AUGMENTED PERCEPTION AND INTERACTION WITH HANDHELD DEVICES

De Tommaso D., Calinon S. - Istituto Italiano di Tecnologia, Genova
{davide.detommaso, sylvain.calinon}@iit.it

ISTITUTO ITALIANO DI TECNOLOGIA

Abstract

We propose a method for enabling handheld devices to share a unique 3D spatial representation of the same environment using mixed reality. This method provides an architecture for collecting environmental data from different kinds of portable devices by their relative perceptual space, but also enables humans to actively interact inside the scene. In our real scenario we consider a RGBD sensor mounted on a robotic arm, used to collect 3D information from the environment, a pico-projector to visualize augmented digital information on real objects and a tablet pc to visualize virtual trajectories of the arm in the real spatial position.

Motivations

As is clear from the developments of the last decade, the electronics research has focused on reducing equipments scale for building embedded systems ever more powerful and small. The combination of their small size, computational capabilities and sensors, enables the user to run a lot of applications everywhere close to hand. Some of these devices can be used: (1) to perceive the sourrounding environment by input sensors such as accelerometers, microphones, RGB or RGBD cameras; (2) to superimpose virtual objects on real environment visible by handheld displays (augmented reality); (3) to superimpose real images on real environments by laser or led lights using pico-projectors. Virtual and real augmented capabilities can be mixed using 3D information about the scene. Points cloud processing and augmented reality enable these devices, with different sensors and computational capabilities, to build a common model of the environment and to extract semantic perceptual information about it. Our research focuses on finding novel methods for robot skill learning by exploring interfaces for both human-computer and human-robot interaction. In this work an infrastructure of interconnected handheld devices, which exchange information about perception and interaction of the same scene, is proposed.

Method and Application

The architecture of the system consists in an infrastructure of interconnected devices which share a common 3D model of the surrounding environment. We consider two class of devices: the *Publisher* and the *Subscriber*. The publisher (the Microsoft Kinect in our case) collects information about the scene, according to its point of view, finds the position and orientation of the *subscribers* and publishes a 3D model of the scene using a cloud of points. The second class of devices (tablet and projector) uses this model to superimpose real or virtual objects on the scene, using geometric transformations of their position and orientation relative to the publisher frame of reference.