On the Automatic Implementation of the Eye Involuntary Reflexes Measurements Involved in the Detection of Human Liveness and Impaired Faculties

François Meunier, Ph. D., ing.

*Université du Québec à Trois-Rivières*

*Canada*

1. Introduction

To this day and in spite aggressive publicity campaigns, impaired driving due to the consumption of legal/illegal drugs such as alcohol and cannabis remains a major cause of road casualties. In order to detect such impairments, police officers first rely on an array of movement coordination tests (Standard Field Sobriety Tests: SFST) which are most of the time applied at road side. If enough suspicions are raised from the first phase of tests, a second set of more reliable tests are conducted at a police station. A driver would then be subjected to an assessment by a Drug Recognition Expert (DRE) in order to determine the type of intoxication (category of drugs) a driver is under as well as his capacity to operate any type of motor vehicle.

In the present chapter as for the law enforcement community we define drugs as substance which, when taken into the human body, can impair the ability of a person to operate a vehicle safely. Drugs are mainly classified into seven categories, the Central Nervous System (CNS) depressants category includes some of the most abused drugs. Alcohol is considered the most familiar CNS depressants drug abused.

The development of the Standardized Field Sobriety Tests (SFST), which was largely done through extensive research efforts performed at the Southern California Research Institute (SCRI) (Burns, 1995); (Burns & Gould, 1997) was a critical step toward the development of the DRE procedures (Page, 2000). The DRE testing process involves twelve distinct components in which the examination of the eyes and dark room examination are parts. The examination of the eyes allows the detection of certain drugs since these drugs can produce very easily observable effects on the eyes. One of the most noticeable of these effects is the horizontal gaze nystagmus. A person under the influence of certain drugs such as alcohol usually will exhibit Horizontal Gaze Nystagmus (HGN) which is an involuntary jerking of the eyes that occurs as the eyes turn toward the side while gazing at an object. The dark room examinations can also allow the detection of certain drugs since it affects how the pupils will appear and how they respond to light stimuli. Some drugs will cause the pupils to dilate and other to constrict. By systematically changing the amount of light entering a person’s eyes, one can also observe the pupil’s appearance and reaction.
In order to detect drivers under the influence of substances, police officers use a standardized set of tests such as the Horizontal Gaze Nystagmus (HGN) test, the Vertical Gaze Nystagmus (VGN) test, the eye convergence test and the pupil’s reaction to light tests. These tests are part of the more complete DRE procedures and are essentially applied manually by law enforcement officers to the eyes of a driver.

However, a study (Booker, 2001) analysing the effectiveness of the HGN test to detect alcohol related intoxication suggested that police officers rarely applied the HGN test according to the standard procedures, when trying to assess the degree of impairment of a driver. These inconsistencies in the administration of a test such as the HGN test could lead to erroneous diagnoses of the intoxication level of drivers.

In order to standardize the administration of the set of tests such as the HGN test involved in the DRE procedures, a first attempt to instrument the HGN test has been reported in experiments done at the Southern California Research Institute (SCRI) (Burns, 1995).

This chapter first describes the main approaches implemented in the literature to extract useful information from the eyes (pupils) of a human to allow the evaluation of his liveness and faculties. These eye-pupil related informations are mainly the involuntary reflexes such as the horizontal gaze nystagmus (HGN), eye divergence, pupil reaction to light, and the naturally occurring change of the pupil’s diameter known as hippus. These involuntary reflexes are detected through the measurement of the gaze direction of the eyes while following specific visual stimuli and the pupil’s diameter change resulting from the modulation of the illumination. Many researches in the field of automatic liveness and impaired faculty detection based on image processing and pattern recognition are outlined. Results of different approaches implemented in these fields are also presented. More specifically, the automatic detection of eye involuntary reflexes such as the hippus (pupil involuntary movement) and eye saccade (eye involuntary movement) are investigated in more details. The automatic detection of such eye reflexes are involved in the liveness detection process of a human which is also an important part of a reliable biometric identification system (Toth, 2005).

The present chapter also introduces existing video-based systems such as the one developed at the Southern California Research Institute (SCRI) (Burns, 1995) that are in general automating the detection of individual or few eye involuntary reflexes.

Moreover, in this chapter, a new video-based image processing system automating the implementation of the eye involuntary reflexes measurements extracted from the tests involved in the DRE procedures is proposed (Meunier, 2006; Meunier & Laperrière, 2008). This system integrates all the eye related involuntary reflexes measurements of the DRE procedures and is build around a video-based image processing system divided into three main modules, a video capturing module to grab videos of the eye movement following automatically generated visual stimuli, an image processing module that essentially extract information about the pupil such as the gaze direction and its size, a diagnosis module assesses the type of impairment of a person from the pupil’s information extracted from the image processing module. This novel system must internally generate visual stimuli and captures video sequences of the eyes following and reacting to these visual stimuli. The video sequences are processed and analyzed using state of the art feature extraction techniques. In this present book chapter, the results outlined are obtained from the last updated prototype of the video-based image processing system described. This automatic human faculty assessment system tested is used in the context of alcohol related intoxication. This system was tested in experiments involving many subjects dosed to a...
blood alcohol concentration (BAC) in a wide range interval of 0.04% to 0.22%, and to show the efficiency of this system to detect alcohol intoxicated individuals at BAC above 0.08%. In order to demonstrate the effects of alcohol on the eye involuntary reflexes comparisons are made between pre-dose and post-dose BAC. This video-based image processing system was experimented with alcohol related intoxication for two reasons. First, the experimental data extracted in the experiments were correlated with Bulk Alcohol Concentration (BAC) measurements obtained with measurement devices such as the Breathalyser 900A or the Intoxilyser 5000 which are easy to use since they only require breath samples. Similar correlations are not so easy to perform for experiments with other drugs since blood or urine samples are often required. Second, the ethical certificates granted by the university authorities are easier to obtain in the case of alcohol workshops but far more difficult to obtain for other drug related workshops.

The present updated prototype of this video-based image processing system is an enhanced version of a previously developed automatic prototype that only implemented the HGN test (Meunier, 2006). The newly updated prototype is also a video-based image processing system but with added automated DRE tests (Meunier & Laperière, 2008). These added automated tests are the convergence test and the pupil’s dark room examinations test. These added functionalities constitute an improvement over previous video-based systems (Burns, 1995); (Meunier, 2006) since these systems were dealing with fewer ocular properties limiting there ability to detect diverse intoxications. This new system also generates visual stimuli and captures video sequences of the eyes following and reacting to these visual stimuli. The ocular properties (signs) extracted from these new tests include convergence of the eyes, the pupil’s reaction in room light, darkness and direct illumination.

During the step four (eye examinations) of the DRE procedures (Page, 2000) two separate eye movement examinations are conducted: the already known HGN (Meunier, 2006) and the eye convergence. During the HGN test a suspected impaired person is instructed to follow an object (pen or a pencil) that is moved horizontally in front of his eyes. During the eye convergence examination a subject is instructed to look at an object until that object is placed on the tip of the subject’s nose. The step seven (Darkroom Examinations) of the DRE procedures enables the extraction of the pupil sizes. This test is administered by changing the amount of light directed to the pupils. With an alive, unimpaired healthy subject, the pupils enlarge in response to darkness and constrict to bright light. The pupils of an impaired person react differently to the light stimuli depending of the type of intoxication.

To implement the HGN and convergence tests, the automated system essentially extracts, the eye-gaze direction (angular position) of the pupils from video sequences of the eyes following visual stimuli. The analysis of these angular positions enables the detection of the saccadic motion of the eyes related to the horizontal gaze nystagmus and also the detection of the eyes convergence which in turn can enable the determination of the type of intoxication of a driver. The pupil’s reaction to light is measured by extracting the radius of the pupils from video sequences of the eyes subjected to controlled illumination.

The new video-based image processing system outlined in this chapter has plenty of practical applications. As a decision support system to assist the law enforcement officers to evaluate the type of impairment of a driver, as an automatic impairment detection system to assess the capacities of an operator to operate a motor vehicle such as a bus, a subway train, a train, a plane, etc., or as a new version of ignition interlock system. This book chapter also provide a description of future research topics in the field of automatic liveness and human faculty detection.
2. Software and hardware requirements for the automatic implementation of eye involuntary reflexes measurements

The implementation of a video-based image processing system for the automation of the eye involuntary reflexes measurements requires three main modules, a video capturing module, an image processing module and a diagnosis module. The image processing module should offer some fundamental functionalities such as the segmentation of the pupils in each image frame of a video of the eye movement captured during the application of a specific DRE symptomatic test. Moreover, the gaze direction and the pupil’s size are computed from the segmented pupils. The video capturing module must also provide some important functionalities. For instance, it must offer a physical environment such as a helmet or set of goggles in which the illumination devices required to test the pupillary reflexes are using the modulation of a visible spectrum light source (White LEDs). Therefore, the capturing device must be sensitive to the near-infrared (NIR) spectrum since the eyes have to be illuminated with NIR LEDs.

2.1 Software requirements: Image processing tools to extract pupil’s information from images

2.1.1 Iris/Pupil region segmentation

In many researches (Ji & Zhu, 2004; Batista, 2005; Ji & Yang, 2002; Zhu & Ji, 2005; Li et al., 2004; Morimoto et al., 2000; Haro et al., 2000) the Dark/Bright pupil effect is used to segment the pupils in video images. This effect is obtained by illuminating the pupils with evenly and symmetrically distributed near-infrared (NIR) light emitting diodes (LED) along the circumference of two coplanar concentric rings. The center of both rings coincides with the camera optical axis. The bright pupils (Fig. 1 left) is produced when the inner ring centered around the camera lens is on and the dark pupil (Fig. 1 center) when the outer ring offset from the optical center is on. The pupil can be easily segmented by image subtraction (Fig. 1 right).

![Fig. 1. Dark/Bright pupil effect and pupil segmentation.](image)

Figure 1 (left and center) also reveals the presence of the first Purkinje image also called glint or corneal reflection. The glint is a high contrasted bright spot corresponding to the reflection of the NIR LED illuminators on the cornea surface. Since the 2D position of the glint is easy to find from frame to frame it can be useful for the estimation of the eye-gaze direction (Batista, 2005; Ohno & Mukawa, 2004; Ji & Bebis, 1999; Ji & Yang, 2002). The glint is easily detected in the dark pupil image (Fig. 1 center) using an image binary thresholding technique in the neighbourhood region of the pupils. The centroid of the segmented region of a glint is computed and gives its 2D position in an image.
The location of the eyes in a video image can also be achieved using a correlation process that maximizes the match between regions of interest (ROI) in the video image corresponding to the eyes and patterns (Fig. 2) that are reference images of an eye with different gaze directions (Matsumoto & Zelinsky, 2000; Zhu & Ji, 2005; Li et al., 2004). The results obtained from the correlation process are the 2D position of the eyes in the video image. Moreover, the patterns maximizing the match with the eye ROIs allow the estimation of the eye-gaze direction. We can also formulate the matching process as a minimization problem where the matching criterion \( E \) is the sum of squared difference errors given by:

\[
E = \sum_{i=1}^{N} (I_p(i) - I_c(i))^2
\]

where \( I_p(i) \) is the pixel value of the \( i \)th pixel of a given pattern image \( I_p \), \( I_c(i) \) the pixel value of an eye ROI in the current image frame, and \( N \) the total number of pixels in each pattern image and the eye ROI in the current image.

Fig. 2. Eye patterns with different eye-gaze directions.

Most of the approaches implemented to locate the eyes and the pupils/irises in video image rely on the notion of image gradient (Sirohey et al., 2002; Peng et al., 2005; Ji & Bebis, 1999; Wang & Sung, 2001; Halir & Flusser, 1998; Smith et al., 2000; Li et al., 2004). The iris/sclera and iris/pupil borders exhibit a noticeable contrast that enables the detection of these borders using methods based on the image gradient. The image gradient can be expressed theoretically and numerically using the Sobel operator by:

\[
\nabla I(x, y) = \left[ \frac{\partial I(x, y)}{\partial x}, \frac{\partial I(x, y)}{\partial y} \right]
\]

\[
I_\theta(x, y) = \tan^{-1} \left( \frac{\partial I(x, y)}{\partial y} \right) \frac{\partial I(x, y)}{\partial x}
\]

\[
\| \nabla I(x, y) \| = \sqrt{I_x^2 + I_y^2}
\]
\[
D_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, \quad D_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}
\]

(5)

where \(\nabla I\) is the gradient of the image \(I(x,y)\) given by the image partial derivative in the x and y coordinate direction, \(I_\theta(x,y)\) is the orientation of the gradient, \(\|\nabla I\|\) is the magnitude of the gradient and \(D_x, D_y\) the numerical implementation of the partial derivative \(\partial I/\partial x\) and \(\partial I/\partial y\) of the Sobel operator. Figure 3 shows the result of the application of the Sobel operator on the dark pupil image depicted in Fig. 1 center that essentially allows the detection of the edge between the pupil/iris region. Figure 3 right shows the elliptic shape corresponding to the pupil contour. With its strong gradient the glint contour is also easy to detect.

Fig. 3. Pupil/Iris border detection (Sobel edge detection). (Left) Horizontal edge detection. (Center) Vertical edge detection. (Right) Gradient image.

The refinement of the location of the edges detected by the application of a Sobel gradient operator can be achieved using the Canny edge detection algorithm. The Canny edge detection starts by computing the gradient of an image and enhance the location of the original edges with a thinning and a thresholding phase to obtain finer edges. These refinement steps operate on the gradient magnitude computed in the linear filtering step and intend to use hysteresis thresholding on edge strength to only retain edge pixels for which the gradient magnitude reaches a maximum. Results of the application of the Canny edge detection algorithm applied to Fig. 1 center is shown in fig. 4. As the threshold value is increased, the preserved edges are stronger. With a threshold value of 60 (Fig. 4 center) the pupil /iris border is clearly observable with its elliptic shape. The glint (corneal reflection) can be also easily detected at a high threshold value and as previously mentioned is the strongest contrasted region (see Fig. 4 right).

Fig. 4. Canny edge detection algorithm. (Left) Edge detected with a threshold value of 20. (Center) Edge detected with a threshold value of 60. (Right) Edge detected with a threshold value of 80.
2.1.2 Circle/Ellipse detection and fitting

Some of the previously introduced techniques to extract iris/pupil edges and contours are used as pre-processing steps to circle/ellipse detection (Ji & Bebis, 1999; Wang & Sung, 2001; Smith et al., 2000; Li et al., 2004).

The circle/ellipse detection process can be done mainly in two different ways. A first approach (Halir & Flusser, 1998; Fitzgibbon et al., 1999) where the edge and contour pixels are fitted by a implicit second order polynomial given by:

\[ F(x, y) = ax^2 + bxy + cy^2 + dx + ey + f = 0 \]  

where \(a, b, c, d, e, f\) are the coefficients of the ellipse and \((x, y)\) the 2D position of the detected pixels at the edge/contour detection phase belonging to an elliptical 2D shape. The fitting of an ellipse to a set of 2D pixels consists in finding the coefficients of the polynomial \(F(x, y)\) by a least-square minimization approach.

A second approach for circle/ellipse detection is based on the Hough transform (HT) (Guil & Zapata, 1996; Zhang & Liu, 2004). The main goal of this approach is to find the ellipse center position \((x_0, y_0)\), the orientation of the ellipse \(\theta\) and the semi-axis \((a, b)\). Therefore, the Hough parameter space is 5-Dimensional with a parameter vector represented by \([x_0, y_0, \theta, a, b]\). The center of the ellipse can be found by searching the maxima of this Hough parameter space satisfying the following equation for each 2D pixel of the searched ellipse:

\[
\frac{((x - x_0) \cos(\theta) + (y - y_0) \sin(\theta))^2}{a^2} + \frac{((x - x_0) \sin(\theta) + (y - y_0) \cos(\theta))^2}{b^2}
\]

where \((x, y)\) are the edge 2D pixel coordinates, \((x_0, y_0)\) is the ellipse center, \(\theta \in [0, 2\pi]\) is the ellipse orientation, \((a, b)\) the semi-axis lengths.

2.1.3 Eye-gaze direction and pupil’s size computation

The previously introduced Circle/Ellipse detection algorithms can be used to measure the pupil’s size (diameter) and furthermore allow the measurement of the pupil’s reaction to light reflexes or the hippus.

The eye-gaze estimation can be achieved by a linear mapping procedure (Ji & Bebis, 1999; Ji & Yang, 2002) using the relative position between the pupil and the glint. This mapping procedure can allow depending of the application, the estimation of the screen coordinates gazed by the eye, or the eye-gaze direction given by the horizontal and vertical pupil displacement angles. The mapping procedure is implemented by:

\[
i = g \cdot w = a\Delta x + b\Delta y + c g_x + d g_y + e
\]

with \(g = [\Delta x, \Delta y, g_x, g_y, 1]\), the pupil-glint vector, where \(\Delta x\) and \(\Delta y\) are the pupil-glint displacements, \(g_x\) and \(g_y\) the glint image coordinates and \(w = [a, b, c, d, e]\) the coefficients of the linear mapping equation and \(i\) the gaze region index corresponding to a given gaze direction.

The vector \(w\) is deduced through a calibration procedure that allows the modeling of the transformation from the pupil-glint matrix \(A\) to a target vector \(B\) given by:
During the calibration procedure a user is asked to gaze at $N$ distinct targets on a screen while each pupil-glint image are captured. The matrix $A$ is obtained from the computation of the pupil-glint vectors $g$ with their corresponding target index allowing the creation of the vector $B$. The coefficient vector $w$ is obtained by solving eq. (9) using a least-squares method.

Another way to compute the eye-gaze direction is through the pose-from-ellipse algorithm (Trucco & Verri, 1998; Ji & Bebis, 1999; Wang & Sung, 2001). The underlying problem consists in deducing the normal vector of a plane containing a circle shape such as an eye pupil which is depending of the eye-gaze direction imaged as an ellipse. The geometry of the pose-from-ellipse algorithm is illustrated in Fig. 5.

The image ellipse defines a 3D cone with vertex in the center of projection of the modeled pinhole camera. The orientation of the circle’s plane (iris/pupil plane) is deduced by rotating the camera reference axis $OXYZ$ in such a way that the original elliptic shape imaged in the image plane becomes a circle shape. This is achieved when the image plane is parallel to the circle plane. The equation of the 3D cone is given by:

$$ax^2 + bxy + cy^2 + dxz + eyz + fz^2 = P^T Q P = 0$$

where $P = [x, y, z]^T$ and $Q$ is the real, symmetric matrix of the ellipse.

\[
A \cdot w = B
\]

\[
A = \begin{bmatrix}
\Delta x_1 & \Delta y_1 & g_{x_1} & g_{y_1} & 1 \\
\Delta x_2 & \Delta y_2 & g_{x_2} & g_{y_2} & 1 \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
\Delta x_{N-1} & \Delta y_{N-1} & g_{x_{N-1}} & g_{y_{N-1}} & 1 \\
\Delta x_N & \Delta y_N & g_{x_N} & g_{y_N} & 1
\end{bmatrix}
\]

(9)

Fig. 5. Geometry of the pose-from-ellipse algorithm.

The normal $n$ of the circle plane corresponding to the eye-gaze direction is given by:
On the Automatic Implementation of the Eye Involuntary Reflexes Measurements Involved in the Detection of Human Liveness and Impaired Faculties

\[
R = \begin{bmatrix}
\cos \theta & 0 & \sin \theta \\
0 & 1 & 0 \\
-\sin \theta & 0 & \cos \theta
\end{bmatrix}
\]

\[
n = R \begin{bmatrix}
0 \\
0 \\
-1
\end{bmatrix} = \begin{bmatrix}
-R_{13} \\
-R_{23} \\
-R_{33}
\end{bmatrix}
\]

\[
\theta = \pm \arctan \left( \frac{\lambda_2 - \lambda_1}{\lambda_3 - \lambda_2} \right)
\]

where \( \lambda_1, \lambda_2, \lambda_3 \) are the eigenvalues of \( Q \) deduced by diagonalizing the ellipse matrix \( Q \), knowing that \( \lambda_1 < \lambda_2 < \lambda_3 \), and \( e_1, e_2, e_3 \) the corresponding eigenvectors.

A Hough based approach can also be used to compute the eye-gaze direction (Takegami et al., 2003). In order to compute the eye-gaze direction, the contour of each pupil is first located from video sequences. Each edge point composing a pupil’s contour is assumed to be isolated. According to the Hough transform theory there could be an infinite number of ellipses that could pass through this edge point. The centers of these ellipses generate another ellipse known as a second layer ellipse. The second layer ellipse is deduced by the following expressions:

\[
x = b_a \cdot \cos \alpha \sqrt{\frac{r_0^2 - x_e^2 - y_e^2}{r_0^2}} - b_a \cdot \sin \alpha \sqrt{\frac{y_e^2}{x_e^2 + y_e^2}} + r_a \frac{x_e}{r_0}
\]

\[
y = b_a \cdot \cos \alpha \sqrt{\frac{r_0^2 - x_e^2 - y_e^2}{r_0^2}} + b_a \cdot \sin \alpha \sqrt{\frac{y_e^2}{x_e^2 + y_e^2}} + r_a \frac{y_e}{r_0}
\]

\[
z = -b_a \cdot \cos \alpha \sqrt{\frac{x_e^2 + y_e^2}{r_0^2}} + r_a \frac{r_0^2 - x_e^2 - y_e^2}{r_0}
\]

\[
b_a = a_0 \sqrt{r_0^2 - a_o^2}
\]

where \((x_e, y_e)\) are the edge point coordinates, \(b_a\) is the current circle radius, \(r_a\) the distance between the cornea curvature center and the second layer ellipse, \(a_0\) the pupil radius and \(r_0\) the distance between the cornea curvature center and the pupil rim.

The application of the Hough transform algorithm allows the determination of the optimal pupill’s center positions \((x_0, y_0, z_0)\). These positions correspond to maximal counts in the \((x, y, z)\) space generated by the application of eq. (12) for each pupil’s contour point given their coordinates \((x_e, y_e)\).

Since the position of the cornea curvature center is already known and the pupil’s center \((x_0, y_0, z_0)\) is deduced by the application of the Hough transform, the pupil’s angular horizontal position (angle \(\theta\)) can be easily computed by (Fig. 6):

\[
\tan \beta = \frac{dy}{dx}
\]
2.2 Hardware requirements: Video-image capturing system

Few commercial and experimental systems that implement automatic detection of the eye involuntary reflexes to assess human liveness and impaired faculties actually exist. One of the first attempts to automatically apply the horizontal gaze nystagmus (HGN) test has been reported in another experiment done at the SCRI by Marceline Burns (Burns, 1995). This system also named ocular motor module (OMM) generates visual stimuli for testing eye movement. Volunteers look into a foam-lined viewport, which allows testing in complete darkness. The illumination is provided by near-infrared (NIR) light emitting diodes (LEDs). The video sequences are captured by a NIR sensitive video camera at a sample rate of 60 samples per second. The captured video sequences are further analyzed using feature extraction algorithms to obtain the eye related measurements required to implement the HGN test. The VideoNystagmoGraphy (VNG) ULMER from Synapsys is another system allowing the evaluation of vestibular functions and the detection of vertigo, dizziness and balance disorders. This system uses a near-infrared sensitive CCD camera mounted on a pair of goggles to capture the eye motion while following externally generated visual stimuli. Since the goggles maintained the eye in a dark environment during the test procedure, near-infrared light emitting diodes (NIR LEDs) are used to illuminate the eye. Another system, the Compact Integrated Pupillograph CIP from AMTech is mainly used to automatically measure the pupil dynamics such as the pupillary light reflex. This system can be useful to analyze sleep disorder and the influence of drugs on drivers. This system essentially captures and stores video images of the eye pupil dynamics during pupil’s reflexes related tests. A forth system (Iijima et al., 2003), consists of a head mounted goggle equipped with a near-infrared sensitive CCD camera, a liquid crystal display allowing the investigation of eye tracking functions of neurological disease patients. The head mounted display showed visual stimuli followed by the eyes that were video captured. Image analysis of the eye movement was performed to detect eye saccades which are known to be symptomatic of Parkinson’s disease.
The automatic impairment detection system implemented in more recent research (Meunier & Laperrière, 2008) is composed of the three main modules required to automatically implement a wider range of eye involuntary reflexes and is an enhanced version of a simpler system developed in previous research (Meunier, 2006). A video capturing module allows the grabbing of video sequences of the eyes following automatically generated visual stimuli. The grabbed video data are subsequently analysed by an image processing module and also a diagnosis module to allow the automatic detection of drugs and alcohol impairment. The newer version of this system is also built around the same modules. However, the video capturing module has been extensively enhanced. In order to eliminate the saccadic eye motion induced by the discrete positioning of the red Light Emitting Diodes (LEDs) (see Fig. 7) used to produce the visual stimulus a LCD flat screen and a reflective surface was installed.

With the new prototype the near-infrared (NIR) illumination system used to allow the visualization of the subject face by the NIR sensitive video-camera was also enhance to increase the coverage of the NIR LEDs beam.

The last two versions of the implemented prototype are also depicted in Fig. 8. In the previous version (Fig. 8 left) a helmet integrates the NIR sensitive camera and the automatic visual stimuli generation system. In the more recent version (Fig. 8 right), a black box integrates a LCD flat screen that projects visual stimuli in continuous motion. This version also increased the amount and coverage of the NIR illumination beam projected to subject’s face which in turn enhanced the contrast of the pupils in images.

Fig. 7. A previous version of the prototype with visual stimulus produced by series of red LEDs (Meunier, 2006).

The image processing module essentially used a Hough based approach algorithm described in the previous section (Takegami et al., 2003), to extract the position (center) of the pupils in each image frame useful in the HGN and convergence tests. Since pupils correspond to darker regions in the images, pupils are easily detected by applying a simple optimal binary thresholding technique (Gonzalez & Woods, 2008).
To avoid pupil false detection, each region segmented in the thresholding phase is validated using biometric properties of the eyes and pupils (Nikolaidis & Pitas, 2000). The contour of each pupil’s segmented region is also extracted. The center of each detected pupil is then computed using eq. (12) which allows the calculation of the eyes horizontal angular position (eq. (13)). These angular positions are stored on disk for further processing by a diagnosis module.

The image processing module is also used to extract the radius of the pupils measured in the pupil’s reaction to light test. These measurements correspond to the radius of the fitting circles on the pupil’s contours extracted from thresholded video images. The approach implemented is similar to the one reported by Toth (Toth, 2005). A Hough transform approach for fitting circles to the captured pupil’s edge points contours has also been investigated. The radius of each detected pupil is also stored on disk for subsequent processing by the diagnosis module.

For the HGN test, the diagnosis module essentially performs the comparison of the pupil’s angular horizontal position curves corresponding to post-dose BAC levels (BAC > 0.00%) with equivalent ideal curves at pre-dose BAC level (BAC ≈ 0.00%), to determine if a subject’s BAC exceeds the tolerance limit of 0.08%.

The summation of the square difference between the pupil’s horizontal angular position curves at post-dose BAC level and the approximate curves corresponding to the pre-dose BAC level are used to detect the lack of smooth pursuit and the nystagmus at maximum deviation which are two out of three visual signs obtained with the HGN test. These visual signs are known to be sufficient to establish the level of impairment of an individual intoxicated by CNS depressant drugs such as alcohol (Citek et al., 2003).

Furthermore, for the HGN test, the diagnosis module also performs a time-frequency analysis of the pupil’s angular horizontal position curves at post-dose BAC levels (BAC > 0.00%), to determine the presence of pupil’s saccadic motions associated with high doses of alcohol (BAC > 0.08%). This time-frequency analysis is based on a short time Fourier transform to allow a more precise localization of the saccadic motions inherent to the
horizontal gaze nystagmus. Such saccadic motions tend to occur when the eyes are gazing at high horizontal deviation.

The diagnosis module also determines from the pupil’s horizontal deviation curves extracted in the convergence test, if the eye-gaze direction of both eyes are converging at the same target. The horizontal deviation is computed as the difference between the pupil’s horizontal position at a given time frame and the pupil’s horizontal position at time 0 when a subject’s eyes are gazing at reference targets placed in front of the eyes. The check for lack of convergence can provide another clue as to the possible presence of CNS depressants such as alcohol.

Ultimately, the diagnosis module assesses the response time of the involuntary reflex of the pupils triggered by changing illumination from the measured radius data extracted by the image processing module. Under ordinary conditions, the pupil should react very quickly, and constrict noticeably when a white light beam is directed to the eyes. Under the influence of certain substances such as alcohol, the pupil’s reaction may be very slow. Experiments show that pupil’s reaction is considered slow if it takes more than one second (roughly 30 images frames) to reach full constriction.

2.3 Alcohol workshops

The enhanced version of the video-based image processing system has been tested in alcohol workshops held at the École Nationale de Police du Québec (ENPQ). Alcohol workshops are used to train new recruits on the use of the breath analysis instruments. Workshops usually last about 4 hours, during which volunteer drinkers are consuming alcoholic beverages and snack food for about an hour and a half. After the consumption period, the BAC of each subject is periodically measured at time intervals of 15 minutes. Pre-dose evaluations were performed at the beginning of each workshop, before the subjects start consuming alcohol.

Many volunteer drinkers were involved in the testing of the recently developed system. Subjects were recruited from the local police academy in Trois-Rivières. Each subject signed an informed consent form previously approved by the research ethical comity of the Université du Québec à Trois-Rivières.

All subjects were of legal drinking age (19-26 years), healthy and reported no sign of fatigue since the workshops were all scheduled at 11h30 in the morning. Blood alcohol levels were assessed at each test time during each workshop using calibrated breath analysis instruments. In the course of each workshop, certified breath analysis specialists trained the recruits on how to performed BAC measurements using a Breathalyzer 900A.

The automated HGN, convergence and pupil’s reaction to light tests were performed with the newly implemented video-bases image processing system at the same rate of the BAC measurements.

3. Automatic eye involuntary reflexes measurements: Experimental results

The video-based image processing system introduced above to automatically assess human’s faculties (Meunier & Laperrière, 2008) (see Fig. 8 right) was used to extract the pupil’s horizontal angular position curves in the HGN test, the pupil’s horizontal deviation curves deduced from the convergence test and the pupil’s radius curves obtained from the pupil’s reaction to light test, from video images sequences.
Figure 9 shows an example of the region segmentation process which reveals the pupils in an image. Figure 9 right exhibits both pupils which correspond to regions with elliptic shape. The pupil’s regions segmentation operation is performed by simple binary thresholding (Gonzalez & Woods, 2008). The segmented bright regions in Fig. 9 right is related to darker regions of the original intensity image (see Fig. 9 left) such as the pupil’s regions. By looking to a typical pupil’s image region (see Fig. 10 left) and its corresponding intensity histogram (see Fig. 10 right) one can easily extract the threshold value used for the segmentation process. This threshold value (dashed line in Fig. 10 right) is actually the intensity value isolating the left most mode of the histogram corresponding to the pupil’s intensity probability distribution. The observation of Fig. 10 left also reveals the presence of the glint (brighter region) and its intensity probability distribution corresponding to the right most mode in the pupil’s image region intensity histogram (Fig. 10 right). The larger regions visible in Fig. 9 right are considered as shadow. The occurrence of shadow is a side effect of the non uniformity of the NIR illumination as well as the non uniformity of the face’s surface. Nevertheless, the enhanced NIR illumination system improved the illumination pattern and allows the decrease of the shadow effect.

Fig. 9. Results of the pupil’s segmentation process. (Left) Typical image of the pupils captured with the video-based image processing system. (Right) Segmented regions with the pupils appearing as circular shapes.

Fig. 10. Pupil’s image region intensity histogram. (Left) Typical pupil’s image region. (Right) Pupil’s image region histogram.

These regions are eliminated at the validation stage since their respective area and shape are not compatible with pupils. Fig. 11 depicts the results of the pupil’s extraction algorithm.
Figure 11 left shows the contours extracted from the segmented image (Fig. 9 right). Figure 11 right presents each pupil located by a circle with corresponding radius.

Prior to performing any comparison between the pupil’s horizontal angular position curves at post-dose BAC level and the approximate curves related to the pre-dose BAC level obtained from the HGN test, we first deduced a confidence interval for the mean value of the square difference (MSD) between the pupil’s horizontal angular position curves at BAC level of 0.00% and their corresponding polynomial approximations.

This confidence interval establishes a base of comparison which represents a sober state (BAC ≈ 0.00%). In this research we used a confidence interval with a confidence level of $1-\alpha = 0.90$ given by $1.722439 \leq \mu_{MSD} \leq 4.681783$. Figure 12 shows a typical horizontal angular position curve of the left pupil of one of the volunteer drinkers at BAC level of 0.00% (pre-dose) following visual stimuli generated in the HGN test.

![Figure 11](image1.png)

Fig. 11. Results of the pupil’s position detection. (Left) Contour extracted from the segmented image. (Right) Located pupils.

![Figure 12](image2.png)

Fig. 12. Left pupil horizontal angular position curves of a subject at a BAC of 0.00% following visual stimuli in the HGN test. The continuous curve represents the extracted positions and the dashed curve its corresponding approximation polynomial.
The curves depicted in Fig. 12 are not revealing great discrepancies which is usual for healthy persons not impaired by alcohol or drugs. Such curves were used to deduce the previously mentioned confidence interval. For BAC values greater than 0.00%, the pupil’s horizontal angular position curves start to show discrepancies with their related approximate polynomials. When the amplitude of the discrepancies increased sharply in saccadic fashion it means that the HGN is present.

Figure 13 shows the same curves depicted in Fig. 12 but for a subject with a BAC level of 0.22%. The differences between the curves are easily noticeable and reveal the presence of two visual signs used to assess the type of impairment of an intoxicated individual, the lack of smooth pursuit also called tracking error and the HGN. The dashed curve shown in Fig. 12 and 13 represents the ideal pupil’s horizontal angular position curve associated with an unimpaired individual and is compared with the continuous pupil’s horizontal angular position curve which allows the detection of the lack of smooth pursuit and also the distinctive nystagmus at maximum deviation (40-45°).

Fig. 13. Left pupil horizontal angular position curves of a subject at a BAC of 0.22% following visual stimuli in the HGN test.

Figure 14 shows the pupil’s horizontal angular position curves corresponding to a BAC level of 0.09% which is within the reach of the criterion BAC level of 0.08% considering the uncertainty (10%) of the BAC measurement taken by the Breathalyzer 900A. From Fig. 14 one can easily notice that the amplitude of the discrepancies between the pupil’s horizontal angular position curves are much smaller at lower BAC levels and the jerking at maximum deviation is less frequent (lower frequency) and also with lower amplitude.

The pupil’s horizontal deviation curves extracted in the convergence test may reveal the lack of convergence where the horizontal deviation of one eye may converge toward the tip of the nose while the other diverges away from the tip of the nose. Examples of eye’s convergence and divergence are shown in Fig. 15 (left and right). Figure 16 shows a typical pupil’s horizontal deviation curves corresponding to a BAC level of 0.00%. One can easily notice by observing the pupil’s horizontal deviation curves in the time interval 30-60 that both eyes converge. When the lack of convergence (see Fig. 17) occurs one of the pupil’s horizontal deviation curves tend to diverge from one and other in the same time interval.
Fig. 14. Left pupil horizontal angular position curves of a subject at a BAC of 0.09% following visual stimuli in the HGN test.

Fig. 15. Examples of eye convergence and divergence. (Left) Convergence: Both eyes are pointing to the tip of the nose. (Right) Divergence: One eye (volunteer’s right eye) is pointing to the tip of the nose while the other is pointing away from the tip of the nose.

(30-60). In the conducted studies this phenomena was not observed for all subjects even with fairly high BAC levels.

The pupil’s reaction to light test assesses the response time of the involuntary reflex of the pupils triggered by changing illumination. A pupil’s slow reaction to changing illumination conditions may reveal impairment caused by certain drugs such as alcohol. Figure 18 depicts the pupil’s radius of the left eye of subjects with two different BAC levels subjected to the reaction to light test. The pupil’s radius is deduced from a circle/ellipse fitting algorithm such as the one exposed previously (see section 2.1.3) and gives results such as the ones showed in Fig. 11 right. The continuous curve was obtained with a subject at a BAC level of 0.138%. From that curve one can deduce the response time of the involuntary reflex of the pupil which is the elapsed time between the time the white LEDs are switched on and the time of the pupil’s full constriction. The response time deduced at this BAC level (BAC: 0.138%) is roughly 1400 ms. The response time deduced from the dashed curved (BAC: 0.09%)
0.00%) depicted in Fig. 18 is about 850 ms. Experiments (Page, 2000) suggest that a response time of more than 1000 ms is considered a slow reaction to light and may be linked to CNS drug related impairments.

Fig. 16. Left and right pupil’s horizontal deviation curves of a subject at a BAC of 0.00% following visual stimuli in the convergence test. The continuous curve represents left pupil horizontal deviation curve positions and the dashed curve the right pupil’s horizontal deviation curve.

Fig. 17. Left and right pupil’s horizontal deviation curves of a subject at a BAC of 0.138% following visual stimuli in the convergence test.
Fig. 18. Left pupil radius measured in the reaction to light test. The continuous curve corresponds to a subject with a BAC of 0.138% and the dashed curve to a subject with a BAC of 0.00%.

The newly developed video-based image processing system has been tested to classify a group of subjects into one of two groups (impaired/unimpaired) based on the detection of the four visual signs (lack of smooth pursuit and the presence of horizontal gaze nystagmus at maximum deviation for both eyes) used in a previous study (Meunier, 2006). Other visual signs (eye convergence, eye reaction to light) were also tested to see if they increase the certainty of the diagnosis on a subject type of impairment. The criteria used in the classification process (human faculty diagnosis) and their reference values are shown in table 1. A person is considered impaired if the measured eye involuntary reflexes exceed their reference criteria. Moreover, the criterion for impairment in the case of alcohol related intoxication is always based on the applied BAC level of 0.08% by law enforcement agencies.

<table>
<thead>
<tr>
<th>Classification criteria</th>
<th>Reference values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eye Tracking error (Mean Squared Difference: degree²)</td>
<td>4,6817</td>
</tr>
<tr>
<td>HGN at maximum deviation (40-45°) (MSD: degree²)</td>
<td>9,3267</td>
</tr>
<tr>
<td>Eye divergence (MSD: pixel²)</td>
<td>5,0322</td>
</tr>
<tr>
<td>Reaction to light (sec)</td>
<td>1,0</td>
</tr>
</tbody>
</table>

Table 1. Classification criteria and their reference values.

The previous version of the video-based image processing system for human’s faculties assessment (Meunier, 2006) reported a classification accuracy of about 80%. The experiment conducted with the new version of this system gives roughly the same classification rate.
One main objective of this recent study was to evaluate the efficiency of the new version of the automatic human impairment detection system especially to cope with the false negative failure rate. This figure was grossly 22% for the previous implemented system prototype and represented the fraction of impaired subjects actually not detected. The false negative failure rate is about the same reported in the previous study (Meunier, 2006). Nevertheless, the success rate of the new system version is almost perfect for a high dose of alcohol (BAC > 0.1%). However, the failure of the classification occurs mainly at BAC levels within about 10% of the criterion (BAC ≈ 0.08%).

4. Discussion and future work

The results obtained in the previously presented experiment confirmed the observations made in the previous study (Meunier, 2006). These results are also consistent with other previously published experiments (Burns, 1995) on the HGN test and confirmed its validity to detect alcohol related impairments. In the more recent experiments we also used other tests such as the eye convergence and the eye reaction to light tests with the HGN test procedures. The results obtained suggest that these other visual signs are not essential for the detection of alcohol related impairments and they are not adding much discriminating power to the classification process. This observation leads us to conclude that the HGN test is the primary tool to accurately detect alcohol related impairments since the use of the visual signs associated to the darkroom examination test had little effect in the classification efficiency reported in Burns (Burn, 1995).

The implemented eye convergence test needs to be reconfigured and redesigned to improve the detection of the lack of convergence that should had been more consistently detected in alcohol related impairments encountered in the study presented here. Some volunteer drinkers mentioned during the execution of the convergence test experiencing focussing problems on the visual stimuli. In order to fully reveal the classification power of the newly implemented eye involuntary reflexes further experiments must be conducted. These experiments must also be held in the context of drug workshops in which subjects are consuming drugs such as cannabis.

Results also suggest that the automated system implemented to detect human’s impaired faculties based on the eyes involuntary reflexes such as the horizontal gaze nystagmus, the eye convergence and the pupil’s reaction to light needs to be improved in order to allow better detection rate at BACs near 0.08% or low level impairment. As previously stated, adding more visual signs (eye’s involuntary reflexes) improved the impairment detection efficiency. Nevertheless, adding other types of information related to the DRE evaluation such as divided attention tests, examination of vital signs, examination of muscle tone to the automated impaired faculties detection process should greatly improved its efficiency. Therefore, a probabilistic framework similar to the one implemented in a research reported in the literature (Ji et al., 2006) that is based on the Bayesian networks for modeling and inferring human’s impaired faculties by fusing information from diverse sources will be developed in future research. This probabilistic framework will first allow the modeling of a Bayesian networks by which the criteria maximizing the certainty of a diagnosis given on a person’s faculties are chosen. Moreover, the Bayesian networks can be used to infer from the fusing information process of multiple sources the probability that a person’s faculties are
impaired by a given category of drugs. Figure 19 shows a first version of the Bayesian networks of the probabilistic framework dedicated to the modeling and inferring of impaired faculties.

Fig. 19. Bayesian networks for modeling and inferring human’s impaired faculties.

5. Conclusion

In this book chapter we outlined a new video-based image processing system implementing the HGN, eye convergence and eye reaction to light tests. This system as the one reported in a previous study (Meunier, 2006) accurately detects impairments with a success rate of about 80%. The present system is also quite efficient to detect impairments for high doses of alcohol but still needs some refinements to increase its success rate for low doses near the BAC criterion of 0.08%, the level to which the false negative failure rate is higher. The use of a probabilistic framework fusing information to improve the classification accuracy will be investigated. Nevertheless, the actual system may also be used to detect other impairments caused by drugs. By using more eye involuntary reflexes (visual signs) the newly developed video-based image processing system prototype is more versatile and can be used to detect a wide array of drug related impairments. It can also be used for the automatic detection of liveness since a living human exhibits eye involuntary reflexes such as the pupillary reflex.
The implemented automatic human faculties assessment system presented in this chapter (Meunier & Laperrière, 2008) is also useful for many other reasons. First, as an educational and interactive tool to assist trainers in alcohol workshops. Second, to document the eye’s involuntary reflexes related test procedures with video sequences which subsequently can be presented in court of law to prove that an individual is under the influence of alcohol or drugs. Third, to standardize the application of these test procedures to avoid inconsistencies in its administration by law enforcement officers.

6. References


www.intechopen.com
There are six sections in this book. The first section presents basic image processing techniques, such as image acquisition, storage, retrieval, transformation, filtering, and parallel computing. Then, some applications, such as road sign recognition, air quality monitoring, remote sensed image analysis, and diagnosis of industrial parts are considered. Subsequently, the application of image processing for the special eye examination and a newly three-dimensional digital camera are introduced. On the other hand, the section of medical imaging will show the applications of nuclear imaging, ultrasound imaging, and biology. The section of neural fuzzy presents the topics of image recognition, self-learning, image restoration, as well as evolutionary. The final section will show how to implement the hardware design based on the SoC or FPGA to accelerate image processing.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:
