A Joint Statistical Approach for Geon Detection [1]

Xiangyu Tang, Kazunori Okada and Christoph von der Malsburg [2]
Computer Science Department, University of Southern California
HNB 228, USC, Los Angeles, CA 90089-2520
{tangx, kazunori, malsburg}@organic.usc.edu

Introduction

Known as *geons*, 3D shape primitives have been proposed as components of object models in the brain to deliver quick and accurate object recognition. Biederman differentiates geons by curvature of axis, shape and size of cross-section, etc. (RBC theory, Biederman 1987). This results in a full family of 24 geons. But, the combination of the geons still remains a large number, especially when they change in orientation, location, and size. The object recognition performance directly depends on how to detect these geons in objects.

We present here a *parametric* method to represent and detect geons. The parameters are extracted from joint statistical constraints defined on *steerable pyramid* decomposition. The choice of the the joint statistical constraints is supported by Marr's four stages of object recognition theory (Marr 1982). We examine the representation in detail when geons rotate, translate, and scale. Constraint-wise similarity is introduced to describe the corresponding parameter variations. Based on our studies of the representation properties, we develop and test a geon detection system on a number of grey level synthetic images containing geon-constructed objects.

Represent and Detect Geons

Our geon representation model is an extension of Portilla and Simoncelli's work on visual texture (Portilla and Simoncelli 2000). In order to represent the geons, first we adopt a set of multi-scale, multi-orientation filters to recursively decompose each single geon image into a steerable pyramid, which consists oriented highpass subbands at every one half scale plus a lowpass residue. Then four statistical constraints are defined on the pyramid to characterize key features of individual geons. For a 4-level 4-orientation decomposition, 710 parameters are extracted to address pixel intensity distribution (*marginal statistics*), periodic or global orientation (*raw coefficient correlation*), fine structures such as edges, corners (*coefficient magnitude statistics*), and lighting effects due to geon's 3D appearance (*cross-scale phase statistics*).

This parametric representation is then investigated as we rotate, translate, or scale the geons. We compute the representations for a few sets of geons under gradual change of orientation, location, or size, and plot the constraint-wise similarity values between every varied geon and the original one. We found that the representation is i) translate-invariant in terms of all four similarity values, ii) approximately rotate-invariant in terms of the average similarity, and iii) sensitive to scaling but approximately scale-invariant within the range of $\pm 10\%$ in terms of the average similarity values. These properties help detect various geons with only a few model geons in memory.

Our geon detection system stores several individual geon models and their associated parameters in memory, then slides a detection window over an object image (composed of known geons), while simultaneously compute the similarity between the current parameter set with all sets in memory in a constraint-wise fashion. The average similarities thus constitute a *saliency map* for each model geon. Each saliency map helps detect a specific geon, and inhibites detection of others by darkening their corresponding areas. A successful detection occurs when the maximum value in the saliency map exceeds a selected *threshold*.

Discussion

The strong invariance properties of our geon representation against rotation and translation comes natually, because the steerable pyramid decomposition is both translate- and rotate- invariant. To cover a wider range of scaling, we either pre-analyze more model geons in memory by every 10% variation in size, or follow a multi-window search technique, with the tradeoff of longer running time.

We tested 40 trials. 34 of them (85%) result in correct detection: recognize all geons in the objects, and give approximate locations (with the help of the saliency maps). The system can also survive slight rotation in depth, local illumination caused by occlusion, and very noisy (up to 50% random noise) background as well as surface marking. Our future research would focus on how to characterize geons with more viewpoint invariant statistics.

---

[1] A more detailed report is available: Xiangyu Tang, Kazunori Okada, Christoph von der Malsburg, Represent and Detect Geons by Joint Statistics of Steerable Pyramid Decomposition, USC Comuter Science Dept., Technical Report, 2002. The authors wish to thank Irving Biederman for all helpful hints and discussions.

[2] Also at Institute für Neuroinformatik, Ruhr-Universität Bochum, D-44870 Germany.