

Fast Content Placement and Alignment in 3D Scenes

Robin Wiethüchter, Saikishore Kalloori, Ribin Chalumattu

Media Technology Center, ETH Zurich
Zürich, Switzerland

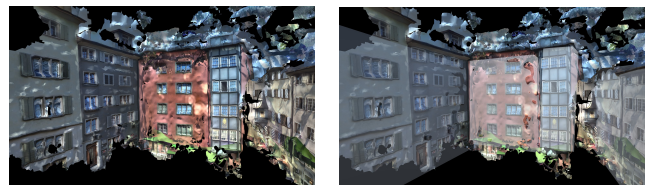
robin.wiethuechter@inf.ethz.ch, ssaikishore@ethz.ch, cribin@inf.ethz.ch

Abstract

We present a system for fast content placement on flat surfaces in 3D scenes that leverages plane detection and computer vision. For content creators, placing and aligning objects is often a time-consuming process. Especially with AR, this problem becomes more prevalent as virtual objects should be aligned to the real world. For certain objects that should be aligned to planar regions, e.g., a virtual painting being placed on a real wall, the scene (wall) implies the rotation of the object. However, placing a virtual painting usually requires manually modifying its 3D coordinates and manually rotating it to exactly match the wall.

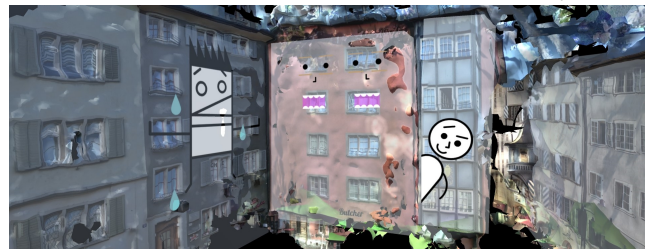
In this work, we propose the use of plane detection to speed up the placement and alignment process for every object whose relation to surfaces has been defined, e.g., images or posters. We do this by first detecting approximately planar regions in our scene and calculating planes for them. Then to place an object, one can simply click a point in the scene to select the location and with a raycast, the actual point in the 3d space is found. The object is then automatically positioned and aligned by moving it to the closest point on the plane of the region and matching the plane's rotation.

Furthermore, we apply object detection and segmentation to the textured planes to gain a semantic understanding of the scene. This enables automatic placement suggestions and assistance when aligning objects to visual markers, e.g., snapping the corners of a virtual painting to those of a real window in AR. The code of our system is available on GitHub¹.



(a) 3D reconstruction of an outdoor scene

(b) Scene with detected planes (shown as white overlays)



(c) Scene with planes and content

Figure 1: Content is placed by choosing a single position on one of the detected plane. The content is then automatically moved to that position and aligned with the plane using its rotation. This workflow allows placing content on flat surfaces with a single click instead of manually modifying the position and rotation of objects.



(a) Window detection applied to textures of the scene

(b) Mapping of detected window markers to the 3D scene

Figure 2: Visual features like windows can be detected in the textures of a scene using object detection models. Mapping this information to the 3D scene enables snapping and aligning content to these visual features as well as automatic semantic placement suggestions, e.g. placing a virtual image on a wall between two windows.