High image rate eye movement measurement

A novel approach using CMOS sensors and dedicated FPGA devices

Andrew H. Clarke, Caspar Steineke and Harald Emanuel

Labor für experimentelle Gleichgewichtsforschung, HNO Klinik
Klinikum Benjamin Franklin, Freie Universität Berlin
Hindenburgdamm 30, 12200 Berlin, FRG
Email: clarke@zedat.fu-berlin.de

Abstract. In contrast to previous devices based on conventional video standards, the present eye tracker is based on programmable CMOS image sensors, interfaced directly to digital processing circuitry to permit real-time image acquisition and processing. This architecture provides a number of important advantages, including image sampling rates of up to 400 /s measurement, direct pixel addressing for pre-processing and AOI acquisition, hard-disk storage of relevant image data. The reconfigurable digital processing circuitry (FPGA) also facilitates inline optimisation of the front-end, time-critical processes. The acquisition algorithm for tracking the pupil and other eye features is designed around the generalised Hough transform

The tracker permits comprehensive measurement of eye movement (three degrees of freedom and head movement (six degrees of freedom), and thus provides the basis for many types of vestibulo-oculomotor and visual research. It is foreseen that the device will be used together with appropriate stimulus generators (rotating or translating devices, visual displays).

Keywords: eye movement measurement; videooculography, CMOS; Image Sensor, FPGA

1 Overview

The use of conventional video techniques for tracking eye movements has proved inadequate for the correct acquisition of saccades with angular velocities of up to 500 deg/s. Nevertheless there is an increasing requirement in both the clinical and research fields for non-invasive measurement of three-dimensional eye movement. The approach presented here side-steps the brickwall of previous video techniques and provides sampling rates of up to 400/s while maintaining high positional resolution.

The eye tracker consists of a head unit, which is individually adjustable, and carries the CMOS cameras for recording eye-in-head images; and the system unit, which accommodates the custom-designed DSP/FPGA architecture for the online, real-time acquisition and pre-processing of image and signal data. This is designed around a standard Windows NT PC with a PCI plug-in board. The head unit is connected to the system unit through high-speed digital data links. These provide the necessary bidirectional data channels for the transfer of high bandwidth image and signal data from the head unit and the command sequences from the system unit.

The image of the eye is reflected by dichroic mirror to the optical lens and projected onto the image sensor. An infrared pass filter (>850 nm) is fitted in front of the image sensing area in order to exclude sporadic incident light from the environment. These optical elements and the cameras are arranged on the head unit to facilitate maximal field-of-view for the test subject. A field-of-view approaching +/-90° horizontal and +40/-60° vertical is attained.

2 Front-end image processing

Significant progress in image processing devices is evident in the most recent CMOS image sensors. Amongst the important features of such smart vision sensors are the configurable acquisition of pixel-defined areas of interest and the on-chip parallel processing of pixel data. For the eye-tracking algorithm, a substantial part of the image pre-processing can be performed by the on-chip circuitry. This has the consequence that only the relevant data must be transferred from the sensor chip to the host computer where the main eye tracking algorithms are performed. This eliminates the bottleneck caused by standard frame-by-frame image acquisition, and thus facilitates considerably higher image sampling rates.

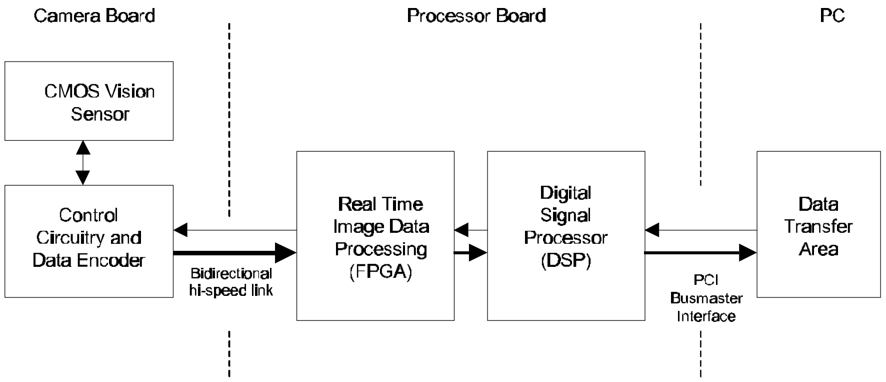

Fig. 1 Layout of front-end processing components used in the eye tracker.

The digital cameras, which have been specially designed around state-of-the-art CMOS image sensors, are interfaced to a dedicated PCI processor board via bi-directional high speed digital transmission links (200 Mbytes/s). This custom-designed PCI plug-in board carries a DSP/FPGA architecture for binocular, online image and signal acquisition. The PC provides the graphic user interface (GUI), data post-processing and storage facilities. The essential element in the System Unit is the Front-end Processor Board. This incorporates a custom-programmed field programmable gate array (FPGA) for each image sensor and a digital signal processor (DSP) with associated storage arrays.

3 Algorithms

The Windows NT software is designed for online acquisition of eye and head movement. Data reduction is performed by the front-end architecture and in the initial NT process permit online storage to hard disk of all relevant data for the subsequent calculation of eye position (3 dof) and head position (5 dof). In addition to online data acquisition and storage, the eye tracker provides online calculation of 2D/ 3D-eye position.

The software package for 3D eye tracking includes the main program running under Windows NT Version 4.0 and a set of front-end firmware modules implemented on the dedicated FPGA and DSP components. During online recording, the critical data acquisition and numerical analysis procedures are performed by algorithms implemented in the FPGA embedded in the CMOS cameras and on the front-end processor board. Here, priority is given to image data acquisition and storage at the selected sampling rate i.e. adequate regions-of-interest around the pupil and the iris markings are stored to facilitate comprehensive off-line processing.

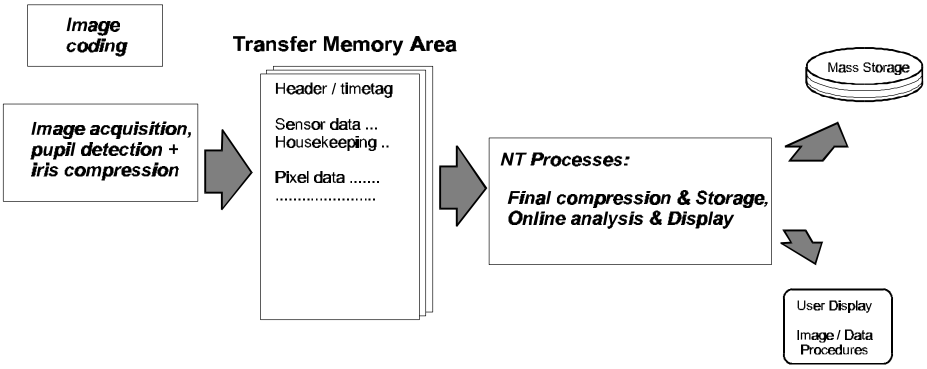

Fig. 2 System structure of eye tracker software illustrating processes and data flow.

A predefined data structure is implemented for the collation, storage and transmission of the sampled image and sensor data streams. Thus, each such data frame includes a header with time stamp, and the related image, sensor (and audio) data. The resulting data files consist of series of data packets (typically 200/s, however it should be noted that with this technique, the limiting factor is the maximum permissible irradiation to the eye by the infrared light source, rather than the performance of the processing capacity). According to the user-selected acquisition parameters i.e. sampling rate, 2D/3D and monocular/binocular acquisition, variable amounts of data may be acquired. For this reason the data type is configured dynamically. Thus, data storage capacity can range from that required for recording of online 3D co-ordinates, through to full grey-level image recording (digital video recorder mode).

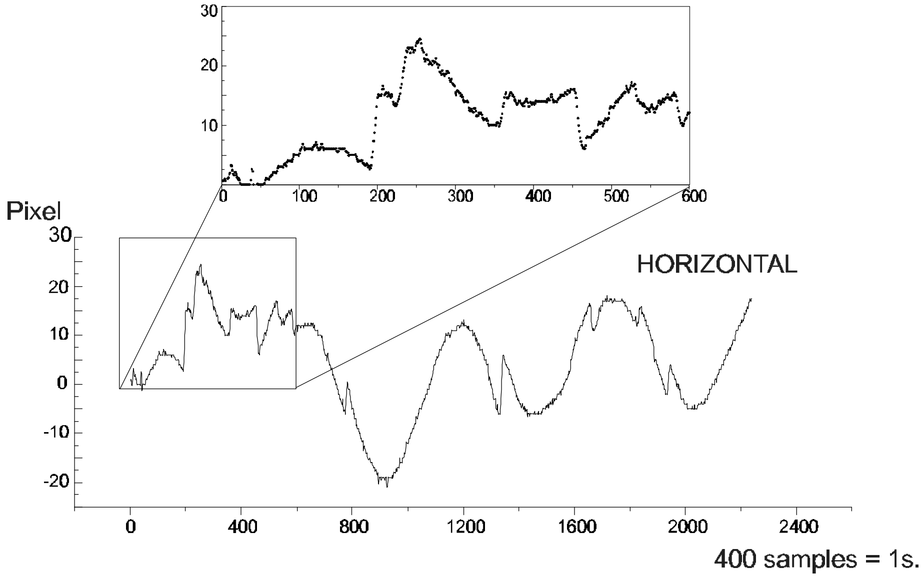

Fig 2. Example of online eye tracking at 400 /s using smart sensors. Eye movement pattern elicited by oscillation of the head (i.e. vestibulo-ocular reflex). Note the complex pattern of slow and rapid phases. The blow-up demonstrates the adequate acquisition, which is achieved with a sampling rate of 400 /s.

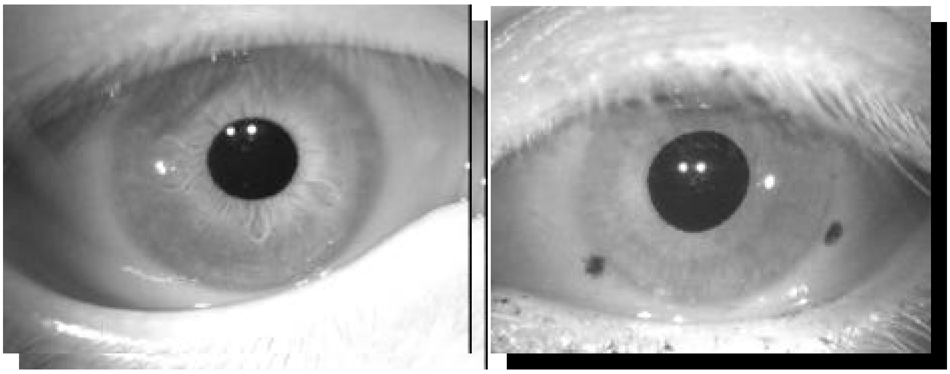

Fig. 3: Eye images recorded under infrared light (930 nm). The iris of example A contains sufficient landmarks to facilitate torsional measurement with iris signatures. In the example B the iris is poorly landmarked, but the tincture markings on the limbus enable measurement of eye movement independently of any natural features.

While the simpler approaches (e.g. pupil centroid) can be rejected for their artefact susceptibility, it appears that the more adequate “circle approximation” techniques for the pupil perimeter (e.g. Barbur et al, 1988) must also be rejected in favour of ellipse-fitting (Pilu et al, 1996) which represents a better approximation to the pupil form of most eyes, and allows for compensation of any geometric distortion during eye rotation (Moore et al, 1996). Besides the problem of pupil form, the greatest sources of error are image artefacts caused by shadowing, reflections from tear fluid, pupil occlusion. To overcome this, an algorithm based on the generalised Hough transform has been implemented (Hough, 1962), which has proved extremely robust against such artefacts. For 3D tracking, the torsional eye position calculated using the polar correlation algorithm (Hatamian & Anderson, 1983).

Considerable improvement to the quality of torsional measurement is facilitated by the application of high-contrast artificial tincture landmarks to the limbus (see example in Fig. 3). Compared to the variable quality of natural iris landmarks, this technique guarantees an ideal contrast profile for the polar correlation algorithm. Of considerably more importance is that the application of such markers permits the complete determination of the rotational position of the eye without having to rely on tracking physiologically variable features such as the pupil or iris patterns. Nakayama (1974) initially demonstrated the necessary numerical routine for this two-marker approach.

4 References

1. Clarke AH Current trends in eye movement measurement techniques. In: Zangemeister WH, Stiehl HS, Freksa C (Eds) Visual Attention and Cognition, Elsevier, Amsterdam, 347-364, 1996.
2. Clarke AH Vestibulo-oculomotor research and measurement technology for the space station era. Brain Research Rev 28: 173-184, 1998
3. Clarke AH, Schücker D, Krzok W Improved three-dimensional eye movement measurement using smart vision sensors. In: Becker W, Deubel H, Mergner T (Eds) Current Oculomotor Research: Physiological and Psychological Aspects. Plenum, NY, 1999.
4. Barbur J, Thomson WD, Forsyth PM A new system for the simultaneous measurement of pupil size and two-dimensional eye movements. Clin Vision Sci 2(2): 131-145, 1987.
5. Hatamian M, Anderson DJ (1983) Design considerations for a realtime ocular counterroll instrument. IEEE Trans Biomed Engg BME-13(2): 65-70.
6. Hough PVC Metods and means for recognising complex patterns. US Patent 3069654, 1962.
7. Moore ST, Haslwanter T, Curthoys IS, Smith ST A geometric basis for measurement of three dimensional eye position using image processing. Vision Res, 36,445-459, 1996.
8. Nakayama K Photographic determination of the rotational state of the eye using matrices. Am J Optom & Physiol Optics 51, 736-741, 1974.
9. Pilu M, Fitzgibbon A, Fisher R Ellipse-specific direct least-square fitting. IEEE Int Conf Image Proc, Lausanne, 1996.