

Chapter 4

Topological Semiotics of Visual Information

ABSTRACT

This chapter explains how active vision and image understanding can be implemented with topological semiotic models, using cognitive architecture with perceptual mechanisms similar to human vision.

INTRODUCTION

There were significant efforts in conversion image data into meaningful informational structures, and also on the usage of context in the processing of visual information. For instance, Geographic Information Systems (GIS) can adequately address problems with geographic and satellite imagery, because geographic knowledge has been well formalized in the form of maps, and maps can be represented well in digital form.

In the field of multimedia, the MPEG-7 standard was an extensive industry effort to address these problems for generic images, converting them into XML structures. MPEG-7 provides a set of image primitives called Descriptors. The MPEG-7 Description Scheme is the structure and semantics of the relationships between image components, which may be both Descriptors and Description Schemes. A MPEG-7 image description consists of a Description Scheme and a set of Descriptor Values.

MPEG-7 supports a range of abstraction levels, from low-level video features, such as are object's shape, size, texture, color, movement, and

DOI: 10.4018/978-1-5225-2431-1.ch004

position, to high-level semantic information. However, the MPEG-7 standard reflects the present state of image/video processing, and it only provides a set of predefined descriptors and schemas. MPEG-7 Visual Descriptors evolve from low-level image processing, which is well understood and formalized. However, Description Schemas relate to mid- and high-level image processing, which has not yet been well formalized.

Neither automatic and semi-automatic feature extraction nor schema creating algorithms is within the scope of the MPEG-7 standard. Although most low-level features can be extracted automatically, high-level features and schemas usually need human supervision and annotation. Only the description format in MPEG-7 is fixed and not the extraction and transformation methodologies. These are the areas that must be addressed.

The highest level of image description is the semantic one, and MPEG-7 standardizes information on these levels. But the problem of transforming primary image structures directly into semantic description has not been solved yet, as processes on the intermediary levels are not well understood and formalized.

Although RDF is better than other schemas in its ability to specify relationships and graphs, the MPEG-7 Group has made a decision to use an easily understandable and readable XML Schema Language as the MPEG-7 DDL. However, neither RDF nor XML Schema has been designed to describe complex dynamic hierarchical structures that constitute most of the real images.

MPEG-7 Visual Descriptors can be used for searching and filtering images and videos based on several visual features such as color, texture, object shape, object motion, and camera motion. This allows measuring the similarity between images and videos. Such a set of descriptors might be sufficient for the entire image.

Similar to MPEG-7 approaches convert images into their structured description that is based on low-level image features and their combinations, which use either top-down or bottom-up flow of processing image data or both types of flow, and attaching linguistic values for semantic querying. Most of them are trying to convert the image into a sort of structural description that can be compared against a similarly described collection of images stored in a database. (See Figure 1)

These approaches might work well for image and multimedia databases as they allow for creating structured collections of images, and querying them on certain similarity criteria, but not for the systems that must perform in the real-time and hostile environments. These approaches are not able to provide the needed level of understanding of the environment.

32 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/topological-semiotics-of-visual-information/189278

Related Content

AI and Blockchain for Industrial Robotics: Ethical, Legal, and Social Implications

Sanjana Prasad and Deepashree Rajendra Prasad (2024). *AI and Blockchain Applications in Industrial Robotics* (pp. 33-57).

www.irma-international.org/chapter/ai-and-blockchain-for-industrial-robotics/336074

Robotic, Grippers, Grasping, and Grasp Planning

Seyed Javad Mousavi and Ellips Masehian (2013). *Engineering Creative Design in Robotics and Mechatronics* (pp. 277-300).

www.irma-international.org/chapter/robotic-grippers-grasping-grasp-planning/78148

Adaptive Self-Organizing Organisms Using A Bio-Inspired Gene Regulatory Network Controller: For the Aggregation of Evolutionary Robots under a Changing Environment

Yao Yao, Kathleen Marchal and Yves Van de Peer (2016). *Handbook of Research on Design, Control, and Modeling of Swarm Robotics* (pp. 68-82).

www.irma-international.org/chapter/adaptive-self-organizing-organisms-using-a-bio-inspired-gene-regulatory-network-controller/141995

The Synergy of AI and Robotics in Modern Education

Froilan Delute Mobo, Ali Sorayyaei Azar and Ana Liza Reclozado Garcia (2025). *Innovations in Educational Robotics: Advancing AI for Sustainable Development* (pp. 69-86).

www.irma-international.org/chapter/the-synergy-of-ai-and-robotics-in-modern-education/372096

Emotion Recognition Approach Using Multilayer Perceptron Network and Motion Estimation

Mohamed Berkane, Kenza Belhouchette and Hacene Belhadeif (2019). *International Journal of Synthetic Emotions* (pp. 38-53).

www.irma-international.org/article/emotion-recognition-approach-using-multilayer-perceptron-network-and-motion-estimation/238081