Robust 2D Tracking for Real-Time Augmented Reality

Shahzad Malik [1], Gerhard Roth [2], Chris McDonald [1]

[1] School of Computer Science, Carleton University, Ottawa, Canada, K1S 5B6

[2] Computational Video Group, IIT, National Research Council, Ottawa, Canada, K1A 0R6

http://www.cv.iit.nrc.ca/research/ar

Abstract

Vision-based registration techniques for augmented reality systems have been the subject of intensive research recently due to their potential to accurately align virtual objects with the real world. The downfall of these vision-based approaches, however, is their high computational cost and lack of robustness.

This paper describes the implementation of a fast, but accurate, vision-based corner tracker that forms the basis of a pattern-based augmented reality system. The tracker predicts corner positions by computing a homography between known corner positions on a planar pattern and potential planar regions in a video sequence. Local search windows are then placed around these predicted locations in order to find the actual subpixel corner positions. Experimental results show the robustness of the corner tracking system with respect to occlusion, scale, orientation, and lighting.

Keywords: Corner tracking, pattern tracking, blob finding, robust, computer vision, augmented reality.

1 Introduction

Unlike virtual reality, which encompasses a user in a completely computer-generated environment, augmented reality is a technology that attempts to enhance a user's view of the real environment by adding virtual objects, such as text, 2D images, or 3D models, to the display in a realistic manner.

Clearly, the realism that a user will experience in an augmented reality environment is directly related to the stability of the registration between the virtual and real-world objects; if the virtual objects shift or jitter, the effectiveness of the augmentation is lost.

Vision-based augmented reality systems rely on extremely accurate optical trackers in order to obtain the required registration stability. Additionally, these accurate trackers must operate in real-time.

One of the most promising vision-based augmented reality techniques involves tracking a planar pattern in real-time and then augmenting virtual objects on top of the pattern based on its pose. In [5, 7, 9, 10, 14], black and white planar patterns are tracked resulting in relatively stable registrations, but the tracking algorithms fail to provide any robustness to partial pattern occlusions. Specially arranged coloured blobs are tracked in [8] that can handle partial occlusions for a brief period of time via Kalman filtering, but the blob centroids are less reliable at different scales or plane orientations. Other techniques address robustness and occlusion, but only in hybrid configurations involving expensive magnetic or inertial trackers and stereo configurations [1, 2, 12].

In this paper we describe the implementation of an optical corner tracker for an augmented reality system that is precise, fast, and robust, and which can be implemented using a standard, consumer-level camera and PC. In Section 2 planar homographies are first reviewed since they form the mathematical core of the 2D tracker. Section 3 then provides a description of a fast and reliable region detector that allows the system to self-identify predetermined planar patterns consisting of black and white corners. Section 3.2 then proposes an accurate corner tracker which uses a robustly computed homography to predict corner positions that are then refined using localized search windows. Experimental results are then presented in Section 4, which show the tracker's stability and robustness to occlusion, scale, orientation, and lighting changes. Additionally, a comparison between corner tracking and commonly used blob tracking techniques is made.

2 Planar Homographies

Tracking planar patterns are advantageous since they define a convenient world coordinate space that can be used for augmentations (Figure 1).

Since we will be tracking such planar patterns, we can make the assumption that they are located on the Z=0 plane in world (pattern) coordinates. Thus we can associate our original pattern with the viewed pattern in a video frame by a 2D-to-2D projective warp. In other words, if $\mathbf{x} = (x, y, 1)$ is a homogeneous coordinate in pattern space, and $\mathbf{x}' = (x', y', 1)$ is the associated coordinate in image space, $\mathbf{x} \leftrightarrow \mathbf{x}'$ defines a correspondence that is related by