

Virtual Billboard

CS 231M Spring 2015 -- Stanford University

Jonathan Noyola (jnoyola@stanford.edu), Sean Stanko (ststanko@stanford.edu)

Abstract

Augmented Reality (AR) has recently become a popular advertising tool; in this vein this paper examines a method to display augmented flyers on a public bulletin board using generalized visual markers. This method differs from existing image and QR-based augmentation methods in its simple marker design and detection methodology, which allow for easy image scaling and efficient mobile computation. In addition, the implementation and usage in augmentation of a simple proportion-based 3D estimation is explored. The methods used are able to quickly detect and track valid sets of markers and produce an augmentation with a realistic 3D effect from a readable distance in real time.

1 Introduction

With the widespread adoption of smartphones, tablets, and other camera-equipped mobile devices, augmented reality has seen a boost in the public eye, in applications ranging from gaming to medicine. The field of advertising has arguably been the most affected by the adoption of this technology, as the ability to enhance physical advertising campaigns with a virtual presence enables companies and individuals to present their message in new, clever and impactful ways. [1]

In this vein, Virtual Billboard looks to provide a framework for small groups, clubs, and individuals to take advantage of AR in their promotions. On university campuses and office buildings, the public notice board is a common place for these groups to promote events, spread ideas, and waste copious amounts of paper. Virtual billboard extends on this idea by creating reusable virtual “flyers” that can be posted on this board, allowing small advertisers to both save paper and extend their message with creative digital and 3D effects.

2.1 Previous Work

Displaying augmented images is generally done via two methods: detecting an existing physical image and augmenting it with virtual effects, or identifying computer specific markers and replacing them with the augmentation.

For many artists and advertisers currently using augmented reality, the former method is common. This approach can aid computer identification by giving the platform a distinct and feature rich image to detect and track, and can yield impressive multi-media results by

interplaying with the physical display. As an example, the 2014 NO AD art application detects New York subway advertisements and replaces them with gallery art. [2] While these applications can be impactful, their nature of being tied to an existing image database limits their widespread use; in the case of NO AD, for example, the need to compare each candidate match against a constantly changing database of advertisements to replace restricts it the New York City area due to size and computation constraints.

The second method, which we emulate, is to use a specialized computer-recognizable image for detection and tracking. While these markings do not have the same effect on viewers as full physical and mixed media images, their more generic and oftentimes high-contrast shapes are easier to generalize than image-specific applications. A common form of machine detection mark is the QR code, a small matrix of black-and white squares used both to encode information and as a detection marker for AR displays.

2.2 Our Contribution

Virtual Billboard provides a generic framework for displaying augmented images; it is compatible with any possible input image and as such provides a lightweight counterpart to image-specific AR applications, which are constrained by the nature of their target image.

In the realm of marker-based augmented reality, Virtual Billboard presents a scalable alternative to QR-code based augmentation. Due to their complex design QR codes can be difficult to read at a distance; [3] Virtual Billboard utilizes a simpler and easier-to-scale design to detect markers at a comfortable reading distance from post boards. Additionally, Virtual Billboard's distinct, generalized markings are improvements upon QR codes by being less obvious (especially for large advertisements) and by being less susceptible to occlusion.

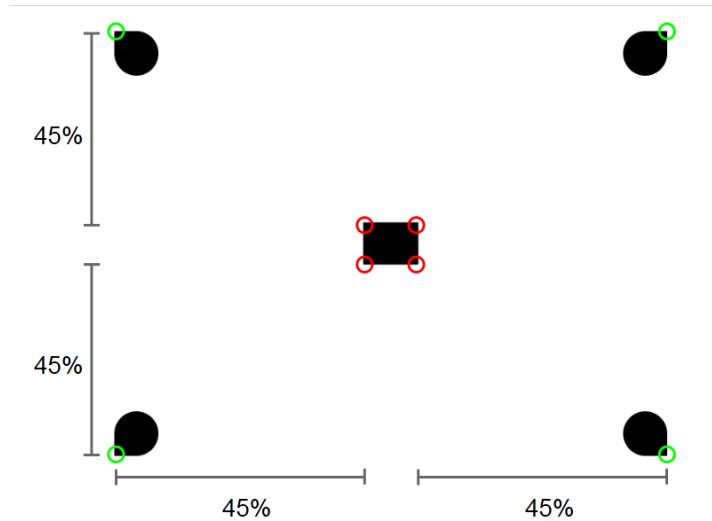
3.1 Technical Summary

Virtual Billboard operates as a three-step detection/tracking/augmentation pipeline. In the detection step the application uses a Harris corner detector to locate and fit keypoints in the detection marker. The tracking stage uses a Kanade-Lucas-Tomasi tracker to then follow the detected keypoints a homography, which is used to compute and apply a simple linear augmentation. As an additional augmentation step, the application can also extract the relative dimensions of the detection marker to find its orientation and spoof a simple 3D on the displayed image.

All three steps--detection, tracking, and augmentation--are run directly on the mobile device, so lightweight and fast operations are preferred in the implementation. Notably, the detection phase is performed without costly filtering operations, and the 3D augmentation is approximated from the marker orientation instead of explicitly calculated. By performing these optimizations, mobile load is reduced and Virtual Billboard is able to run more efficiently.

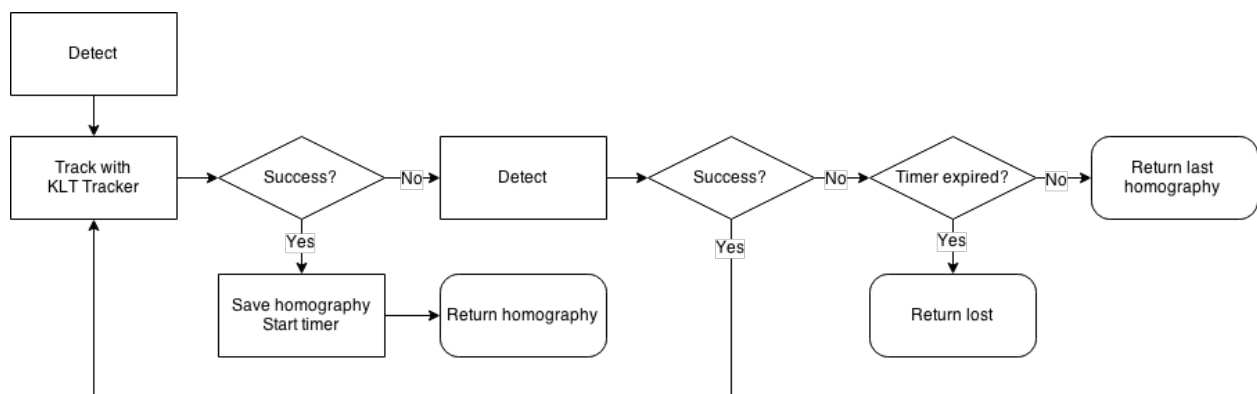
3.2 Technical Implementation

3.2.1 Detection



In the detection phase, a Harris detector is used to extract all corners in the frame. The corner detector has been modified for these purposes to not consider points in clusters of more than four. Once the corners are extracted, the application iterates over all sets of four corners at least a specified distance apart and computes a homography between them and the unit square. Once once a homography between four points (green) is computed, a series of test points (red) is fed into the homography; if their transformed location lies within a small tolerance of detected corners, all eight points are saved and the marker is considered “found.” If not, the homography is rejected and the application continues to iterate over point pairs. While this brute force procedure could potentially take many iterations, the correct set of corners are quite distinct and spread apart, so increasing the corner quality and minimum distance can greatly reduce the search space.

3.2.2 Tracking

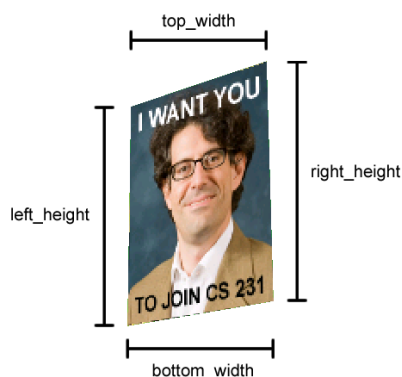


Once the application detects a valid marker, it stores all eight points used in detection for tracking. Tracking is performed with a standard KLT optical flow tracker as described in the above flowchart. [4][5]

In the case that the tracker loses the marker, the app remembers the last location and orientation of the augmented image and displays it there for up to one second while the detector attempts to reacquire the marker. If the marker is rediscovered in this window, tracking resumes as usual; otherwise the display is cleared and detection continues to loop until a marker is discovered.

Memorizing and displaying the last known location gives Virtual Billboard the ability to gracefully handle the temporary occlusion of corners or loss of tracking due to motion blur; as the application is displaying virtual flyers, the user can be expected to be (mostly) stationary while reading the image, and small errors in displayed location from using a stored image are preferable to the flyer flickering in and out due to tracking errors.

3.2.3 Pseudo-3D augmentation



$$x_offset = depth \times \frac{(right_height - left_height)}{\max(right_height, left_height)}$$

$$y_offset = depth \times \frac{(bottom_width - top_width)}{\max(bottom_width, top_width)}$$



Under standard circumstances, the outermost detected keypoints are sorted by location in the frame to determine orientation, and a homography between them and the target image is computed. After this, simple masking and addition is used to overlay the desired image over the marker in the frame.

In addition to this, Virtual Billboard supports the display of layered three-dimensional images. In this feature, the length of each edge of the marker (top, left, bottom, and right) is computed and the relative size difference between opposing edges is used to approximate the orientation of the marker. Using this orientation approximation, the display image is then

translated along a supplied “depth” value from its initial transform. To further improve this illusion of depth, a “drop shadow” effect is placed behind the image.

While the above calculations do not actually yield the 3D orientation of the display, they do produce a visually believable approximation of it for these flat objects, without the need of expending computational resources on expensive 3D operations.

3.2.4 Mobile Optimizations

As touched on in the above sections, several optimizations have been made to each phase of the augmentation pipeline in order to improve Virtual Billboard’s mobile-specific runtime.

In the detection phase, a slightly restrictive distance threshold is enforced in order to reduce the number of candidate homographies to iterate over. In corner detection, minimum distance and cluster sizes are utilized to reduce the number of points considered. Additionally, detection is only run once per four frames in order to reduce the number of costly calculations per second. This design choice is supported by the memorization step in the tracker, which maintains the display through the delayed detection phase. Orientation estimation, as used in the pseudo-3D step, leverages existing metrics already calculated for 2D augmentation and requires minimal additional calculation, providing a notable boost in performance and simplicity as compared to homography decomposition or other true 3D calculations.

4 Experimentation

Tracking & Detection

A series of trials simulating real-world conditions the app might face was performed to judge Virtual Billboard’s ability to detect, track, and render over the AR detection marks. In all cases a 8.5”x11” printed detection symbol was placed on a communal bulletin board or wall, with additional pre-existing flyers adjusted in order to frame performance tests. The application itself was run on a Nvidia SHIELD tablet running Android 5.0. Trials under each condition were run several times, with the average performance reported.

Trial A: Single image



Detection range	Tracking Range	Detection time (average)
38 inches	115 inches	48ms

As an initial performance test, Virtual Billboard was tasked with detecting and tracking a single marker on an otherwise empty bulletin board. Under these isolated conditions the application performed very well, picking up the image quickly and tracking it beyond readable distance.

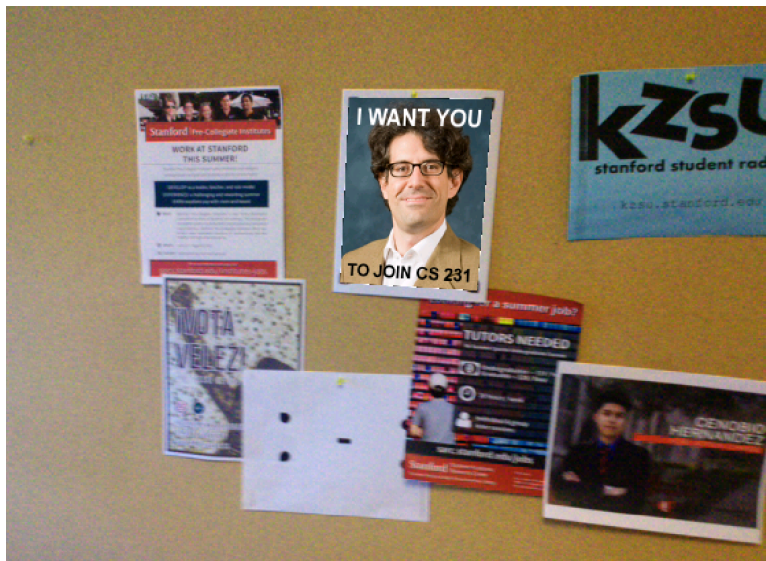
Trial B: Overlap



Detection range	Tracking Range	Detection time (average)
35 inches	95 inches	54ms

As a primary robustness test, the app was then tasked to identify a marker partially covered by other flyers. So long as the five detection markers were not occluded, detection quality and length were both mostly unaffected. Augmentation similarly proceeded unaffected, neatly overlaying the augmented image over the offending flyers.

Trial C: Crowded board



Detection range	Tracking Range	Detection time (average)
35 inches	100 inches	145ms

Building on the above trial, the bulletin board was then fully populated and the trial repeated. Detection and tracking performance were largely unaffected, as with the overlap trial, although the additional corners did cause a noticeable drop in framerate during the detection period.

Performance of additional Features

Orientation & Scaling



Under all conditions, the application was able to correctly identify the top, bottom, left, and right edges of the marker and use this information to properly scale and display the augmented image. If the marker was rotated during tracking the application would successfully track this shift and, if the marker was rotated ninety degrees, the tracker immediately flipped the image to respond to the new top edge.

Pseudo-3D Projection

Thanks to the simplified 3D approximation, the addition of 3D effects had little to no effect on the runtime or quality of tracking. The 3D layering output was very overall believable, with the shifting between layers occurring at rates visually consistent with their simulated depths.

Mobile Performance

Under most testing conditions, the mobile performance was rather good. Tracking maintained a steady 15-20 frames per second, as did detection for simple cases (the crowded board tests dropped to about ten frames during detection, but this effect was mitigated somewhat by only detecting every four frames). Mobile camera blur and artifacts had some effect on the quality of tracking, although the memorization display location usually helped mask this effect long enough for the frame to stabilize and the detector to recover.

5 Conclusions

For the described use of Virtual Billboard--displaying augmented flyers on public bulletin boards--this application performs very well. It is able to detect and track the marker throughout most angles and distances one can expect to view a notice board from, even if the board is full of other flyers. The application's ability to handle variations in scale and rotation make it quite robust.

The 3D estimation performed very well and produced believable results throughout testing. While not as applicable as traditional methods--it requires a planar shape, and can only

augment planar overlays, instead of full 3D objects--it provides a very lightweight method for applying a sense of depth to simple image and text overlays.

In conclusion, Virtual Billboard and member routines provides a robust platform for displaying flyer size augmentations. While it lacks the interactivity of QR codes or mixed media augmentations, it makes up this deficit in its generality and detection range. When synced with a database of flyers, Virtual Billboard has the potential to be an efficient way for individuals and groups to communicate data in an augmented world.

6 References

[1] Russell, Mallory. "11 Amazing Augmented Reality Ads." Business Insider. Business Insider, Inc, 28 Jan. 2012.

[2] Romano, Jowy. "NO AD." NO AD. <http://noad-app.com/>

[3] "QR Codes Scanning Distance." Qrworld. 16 July 2011.

[4] Bruce D. Lucas and Takeo Kanade. An Iterative Image Registration Technique with an Application to Stereo Vision. International Joint Conference on Artificial Intelligence, pages 674-679, 1981.

[5] Carlo Tomasi and Takeo Kanade. Detection and Tracking of Point Features. Carnegie Mellon University Technical Report CMU-CS-91-132, April 1991.