Duo: A Human/Wearable Hybrid for Learning About Everyday Human Activities

Charles C. Kemp

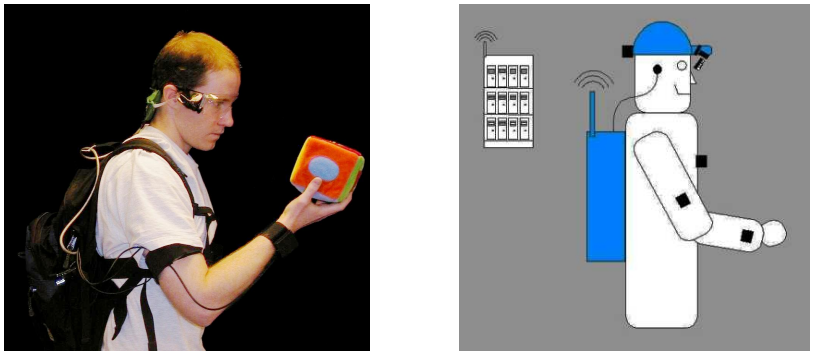

Figure 1: *A picture and diagram of Duo. The diagram depicts the head-mounted camera, the 4 body-mounted orientation sensors, the ear phones, and the wirelessly linked computer cluster.*

What: We are investigating wearable computing as a way for machines to learn about everyday human activities from a first-person perspective [2] [4]. Our wearable system, Duo, uses body-mounted orientation sensors and a head-mounted camera to sense the kinematic configuration of the wearer's body while watching the world from the wearer's point of view. Duo both passively monitors the wearer's behavior and actively requests help from the wearer in order to better learn from the wearer's activities.

Why: Human intelligence relies on a wealth of common-sense acquired from a lifetime of experience. In order to achieve the long term goals of artificial human intelligence, researchers must find ways to endow machines with this type of common-sense.

A wearable system with sufficiently sophisticated sensors could experience the world in approximately the same way as the person wearing the system. If this person is also willing to help the wearable creature by cooperating with requests, then the person can, in a sense, serve as the wearable creature's body. Through requests, the wearable creature can coarsely control the actions and behaviors of the cooperative human. In this way, we can think of the human and wearable as being part of a subsumption architecture in which the wearable subsumes some of the physical and intellectual abilities of the human [1]. In addition, a cooperative wearer can play a role similar to a teacher for the wearable creature.

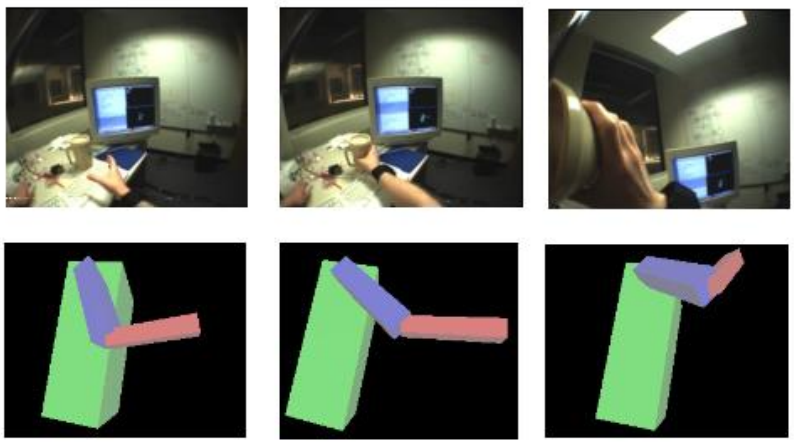

Figure 2: *This figure shows three snapshots of data from a sequence of activity monitored by the system. The wearer reaches for a cup and drinks from it. The top row consists of images from the head-mounted camera. The bottom row shows a kinematic model using the corresponding orientation measurements.*

How: As shown in figure 1, Duo currently consists of a head-mounted camera from which the creature watches the world and 4 orientation sensors with which the creature measures the kinematic configuration of the wearer's head, torso, and dominant arm. Duo can also make requests of the wearer via headphones using text-to-speech software, and can make use of an LED array mounted around the camera to actively illuminate objects held by the wearer.

Progress: We are initially focusing on having Duo learn about the common manipulable objects with which the wearer interacts. By observing the kinematic activity of the wearer's dominant arm, Duo will have the opportunity to learn about the ways in which these common objects are used. The actions applied to an object should be closely related to the object's function, identity, and importance to people. For example, cups can look very different from one another, but the drinking actions applied to cups are very similar to one another. See figure 2.

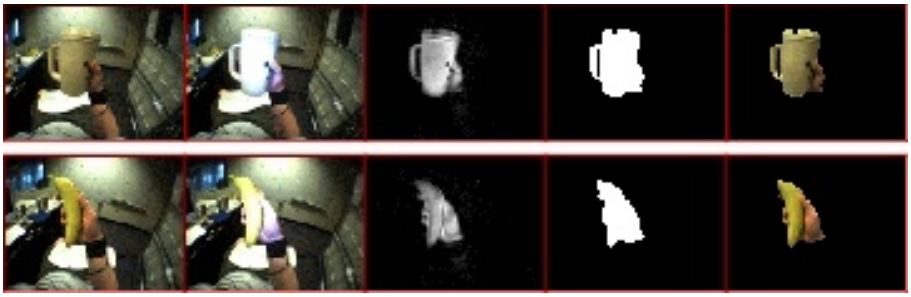

Figure 3: *This figure shows two segmentations of common manipulable objects by Duo. The first column shows Duo's view before the LED flash and the second column shows the view during the LED flash. The third column shows the difference between the flashed and non-flashed images. The fourth column shows the object and hand mask produced by thresholding the difference image from the third column. The final column shows the masks applied to the image to segment the hands holding the objects in the images.*

Duo's first behaviors work together with a cooperative human to acquire high-quality segmentations of everyday manipulable objects. When Duo detects that the wearer has reached for an object, Duo uses speech through the headphones to ask the wearer to look at the object. While looking at it Duo flashes the LED array in order to segment the hand and object, which are in the foreground, from the rest of the world in the background. See figure 3. While looking at the object, Duo also monitors the wearer's head movements. If the wearer's head moves significantly, Duo requests that the wearer keep his head still. While the wearer interacts with the object, Duo records the kinematic actions of his dominant arm [3].

We are currently working on kinematic action recognition, visual object segmentation, and visual object recognition. We are also developing methods to relate observed kinematic actions with the object to which the actions were directed, as well as methods by which Duo can more effectively benefit from the user's expertise and help. For example, by asking the user to "do that again" when an action of interest is observed. Ideally, Duo will eventually be able to request that the wearer perform actions that Duo has previously learned and the wearer has labeled.

Research Support: Funds for this project are provided by DARPA as part of the "Natural Tasking of Robots Based on Human Interaction Cues" project under contract number DABT 63-00-C-10102.

References:

[1] R. A. Brooks. A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation*, 2(1):14–23, March 1986.

[2] Brian Patrick Clarkson. *Life Patterns: structure from wearable sensors*. PhD thesis, MIT Media Laboratory, September 2002.

[3] Charles C. Kemp. Duo: A human/wearable hybrid for learning about common manipulable objects. In *Proceedings of the 3rd International IEEE/RAS Conference on Humanoid Robots*, Karlsruhe, Germany, October 2003.

[4] Taro Maeda, Hideyuki Ando, Maki Sugimoto, Junji Watanabe, and Takeshi Miki. Wearable robotics as a behavioral interface - the study of the parasitic humanoid -. In *Proceedings of the 6th International Symposium on Wearable Computers (ISWC'02)*, October 2002.