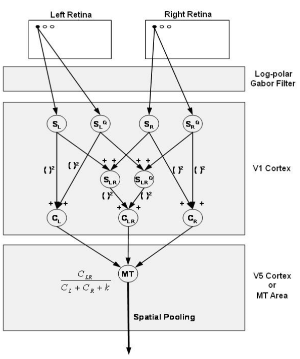
Implementation of a Biologically Inspired Stereoscopic Vision Model in C+CUDA

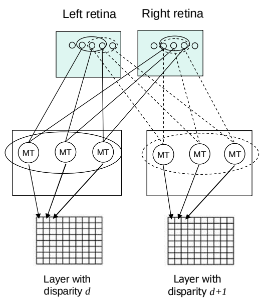
Camilo A. Carvalho, Alberto F. De Souza, Lucas Veronese, and Hallysson Oliveira
Universidade Federal do Espírito Santo, 29075-910-Vitória-ES, Brazil
{camilo, alberto, hallysson}@lcad.inf.ufes.br

The images formed on our retinae are bidimensional; however, from them our brain is capable of synthesizing a 3D representation with color, shape and depth information about the objects in the surrounding environment. For that, after choosing a point in 3D space, our eyes verge to this point and, at the same time, the visual system is fed back with the eyes position information, interpreting it as the distance of this point to the observer. Depth perception around the vergence point is obtained using visual disparity, i.e., the difference between the positions in the retinae of the two projections of a given point in 3D space caused by the horizontal separation of the eyes. Most of the depth perception processing is done in the visual cortex, mainly in the primary (V1) and medial temporal (MT) areas. In this work, we modeled the neural architecture of the V1 and MT cortices using as building blocks previous models of cortical cells and log-polar mapping. A sequential implementation of our model can build a tridimensional representation of the external world using stereoscopic image pairs obtained from a pair of fronto-parallel cameras. A C+CUDA parallel implementation is almost 60 times faster and allows real-time 3D reconstruction.

In our model (Figure 1(a)), we use: (i) two types of simple cells (see simple cells models in [1]) – monocular simple cells (S_L, S_L^Q, S_R and S_R^Q) and binocular simple cells (S_{LR} and S_{LR}^Q); (ii) two types of complex cells – monocular complex cell, build from two monocular simple cells in quadrature phase (90º) (C_L and C_R), and binocular complex cell build from four monocular simple cells (C_{LR} – energy model [2]); and our medial temporal (MT) cell. Our MT cell divides the output of the binocular complex cell by the sum of the output of two monocular complex cells and constant (see equation in Figure 1(a)), which can be adjusted to make the MT cell behave like a tuned inhibitory cell [3].



(a)

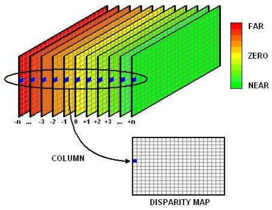


(b)

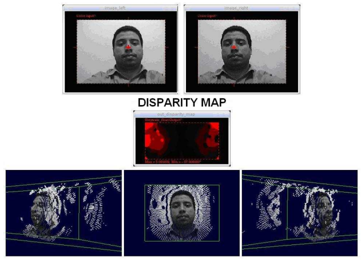
Figure 1: (a) MT cell model. (b) MT neural layers.

In our model, a MT "layer" is a group of log-polar-retinotopically-arranged MT cells with the same disparity, *d* (see Figure 1(b)). A group of MT layers codes different degrees of disparity, as shown in Figure 2(a). By summing up the output of all neurons of each MT layer we can perform vergence, i.e., we can find the point in the left image that corresponds to the center of the log-polar mapping of the right image onto MT – we just need to take de disparity of the least active MT layer (the MT cells are tuned inhibitory cells).

Once we have vergence, we can build a disparity map by taking the disparity of the least active neuron from each column of MT layers as the disparity of the corresponding point of the right image (Figure 2(a)). By performing the inverse of our log-polar transformation, we can map each point of the right image back onto 3D space (Figure 2(b)).



(a)



(b)

Figure 2: (a) From MT to disparity map. (b) 3D reconstruction.

We ran the sequential and C+CUDA versions of our model in an AMD Athlon 64 X2 (Dual Core) 5,200+ of 2.7 GHz, with 3GB of 800 MHz DRAM DDR2, and video card NVIDIA GeForce GTX 285, with 1GB of DRAM GDDR3. Table 1 shows the execution times of each implementation (columns), in addition to the speed-up over the sequential implementation (last column). As the table shows, the speed-up achieved (about 60) with C+CUDA allows real-time.

Table 1: One stereo frame 3D reconstruction: experimental results.

C (s)	C+CUDA (s)	*Speedup*
16,8806	0,2942	57,38

[1] ANZAI, A.; OHZAWA, I.; FREEMAN, R. D. Neural Mechanisms for Processing Binocular Information I. Simple Cells. The Journal of Neurophysiology, Vol. 82 No. 2, August 1999, pp. 891-908.

[2] OHZAWA, I.; DeANGELIS, G. C.; FREEMAN, R. D. Encoding of Binocular Disparity by Complex Cells in the Cat's Visual Cortex. The Journal of Neurophysiology, Vol. 77, No. 6, June 1997, pp. 2879-2909.

[3] GONZALEZ, F.; PEREZ, R. Neural mechanisms underlying stereoscopic vision. Progress in Neurobiology, Vol. 55, June 1998, pp. 191-224.