

Virtual Postcards:

An Augmented Reality System for Advertising and Entertainment

1.0 Overview

Virtual Postcards is a real-time augmented reality system implemented using dedicated hardware. Augmented reality has been a growing research field interested in supplementing visual input from the world rather than creating it. AR systems can involve object detection and labeling, motion tracking, and real time special effects. Recently, AR has been adopted by the commercial sports broadcasting industry to superimpose “virtual” 10-yard lines on the football field, replace blank billboards with dynamic advertisements in NASCAR racing, and highlight the always difficult to see puck in hockey.

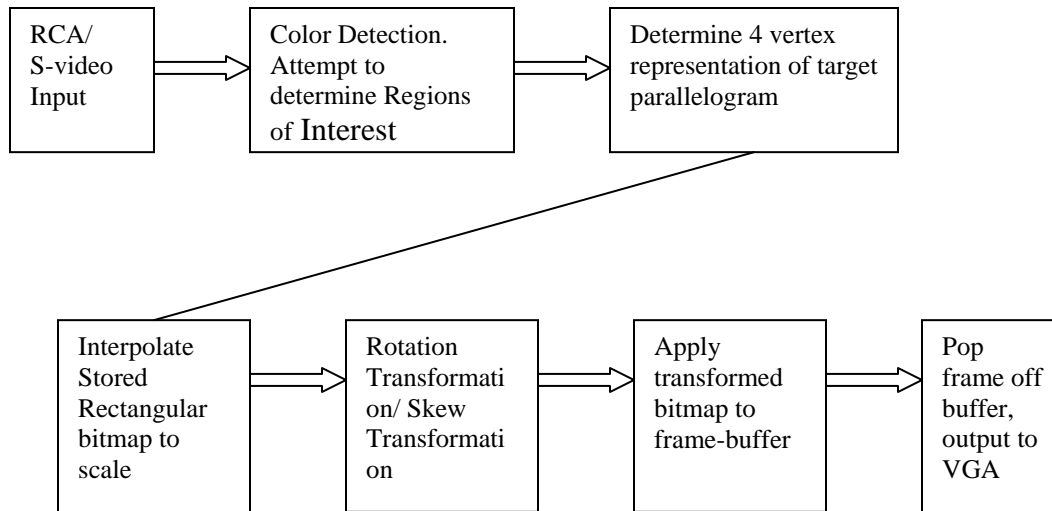
We intend to build a system in a similar vein to these commercial products, but engage our own creative solutions to many of the engineering challenges related to augmented reality. Our system should be able to identify blank index cards in a real video feed and superimpose arbitrary media onto them. An ideal instantiation of our system should be accurately estimate the pose of the postcard “blanks” regardless of lighting or background conditions. This means that our system must scan real images, detect blank postcards, estimate their rotation and skew relative to the camera, appropriately transform stored media to conform to the detected pose, and finally superimpose the media onto the video feed.

This ideal construction is certainly an ambitious task. Therefore, we have designed our project to be fully modular in both function as well as implementation. At the very least, we would like our system to be capable of detecting a brightly outlined rectangle over a dark background and filling it in with arbitrary color. At the very most, our system will be robust to light and dark backgrounds, capable of handling simple uncolored index cards (as ‘blanks’), and elegantly handle occlusion between multiple cards. A crowning feature of the ideal system may be to superimpose simple animated advertisements rather than simple bitmaps onto the cards.

2.0 Components

The core of our design is the minimalist implementation described above. In this case, we require a pipelined system which outputs one frame of augmented video data per one frame inputted from the camera with acceptable latency. The input stage must first interface with s-video or RCA based video input. Next, our base framework assumes only a single index blank is present and attempts to identify its corners based on simple logic and color. With this information (rectangle vertices) the output stage takes over. A solid block of color is drawn onto the frame buffer to fill the parallelogram formed by the

vertices specified in the first stage. A more sophisticated system would attempt to rotate, skew, and scale a simple bitmap to fit the desired parallelogram. The final output module generates a VGA signal whose frames appear exactly as the camera captured the scene with the addition of our superimposed color or image.



3.0 Evolution

We feel that starting with a simple instantiation is ideal, however we anticipate being able to create a much more effective system than the one described above. We feel some realistic goals for the project should be to effectively approximate index-card pose in real images with real backgrounds (not bright colors or dark backgrounds) and handle multiple virtual postcards at once. This latter goal comes with the extreme complication introduced by occlusion. Our system must then be able to handle depth of cards, and maintain robust pose estimation despite significant occlusion. This will likely be the most challenging aspect of the design of our more advanced system.

Our ambitious functionality ideas for the system are clearly not represented by the simple diagram above. We do have several ideas which should be useful first approaches to tackling robust pose estimation and occlusion. Including high pass or other filtering stages may aid in detecting card edges. A robust mechanism for pose estimation might begin with color-based region of interest filtering (ROI, this means to only consider off-white regions of the image to be viable postcards). We might continue with simple edge detection, then corner detection. Finally, grouping opposing corners might allow us to approximate parallelogram vertices. To handle occlusion, postcards might have some kind of “instantiation protocol” such as holding up a postcard so that it is clearly identifiable and pressing a button. From here, postcards are tracked rather than detected at each step. That is, at each step we need not detect pose and placement of an unknown number of postcards, rather we need to determine the change in pose of N known postcards in the field. While conventional vision mechanisms such as optical flow could

be used for computing such a delta, our restriction to dealing with simple rectangles should allow us to design something far simpler.