Cognitive Computer Vision – Past, Present, and Future

David Vernon
Etisalat University
UAE

It is now just over forty years since Roberts first published the results of his seminal attempts to construct a computer vision system. Since then, computer vision has matured and undergone many stages in its evolution. From the blocks world approaches of the sixties, to the knowledge-based approaches of the seventies, the hierarchical modular information processing approaches of the eighties, the creation of mathematically-sound robust early vision of the nineties and the associated expansion of vision based on computational geometry, to the more recent probabilistic techniques. On the way, computer vision has spawned a number of successful offshoots, such as the machine vision of industrial inspection, the analysis of video data for remote monitoring of events, and the use of image analysis to facilitate special effects in the film industry. However, to date, the ultimate goal of creating a general-purpose vision system with anything close to the robustness and resilience of the human visual system remains as elusive as ever.

One of the more recent trends in computer vision research in the pursuit of human-like capability is the coupling of cognition and vision into cognitive computer vision. Unfortunately, it is apparent that the term cognitive computer vision means very different things to different people. For some, it means the explicit use of knowledge and reasoning together with sensory abstraction of data from a perceived environment; for others it implies the emergent behaviour of a physically-active system that learns to make perceptual sense of its environment as it interacts within that environment and as a consequence of that interaction. For others yet, it is a meaningless term in its own right and cannot be treated except as an embedded component of the process of cognition that, in turn, is an inherent feature of autonomous systems.

In this talk, we will trace the lineage of cognitive computer vision and place it in the context of the many relevant contributing disciplines, such as cognitive psychology, cognitive science, artificial intelligence, robotics, developmental psychology, dynamical systems theory, machine learning, as well as computer and human vision. In doing so, we will draw out the various distinctions that people make (and don't make but probably should) in referring to cognitive vision and, in particular, we will set out the various underlying assumptions that each approach makes in its approach to vision, perception, cognition, and intelligence. The ultimate goal in this talk is to sketch a map of the overall domain so that one can see what choices one is actually making when one adopts a particular approach and so that, hopefully, one can identify some opportunities to develop and to reinforce the linkages between complementary but mutually-relevant approaches.