Foreword

This volume, “Developing and Applying Biologically-Inspired Vision Systems: Interdisciplinary Concepts,” highlights many significant contributions to vision systems research that take an interdisciplinary approach, effectively using our knowledge of biological visual processing to build better computer vision systems. As such it represents one more important milestone on a path that was started quite some time ago. In 1962 Harry Blum wrote a report titled “An Associative Machine for Dealing with the Visual Field and Some of its Biological Implications.” The title reveals that he was not only inspired by, but also attempted to have impact on, biological vision. Blum also received inspiration from the Gestalt psychologists in developing algorithms for extracting shape descriptors (1967) and even tried to map his algorithm onto the results of Hubel and Wiesel’s (1962) study of visual cortical neurons. Reading what is widely acknowledged as the first PhD thesis in computer vision by Larry Roberts (1963), one sees a strong influence from J.J. Gibson, and to a lesser extent Attneave. As Roberts writes, the requirement that perceived shapes be invariant to perspective projection was derived from Gibson’s work. Azriel Rosenfeld and John Pflatz (1966) certainly cared about computational operations on images that had perceptual relevance in defining algorithms for computing connected components and distance functions in digital images. The character of human vision played a huge role in the very important “General Purpose Models: Expectations about the Unexpected” paper by Steven Zucker, Azriel Rosenfeld, and Larry Davis (1975). There, they urged a broader view of the problem of computer vision using the generality of human vision as the guide. It was this paper, as I recall, that played a major role in my own formative years. Rosenfeld went on to found and organize a highly influential series of workshops on how human vision might inspire machine vision and Zucker continues to this day to inform his work on shape and curvature interpretation by the neurobiology of the early visual cortex and provide both effective new algorithms for computer vision as well as predictions for new human experimental work. Since those very early days of computer vision, several others have also carried the baton arguing for why the characteristics of biological vision must play a deeper role in the development of computational algorithms (David Marr, Olivier Faugeras, Jan Kooenderink, Dana Ballard, Jan-Olof Eklundh come to mind prominently, among others). If one carefully looks at the development of computer vision from those early years to the present, it is easy to see specific examples of how our knowledge of how to build effective computer vision systems has increased as a result.

It is important to note that this enterprise is not an easy one. It is not the case that we fully understand biological vision, and our task is to only determine how to best use that knowledge to further computer vision. Our knowledge of visual processes in the brain is evolving, quite extensive for some aspects (such as early processing) but quite inadequate for others (such as general object or event recognition, or attentional control). Many have even abandoned this connection saying that our understanding is too meager to enable its effective use. This point of view was common perhaps 20 years ago or so, and likely justifiably so. However, I would argue that the explosion of research on human and non-human primate visual systems in the past 20-30 years has been quite successful at filling in gaps and creating a body of knowledge that is far more appealing to computational researchers. As a result, one sees in recent years more and more papers in mainstream computer vision conferences and journals that reveal biological inspiration.

Consider only one example. Blum was influenced by the Gestalt studies of shape. He points out that the Gestaltists (citing Koffka, Deustch, Kazmierczak, and others) used field theoretic concepts and proposed diffusion/propagation models. These ideas motivated Blum, but he found them unsatisfactory as presented due to their lack of precision and detail. This is a characteristic of functional conceptualization that persists even in the best of current work. This is not a criticism; rather, it is a justification for interdisciplinary efforts where each discipline works on the areas of their expertise. Blum thus took those ideas and developed the now well-known Medial Axis Transform (MAT or 'grass fire' algorithm). This path, in its general sense, is the one followed by the authors represented in this book. One of the papers, by Rodriguez-Sanchez and myself (chosen only because I know it well and for its conceptual connection with Blum), looks at the detection and description of single object 2D silhouettes, the same kind of silhouettes on which MAT might operate. In this case however, the quest is to develop a formalization of the stages of processing the primate visual cortex uses for this task and to show the correspondence between the computational result and the responses of single neurons to the same stimuli.

This volume, well-organized and presented by Marc Pomplun and Jun Suzuki, gives us a detailed view of many current attempts to continue the interdisciplinary inspiration. Topics such as visual representations, visual attention, motion, robot behavior, shape and object recognition, and more are nicely represented, among others. I hope that both current and next generation of scientists and engineers receive the same inspiration for this research enterprise by reading the papers in this book as I did from those very early papers.

ABOUT THE FOREWORD AUTHOR

John K. Tsotsos
York University, Canada

John Tsotsos *is Distinguished Research Professor of Vision Science at York University and hold the Canada Research Chair in Computational Vision. He received York's Inaugural President's Research Excellence Award, is also a Fellow of the Royal Society of Canada, has been a Fellow of the Canadian Institute of Advanced Research and received several other awards. Born in Windsor, Ontario, he holds a doctoral degree from Computer Science at the University of Toronto where he is cross-appointed in Computer Science and Ophthalmology and Vision Sciences. His research spans computer science, psychology and neuroscience and is focused on visual attention and visual object recognition in both man and robots as well as mobile robotics and assistive technology.*

REFERENCES

Blum, H. (1962). *An associative machine for dealing with the visual field and some of its biological implications*. Air Force Cambridge Research Labs. doi:10.1007/978-1-4684-1716-6_34

Blum, H. (1967). A transformation for extracting descriptors of shape . In Wathen-Dunn, W. (Ed.), *Models for the perception of speech and visual forms* (pp. 362–380). MIT Press.

Roberts, L. G. (1963). *Machine perception of 3-dimensional solids.* Doctoral Dissertation, Dept. of Electrical Engineering, MIT.

Rosenfeld, A., & Pfaltz, J. L. (1966). Sequential operations in digital picture processing. *Journal of the ACM*, *13*(4), 471–494. doi:10.1145/321356.321357

Zucker, S. W., Rosenfeld, A., & Davis, L. S. (1975). General-purpose models: Expectations about the unexpected. *ACM SIGART Bulletin*, *54*, 7–11. doi:10.1145/1045247.1045249