Interconnected Representations for Visual Cognition

Rajesh Kasturirangan

Artificial Intelligence Laboratory
Massachusetts Institue Of Technology
Cambridge, Massachusetts 02139

http://www.ai.mit.edu

The Problem: Human Visual Cognition is characterized by the flexibility and ease with which s patial relations and shape properties are computed. How is it possible for a sys tem to be fast as well as being capable of computing a possibly infinite number of different spatial relations? What is the underlying computational structure? Human Visual Cognition is characterized by the flexibility and ease with which s patial relations and shape properties are computed. How is it possible for a sys tem to be fast as well as being capable of computing a possibly infinite number of different spatial relations? What is the underlying computational structure?

Motivation: The human visual system computes qualitatively different proper ties with great ease. For example, topological properties like closure and conta inment are quite different from metric properties like reflectional symmetries a nd pseudo symmetries. Yet, both are computed effortlessly by the human visual s ystem in most typical situations. There is a need for a computational framework that captures the intuition that there are qualitatively different, but intercon nected levels in visual cognition.The human visual system computes qualitatively different properties with great ease. For example, topological properties like closure and containment are quite different from metric properties like reflecti onal symmetries and pseudo symmetries. Yet, both are computed effortlessly by t he human visual system in most typical situations. There is a need for a computa tional framework that captures the intuition that there are qualitatively differ ent, but interconnected levels in visual cognition.

Previous Work: Shimon Ullman has argued for a high-level, context-sensiti ve visual representation that he calls visual routines. In his work, no attempt was made to see if visual routines fall into qualitatively different levels. We hope to further the study of visual cognition by introducing the idea of interconnected qualitative levels and putting these different levels into a computational framework.

Approach: The form of a representation affects the ease of use in a given context. Furthermore, there are many complimentary directions of processing. In particular, we have focussed on two complimentary directions, namely, one in whi ch the representational levels go from concrete to abstract and the other in whi ch the representational levels go from abstract to concrete. All of the qualitat ive levels whether concrete or abstract, are high-level representations that lie beyond low-level visual representations. In our attempt to understand these rep resentations, we focus on three topics: (a) The computational structure of the q ualitative levels, (b) The interconnections between the qualitative levels, (c) The extraction of high-level, qualitative, representations from lower-level visu al representations.

Impact: Understanding visual cognition is a key aspect of understanding th e fluidity and creativity that characterize human intelligence. Putting visual c ognition into a computational framework will help us to understand human intelli gence as well as to help build more flexible artificial systems.

Future Work: Two lines of research emerge naturally from our current work. The first is to discover how the computational structures represented in the var ious interconnected representations are combined in a given visual context. The other is to find out how these visual representations interface with other high- level mental functions, for example, spatial representation in natural language.

Research Support: Gerald J. and Marjorie J. Burnett Fellowship, MIT Endowed Fellowship.

References:

[1] S.Ullman. High Level Vision. MIT Press, 1997.