

Activity Maps for Location-Aware Computing

David Demirdjian, Konrad Tollmar, Kimberle Koile, Neal Checka & Trevor Darrell

Artificial Intelligence Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

<http://www.ai.mit.edu>



The Problem: Location-based context is important for many applications. Previous systems offered only coarse room-level features or used manually specified room regions to determine fine-scale features. We propose a location context mechanism based on activity maps, which define regions of similar context based on observations of 3-D patterns of location and motion in an environment. We describe an algorithm for obtaining activity maps in real time using the spatio-temporal clustering of visual tracking data.

Motivation: In many cases, fine grain location based information is preferred. One example would be to control lights and air conditioning, e.g. the desk lamp might light up and the air conditioning starts whenever a user is sitting at his desk. In addition the phone might become activated and the computer screen get invoked from stand-by mode. Similarly in a small group meeting the system could know where and how many people are in the room and could make appropriate settings for lights, air conditioning, and computer tools. For each of these tasks, location context information is important [3].

Simply considering the instantaneous 3-D location of users is useful, but alone is insufficient as context information. Applications have to generalize context information from previous experience, and an application writer would like to access categorical context information, such as what activity a user is performing. In addition, other features such as motion and shape (configuration) of the user are often important to distinguish activity: contrast a person walking past a desk with a person sitting at that desk.

Previous Work: Previous approaches have partitioned space based on simple proximity or relied on user specified maps for regions. In contrast, we argue that location regions should be learned from observed activity, including motion and shape cues as well as position. Regions can overlap in space, since motion or shape can indicate a different activity.

In this paper we describe an algorithm for computing location context based on 3-D person tracking techniques and the use of automatically generated activity maps. Our system is robust to many of the issues that often plague computer vision systems, such as dynamic illumination or fast motions. We form activity regions using a spatio-temporal clustering method and use the resulting regions to define an activity map. This map is used at run time to contextualize user preferences, e.g., allowing “location-sticky” settings for messaging, environmental controls, and/or media delivery.

Approach: A multi-camera stereo-based tracking system developed in our group [2] is used in order to track people in indoor environments. The person tracker provides a history of 3-D information of every person in the observed space. The 3-D information consists of (x, y, h) where x, y is the coordinates of the person in the ground plane and h is the relative height of the person with respect to the floor.

Since tracking data are time-stamped, the instantaneous velocity (v_x, v_y, v_h) can be derived. We determine a persons features' from the history of spatio-temporal tracking data of a person. We characterize a person at location (x, y) are: $f(x, y) = (h, v, v_{\text{lt}})$ where $v = \sqrt{v_x^2 + v_y^2}$ is the instant ground plane velocity norm and v_{lt} , the average ground plane velocity norm over a certain amount of time. By using the features $f(x, y)$, we can capture the configuration (sitting, standing) and motion of a person over both short and long period of time.

By tracking people in a space for a long period of time, a dense set of observed location features $f_i(x, y)$ can be gathered. We define an activity zone as a connected region where observed location features $f_i(x, y)$ have similar values. An activity zone Z_k is defined by a connected region R_k in the 2-D space defined by (x, y) and a characteristic feature $F_k = (h, v, v_{\text{lt}})$ representing the typical activity in this area. An example of segmentation result is shown



Figure 1: A one-person office. The estimated activity map (right) contains 4 zones. Zone 1 corresponds to the “walking context”, zone 2 corresponds to the “working context” (desk), zones 3 and 4 correspond to the “resting context” (chair on the bottom right of the picture). Zone 3 could be associated to the transition between zone 1 and zone 4 (chair).

on Figure .

Impact: Our experiments on different environments show that our system is able to generate activity maps that give an improved understanding of a person’s context over previous approaches to sub-room location modeling which required that users explicitly define physical regions [1]. The improvement is due to our system’s ability to use fine-scaled features that include position, motion and height, making the identification of a person’s context more accurate than one based on position along.

Future Work: There are many avenues of future work planned for our system. In addition to making the system more accurate and fast, we plan to add a statistical estimation formulation to the region estimation process. This will make the estimated regions more stable to noise in the sensing process. We also wish to include higher-level information about the tasks users are performing in the environment and the objects they are manipulating to aid in determining location. We speculate that more complex application behavior can be achieved by augmenting the system with knowledge of objects (e.g. desk, computer) and human behavior (e.g. people generally read, write at a desk). Adding a simple object recognition system and task knowledge base is planned future research.

References:

- [1] B. Brumitt, J. Krumm, B B. Meyers, and S. Shafer. Ubiquitous computing and the role of geometry. *IEEE Personal Communications*, 7-5:41–43, August 2000.
- [2] T. Darrell, D. Demirdjian, N. Checka, and P. Felzenszwalb. Plan-view trajectory estimation with dense stereo background models. In *2001 International Conference on Computer Vision*, 2001.
- [3] Bill Schilit Roy Want. Special issue on location-aware computing. *IEEE Computer*, 34(8), 2001.