Thesis: Literature Review: Context Understanding in Computer Vision

Abstract: Contextual information plays an important role in many computer vision tasks, such as object detection, video action detection, image classification, etc. Recognizing a single object or action out of context would be challenging sometimes, but context information may help improve the understanding of a scene or an event. Appearance context information, e.g., colors or shapes of the background of an object can improve the recognition accuracy of the object in the scene. Semantic context (e.g. a keyboard on an empty desk vs. a keyboard next to a desktop computer) will improve accuracy and exclude unrelated events. Context information, such as the time or location of an images captured, that are not in the image itself can also help to decide whether certain event or action should occur. Other types of context (e.g. 3D structure of a building) will also provide additional information to improve the accuracy. In this survey, different context information that has been used in computer vision tasks is reviewed. We categorize context into different types and different levels. We also review available machine learning models and image/video datasets that can employ context information. Furthermore, we compare context based integration and context-free integration in mainly two classes of tasks: image-based and video-based. Finally, this survey is concluded by a set of promising future directions of context learning.

Committee:

- Professor Zhigang Zhu, Mentor, The City College
- Professor Jie Wei, The City College
- Professor Hao Tang, Borough of Manhattan Community College
- Professor Jie Gong, Outside Member, Rutgers University-New Brunswick
- Dr. William H. Seiple, Outside Member, Lighthouse Guild