Augmented Reality: Visual Tracking

Author: Rohith Perumalla

Date: 12/2/16

Subject: Computer Engineering

Citations:

Fua, Pascal, and Vincent Lepetit. "Vision Based 3D Tracking and Pose Estimation for Mixed Reality." Emerging Technologies of Augmented Reality (2006): n. pag. Web.

Summary:

Augmented Reality applications require exact information about the relative position of an object and their scenes, and if the object moves or turns in any degree. This means keeping real time tracking of all 6 directions of movement of an object and the three dimensional displacement of an object relative to the camera. There are many methods currently used to track objects in real time like Fiducials, Natural Features, and Detection. There are many pro's and con's to each of these.

Analysis:

Augmented Reality is the concept of enhancing a user's environment by superimposing objects either through wearable or handheld devices. A key part of Augmented Reality is having the augmented elements to properly react to the movement of the user and have the computer automatically reposition the object. For this repositioning and reacting to occur fluidly and realistically enough for the user to have a smooth experience Augmented Reality requires tracking of the object (especially in marker based Augmented Reality). Some methods that attempt to satisfy this need include mechanical trackers, magnetic trackers, Ultrasonic trackers, and Inertial trackers. All have their flaws from limited working environments to distortions from the metal in or temperature in their environments. Computer and camera based tracking seems to provide the most benefits, computer based tracking works using a variety of methods like Fiducials and Natural Features.

Vision based tracking can be broken down to two principal steps: first image processing to extract information from the environment and then analyzing the gathered by the scans. Fiducials are elements added the real environment, like markers or labels, that can be recognized by the camera to help with scanning. When scanning the environment, the addition of fiducials like markers help gather precise data from the environment. These markers can be designed to work with an ad hoc method to maximize its efficiency. Circle fiducials are also recommended due to the relatively invariant perspective distortion. While some fiducial systems run off ad hoc methods some systems are implementing machine learning to maximize accuracy and efficiency. Some fiducials use rectangular planes, making elements adapt or change depending on the distortion of the plane. Under good lighting situations Fiducials can be a low cost real time 3D tracking system satisfying Augmented Reality application needs.

Another method often used is the use of natural features, in an environment there are many features that naturally appear and can be mapped and used for visual tracking. A form of this method commonly used is edge detection, in a natural environment there are many different edges. Two surfaces that meet form strong gradients with shadows and with gradient detection edges can be distinguished. With edge detection, it is naturally stable to lighting changes allowing it to be extremely flexible. Another form of natural features is dependent on information gathered by pixels from an object's projection. A specific type of this tracking system is RAPid which was one of the first real time 3D tracking systems. It's core function is based off the idea that in an image there are "control points" that lie on high contrast points indicating an edge from pixel data. These "control points" can be detected and predicted at real time allowing the fluid manipulation of the augmented element. However, RAPid can be tricked by weak contours making this form of pixel detection slightly unreliable. Another type of natural feature detection is texture based detection, a method that uses the camera input to track and map item to specific textures. Texture detection can be derived from optical flow, template matching, or interest point correspondences. The most effective form of texture tracking would be interest point detection and 2D matching. Interest point detection works by picking a few pixels -only a few to minimize the number of computations required- the selected points are different from their neighbors which allows easy recognition regardless of direction, noise, or distortion. However, when using any type of natural feature tracking if any point is not trackable all methods encounter errors and results in tracking failure, to avoid this, keyframes are introduced. Keyframes are registered beforehand, so when any new frame comes into the system the keyframe is tracked and the drifting of the points doesn't affect the functionality of the system; but, in order for keyframes to work in real time it requires a quick implementation that is relatively insensitive to some distortions. There are other methods of Natural Feature tracking, however, edge detection and pixel analyzing have the most potential for real time tracking even though texture based detection could be more reliable.

Visual tracking is imperative for Augmented Reality application, especially to have smooth immersive experiences for users. In order to have real time results with minimal restrictions and distortions Augmented Reality seems to work best with computer analyzed visual scenes. For the computer tracking there are many methods each with their own benefits and downfalls. Fiducials allow real time tracking with three dimensional displacements but require a physical object and proper lighting to work properly. Unlike Fiducials, that require markers, Natural Feature detectors work purely through algorithms either finding edges, with gradient detection or pixel data, or detecting textures and mapping an element to that unique texture. These Natural Feature detectors have the benefit or having increased stability with lighting changes, but are not always as accurate as Fiducials. Currently these methods are being improved and new methods are being developed for real time visual tracking. These visual tracking methods are vital for Augmented Reality applications, but it is safe to assume with the improvement of hardware and increased interest in augmented reality that a stable, fast, and efficient solution will be available in the near future.