Rectangular Cortical Magnification

Harald Ruda (h.ruda@neu.edu)
Computational Vision Laboratory, Northeastern University, Boston, MA

The primate visual system is strongly space-variant. Starting with the retina and continuing through higher visual areas, the fovea and para-fovea are highly over-represented. Generally this is thought to be an efficient allocation of neural resources.

The most common representation used to perform similar resource allocation in computer vision is a log-polar mapping defined by $w=log(a+z)$. While useful, this mapping is costly and complicated. The computations involve complex logarithms, and cannot be done very efficiently. A more intractable issue is that the resulting map is not strictly rectangular (monopole or dipole). It is therefore difficult to incorporate the log-polar mapping into traditional computer vision algorithms. Still, this mapping has been found to be useful in many robotics applications (Traver & Bernardino, 2010).

I suggest that a better approach is to use a *rectangular mapping* (mapping a rectangle to a rectangle) that incorporates *non-linear cortical magnification*. A *moveable rectangular fovea* allows for additional benefits. Because of the *separability* of the horizontal and vertical dimensions, *very fast algorithms* can be used. While none of these ideas are new in themselves, by putting them together, it is possible to achieve tremendous gains in speed and usability. Rather than preserving angles (conformal) this mapping preserves relative position (right-left, up-down).

The figure on the right shows an example of how the original input would be sampled and turned into a usable matrix for computations. This scheme can advantageously be used for (1) robotics, (2) very large sensors, and (3) modeling of primate vision. For robotics and sensors the simplified sampling can be put into embedded processors. When modeling primate vision this mapping can be combined with location jitter to create a realistic sampling structure of the optic array; the same structure, with jitter, can be used to index the locations of higher visual nodes.