable work.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## References

1. Al-Rahayfeh A, Faezipour M (2013) Eye tracking and head movement detection: a state-of-art survey. *IEEE J Transl Eng Health Med* 1. <https://doi.org/10.1109/JTEHM.2013.2289879>
2. Bacivarov I, Ionita M, Corcoran P (2008) Statistical models of appearance for eye tracking and eye-blink detection and measurement. *IEEE Trans Consum Electron* 54(3):1312–1320. <https://doi.org/10.1109/TCE.2008.4637622>
3. Baranoski GVG, Krishnaswamy A (2010) Light & skin interactions: simulations for computer graphics applications. Morgan Kaufmann Publishers Inc., San Francisco
4. Bartkowiak G, Baszczyński K, Bogdan A, Brochocka A, Hrynyk R, Irzmańska E, Koradecka D, Majchrzycka K, Makowski K, Marszałek A, Owczarek G, Żera J (2012) Use of personal protective equipment in the workplace. In: Salvendy G (ed) *Handbook of human factors and ergonomics*, 4th edn. John Wiley & Sons, Inc. <https://doi.org/10.1002/9781118131350.ch30>
5. Bastos-Filho T, Ferreira A, Cavalieri D, Silva R, Muller S, Perez E (2013) Multi-modal interface for communication operated by eye blinks, eye movements, head movements, blowing/sucking and brain waves. In: *Proc. of ISSNIP biosignals and biorobotics conference: biosignals and robotics for better and safer living*, Rio de Janeiro, Brazil, pp 1–6. <https://doi.org/10.1109/BRC.2013.6487458>
6. Calvo AA, Perugini S (2014) Pointing devices for wearable computers. *AHCI* 2014:10. <https://doi.org/10.1155/2014/527320>
7. Cui J, Liu Y, Xu Y, Zhao H, Zha H (2013) Tracking generic human motion via fusion of low- and high-dimensional approaches. *IEEE Trans Syst Man Cybern Syst Hum* 43(4):996–1002. <https://doi.org/10.1109/TSMCA.2012.2223670>
8. Dorsey J, Rushmeier H, Sillion F (2007). *Digital modeling of material appearance*. Morgan Kaufmann Publishers Inc., San Francisco

9. Evinger C, Manning KA, Sibony PA (1991) Eyelid movements. Mechanisms and normal data. *Invest Ophthalmol Vis Sci* 32(2):387–400
10. Grauman K, Betke M, Gips J, Bradski G (2001) Communication via eye blinks - detection and duration analysis in real time. In: Proc. of IEEE CVPR, Kauai, HI, USA, pp 1010–1017. <https://doi.org/10.1109/CVPR.2001.990641>
11. Grauman K, Betke M, Lombardi J, Gips J, Bradski G (2003) Communication via eye blinks and eyebrow raises: videobased human-computer interfaces. *Univ Access Inf Soc* 2(4):359–373
12. ISO Standard: ISO 9241-411:2012 (2012) Ergonomics of human-system interaction – Part 411: Evaluation methods for the design of physical input devices
13. Jaimes A, Sebe N (2007) Multimodal human computer interaction: a survey. *Comput Vis Image Underst* 108(1–2, (Special Issue on Vision for Human-Computer Interaction)):116–134. <https://doi.org/10.1016/j.cviu.2006.10.019>
14. Jian-Zheng L, Zheng Z (2011) Head movement recognition based on LK algorithm and Gentleboost. In: Proc. 7th international conference on networked computing and advanced information management (NCM), pp 232–236
15. Kapoor A, Picard RW (2001) A real-time head nod and shake detector. In: Proc. 2001 Workshop on perceptive user interfaces, pp 1–5. <https://doi.org/10.1145/971478.971509>
16. Karson CN (1988) Physiology of normal and abnormal blinking. *Adv Neurol* 49:25–37
17. Kim S, Park M, Anumas S, Yoo J (2010) Head mouse system based on gyro- and opto-sensors. In: Proc. 3rd international conference on biomedical engineering and informatics (BMEI), vol 4, pp 1503–1506. <https://doi.org/10.1109/BMEI.2010.5639399>
18. Kim D, Choi S, Choi J, Shin H, Sohn K (2011) Visual fatigue monitoring system based on eye-movement and eye-blink detection. In: Proc. SPIE 7863, stereoscopic displays and applications XXII, 786303. <https://doi.org/10.1117/12.873354>
19. Kojima N, Kozuka K, Nakano T, Yamamoto S (2001) Detection of consciousness degradation and concentration of a driver for friendly information service. In: Proc. of the IEEE international vehicle electronics conference, Tottori, Japan, pp 31–36. <https://doi.org/10.1109/IVEC.2001.961722>
20. Krishna, GVS, Amarnath K (2013) A novel approach of eye tracking and blink detection with a human machine. *IJOART* 2(7):289–297
21. Krishnasree V, Balaji N, Rao PS (2014) A real time improved driver fatigue monitoring system. *WSEAS Transactions on Signal Processing* 10:146–155
22. Le H, Dang T, Liu F (2013) Eye blink detection for smart glasses. In: Proc. of 2013 IEEE international symposium on multimedia, pp 305–308. <https://doi.org/10.1109/ISM.2013.59>
23. Liu K, Luo YP, Tei G, Yang SY (2008) Attention recognition of drivers based on head pose estimation. In: Proc. IEEE vehicle power and propulsion conference, pp 1–5. <https://doi.org/10.1109/VPPC.2008.4677536>
24. Liu Y, Nie L, Han L, Zhang L, Rosenblum DS (2015) Action2Activity: recognizing complex activities from sensor data. In: Proc. of the international conference on artificial intelligence, IJCAI'15, pp 1617–1623
25. Liu Y, Nie L, Liu L, Rosenblum DS (2016) From action to activity: sensor-based activity recognition. *Neurocomputing* 181:108–115. <https://doi.org/10.1016/j.neucom.2015.08.096>
26. Mandal B, Eng H-L, Lu H, Chan DWS, Ng Y-L (2012) Non-intrusive head movement analysis of videotaped seizures of epileptic origin. In: Proc. 34th annual international conference of the IEEE engineering in medicine and biology society, pp 6060–6063. <https://doi.org/10.1109/EMBC.2012.6347376>
27. Manogna S, Vaishnavi S, Geethanjali B (2010). Head movement based assist system for physically challenged. In: Proc. 4th International conference on bioinformatics and biomedical engineering (iCBBE), pp 1–4. <https://doi.org/10.1109/ICBBE.2010.5517790>
28. Nakano T, Kato M, Morito Y, Itoi S, Kitazawa S (2012) Blink-related momentary activation of the default mode network while viewing videos. *Proc Natl Acad Sci* 110(2):702–706. <https://doi.org/10.1073/pnas.1214804110>
29. Patel RA, Panchal SR (2015) Detected eye tracking techniques: and method analysis survey. *IJEDR* 3(1): 168–175
30. Ponder E, Kennedy WP (1927) On the act of blinking. *Q J Exp Physiol* 18:99–110
31. Sawicki D, Miziolek W (2015) Human colour skin detection in CMYK colour space. *IET Image Process* 9(9):751–757
32. Schiffman HR (2001) Sensation and perception. An integrated approach. Wiley, New York
33. Singh H, Singh J (2012) Human eye tracking and related issues: a review. *IJSRP* 2(9):1–9
34. Smith S (2010) Workers are risking injury by not wearing safety equipment. <http://www.ehstoday.com/ppe/hand-protection/workers-risking-injury-safety-equipment-6332>. Accessed 20 Dec 2017
35. Strumillo P, Pajor T (2012) A vision-based head movement tracking system for human-computer interfacing. In: Proc. new trends in audio and video/signal processing algorithms, architectures, arrangements and applications (NTAV/SPA), pp 143–147
36. Tada H, Omori Y, Hirokawa K, Ohira H, Tomonaga M (2013) Eye-blink behaviors in 71 species of primates. *PLoS One* 8(5):e66018. <https://doi.org/10.1371/journal.pone.0066018>

37. Wolska A (2013) Artificial optical radiation – principles of occupational risk assessment (in polish: Sztuczne promieniowanie optyczne - zasady oceny ryzyka zawodowego). Central Institute for Labour Protection - National Research Institute, Warsaw
38. Workers fail to wear required personal protection equipment. In: Sustainable plant, by objective. <http://www.sustainableplant.com/2011/07/workers-fail-to-wear-required-personal-protection-equipment/>. Accessed 20 Dec 2017



**Piotr Kowalczyk** was born in 1986. He received M.S. in Computer Science in 2010 and Ph.D. in 2015 all from Warsaw University of Technology, Warsaw, Poland. From 2015, he is working as Technology Delivery Analyst in Accenture - consulting services company. He implements projects in the field of banking systems. He continues also cooperation with the Warsaw University of Technology. Main interests: computer graphics, pattern recognition, HCI, banking systems.



**Dariusz J. Sawicki** was born in 1957. He received M.S. in Electrical Engineering in 1981, Ph.D. in 1986 and D.Sc. (habilitation) in 2008, all from Warsaw University of Technology, Warsaw, Poland. From 1986 to 2011, he was a Research Assistant and an Assistant Professor with the Warsaw University of Technology. Since 2011 he is an Associate Professor at the Institute of Theory of Electrical Engineering, Measurement and Information Systems, at the Warsaw University of Technology. Main interests: computer graphics, pattern recognition, HCI, simulation in lighting technology. Dr. Sawicki is a Member of IEEE since 1992 (Senior Member since 2012), Member of ACM since 2001 (Senior Member since 2011) and Member of Polish Information Processing Society since 1987.