Real Time Visual Attention on ProtoEye

H. Hügli*), N. Ouerhani*), P.-Y. Burgi, P.-F. Ruedi

The most important function of visual attention is to rapidly detect visually salient objects in complex visual scenes, on which higher level vision tasks can focus. Through targeting the analysis process of scene images, the amount of information to be processed further can be drastically reduced.
In recent works, we showed the usefulness of visual attention in some computer vision fields, such as image segmentation[1] and image compression[2].
The complex structure of computational models of visual attention represents, however, an obstacle to the real time implementation of the attention process on conventional computers. Recently implemented on a high performance personal computer (950 MHz), the saliency-based algorithm of visual attention runs at a frequency of only 3 Hz.

In order to overcome this performance limits, the visual attention model is implemented on a highly parallel Single Instruction Multiple Data (SIMD) architecture. Generally a SIMD machine consists of a 2D array of identical processing units which perform the same operation on a 2D array of data. Unlike conventional sequential processors, the computation time of such a parallel architecture is independent of the size of images.

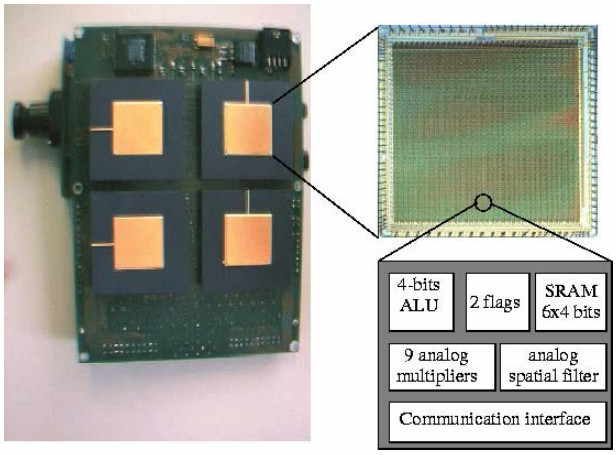
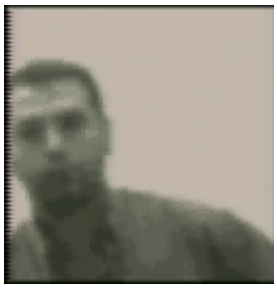


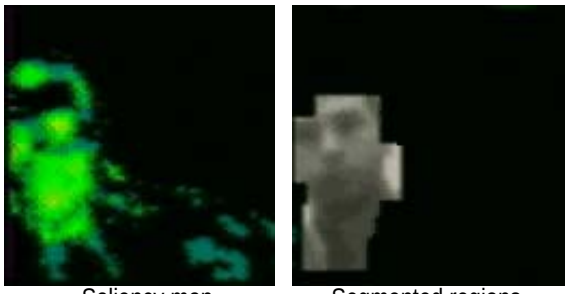
Figure 1
ProtoEye: A SIMD machine tailored for image processing.

ProtoEye is a CSEM image processing ASIC based on the principle of SIMD. It consists of a 35 x 35 array of mixed analog-digital cells. The digital part of a cell, working on 4-bits words, performs all operations needed to transform single images and to combine pairs of images. The analog part is composed essentially of a diffusion network which efficiently performs the time consuming task of low and high-pass spatial filtering of images. Thus, a processor is assigned to each pixel of the image. Four ProtoEye chips are connected together to process 64 x 64 gray level images, provided by a CMOS camera. The whole architecture is controlled by a general purpose microcontroller (sequencer) running at a frequency of 4 MHz, yielding an effective performance of over 8 Giga operations per second. In addition to its high performance, the image processing platform is fully programmable.

The high performance on one hand and the flexibility on the other hand allow the real time implementation of complex computer vision applications, such as visual attention, on the presented image processing platform. The most time consuming part of the visual attention algorithm, related to spatial filtering, is performed by the analog spatial filter which is based on a diffusion network. The implemented system provides, in real time, saliency maps related to intensity at a frequency of 16 Hz.
A simple segmentation algorithm based on visual attention has been also implemented. Unlike conventional segmentation algorithms, the implemented method segments only visually salient regions.



Original image



Saliency map | Segmented regions

Figure 2
Computing visual attention from intensity and segmenting salient regions.

Future works aim at considering also color components to compute visual attention on the SIMD-based platform. Furthermore, more sophisticated segmentation algorithms based on visual attention will be implemented.

*) Institute of Microtechnology, University of Neuchâtel.

[1] Nabil Ouerhani, Neculai Archip, Heinz Hügli & Pierre-Jean Erard, "Visual Attention Guided Seed Selection for Color Image Segmentation", Proc. Conf. Computer Analysis of Images and Patterns, CAIP'2001, September 5-7, 2001, Warsaw, Poland, Lecture Notes on Computer Science, Springer Verlag, LNCS-2124, pp. 630-637

[2] Nabil Ouerhani, J. Bracamonte, H. Hügli, M. Ansorge & F. Pellandini, "Adaptive Color Image Compression Based on Visual Attention", Proc. 11th Int. Conf. on Image Analysis and Processing, ICIAP 2001, 26-28 Sept. 2001, IEEE Computer Society Press, 2001, pp 416-421