

Arnold W.M. Smeulders
"Object Class Recognition"

March 26, 2009
2 pm
HEC 101

Abstract:

Since the turn of the century, machine learning has entered computer vision to its core much to the benefit of the state of the art. In fact, many computer vision solutions are reached by connecting a good set of features to learning from selected examples. Objects in the world have a predictable repertoire of shapes and appearances. Textures, soil, sky and the like, only share a common appearance. In either case we have no other choice than to use the regularities in the appearance for recognition. When analyzing the actually recorded data it is far from trivial to decide what component is intrinsic to the object and what part in the recorded data is accidental due to this particular recording.

We argue that it is important to build a representation of objects by an array of features ranging from a set of features holding the basic spectral values, to sets of invariants, to regional, textural and contextual features. The first set of features is ruled by the laws of image formation, the second set by the laws of object and texture reflection, the third one by spatial coherence, and the latter types are ruled by image statistics. Under this paradigm we have treated tracking as an instant machine-learning problem: how can we recognize the object learned from one image in the next one? And we have built a video search engine classifying objects into a large vocabulary of concepts. The system has consistently performed well in the TRECvid competitions over the last 5 years amidst participations from IBM research, Columbia, Oxford, and Tsinghua University. Last year, we have also performed well in the PASCAL object recognition competition, with the same system that is. We will discuss the contributions on video search and object classification.

Arnold W.M. Smeulders

Arnold W.M. Smeulders graduated from Technical University of Delft in physics in 1977 (M.Sc.) and in 1982 from Leyden University in medicine (Ph.D.) on the topic of visual pattern analysis, and turned professor in bio-informatics in 1990 working on the topic of 3D image processing. Since 1996, I am professor in multimedia information analysis at the University of Amsterdam and head of the Intelligent Systems Lab Amsterdam. My favorite research topics these days are video search engines, learning to recognize object categories, and the picture to language barrier in general. In ISLA, color vision invariant features and natural image statistics are important features. In 2008 the same system of ours landed on top of both the video search engine competition TRECvid, run by NIST, as well as the PASCAL challenge for object class recognition.

The lab has an extensive record in co-operations with Dutch institutions and industry on the topics of multimedia search engines, machine learning and Internet content exploration. He is scientific director of the MultimediaN public-private national partnership, and of the ASCI national research school. He leads the EU Vidi-Video on video search engines. He is fellow of International Association of Pattern Recognition and associated editor of the International Journal for Computer Vision.