

TOWARDS A SCIENCE OF IMAGE UNDERSTANDING

Berthold K. P. Horn
M.I.T. Artificial Intelligence Laboratory
545 Technology Square
Cambridge, MA 02139

of image understanding we could benefit tremendously from the resultant cross-fertilization amongst experts in different domains.

A great deal of work in machine vision has brought us to the point where we can tell that:

- (a) image understanding is difficult, and
- (b) image understanding is very useful.

In order to conserve our limited resources, we should pause briefly and consider what new problems to approach and how to tackle them in order to maximize the pay-off. In particular, we might hope to pull together results of efforts in many fragmented areas in order to build the foundations for future work. Perhaps, too, we can see how to structure investigations so that they contribute to basic understanding, rather than just developing special solutions. This may not always be practical under real-world pressures. Nevertheless, I will list four ideas I consider germane to this topic:

- (1) We should pursue vision problems which lead to theories that can be tested by programming a digital computer.

The proof of the pudding is in the eating of it. It is the digital computer that enforces a certain rigour in our discipline and we ought to avoid spending much effort pursuing theories that cannot be tested.

- (2) We should look at real images -- in all their natural splendor.

It is unusual not to be surprised by the actual data -- the real world rarely conforms to our expectations. Often the tools for doing this are not available -- an interactive system capable of displaying arbitrary intensity profiles and calculating various simple measures and histograms is the minimum needed. One should be able to browse around in the image in a natural way.

- (3) We ought to understand the imaging process -- the way surfaces reflect light in particular.

It is hard to imagine any serious computer scientist processing information in some interesting fashion when he has no understanding of its origins. We are only now beginning to understand how the physics of image-formation can be used to guide and constrain the process of obtaining a symbolic description of an image.

- (4) We need the equivalent of the molecular biologist's *Escherichia coli* or the geneticist's *Drosophila melanogaster*.

Research in machine vision, scene analysis and image understanding has been fragmented because of the enormous domain of potentially interesting problems. If we could agree that one, or a small number, of vision problems or image types can provide enough stimulation for work on several aspects