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 (SL, SLQ, SR and SRQ) and binocular simple cells (SLR and SLRQ); (ii) two types of complex cells – monocular complex cell, build from two monocular simple cells in quadrature phase (90°) (CL and CR), and binocular complex cell build from four monocular simple cells (CLR – 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].

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

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.

<table>
<thead>
<tr>
<th>C (s)</th>
<th>C+CUDA (s)</th>
<th>Speedup</th>
</tr>
</thead>
<tbody>
<tr>
<td>16,880</td>
<td>0.2942</td>
<td>57.38</td>
</tr>
</tbody>
</table>

