

Where Am I?

From Spatiotemporal Descriptions To a Sketch To a Geospatially Grounded Map

JAMES M. KELLER

*Electrical and Computer Engineering Department
University of Missouri-Columbia
Columbia, MO 65211
KellerJ@missouri.edu*

ABSTRACT

With the collaboration of several faculty colleagues and many students, I have been studying the creation and utilization of spatial relations in various sensor-related domains for many years. Scene description, involving linguistic expressions of the spatial relationships between image objects, is a major goal of high-level computer vision. In a series of papers, we introduced the use of histograms of forces to produce evidence for the description of relative position of objects in a digital image. There is a parameterized family of such histograms, for example, the histogram of constant forces (much like the earlier histogram of angles) and the histogram of gravitational forces that highlights areas that are close between the two objects. Utilizing the fuzzy directional membership information extracted from these histograms within fuzzy logic rule-based systems, we have produced high-level linguistic descriptions of natural scenes as viewed by an external observer. Additionally, we have exploited the theoretical properties of the histograms to match images that may be the same scene viewed under different pose conditions. In fact, we can even recover estimates of the pose parameters. These linguistic descriptions have then been brought into an ego-centered viewpoint for application to robotics, i.e., the production of linguistic scene description from a mobile robot standpoint, spatial language for human/robot communication and navigation, and understanding of a sketched route map for communicating navigation routes to robots. This last activity can be labeled as Sketch-to-Text.

Recently, with a grant from the National Geospatial Intelligence Agency and collaboration with personnel at the Institute for Human and Machine Cognition, we are tackling the inverse problem: given one or more text descriptions of a temporal and spatial event, construct a sketch of the event for subsequent reasoning. This is called Text-to-Sketch by the NGA. The idea is that the person or persons providing the linguistic descriptions either may not know where they are exactly or they may not use referenced landmarks in the descriptions. Hence, the input may only be a temporal sequence of objects and their relationships. The produced graphics-based sketch must be grounded in reality by matching it to a satellite image or geospatial database. The techniques involve natural language understanding, fuzzy approaches and force histogram matching for intelligent sketch production, and subgraph isomorphism algorithms and/or genetic algorithms for sketch to geospatial database matching. This talk will quickly survey the early applications and focus on our approach to the new problem.