Transactions on Biomedicine and Health vol 1, © 1993 WIT Press, www.witpress.com, ISSN 1743-3525

Spatial resolution enhancement and image restoration for high speed video-oculographic systems

F. Fioravanti, P. Bruno & P. Inchingolo
Dipartimento di Elettronica, Elettrotecnica ed Informatica, University of Trieste, 34100 Trieste, Italy

ABSTRACT

Current Video-Oculography (VOG) permits multidimensional, drift-free, non-invasive measurement of eye movements, but cannot be used for fine recording of fast eye movements (*saccades*), which would require high spatial and high temporal resolutions, corresponding to global data rates in the range of 300Mb/s ÷10Gb/s. A new technique is proposed for achieving high accuracy using very low spatial frequency sampling and reconstructing a high-resolution binary image of the pupil from a low-resolution grey-level frame, acquired with a solid-state sensor. A recursive fitting algorithm for eye position detection is also introduced, allowing correct measurements also when the pupil contour is only partially available.

INTRODUCTION

Analysis of eye movements is very common in a wide spectrum of research and applicative fields, from neurophysiology to clinics or ergonomics.
For this reason many eye position measurement systems (*oculometers*) have been proposed, in view of the disparate requirements typical of each application. In particular, for a precise recording of fast eye movements (*saccades*) a large bandwidth (at least 75Hz [1]) and high resolution (better than 0.2deg) are needed. Moreover, the study of small saccadic oscillations requires a double bandwidth and a resolution ten times better. Only the *search coil in a magnetic field technique* [2] satisfies nowadays these specifications over a wide range of eye rotations; unfortunately it cannot be employed extensively in humans, and particularly in children, because of its invasivity.

In recent years, some video-based oculometers have been developed. The video-oculography (VOG) is based on a temporal sampling of the eye image by means of a bi-dimensional sensing device (usually a solid-state sensor); the pupil image edge is related to the eye position (referred to a craniotopic co-

ordinate system) and to the displacement of the sensing device. In most VOG realisations these relationships are not well explained, and some simplifications are often used, introducing errors in the evaluation of eye position. The most popular methods are based on the linearisation of the function describing the position of the pupil centre, which yields to relevant errors for large rotation angles; the pupil centre is usually found by averaging the co-ordinates of the black pixels [3-6]: this method is correct only if the pupil image is completely available, i.e. when there are no light reflexes on the pupil edge nor eye-lids covering it. Another limit of all current VOG realisations is that either temporal bandwidth or resolution is sacrified, so that saccades cannot be detected correctly.

NEW APPROACH FOR VOG EYE POSITION MEASUREMENTS

We developed an accurate analytical model of the whole recording environment (Figure 1), considering the patient head (*PH*) and eyes (*PE*), the optical system (*OS*) and the sensing device (*SD*). Making some assumptions about the image formation process, it is possible to show that the pupil image edge is well described by an ellipse for gaze deviations from the frontal position with respect to the sensing surface within a *tolerance field* (*TF*), whose amplitude depends on the aperture diameter of the optical system and the eye-objective distance.

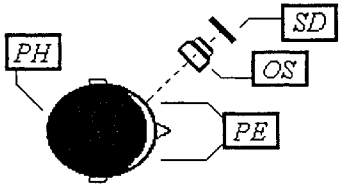

Figure 1: VOG Measurement System

For a diameter of 3mm and a distance of 20mm between *PE* and *OS*, *TF* is ±60deg, which is large enough for the requested recording range. We proved that within this *TF* defocusing effects can be neglected and that it is possible to relate the ellipse parameters to the eye position by solving a non-linear equation system. However, some conditioning problems in the algorithm introduce the need for a calibration procedure, in order to reduce the error propagation. For each eye, the complete procedure we propose for a movement recording session is described by the following two phases: 1) *Instrument Calibration (IC)*, divided in two steps: objective detection of the craniotopic co-ordinates of the sensor plane and error-reduction calibration (for high precision measurements only); 2) *Eye Position Measurement (EPM)*, obtained in three steps: recognition of the pupil-image edge, estimation of the ellipse parameters (*fitting procedure*) and evaluation of eye rotation angles.

Introducing the effects of the spatial sampling due to the discrete nature of a solid-state sensor and a 256-grey-level quantisation, it is possible to evaluate the Maximum Absolute Error (*MAE*) affecting the estimated rotation angles. The simulation of the whole system has shown that image resolution effects are more significant during the *IC* phase than during the *EPM* one: a high spatial resolution during *EPM* has few effects on instrument precision if the calibration procedure has been executed with a lower resolution. A high spatial resolution during *IC* might instead allow to use a lower resolution in the measurement phase. Figure 2 shows the *MAE* bounding obtainable in a measurement as a function of the spatial resolution of the *IC* and *EPM* phases.

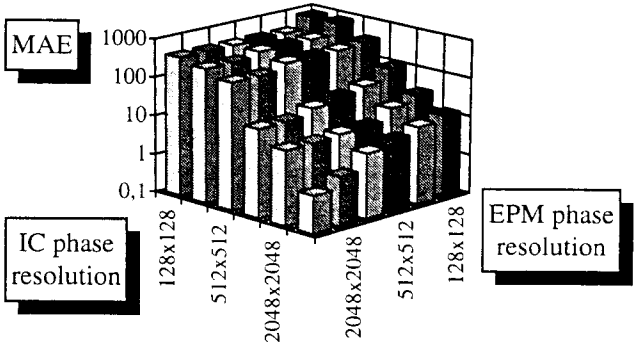

Figure 2: Maximum Absolute Error as a function of image resolution during IC and EPM phases. MAE is expressed in hundredths of degree and log-scale.

The relationship between frame spatial resolution during acquisition and maximum absolute error during the above discussed evaluation of eye position shows that an image sampling of 512x512 pixels is required in order to achieve a 0.2deg spatial resolution. Since at least 256 grey-scale levels are needed to detect the eye movement torsional component using cross-correlation algorithms [3], a bit-rate of about 300Mb/s is necessary, considering that a minimum sampling rate of 150Hz is required for a 75Hz bandwidth. Moreover, when small saccadic oscillations must be recorded, the required bit-rate raises to about 10Gb/s, since 2048x2048 pixel frames at a sampling rate of 300Hz should be acquired. Evidently, a so high throughput is not practically reachable with common solid-state image sensors and acquisition electronics. For this reason, existing VOG systems cannot perform simultaneously both high-spatial and high-temporal resolutions.

EYE IMAGE BINARISATION AND RESOLUTION ENHANCEMENT

To overcome the above described throughput limits, we propose to build up a binary, high-resolution pupil image using a reconstruction algorithm which

realises the decoding of pupil-edge position stored in a low-resolution, grey-level image. This solution permits to elude typical VOG high-rate data-transfer problems by implementing a wide bandwidth, low spatial resolution, 256-grey-level acquisition system an then increasing pupil image resolution.

Our pupil-image reconstruction method is based on the integration properties of solid-state image sensors and, in particular, of diode arrays. These devices are made up of a NxN matrix of photodiodes, each one of them giving an output voltage proportional to the incident field intensity. This characteristic allows to express the grey level of the generic pixel as follows:

$$G(m,n) = \iint K f(x,y)\, w(m\delta - x, n\delta - y)\, dx\, dy \qquad (1)$$

where δ is the sampling step, K is a conversion constant, $w(x,y)$ is a weight function and $f(x,y)$ is the intensity of the incident field. For the considered diode arrays, $w(x,y)$ is a unitary square pulse of width δ; therefore the grey level of each pixel depends on the field over a single photodiode sensing area, i.e. no overlapping occurs. Let us divide the sensor surface into NxN meshes of width δ representing the NxN photodiodes sensing areas (Figure 3). Let us divide now more finely each mesh into $N_r xN_r$ smaller ones of width δ/N_r (both N and N_r are powers of 2), and assume that each small mesh is the sensing area of one photodiode in a higher-resolution device. For each pixel (m,n) in the low-resolution image (called macropixel, MAP), a pattern $\boldsymbol{PAT}[m,n]$ of $N_r xN_r$ pixels in the high-resolution image (called micropixels, MIPs) can be defined.

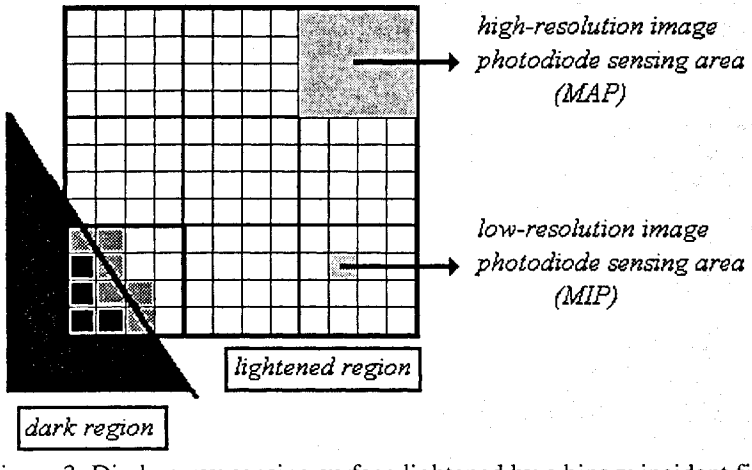

Figure 3: Diode-array sensing surface lightened by a binary incident field

If $G(m,n)$ is the generic MAP grey level and $g(m,n,i,j)$ is the grey level of the generic MIP inside the pattern $\boldsymbol{PAT}[m,n]$, from equation 1 comes:

$$G(m,n) = \sum_{(i,j)\in \boldsymbol{P}} g(m,n,i,j) \qquad (2)$$

where $\boldsymbol{P}=\{i,j|1\le i\le Nr, 1\le j\le Nr\}$. If we consider now a binary incident field, dividing with a straight-line edge the sensing surface into a lightened region and a dark one, the grey level scale in the high-resolution image will be bounded by the upper limit g_{max} relative to those MIPs whose diode is completely lightened. Let us discriminate three sets of MIPs as follows (Figure 3): $\boldsymbol{L}$ (MIPs having grey level g_{max}), $\boldsymbol{B}$ (MIPs having grey level 0) and $\boldsymbol{T}$ (MIPs having grey level between 0 and g_{max}). We showed that if $N_r \ge 4$ the contribute of the terms $g(m,n,i,j)$ in Equation 2 is negligible for $(i,j)\in \boldsymbol{P}\cap\boldsymbol{T}$ and is null for $(i,j)\in \boldsymbol{P}\cap\boldsymbol{T}$. With this assumption, it is possible to evaluate the number N_L of MIPs lightened (i.e. belonging to class $\boldsymbol{L}$) in the pattern $\boldsymbol{PAT}[m,n]$ as follows:

$$N_L = G(m,n) / g_{max} \qquad (3)$$

Moreover, since the maximum grey level in the low-resolution image G_{max} is N_r^2 times larger than g_{max}, we obtain:

$$N_L = G(m,n)\ N_r^2 / G_{max} \qquad (4)$$

The displacement of these lightened pixels inside the pattern $\boldsymbol{PAT}[m,n]$ must define two regions and a rectilinear edge between them having the same direction of that one dividing the dark and the lightened regions of the incident field. If this edge is a straight line at least inside a 3x3 MAP lattice (condition that can be satisfied for the pupil edge), its direction is easily found by means of the local-gradient vector evaluated on the low-resolution image using the finite-differences approximation:

$$\boldsymbol{Grad}(m,n) = 0.5\,[G(m+1,n) - G(m-1,n)\,,\ G(m,n+1) - G(m,n-1)] \qquad (5)$$

Because of the discrete domain, the number of distinct edge directions inside an $N_r x N_r$ MIP pattern is $4(N_r-1)$. We conclude that each distinct pattern $\boldsymbol{PAT}[m,n]$ is finally identified by two conditions: an integrative one (the number of lightened MIPs) and a derivative one (the direction of the edge, i.e. the local-gradient vector). When a grey-level low-resolution image is available, it is then possible to build up a high-resolution binary image by selecting, for each MAP, an $N_r x N_r$ MIP pattern as follows:

$$\boldsymbol{PAT}[m,n] = \boldsymbol{PAT}[N_L, \boldsymbol{Grad}(m,n)] \qquad (6)$$

This high-resolution binary-image reconstruction method can be implemented in real time by storing all distinct $N_r x N_r$ MIP patterns in a RAM device and addressing the selected one with the number of lightened MIPs and

the edge direction. The storage capacity required depends on the *reconstruction factor* N_r: 29Kbytes for N_r=8 and 983Kbytes for Nr=16. An alternative implementation is possible if only the co-ordinates of the edge-points are required: in this way the amount of required memory decreases of 25% for N_r=8 and 50% for N_r=16.

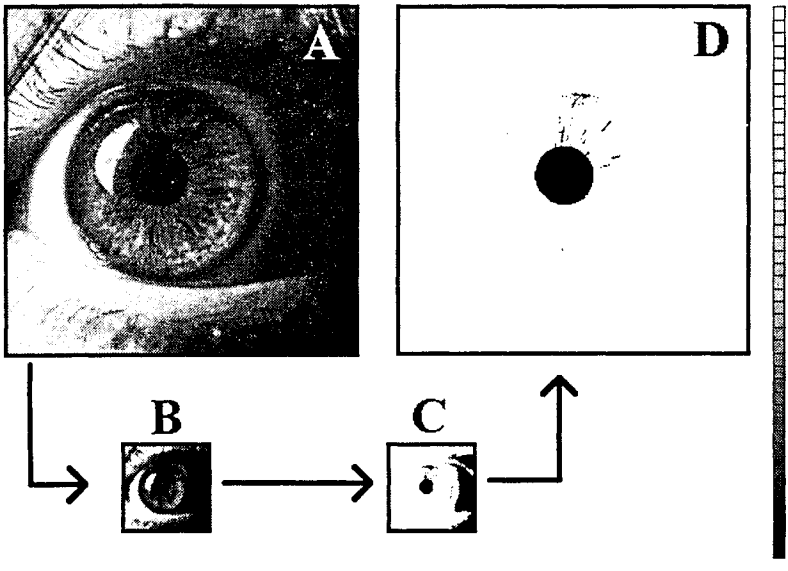

Figure 4: A) High-resolution eye image; B) low-resolution resampling; C) segmentation and grey-level local stretching; D) binary high-resolution pupil-image reconstruction.

The results of simulation using a binary incident field are very good, showing more than 97% of exactly reconstructed points for Nr = 4, 8 and 16. However, in the case of the pupil image, the incident field is not binary because of the non-uniform iris reflectivity. Therefore, the eye image must be segmented, extracting black MAPs (pupil region) and their adiacent ones (pupil-edge transition region). The reconstruction method is then applied to these MAPs, by redefining G_{max} (see Equation 4) as the local-maximum grey level inside a 3x3 lattice centred on the current MAP (*local stretching*). In the example of Figure 4 a high-resolution (512x512) eye image is resampled with a lower resolution (128x128) and then reconstructed to a high-resolution (512x512) binary image. The segmentation procedure includes some iris-texture pixels, which appear also in the reconstructed binary image. These undesired MIPs can be easily removed with a median filter or with some image compression algorithm.

ESTIMATION OF THE PUPIL-IMAGE CONTOUR

As introduced before, our method of eye position evaluation is based on the best-fitting of an ellipse to the edge of the pupil image. If the entire contour of the pupil is available, the best-fitting ellipse

$$g(x,y) = a_{11}\,x^2 + 2\,a_{12}\,xy + a_{22}\,y^2 + 2\,a_{13}\,x + 2\,a_{23}\,y + 1 = 0 \qquad (7)$$

can be found by minimising with a recursive algorithm the error function E, evaluated as the sum of the square distances of the generic pupil edge point (x_k, y_k) from the ellipse, extended to all the N_p edge points:

$$E = \sum_{k=1}^{Np} \frac{g(x_k, y_k)^2}{(\partial g(x_k, y_k)/\partial x)^2 + (\partial g(x_k, y_k)/\partial y)^2} \qquad (8)$$

Unfortunately, in most cases the pupil contour is only partially available on the acquired image, because of the light corneal reflection and the eye-lid occlusion. The latter effect is particularly evident in strabismic subjects, on which current VOGs cannot be used.

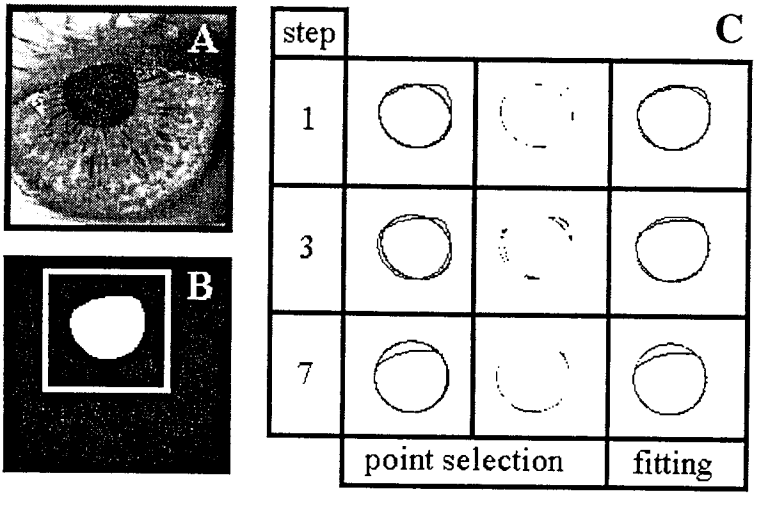

Figure 5: A) Eye-lid partially occludes pupil image; B) Pupil image after binarisation; C) Iterative estimation of the pupil-image contour

This problem can be solved by adopting a recursive point-selection algorithm able to discriminate the edge points actually belonging to the pupil contour.

The selection is possible by neglecting, in the evaluation of the error function, the terms due to those points whose distance from a reference ellipse is larger than a certain threshold.
The reference ellipse is chosen initially by means of a pseudo-inverse method and updated for each iteration with the fitting ellipse found in the previous step. The iteration is stopped when the parameter variations between two steps are negligible. The number of steps required for convergence depends on the ratio ρ between the number of points belonging to the pupil contour and the total number of edge points: 2 steps for $\rho=1$, 5 steps for $\rho=0.5$ and 9 steps for $\rho=0.35$. Convergence is ensured for ρ greater than 0.25. This limit results low enough for an accurate eye position identification of most eye deviations also for strabismic subjects.

CONCLUSIONS

The methods described in this paper permit to realise a non invasive, precise, high-resolution and large-bandwidth oculometer, satisfying quite well all the desired characteristics for an eye movement measurement system. They will be implemented on a hardware prototype we have designed and which is very close to be completed. This work was particularly motivated by the need to study accurately saccadic eye movements in strabismic and amblyopic children, within a project of the European Community aimed to help functional recovery from these pathologies.
Acknowledgements: work supported by the European Community, contract #SCI*-CT91-0747 (TSTS)

BIBLIOGRAPHY

1. Inchingolo, P. and Spanio, M. 'On the Identification and Analysis of Saccadic Eye Movements - A quantitative study of the processing procedures' *IEEE Transactions Biomedical Engeneering,* Vol. BME-32, pp.683-695, 1985
2. Robinson, D. A. 'A method of measuring eye movement using a scleral search coil in a magnetic field' *IEEE Transactions Bio-Medical Electronic*, Vol. BME-10, pp.137-145, 1963
3. Hatamian, M. and Anderson, D. J. 'Design Considerations for a Real-Time Counterroll Instrument' *IEEE Transactions Biomedical Engeneering*, Vol. BME-30, pp.278-288, 1983
4. Clarke, A. H., Teiwes, W. and Scherer, H. 'Video-Oculography - An alternative method for measurement of three-dimensional eye movements', in *Oculomotor Control and Cognitive Processes* (Ed. Schmid, R. and Zambarbieri, D.), pp.431-443, 1991
5. Barbur, J. L., Thomson, D. W. and Forsyth, P. M. 'A new system for the simultaneous measurement of pupil size and two-dimensional eye movements' *Clinical Vision Sci.*, Vol.2, No.2, pp.131-142, 1986
6. Li, H. F., Chan, F. H. Y., Poon, P. W. F., Yan, W. F., Womg, W.Y., Hwang, J. C. 'Realtime online pattern recognition of eye position and movement' *IEEE Proceedings*, Vol.132E, No.6, pp.293-308, 1985