MOTION ESTIMATION FOR MEDICAL APPLICATION IN PSYCHIATRY

[1]*Olivier Laligant,* [1]*Remy Leconge,* [1]*Fred Truchetet,* [2]*Olivier Mahu and* [1]*Alain Diou*

[1]o.laligant@iutlecreusot.u-bourgogne.fr

[1]Le2i Lab., 12 rue de la Fonderie, 71200 Le Creusot

[2]Centre Hospitalier Spécialisé, 58400 La Charité sur Loire

ABSTRACT

Some systems have been developped to help psychiatrist to do diagnoses. Among the tried approaches to characterize disease by systems, the study of face movements seems very informative concerning the state of the patient. We present in this work a set of methods allowing to characterize 3D head and eyes movements in the following shooting condition: mobility of the subject in sitting position, background lighting of the room. Our approach yields in a simple way the localization of iris and the characterization of their movement in the three dimensional space. The 3D data are derived from stereoscopic system. The absolute 3D movement of eyeballs and their relative movement with regard to the head are obtained, even if this one are moving. A multiresolution method to deal with movements of great differences in amplitude, like eyes closure, is also presented. No specific lighting has to be used.

1. INTRODUCTION

This research was carried out within the framework of psychiatry, where we are interested in the movement of patient's face and particularly the eyes and the eyelids. The study of face movements seems more interesting and informative to characterize disease than walk analysis for example. We have developped methods for head and eyes motion tracking. A prototype has been built on a steroscopic configuration of cameras. Another prototype is based on a high speed camera (Kodak). A screen and a computer are used to create stimuli in front of the patient.

Head, eyes and eyelids movements present great differences in amplitude. Moreover, area of eyelids is not entirely structured and the speed is not homogenous. The motion characterization must be sufficiently accurate to take into account these properties. Numerous domains are interested in the characterization of human movements and in particular the eye movements (customer attention, driver attention, ...). Most of proposed methods allow to track eyes from data 2D (typically a sequence of image). These techniques estimate the apparent eyes movement between two images. But optical flow is different from the real movement when motion is a rotation. Eyes movement is composite and consists of two components. The first one is the absolute movement of the head and second one is the relative eyes movement with regard to the head. The component of relative movement is the only one leading to the real eyes motion.

In this work, we have developped specific methods to detect head movement, eyes movement and eyelids movement. Our approach leads to an estimate of these movements in normal conditions of acquisition (background lighting, freedom of movement of the subject in sitting position). The second section explains the localization method of the iris which

allows to estimate their real motion. The case of the eyelids is considered in the third section. Fourth section exposes the stereoscopic acquisition which gives three dimensional data. These data allow to compute the real movement of the observed scene. Then we present the method to compute the relative eyes movement with regard to the head. Finally, medical application in progress of this work is described.

2. DETECTION AND LOCALIZATION OF IRIS

Most algorithms use a specific lighting to localize eyes in front of a camera. This ligthing, generally infrared [1] [2] [3], produces a reflection on the cornea. In this way eyes are localized by detection of maximum of luminance. Unfortunately, some applications or particular lighting conditions do not allow any additional lighting. Consequently, it is necessary to use a method suitable to detect and localize iris whatever the background lighting conditions. The neural networks [4] can solve this constraint but require a learning. Algorithms based on deformable models [5] use a generic shape of an eye. This shape is deformed in a iterative process to match the eyes in the image. These methods present an important number of iterations which penalize the computing time. Consequently, this one is often reduced by using a localization process based on a specific lighting. Below, a simple and fast method is proposed to localize iris without any specific ligthing.

2.1. Detection and localization mask

Iris are the only circular parts of the human face. A detector of circular shape would localize the center of the two eyes. However, the precision of the localization depends on the person in the front of the camera. Precision varies according to the degree of opening of the superior eyelid. Figure 4 present some examples of subimages of eyes. Persons have naturally the superior eyelid which covers the high part of the iris. This mask effect disrupts the detector and leads to bad results. The covering of the lower eyelid is negligible and does not influence results of localization. This is why the operator of detection and localization is based on a detector of half circular shape rather than a detector of circular shape.

2.2. Algorithm

The half circular mask C_n of radius R_n is shown figure 1 where several masks are presented. The mask can be defined as following:

$$C_n(i,j) = \frac{1}{p_{R_n}} \quad (1)$$

with

$$i = round(R_n.cos\alpha)$$

$$j = round(R_n.sin\alpha)$$

$$\alpha \in [\pi, 2\pi]$$

where p_{R_n} is the length of the perimeter of the half circle of radius R_n and i and j are local pixel coordinates.

The algorithm is based on the Hough transform. N masks $\{C_n\}_{n\in[0,N-1]}$ are constructed where R_n varies in the range $[R_{min}, R_{max}]$. R_{min} and R_{max} are chosen according to shooting condition in order to detect iris of radius R_{iris}: $R_{iris} \in [R_{min}, R_{max}]$.

[Figure 1 about here.]

From the centre of the iris, the luminance profile (S) of the eye in the direction α can be considered as an Heaviside function (Y). In the same way, the mask C_n will be considered as the response signal to the iris detector. In the direction α, on the r axis (on the radius):

$$C_n(r) = M_n * S(r) \tag{2}$$

where $M_n(r)$ is the unknow detector and $S(r)$ the eye luminance profile. Then:

$$\delta(r - R_n) = M_n(r) * Y(r - R_n) \tag{3}$$

We deduce easily the theoretical result ($M_n = \frac{d}{dr}\delta$) and a regularized expression for M_n:

$$M_n(r) = [-1 \;\; 0 \;\; 1] * h(r - R_n) \tag{4}$$

with h a regularization function. In this work, h is defined as the Parzen's window:

$$h = \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{bmatrix} \tag{5}$$

Finally, the detector mask is contructed as following:

$$M_n = \begin{cases} -\frac{1}{p_{R-2}} & \text{for the radius} \quad R-2 \\ -\frac{1}{p_{R-1}} & \qquad\qquad\quad R-1 \\ 0 & \qquad\qquad\quad R \\ \frac{1}{p_{R+1}} & \qquad\qquad\quad R+1 \\ \frac{1}{p_{R+2}} & \qquad\qquad\quad R+2 \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

The convolution of the N masks with one image of the sequence generates N images results. The pixel contrast value is then:

$$\{F_n[x, y] = (M_n * I_k)[x, y]\}_{n\in[0,N-1]} \tag{7}$$

The iris localization correspond to the contrast maximum detection of the border of the iris. The localization of the eyes is obtained in the following way:

- for each image I_k, where k is the index of the images sequence:

- compute N images $\{F_n[x,y]\}_{n\in[0,N-1]}$, n denotes the mask index
- for each of the N image $\{F_n[x,y]\}_{n\in[0,N-1]}$:
  - $*$ search for the two highest absolute maximums of the gradient which can represent iris localizations: $m_{k,n,1}$ and $m_{k,n,2}$. The curves in figure 2 represents $m_{k,n,1}$ and $m_{k,n,2}$.
- extract the two highest maximums among the two lists above: $m_{k,1}$ and $m_{k,2}$. $m_{k,1}$ and $m_{k,2}$ represent the localization of the eyes for the image I_k.

[Figure 2 about here.]

2.3. Implementation

[Figure 3 about here.]

All operations are processed in the Fourier space to save computating time. The Fourier transform of the masks are first calculated (offline computing). Then processing of $\{F_n[x,y]\}_{n\in[0,N-1]}$ is carried out in Fourier space (see figure 3). The rest of the algorithm is performed as described in the above subsection.

2.4. Results

[Figure 4 about here.]

The robustness of the algorithm has been tested with persons having more or less opened eyes and with or without glasses. The results show that localization can be considered as accurate as long as the eyelid covering does not exceed 50% of the iris surface. In this case the localization of the iris is guaranted with a precision of 0.5 pixel. The figure 4 shows some subimages results among the realized tests. These images show clearly that no specific lighting has been used.

3. EYELIDS MOVEMENT

Knowing the localization of the iris, the areas corresponding to the eyelids can be deduced. During the eyelid closure, areas tracking can be assisted by head movement estimation (presented in the next section). The movements of eyelids present great differences in amplitude and their area is not well defined. A motion is considered as a motion of large amplitude when it exceeds three or four pixels per frame. There are a lot of techniques of movement estimation, but generally they are classified in three main categories: differential methods (methods of Horn [6], Schunk and Nagel [7],), mapping methods [8, 9] (block matching), and transform method (Gabor transform, Fourier trans-form) [10, 11]. The first methods give good results, provided that only movements of small amplitude are considered. The second techniques operate correctly on structured elements and motions of different amplitudes, nevertheless the precision of this process is not a subpixel one. The last methods are exploitable only for elementary movements. In our case, movements have various

amplitudes. So a simple and efficient method to compute motion estimation for a large range of speed is proposed. The complement of differential methods and block matching allows us to characterize all the movements in image sequences. The cooperation of these two methods is achieved by a multiresolution strategy. In the next subsections are presented the two used methods and the algorithm mixing these methods. This section is closed witht he obtained results.

3.1. Markov Random Fileds (MRF) and Block Matching Process (BMP)

MRF make assumption that the luminance of a pixel representing the same information is constant in two successive images. MRF modeling is equivalent to an energy function, and the motion estimation problem is reduced to the minimization of this energy function. Classically, the energy function is composed of two parts .The first one is the data energy and the later is the Tikhonov regularization energy:

$$E(s,t) = \left[f\left(s + p_{s,t}, t + dt\right)\right]^2 + \sum_{\{s,s_j\} \in C_2} \beta \left\| p_{s,t} - p_{s_j,t} \right\|^2 \tag{8}$$

where :

- s is the current site (pixel)
- $p_{s,t}$ is the motion vector on the site s at the time t. This vector has two components: $p_{s,t} = (u_s, v_s)$
- s_j is in a set (C_2) of sites in the neighborhood of s
- β is the parameter of regularization ($\beta \in [0, \infty[$). Small values lead to an accurate solution for image data, but results are sensitive to noise. High value implies smooth solution, but results do not match the luminance data.

The solution $p_{s,t}$ is obtained by an iterative process that converges toward the final value. Unfortunately, this method leads to incoherent results for great amplitude movements.

The estimation of motion by block matching consists (choosing a block in an image) in finding the best similar block with respect to criterion in another reference image. The criterion C of similarity is defined as following:

$$C = \sum_{B} |(I_1 - \bar{I}_1) - (I_2 - \bar{I}_2)| \tag{9}$$

where I_1 and I_2 are the luminances of blocks, respectively, in the previous image and in the current image. $\bar{I}_1$ and $\bar{I}_2$ are the average pixel luminances of blocks, respectively, in the previous image and in the current image. B is the subset of pixels of the block.

The block matching allows the estimation of great amplitude movement if initialization and correlation are correctly chosen.

3.2. Multiresolution motion estimation

A multiresolution approach is used to merge these two methods. Indeed, if a movement is significant on scale 0 (initial image), it will become a movement of low amplitude in scales of coarser resolution. The MRF is used to characterize this small displacement, and motion tracking across the scales is achieved by BMP. In this way, BMP initialization is obtained with MRF.

Let there be a temporal sequence of $n+1$ images:

$$\{I_0, I_1, ..., I_k, ...\}_{k\in[0,n]} \tag{10}$$

where k denotes the time.

First of all, a multiresolution pyramid of each image I_k of the sequence is built:

$$P_k = \{I_{k,0}, I_{k,1}, I_{k,2}, ..., I_{k,j}, ...I_{k,J}\} \tag{11}$$

where j denotes the scale and $J+1$ the number of scales.

Starting on the lowest scale (J), the algorithm consists in building gradually the pyramid of movement:

$$(U_{k,J}, V_{k,J}), (U_{k,J-1}, V_{k,J-1}), ..., (U_{k,J}, V_{k,J}), ..., (U_{k,0}, V_{k,0}), \tag{12}$$

where $(U_{k,j}, V_{k,j})$ depicts the movement in the image I_k at scale j. The movement image $U_{k,0}, V_{k,0}$ is the expected result. All calculations will be carried out on a pair of successive pyramids noted (P_{k-1}, P_k). The algorithm is organized as following:

for each image I_k of the sequence $\{I_0, I_1, ..., I_k, ...\}_{k\in[0,n]}$

let there be (P_{k-1}, P_k), initialization is processed by MRF on the coarsest (J) image and produces vector $(U_{k,j}, V_{k,j})$ with $j = J$

for each scale $j = j-1$ of the pyramid until scale 0

- vector are propagated on the next finer scale to initalize $(U_{k,j}, V_{k,j})$
- at the finer scale, two configurations are possible to refine $(U_{k,j}, V_{k,j})$:
  - the initial vector is null, and in this case the motion vector is determined by MRF,
  - or a nonnull initial vector is found and used to initialize the BMP. Remark: motion detected for a given point at scale $j+1$ involve initial nonnull vectors at scale j in the 4-neighborhood. Consequently, all these points (scale j) are checked to know if they correspond to real movement. So, before estimation by the block matching, the propagation of movement is refined by MRF. If a point does not correspond to movement, its initialization is forced to 0.

the final result is $(U_{k,0}, V_{k,0})$

More details can be found in [12].

3.3. Speed estimate

[Figure 5 about here.]

[Figure 6 about here.]

For precision evaluation the method has been applied to estimate the local speed (see figure 5) of a rotating disk. A white disk with a black sector has been designed. The disk put in uniform rotation with angular velocity of $\pi/3$ rd/s has been shot at 25 images/s. Figure 6 presents the precision of the MRF and BMP cooperation. Clearly this method takes the advantage of each process. Indeed the curve is composed of two parts corresponding to the impact of the Markov method for the lowest speed and the block matching process for the highest speed. The speed increases by 1 pixel/s when the radius increases by 23 pixels. Motion estimate by Markov random fields is subpixel, consequently the experimental curve is similar to the theoretical one. The motion estimate using the block matching process is constant at intervals of 23 pixels, because the block matching process has a resolution of one pixel.

3.4. Results

[Figure 7 about here.]

[Figure 8 about here.]

[Figure 9 about here.]

Application to the volontary eyelid closure is presented in figure 7. The algorithm seems reliable for the determination of the direction and norm of the motion vectors. Some estimation errors are due to noise and aliasing and they will be eliminated in future experimentation with a full frame camera. Figure 8 shows two image pyramids where the complement of MRF and BMP is clear. The curve in figure 9 depicts the eyelid velocity maximum as a function of the time. The experiment seems to point out two phases in this movement: beginning and middle courses of eyelid movement, where acceleration and then decceleration are regular; and the end of eyelid movement, which present a different speed evolution. Further experiments will be made to confirm these observations.

4. 3D MOVEMENT ESTIMATE

The method presented in the second section allows to localize iris and to estimate apparent motion. We present in this section the method to estimate the 3D movement of eyes and head by using a classical stereo vision system [13] [14]. The algorithm of localization of iris applied to the stereo images allows easily to determine 3D position of the two eyes. The difficult task of the stereo system is to estimate the movement of the head. Knowing this, composite movement of the eyes can be deduced. This movement has two components : the movement of the eyes relative to the head and the movement of the head.

4.1. Head motion

Mapping points correspond to the matching of the two images of the stereo system. The matching is processed on the base of the Fua criteria [15]. The 3D image is then obtained by triangulation from the mapping points.

[Figure 10 about here.]

The estimate of head movement is carried out with a system of spherical coordinates. In the conditions of acquisition presented in the introduction, the person who is filmed must read a text or follows a moving target on a screen. In such a situation, head movements are assumed as rotations (according to 3 axes). In the spherical system of coordinates, the rotation of the head between two images is mathematically expressed as a translation. The transform of cartesian coordinates to spherical coordinates is achieved by estimate the position of the occiput hole of the head. This is the pivot between the skull and the vertebras. This point will be the origin of the new spherical coordinates system (figure 10).

[Figure 11 about here.]

The depth images are constructed from these spherical coordinates. The grey level of the image is proportional to the depth ρ and the pixel coordinates are varying according to φ and θ. Due to the big size, the data are sub-sampled version of cartesian images. Only surfaces with normal near the normal of the cameras are taken into account because the reconstruction of the other surfaces is less precise. The estimate of rotations consists in identifying translations in these images. This is achieved by correlation between two images (figure 11) which gives the translation (rotations $\Delta\varphi, \Delta\theta$). The rotation on the third axis is deduced from the iris position in cartesian coordinates.
This three angles characterize the head movement in the spherical coordinate system and allow to determine the relative movement of eyes. A second transform allows to formulate the head motion in respect of the cameras coordinate system (absolute eyes motion is already known in this coordinate system). In this way the movement of eyes with regard to the head is obtained by difference between the absolute movement of the eyes and head movement.

4.2. Temporal improvement

[Figure 12 about here.]

Our approach leads to an estimate of movements in normal conditions of acquisition : background lighting, freedom of movement of the subject in sitting position. These conditions make work difficult. Nevertheless, not homogeneous lighting is used in this work to improve 3D reconstruction along the temporal images sequence. Head movements involve different and complementary results at correlation stage. New results are merged with current reconstruction according to a quality criterion. This criterion is the confidence one computed at the matching stage. The figure 12 present a depth image (twentieth in the sequence) of raw 3D reconstruction and the improved one (due to a lateral movement of the head).

4.3. Accuracy

[Figure 13 about here.]

Considering the system parameters and the precision of the iris localization, the 3D components of the iris are given with a relative error of 1%. Concerning the head motion, precision estimate is achieved by studying rotating movements of a synthetic head (see figure 13) on a calibration bench. The rotation measurements obey the Gaussian law with a standard deviation $\sigma = 0.36^o$.

4.4. Results

The results are obtained in the following experimental conditions. The stimulus is a virtual object moving from left to right and from right to left on a screen. The person placed in front of cameras follows this object. In these experimental conditions, the motion estimation of the head is shown in figure 14. The vector speed of eyes with regard to the head is deduced by difference between global eyes speed and head speed. This velocity is presented in figures 15 and 16.

[Figure 14 about here.]

[Figure 15 about here.]

[Figure 16 about here.]

The curves of global speed of head according to components x and z have sinusoidal shapes (due to the alternative movement of the head). Eyes curves have *peaks* which represent the jerk movements. Seeing that eyes follow an object moving horizontally, the component y is weak. These experiments reveal the pursuit mechanisms occurring when a person observe a moving object. This mechanism consists in a conjugated, slow and continuous movement, called eye pursuit. Eyes move generally in a proportional speed to the target. If this speed exceeds the maximum speed of eyes pursuit, corrective jerks (very fast movements) occur and break up the pursuit.

4.5. Medical application

[Figure 17 about here.]

[Figure 18 about here.]

A 2D eyes motion estimate system is currently used by a medical team at *Centre Hospitalier Spécialisé de la La Charité sur Loire*. They undertake to characterize eyes motion of some volunteers at different time and under different (sedative) treatments. We emphasize that this system part works at the rate of 200 images per seconds. Example of preliminaries results is shown in figures 17, 18. First measures confirm that eyes reflex (delay and then speed) decreases as sedative treatment increases. In such application, precision is validated by relevance of results. Presently, we are developping a

fast version of the prototype based on a ROI-camera (Region Of Interest). The team has in perspective to find additional characteristic variations of eyes movements according to the physiology state.

5. CONCLUSION

This work leads to eyes and head movements estimate without any specific lighting. The efficiency of the simple 2D eyes localization method allows to characterize eyes motion. A multiresolution scheme for the cooperation of two different approaches in motion estimation has been developped to estimate eyelids movements. The movement of the head is estimated by a classical stereo vision system. Computing strategy based on the assumption that movements are rotations yield precise results for head motion estimate [16]. Finally, the 3D eyes movement can be characterized.

While the precision and the robustness of each stage of the process have been assessed by various experiments, the accuracy of the overall approach cannot be easily estimated. Two ways could be considered: - make use of an artificial head with realistic eyelid, iris and fully controllable movements; - compare with results given by existing devices. The former approach is still to be performed in our laboratory. The latter has a serious flaw. Indeed existing commercial devices dedicated to eye movement estimates are based on the detection and tracking of the reflexion of a special lighting. And therefore it is the reflexion spot movement that is measured and it can be quite different form the eye one.

We have the perspectives to study the other movements of the face such as the movements of the jaw. Automation of the developments for medical prototype are in progress.

6. REFERENCES

[1] T. E. Hutchinson, K Preston White, J.R. Worthy, N. Martin, and K. C. Reichert, "Human computer interaction using eye gaze imput," *IEEE trans on systems, man and cybernetics*, pp. 1527–1534, 1998.

[2] Y. Ebisawa, "Improved video based eye gaze detection method," in *IEEE IMTC*, Hamamatsu, May 1998, pp. 963–966.

[3] C. Collet, "Capture et suivi du regard par un système de vision dans le contexte de communication homme-machine," *Thèse de doctorat, Ecole Normale Supérieure de Cachan*, 1999.

[4] S. Baluja and D. Pomperleau, "On intrusive gaze tracking using artificial neural networks," in *CMU technical report*, Carnegie Mellon University, Jan. 1994, pp. 94–102.

[5] A.L. Yuille, D.S. Cohen, and P.W. Hallian, "Feature extraction from faces using deformable templates," *Journal of Computer vision*, pp. 99–111, 1992.

[6] B.K. Horn and B.G. Schunck, "Determining optical flow," *Artif. Intel. 17*, pp. 183–185, 1981.

[7] H.N. Nagel, "On the estimation of optical flow: relation between different approaches and some new results," *Artif. Intel. 33*, pp. 299–324, 1987.

[8] B.C. Song and J.B. Ra, "A fast multiresolution block matching algorithm for motion estimation," *Signal Process. Commun. 15(9)*, pp. 799–810, 2000.

[9] C. Alexandre and H. Vu Thien, "Fast motion estimation algorithm," *Traitement du signal 13(4)*, pp. 351–359, 1996.

[10] A. Spinei, D. Pellerin, and J. Herault, "Spatiotemporal energy-based method for velocity estimation," *Signal Process. 65(3)*, pp. 347–362, 1999.

[11] D. Vernon, "Computation of instantaneous optical flow using the phase fourier components," *Image Vis. Comput. 17(3-4)*, pp. 189–199, 1999.

[12] R. Leconge, O. Laligant, F. Truchetet, and A. Diou, "Methods cooperation for multiresolution motion estimation," *Optical Engineering*, pp. 435–442, 2002.

[13] O. Faugeras and R. Keriven, "Complete dense stereovision using level set methods," in *Proc EECV*, Carnegie Mellon University, 1998, pp. 379–393.

[14] O. Faugeras, B. Hotz, H. Mathieu, and T. Vieville, "Real time correlation based stereo: algorithm, implementations and applications," in *Rapport de recherche INRIA*, 1993.

[15] P. Fua, "Combining stereo and monocular information to compute dense depth maps that preserve depth disconuities," in *International joint conference on artificial intelligence*, 1991, pp. 1292–1298.

[16] R. Leconge, "Estimation de la cinétique de certains éléments du visage," *Thèse de l'Université de Bourgogne*, 2002.

List of Figures

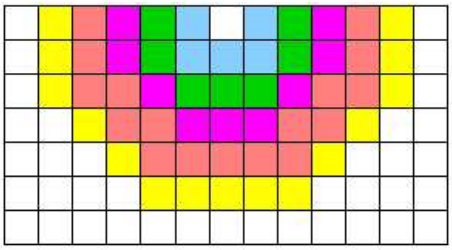

Figure 1: *Five half circular masks presented on the same illustration.* R_n *varies in the range* $[R_{min}, R_{max}] = [1, 5]$.

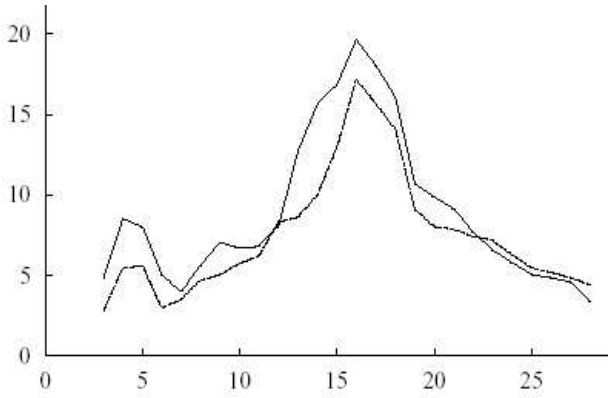

Figure 2: *Curves depicting iris contrast detection maximums as a function of masks radius. The highest value on each curve localize one iris.*

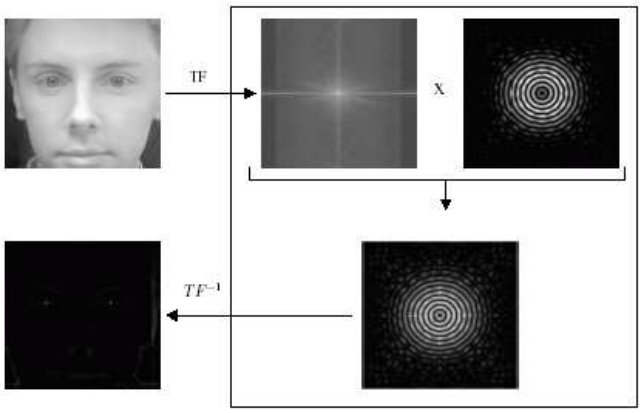

Figure 3: *Processing of F_n in the Fourier space with the mask of radius $R_n = 15$.*

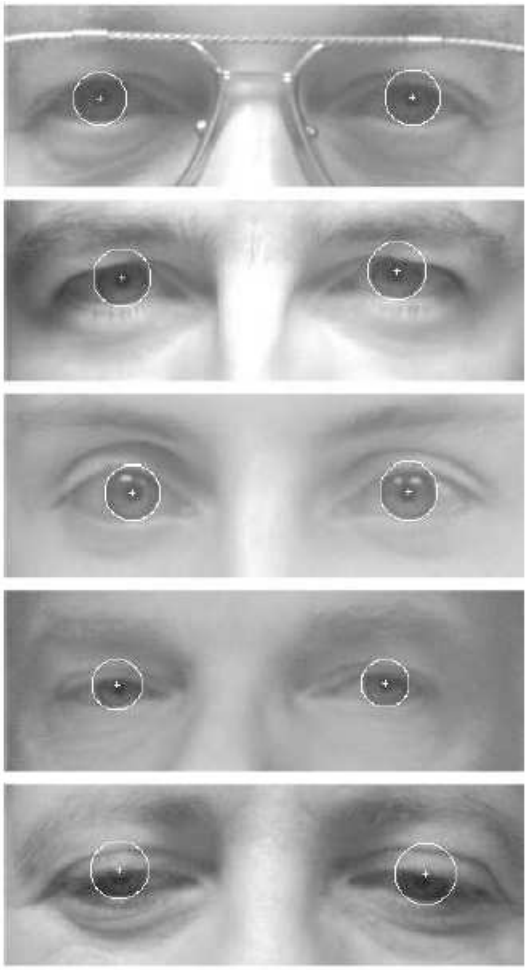

Figure 4: *Localization of the iris on original subimages. Images luminance shows clearly that no specific lighting has been used.*

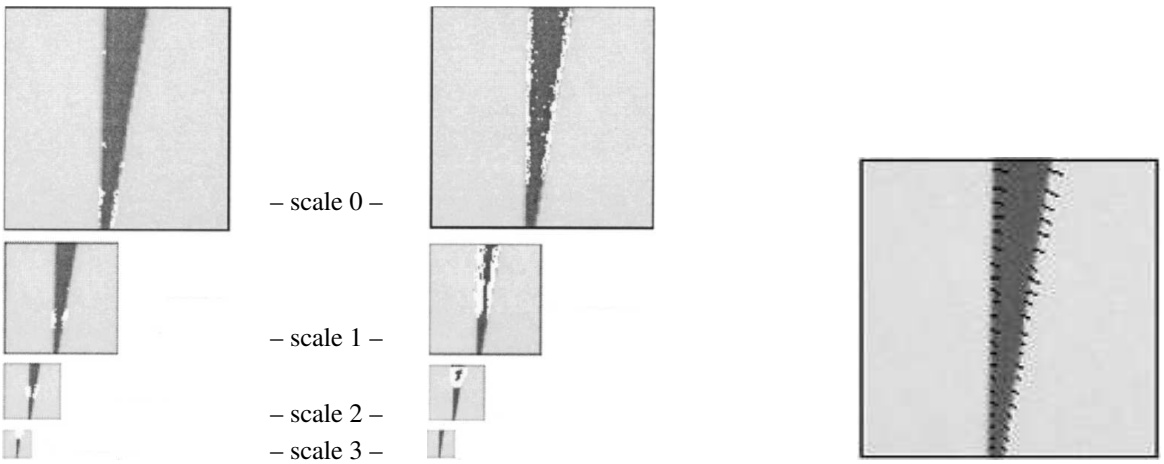

Figure 5: *Speed estimate (right, one vector out of four is shown) of the disk in uniform rotation: contribution at different scales of the Markovian approach (left), and of the block matching (center). The white points show pixels over which motion is detected (video rate: 25 images/s).*

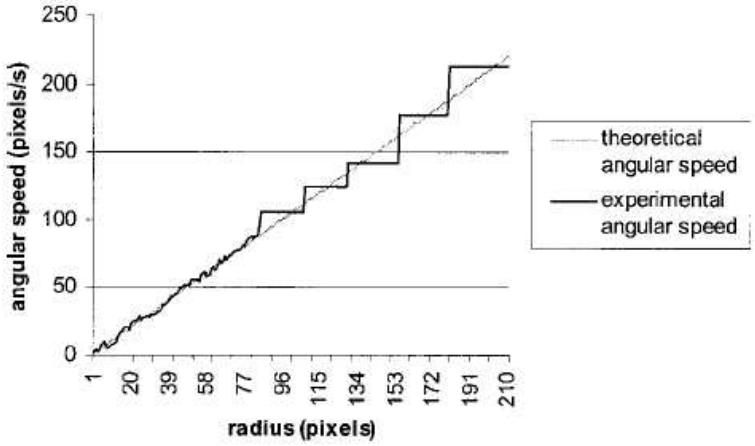

Figure 6: *Comparison between theoretical and experimental angular speed (video rate: 25 images/s).*

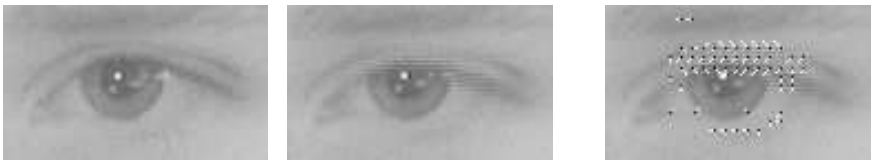

Figure 7: *Images and results for voluntary eyelid closure. For visibility, Only one vector out of four is shown. Horizontal aliasing is due to interlacing.*

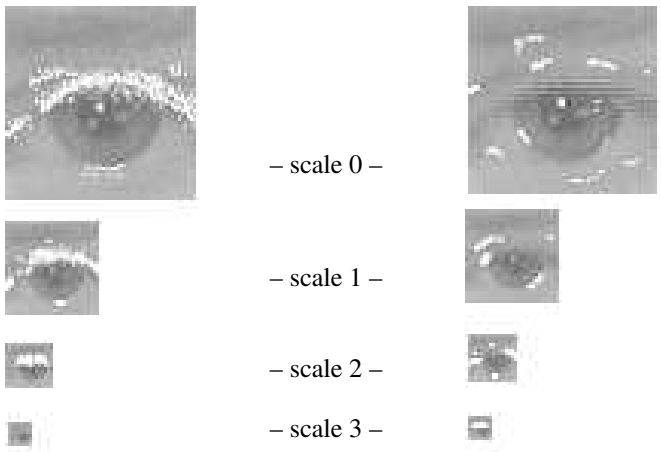

Figure 8: *Contributions of (left) the Markovian (MRF) approach and (right) of block matching (BMP). The white points show the pixels where motion is detected.*

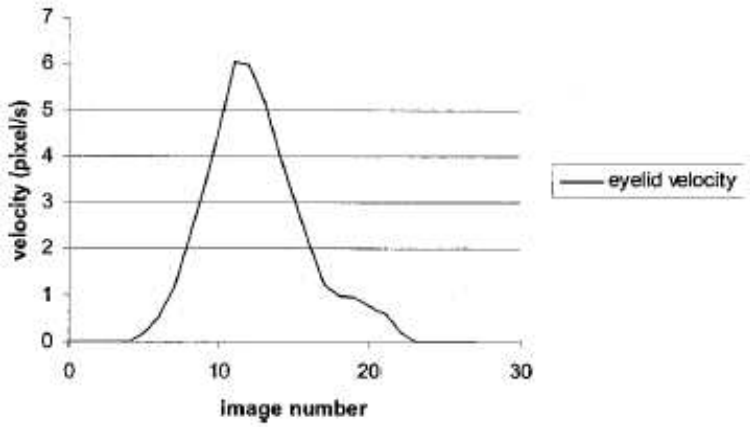

Figure 9: *Curves depicting the eyelid velocity (video rate: 200images/s). The movement seems to feature two phases.*

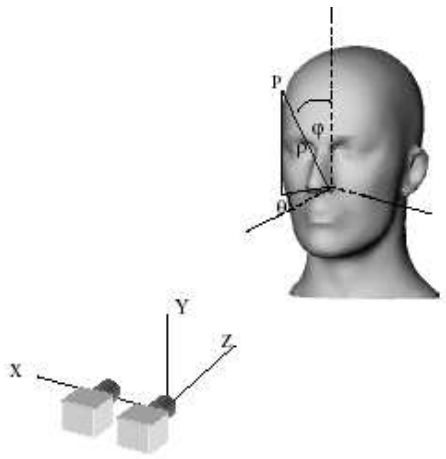

Figure 10: *Spherical coordinates.*

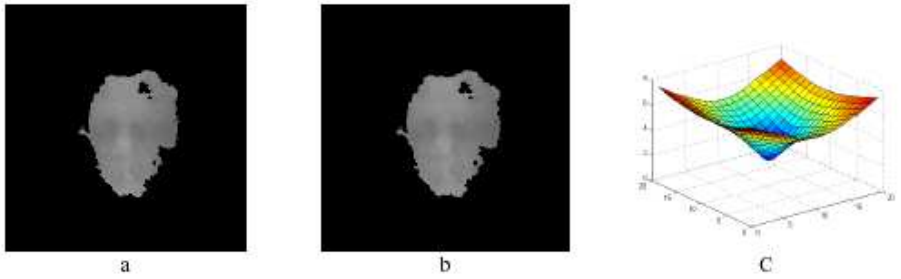

Figure 11: *Depth images in spherical coordinates and correlation function of these images.*

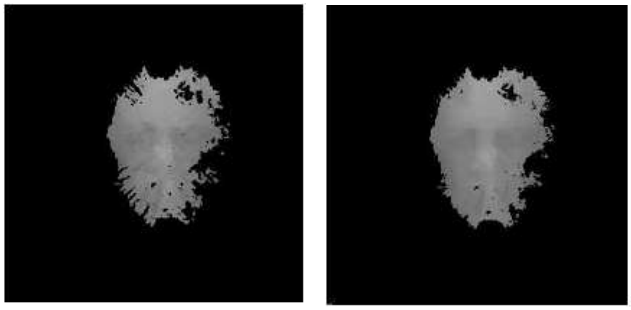

Figure 12: *Depth image (twentieth in the sequence) of raw 3D reconstruction and the refined one due to a lateral movement of the head.*

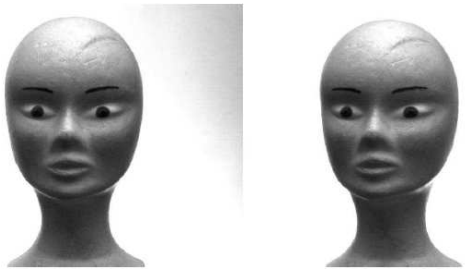

Figure 13: *Stereoscopic images of the synthetic head.*

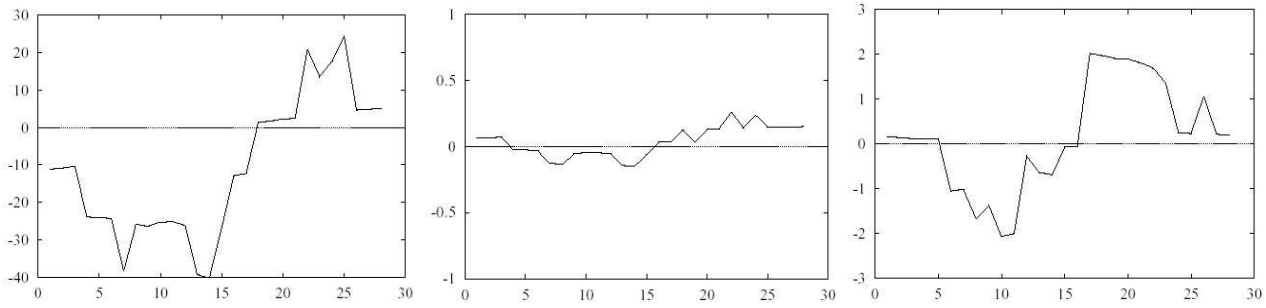

Figure 14: *Head motion: components x, y and z. Velocity (mm/s) curves as a function of the image in the sequence.*

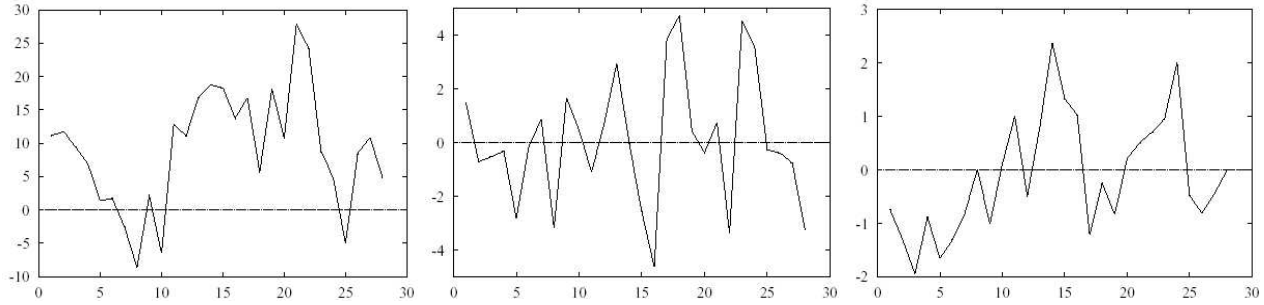

Figure 15: *Left eye motion: components x, y and z. Velocity (mm/s) curves as a function of the image in the sequence.*

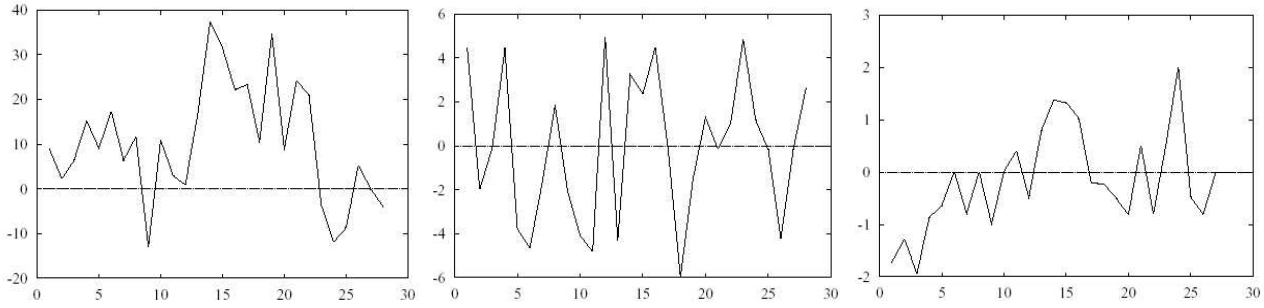

Figure 16: *Right eye motion: components x, y and z. Velocity (mm/s) curves as a function of the image in the sequence.*

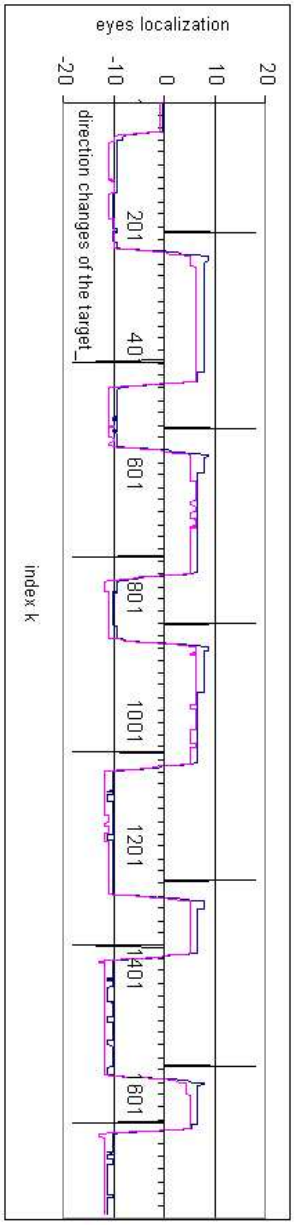

Figure 17: *Example of test for normal physiology of the subject. The eyes follow a target moving alternatively on a horizontal line. Each curve shows the localization of each eye in respect of sequence images (acquisition rate: 200i/s). The vertical segments represent direction changes of the target.*

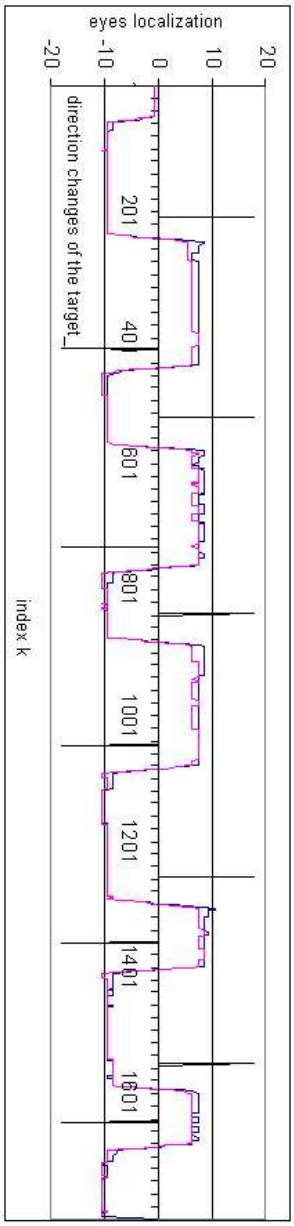

Figure 18: *The same subject under light sedative treatment, two hours later. The eyes follow a target moving alternatively on a horizontal line. Each curve shows the localization of each eye in respect of sequence images (acquisition rate: 200i/s). The vertical segments represent direction changes of the target. This result demonstrates that reflex delay of pursuit increases (compare with figure 17).*