

Vision: ideas for enabling Ambient Intelligence and serving Social Networks

Hamid Aghajan, Stanford University

aghajan@stanford.edu

Vision offers rich information about events involving human activities in user-centric applications from gesture recognition to occupancy reasoning. Multi-camera vision allows for applications based on 3D perception and reconstruction, offers opportunities for collaborative decision making, and enables hybrid processing through task assignment to different cameras based on their views.

In addition to the inherent complexities in vision processing stemming from perspective view and occlusions, setup and calibration requirements have challenged the creation of meaningful applications that can operate in uncontrolled environments. Moreover, the task of studying user acceptance criteria such as privacy management and the implications in visual ambient communication has for the most part stayed out of the realm of technology design, further hindering the roll-out of vision-based applications in spite of the available sensing, processing, and networking technologies.

The output of visual processing often consists of instantaneous measurements such as location and pose, enabling the vision module to yield *quantitative knowledge* to higher levels of reasoning. While such knowledge is essential in many smart environments applications such as gesture control and accident detection, most ambient intelligence applications need to also depend on *qualitative knowledge* accumulated over time in order to learn user's behavior models and adapt services to the preferences explicitly or implicitly stated by the user.

Proper interfacing of vision to high-level reasoning allows for integration of information arriving at different times and from different cameras according to: associated confidence levels, available contextual data, as well as the accumulated knowledge base.

This talk presents ideas for interfacing vision to other modules, and overviews a number of potential applications involving the recognition of human activities, namely environment discovery based on user interaction, context-based ambience control services, and experience sharing using avatars.