Is Computer Vision Still AI?

Robert B. Fisher
Department of Artificial Intelligence
University of Edinburgh

Abstract

Recent general AI conferences show a decline in both the number and the quality of vision papers, whereas there is a great growth and specialization of computer vision conferences. Hence, one might conclude that CV is parting, or has parted, company from AI. This essay proposes that the divorce of CV and AI suggested above is actually an "open marriage", and while CV is developing through its own research agenda, there are many shared areas of interest and many of its key goals, assumptions, and characteristics are also clearly found in AI.

Keywords: Computer Vision, Image Processing, subfields of AI

It is not easy to infer the relationship between the field of Computer Vision (hereafter abbreviated CV) and Artificial Intelligence (hereafter abbreviated AI) from the external appearances of both fields. Recent general AI conferences show a decline in both the number and the quality of vision papers, whereas there is a great growth and specialization of CV conferences. Some CV or robotics researchers even claim that AI is unnecessary or irrelevant – special purpose, dedicated, well-engineered, mathematics based, processes will lead to success.

Alternatively, we could consider the interests and directions of the two fields: an examination of many recent CV conferences and journals shows a marked inclination, especially in the more theoretical papers, towards complex mathematics (e.g. geometric invariance, differential geometry, functional analysis, control theory), models of the physics of light, color, shape, motion appearance, texture, etc., statistical models of the scene and other properties (e.g. fractal, Markov random fields, Bayesian), and non-symbolic image–to–image transformations. Most successful practical vision systems are well-engineered combinations of special purpose sensors, hardware, and algorithms, tailored to solve specific visual problems (and are generally more successful in proportion to the narrowness of the task). Some texts (e.g. [Batchelor 1985]) don't even mention AI nor it's methods. Example practical applications include assembly-line robot welding guidance, inspection of computer keyboards and integrated circuits, counting of teabags, tomato grading, optical tracking systems, etc.

We can contrast this with the outward appearance of AI, which has spent much time investigating more general methods, such as search controlled by domain–dependent constraints, models of (formal) logical and non–logical reasoning, heuristic and uncertain reasoning, representation of general physical and world knowledge (e.g. surface shape, elasticity, gravity) and human knowledge (e.g. conceptual, belief), and learning (i.e. inference of new relationships, self–organization).

Hence, one might conclude that CV is parting, or has parted, company from AI. This phenomenon is a more general problem of AI, and most subfields are specializing to the point that wasteful duplication is occurring and loss of sight the grand goal of re-integration of the different aspects of intelligence is evident. But, though there are many specializations of AI (e.g. planning, natural language understanding, knowledge representation, etc.), no one questions the separateness of these fields from AI. However, the technology of most CV research is so non-mainstream AI (i.e. it looks more like manufacturing engineering or applied physics) that the connection is no longer obvious.

Phrased in these terms, the differences seem extreme; however, I claim the differences are largely illusory and that CV still has and will continue to have a strong relation to AI. CV's specialization and unique preoccupations define it as a field of study, but no more exclude it from the community of AI than do the distinctive formalisms of natural language grammars, the logics underpinning theorem proving and formal reasoning or the numerical calculus of neural networks (and its learning algorithms) exclude their subfields.

This essay proposes that the divorce of CV and AI suggested above is actually an "open marriage", and while CV is developing through its own research agenda, there are many shared areas of interest and many of its key goals, assumptions, and characteristics are also clearly found in AI. We will take several views on this question, not from insecurity, but because many strands of connection imply a close coupling in the space of cross-disciplinary linkages — as contrasted with a sparser coupling (in the sense of few links between concepts) between CV and, for example, architecture (coupled mainly through issues of shape, appearance, and visual aesthetics). I hope to convince the reader that the relation will become even stronger in the future.

1 AI and CV Share Methodological Approaches

There are many eloquent and intelligent discussions of the nature of AI [Boden 1988, Haugeland 1985] (and if it can even exist [Penrose 1989]), but lately most arguments seem to center on the How of AI, not the What. That is, whether true AI is realizable through classical methods (i.e. logic, representation, and search), pattern classification (e.g. case and frame unification), behaviorist methods (i.e. an organized conglomeration of individual goal–pursuing/achieving behaviors), classical emergent-behavior methods (i.e. global, epiphenomenal, behavior arising from the aggregation of distributed actions of separate agents), or neural networks (i.e. distributed, fault–tolerant, connectionist, emergent, numerical computations).

Loosely speaking, CV mirrors the methodological division of AI — there are certainly classical symbol manipulation [Brooks 1981], statistical pattern classification, distributed competence, and connectionist [Hinton 1981] paradigms. In addition, there are at least two other paradigms. The first is the numerical processing of images, where algorithms are linked to the geometric structure of the image, and are constrained by underlying theories of the physical processes that gave rise to the sense data (as in a theory of surface shading [Horn 1975]). The image-to-image transformations typical of this class of processes are not "intelligent" in themselves (i.e. they usually only do a limited amount of interpretation, as in suggesting possible edges). However, they can be considered a product of AI [Schank 1991]. The processes may have originally been intelligent — i.e. occurred as a result of explicit rule-based reasoning, but were later were "compiled" into pure algorithmic form, or they may have never been such, but the actions of the algorithms may be controlled by higher-level control inputs (such as focus of attention).

The second paradigm is that of active perception [Aloimonos 1989], wherein the observer manipulates or maneuvers within the environment in order to produce perceptual effects, such as the kinetic-depth process.

CV also encompasses a large amount of special purpose applications processes, dedicated to extracting a single, perhaps obscure, piece of information from a specialized type of image (e.g. the parametric shape of a range image surface, the average inter-cellular distance in a particular tissue section or the postal code on a letter). These specializations do not fit clearly into the "general vision architecture", in the same way that an expert chess player is not generally considered a component of a general AI program. Even so, the class of transformation that these represent does not differ significantly from the class of competences expected to be found in a general vision system — such as an optical flow from intensity data computation.

2 AI and CV Share Domain Assumptions

Broadly speaking, the following assumptions underpin AI and CV perspectives. This list is undoubtedly incomplete, but I have tried to elucidate beliefs that underlie most research in the general sense — not the individual research areas, but rather the themes that they share. I have listed the key topics from an AI perspective and then discuss how CV fits within this perspective.

As will be seen, part of what influences the distinctiveness of CV is the nature of raw sense data, which is often under-constrained, is always shaped by sensor characteristics and is corrupted by noise. It has a regular geometric structure (e.g. images) that is closely linked to with the geometric ordering in the sensed domain. This does not imply isomorphism, as 2D images arise from the projection of the 3D world. In fact, the relationship may be obscure (as in a holographic image), or inexact, as in a cartoon sketch.

1. **Knowledge Representation:**

   (a) **Representable:** Knowledge is representable, usable, and communicable, although the details of how are unresolved as yet.

   CV programs embody knowledge (e.g. about shape, reflectance, apparent structure), and increasingly make the knowledge explicit in terms of geometric object model-bases and rule bases for how to recognize objects, or when to apply various operators.

   (b) **Representation schemas:** Symbols can be used to represent some concepts and the manipulation of some symbols can be used for modeling or reasoning about a domain.

   Most high-level vision programs manipulate symbolic representations of model and data features, usually attempting some form of search for a correspondence between the two. Intermediate level image interpretation processes more often have numerical image data as inputs, but produce symbolic descriptions as their outputs.

   (c) **Multiple descriptions:** The characterization of real domains (or environments) requires many different descriptors and points of view, and multiple levels of representation.

   The CV community has believed for at least a decade that the description of the visual world requires multiple descriptions, particularly through the concept of "sketches" [Marr 1982] and intrinsic images [Barrow and Tenenbaum 1978]. These are focused, special-purpose representations of the world from the viewer's perspective that describes how it is moving, how it is illuminated, where its significant features lie, etc.

   (d) **Underlying theories:** Mathematical theories underlie some aspects of knowledge and some domains (e.g. theory of mechanics, physics, dynamics of a steam engine, etc.). The theories can be used to model and test hypotheses, interpret sense data or predict effects.

   Mathematical theories of surface shape, motion and image flow, geometry and geometric invariance, image formation, noise processes, combinatorics, and physical theories of shading, mutual illumination, color, and texture underlie much of the successful recent vision research. On the other hand, these theories tend to be "autonomous", that is, require no insight, feedback or control from later, or other, stages of the vision system. Expressing the complexities of algorithms that allow these external interactions in mathematical form may be quite difficult.

(e) **Common sense:** Large amounts of common sense knowledge are required for intelligent and effective behavior in the real world.

A machine that can see as we see the world will require a visual memory of thousands of objects and fragmentary shapes. To exploit this memory, a CV system will need to be able to reason about how appearance can vary by position and illumination, depth ordering, interaction between various objects in a scene, how the members of a class of shapes can appear and deform (e.g. clothing, trees, etc.).

2. **Reasoning:**

(a) **Mechanizable:** Reasoning can be mechanized, modeled, replicated, and experimented with, particularly through the use of computer systems.

Any working CV system exemplifies the mechanization of visual reasoning, in that it is clearly an algorithm implemented on a computer system (in the broad sense – allowing dedicated or hard-wired electronic implementations) [Brooks 1991, Rosenschein 1986].

(b) **Complexity:** Reasoning is complex, and may require non-deterministic decisions.

Because of the difficulty of interpreting data (due to noise, low resolution or under-constrainedness), most real image interpretation systems (i.e. systems that label image features) embody some form of expert system reasoner [Draper et al 1988] that pursues alternative hypotheses and quantifies the degree of verification.

(c) **Heuristics:** Heuristics are often required to: a) model incompletely understood phenomena, b) simplify computationally intractable algorithms, or c) provide a simple and reliable tool when exact methods are unnecessary or expensive.

Most object recognition is still in the heuristic stage, except for simple geometric solids (e.g. polyhedra). Even most edge detectors are based on the heuristic that all edges are intensity step edges, and most uses of the edge information assume that image edges correspond to object edges (ignoring lighting, shadows, specularities, and changes in reflectance).

(d) **Uncertainty:** Reasoning involves uncertainty by virtue of incomplete knowledge, perceptual "noise" and imperfect heuristics.

Uncertain reasoning is needed to quantify the belief in a sensor measurement, and to characterize the certainty of a hypothesis [Durrant-Whyte 1987]. Recently, much CV research has been using statistical methods to represent measurement and hypothesis uncertainty, particularly through the use of the Kalman filter.

3. **Behavior:**

(a) **Humans:** Human and other animal behavior is studyable and is underpinned by intelligible, but as yet unknown, computational processes.

Results in visual neurophysiology over the last 40 years have shown that a large number of neural structures exist for the purpose of extracting visual information, and testable theories for some of these have been developed [Marr 1982].

As we are only well-informed regarding the complexity of the human visual perception, we cannot be sure of the sophistication of other intelligent systems. However, there is no doubting the visual intelligence of humans through observing our arts — particularly the abstract, surreal or cartoon forms. Through these we move from a literal description of shape and reflectance to a reductionist symbolic representation of the world expressed in a functionally useful concepts and then to a cultural dialogue, in which the form and content of the image (or sculpture) is as much a response to current historical, social, and artistic context as a description of a possible reality. As for accessibility, we know a little about human perception from observers' reports and agnosia [Farah 1990] studies; about other species we know almost nothing. For the early stages of animal vision, the knowledge of the active neurology in other species may be greater because of the information obtained from live animal experiments.

(b) **Complex behavior:** Complex, intelligent behavior is a consequence of the complexity of the domain in which that behavior occurs (i.e. the richness of human experience and imagination generates linguistic complexity).

The complexity of visual behavior is reflected in the complexity of the programs that implement that behavior. Almost any vision system that does anything of consequence has thousands of lines of code at a minimum. Systems that aspire to even limited degrees of "general-purpose" capability [Draper et al 1988] have more like hundreds of thousands of lines of code and often involve the use of complex reasoning mechanisms, such as blackboards. Even extracting simple information from a real image (e.g. counting cells in a microscope slide view) is complex because of the detail and variability of the real world (e.g. different cell sizes, optical constraints, placements, debris, adjacencies, variations in boundary appearance and shape, etc).

(c) **Intelligent perception:** Intelligent perception requires integrating many different sources of information, plus the use of knowledge about what is being perceived.

Intelligent vision particularly requires integrating many different sources of visual information, as generating a full understanding of complex scenes seems to require different representations of shape, position, color, motion, etc. For example, while stereo gives information about distance to and shape of regions containing a lot of visual texture, extending the understanding to nearby non-textured regions needs other information, from, e.g. the shading. We can also actively use knowledge about the nature of the world, as in when we reason about how a pile of books that partly hides a water glass affects the appearance of that glass, and how the optics of the light passing through the glass affects what we see through the glass itself.

(d) **Multiple theories:** No single theory explains all intelligent behaviors, and an agent may utilize different behaviors as is appropriate.

CV is a perfect example of where multiple types of behaviors are needed, as much low-level image work is very data-driven, whereas most high-level systems embody both data and model driven reasoning. Most low-level programs (i.e. image-to-image processing) are numeric, whereas most high-level are purely symbolic. Classical symbol manipulation high-level vision systems can interface to neural network low-end modules. Active or multi-camera vision systems can sometimes acquire data far more easily or reliably than passive monocular vision systems.

(e) **Multiple skills:** Truly intelligent systems are capable of: perception, communication, memory, learning, self-analysis, self-knowledge, decision making, acting, planning, attention focusing, and undoubtedly other skills (in varying degrees). These skills will needed to be integrated into a cohesive system to be generally useful.

In CV, perceptual results need to be encoded for use, the shape and appearance of objects must be known and those of new objects must be learned; a vision system that can move to obtain a better viewpoint is more effective; knowing what results can be trusted can guide when to proceed; knowing how well a set of CV processes works when applied to different domains can guide the selection of which to apply.

(f) **Learning:** Learning is required, because: a) the amount of knowledge available is too immense for explicit encoding by human designers, b) domains change over time, c) new concepts enter discourse, and d) agents enter new domains.

Model-based vision needs models; mostly these are constructed by humans and, for complex objects, rather slowly. Self-acquisition is the only possibility for acquiring the description of large numbers of objects that must be known by competent active agents. Visual descriptions must change over time — as a consequence of seasons, aging, wear, growth or encountering new objects.

Some models might be learned using classical structural learning methods, adapted for visual representations. However, experience with medical imaging suggests that there are patterns (e.g. biological) that are too complex for compact description — and hence the communication of these patterns by other than example will be nearly impossible. And so, much as a medical specialist learns to interpret a class of image data by repeated exposure over six months, so too will machine vision systems learn, perhaps in a connectionist manner.

3 AI and CV Share Share Goals

The main goals of AI research are:

1. To characterize intelligence and intelligent behavior in general. This will include theories of the architectures for integrating the various intelligent skills into single autonomous agents and cooperating systems of agents.

2. To understand human competences and computational processes, in part by providing a methodology for developing testable theories.

3. To develop tools that need less human attention, embody greater capability and compiled experience, and extend human control over our environment (i.e. "environment" in the sense of our social, biological, physical and intellectual context). The "intelligence" of the tools is in part a consequence of their complexity and reactive flexibility; however, the real breakthrough is in the embedding of intent into the tool.

4. To develop epistemologies suitable for representing different knowledge domains.

5. To develop tools for advanced computer science and engineering (e.g. distributed processing, programming languages, correctness proving, automatic programming).

6. To extend the philosophical view of humanness in order to reconcile the physical operation of the brain and its computational structure with the concept of the mind and the mental universe (e.g. [Dennett 1992]).

The goals of CV are:

1. To understand the human and other biological vision systems through building testable models. This must also include understanding the purposes of vision [Gibson 1979, Sloman 1989], which consists of multiple information extraction modules, providing visual control (e.g. motion) feedback as well as information extraction.

2. To provide machines that extend human perceptual abilities into new domains or heighten it in normal domains.

3. To provide tools that embody autonomous information extraction abilities (e.g. need not employ a human).

4. To determine the key representations needed for a visual and spatial description of the real world, and to discover computational processes that can reliably infer them.

Hence, it can be seen that CV's goals are largely specializations of AI's goals. (CV's goal 1 is a specialization of AI's goal 2, CV's goals 2 and 3 are specializations of AI's goal 3, and CV's goal 4 is a specialization of AI's goal 4.)

4 AI and CV Share Share a Common Intellectual Context

We also consider the relation of AI and CV from the perspective of three aspects of their shared intellectual context.

4.1 Philosophical Support

There is no issue more central to AI than that of world representation, and perception, through its direct external input from the world, provides the basis for constructing those first, basic representations. The philosophy of perception and mind shows strong connections between AI and CV, through the attempts to understand the relationship between a physical reality and our perceived understanding of that world. Fundamental to the perceptual view is that what we embody is not a literal description of the world as it is, but an internalized and abstracted representation of that world, encoded into terms forming part of the internal states of the agent. Without such an encoding the description would be of no use - we would still need to analyze the new description to extract useful information [Sloman 1989]. The full nature of the actual representations is unclear and is perhaps incompletely communicable between different humans. What is clear is that it is a "representation", as compared to the "real thing" and hence impoverished.

These abstractions are necessarily interpretations. As sensors are tuned to sense particular modalities, then the results they report are biased by what they expect to see. A sensor that detects "green" must decide what is or is not green, whereas the physical spectrum is dense and the light coming off objects usually has components at all frequencies. This complication is particularly acute with the color "brown", whose perception is affected by the relative lightness of the surrounding colors. Hence any conceptual representation of the world must have a relation to the world, but is not the world.

Where AI has its strongest linkage with CV is in those aspects of vision that are distinctly human. This is the symbolic partitioning of our environment into conceptual, named entities. For example, an apple on a table is a relatively distinct entity, yet when still attached to a tree it is part of a larger entity. We decide which piece is the apple. Considering some subset of the world a distinct nameable entity is often simply a human convention and thus suitable for AI. Of course, not all entities are defined purely by convention, as biological and physical processes clearly play a major role in physically defining the world.

Conceptual entities need not be simply objects, but also actions (e.g. when does a jog become a run or a sprint) and attributes (blue versus cyan), etc. And, of course, names are linked to function — so one physical object may be referenced by different names according to the use that the object has (e.g. as a bowl, cup, ashtray, paper clip holder, etc.). It is clear that one must also know a lot about human society and its conceptual structures in order to interpret the meaning of the visual input at any level deeper than a largely empirical, physical description.

4.2 Biological and Psychological Support

Vision is often naively considered a monolithic sense, whereas there are actually many different information extraction and interpretation processes involved. Neurophysiologists have identified separate regions of the visual cortex that appear to extract shape, motion, color and edges from our visual input. Altogether, about 9+ (including the retina) distinct regions of visual processing have been identified, many with several processing layers — so something like 20+ different, but as yet unknown, representations of the world may be extracted. The neural connectivity of some regions in the early stages of the processing are broadly known [Hubel 1988, Zeki 1993], but on the whole, their function is not. Some activities are fully autonomous (e.g. processing the full visual field input), whereas others involve input from other centers of the brain — e.g. attention focusing.

Outputs of visual processing are used in many places — for helping maintain balance, tracking moving objects, or stabilizing our representation of the world while in motion, ducking or blinking when danger approaches, making local maps of our environment, discriminating between alternatives, as well as the obvious labeling of the world.

From this list of visual functions, it is clear that there are many biological visual processes that contribute to and define the behavior of intelligent beings. The linkage also goes the reverse direction, in that intelligent processes are needed to select what to attend to, to record and extract visual memories, to provide the motion needed for active vision (e.g. through motion parallax), to do spatial reasoning, to make visual aesthetic judgements, etc.

4.3 Shared Computational Methodology

The foundation behind the experimental methodology of both AI and CV is the computer. However, many fields use computers, so we must look deeper to the more fundamental notion of the computer as a tool for theory testing, as a tool for empirical explorations and as a vehicle for embodying theories to create usable artifacts (as compared to more conventional numerical calculating, text processing and database transaction machines). While theories of mechanics or energy flow, for example, lie behind the function of other machines, the theory of computational information processing lies behind machines built for AI and CV [Marr 1982]. This entails internal representations, reasoning based on and transformations between representations, constructions of those representations and actions based on them.

The use of the computer is so pervasive in AI and CV (and now in most other fields) that it is taken for granted. What is more important are the many shared tools and techniques that exploit the computer capabilities to achieve both AI and CV. A non-exhaustive list must include: frame representation techniques [Brooks 1982], expert systems machinery [Matsuyama 1990], probabilistic and uncertain reasoning, search tree exploration [Grimson 1990], generate and test algorithms, constraint satisfaction systems [Brooks 1982, Waltz 1975], hierarchical representations and reasoning methods [Fisher 1989], both symbolic and neural network [Hinton 1981] perspectives, and the embedding of much domain specific knowledge. These shared techniques are used mostly for what is loosely called "high level vision", which is preoccupied with "symbol-to-symbol" transformations.

Also shared by both domains is the methodology of experimental programming, in that the computer is the ideal tool for performing experiments to validate or explore intelligent information processing processes.

5 Conclusion

It was notable how easily the more general characteristics of AI relate to the more specific characteristics of CV. Such a straightforward characterization surely suggests that there is still a strong relationship. The largely shared goals, plus the additional support of the shared interests in philosophical underpinnings and neurophysiological mechanisms and processes also strengthen the conclusion.

To complete this essay, I look at a few aspects of CV that are intertwined with some of the central issues of AI. To start with we consider nameability.

CV, at the highest levels, addresses recognition of objects and actions. Since these entities are not distinguishable merely by appearance, this activity necessarily links into other areas of AI, such as natural language for naming conventions, common sense reasoning, and robotics (behavior). As the nameability of the human world depends ultimately on the humans that inhabit it, CV must be based on the modes of intelligence that provide the names.

However, once we have names, we also need some way to invoke mental structures connected with the names, that is, to select their model from the visual description base that potentially explains a set of visual data. The details of how this is done are not clear, but it is interesting that this selection process regularly appears in other subfields of AI, as in invoking schema for dialogue understanding, or cases in case-based reasoning or appropriate meta-level search heuristics.

CV also requires the ability to generalize and reason about similarity. Suppose we encounter a person that we have never seen before. We don't have to go sequentially: "Now, is this a house, or a dog, or an apple, or …?". We can directly generalize from the specific person's appearance to the general nature of humanness. We are not troubled (much) by new haircuts, or missing limbs. Somehow we abstract into a space that compares the generalizations.

On the other hand, CV will need to explore more "case-based" and opportunistic reasoning. It is well-known that people need a period of training before they can achieve expert-level performance at new, non-intuitive, visual interpretation tasks, such as x-ray interpretation, radar display interpretation, etc. It is also clear that people sometimes use specific features for search or identification, cues (e.g. color of a book, or a scar distinguishing "identical" twins). Some visual learning seems to be largely iconic (e.g. alphabetic letters, word groups). These examples suggest the use of case methods, with problem specific, compiled processes rather than use of generic high-level visual processes. And, as it is well-known that people can "compile" explicit reasoning into "intuitive" procedures, this suggests that future vision systems may have to apply their case-based reasoning in both forms.

Once our intelligent AI system is capable of performing, it will need to have a large visual knowledge base (a specialization of the general knowledge-base search [Lenat and Feigenbaum 1991]). For general expert performance, a large corpus of "common sense" visual knowledge will be needed — about how classes of objects typically appear, how to discriminate between specific classes, object's material properties, likely contexts and associations, etc. The vision system will need to build, extend and generalize the database from new examples. Models of time, causality, general physical principles and uncertainty will be needed to interpret the observations.

Through these themes, I am trying to second-guess the future directions for CV research. One common thread is that there will be a greater dependence on the methods being developed for general AI systems: case/frame matching, case/frame invocation, truth-maintenance systems, generalization, learning, control of combinatorial search, etc. Some form of self-understanding will be needed for feedback on performance. Planning and focus of attention mechanisms will be needed to focus computational resources. At the same time, CV will become essential for a truly intelligent autonomous AI machine, if only for the ability to learn for itself (let alone the philosophical contention that an artifact cannot know the world if it is incapable of acting in it, and hence sensing it). Hence, I must conclude that the connections can only grow stronger as the two fields develop.

What is central to Computer Vision are issues of how to represent what is known and observable, how to reason with this represented information and how to act on that knowledge (controlling both internal and external behavior). As these will always be three of the main foundations of Artificial Intelligence, there is no chance that Computer Vision will ever drift far. In the past decade, much Computer Vision research has concentrated on developing competences that reliably extract useful low-level descriptions of the world. As this research phase matures, there will then be a major increase in research that relates those descriptions to stored representations of objects and situations, and this will again make clear the association with Artificial Intelligence.

Acknowledgements

I'd like to thank many people for their thoughts on this theme, but particularly A. Bundy, H. Hughes, C. Malcolm, A. Sloman, M. Trucco and M. Uschold.

Biographical Sketch

Dr. Robert Fisher is a Senior Lecturer in the Department of Artificial Intelligence at the University of Edinburgh (member of staff since 1984). He received his PhD from University of Edinburgh (1987) and has been researching 3D scene understanding and other topics in computer vision for over 10 years.

References

[1] Aloimonos, J. 1989. Integration of visual modules : an extension of the Marr paradigm. Boston, Mass: Academic Press.

[2] Barrow, H. G., Tenenbaum, J. M. 1978. Recovering Intrinsic Scene Characteristics from Images. In Computer Vision Systems, eds. Hanson and Riseman, 3-26, New York: Academic Press.

[3] Boden, M. A. 1988. Computer Models of Mind. Cambridge,UK: Cambridge Univ. Press.

[4] Brooks, R. A. 1981. Symbolic reasoning among 3-D models and 2-D images. Artificial Intelligence, 17:285-348.

[5] Batchelor, B. G., Hill, D. A. and Hodgson, D. C. 1985. Automated Visual Inspection. Kempston,Bedford,UK:IFS (Publications) Ltd.

[6] Brooks, R. A. 1991. Intelligence Without Reason. Proc. 12th Int. Joint Conf. on Artificial Intelligence, 569-595. Menlo Park: Int. Joint Conferences on Artificial Intelligence.

[7] Dennett, D. C. 1992. Consciousness Explained. London: Allen Lane.

[8] Draper, B., Collins, R., Brolio, J., Hanson, A. and Riseman, E. 1988. Issues in the Development of a Blackboard-Based Schema System for Image Understanding. In Blackboard Systems, eds Engelmore and Morgan, 189-218, Addison Wesley.

[9] Durrant-Whyte, H. F. 1987. Uncertain geometry in robotics. Proceedings IEEE Conference on Robotics and Automation, Vol. 2, 851.

[10] Farah, M. J. 1990. Visual Agnosia. Cambridge, Massachusetts: MIT Press.

[11] Fisher, R. B. 1989. From Surfaces to Objects: Computer Vision and Three Dimensional Scene Analysis. Chichester, England: John Wiley and Sons.

[12] Gibson, J. J. 1979. The Ecological Approach to Visual Perception. Boston: Houghton Mifflin.

[13] Grimson, W. E. L. 1990. Object Recognition by computer: The role of Geometric Constraints. Cambridge, Massachusetts: MIT Press.

[14] Harnad, S. (Editor). 1987. Categorical perception : the groundwork of cognition. Cambridge, UK: Cambridge University Press.

[15] Haugeland, J. 1985. Artificial intelligence : the very idea. Cambridge, Mass: MIT Press.

[16] Hinton, G. 1981. A Parallel Computation that Assigns Canonical Object-based Frames of Reference. Proceedings 7th Int. Joint Conf. on Artificial Intelligence, 683-685. Menlo Park: Int. Joint Conferences on Artificial Intelligence.

[17] Horn, B. 1975. Obtaining Shape from Shading Information. In The Psychology of Computer Vision, ed. P. Winston, 115-155. New York: McGraw-Hill.

[18] Hubel, D. H. 1988. Eye, brain, and vision. New York: Scientific American Library/W. H. Freeman.

[19] Lenat, D. B. and Feigenbaum, E. 1991. On the Thresholds of Knowledge. Artificial Intelligence 47:185-250.

[20] Marr, D. 1982. Vision. San Francisco: W.H. Freeman and Co.

[21] Matsuyama, T. 1990. SIGMA : a knowledge-based aerial image understanding system. New York; Plenum Press.

[22] Penrose, R. 1989. The emperor's new mind : concerning computers, minds, and the laws of physics. Oxford: Oxford University Press.

[23] Rosenschein, S. J., Kaelbling, L. P. 1986. The Synthesis of Machines with Provable Epistemic Properties. Proc. Conf. on Theoretical Aspects of Reasoning about Knowledge, J. Halpern (Ed), 83-98. Los Altos, CA: Morgan Kaufmann.

[24] Schank, R. C. 1991. Where's the AI? AI Magazine, 12(4):38-49.

[25] Sloman, A. 1989. On designing a visual system (Towards a Gibsonian computational model of vision). J. Expt. Theor. Artif. Intell. 1:289-337

[26] Waltz, D. 1975. Understanding line drawings of scenes with shadows. In The Psychology of Computer Vision, ed P. Winston, 19-91. New York: McGraw-Hill.

[27] Zeki, S. 1993. A Vision of the Brain. Oxford: Blackwell Scientific Publications.