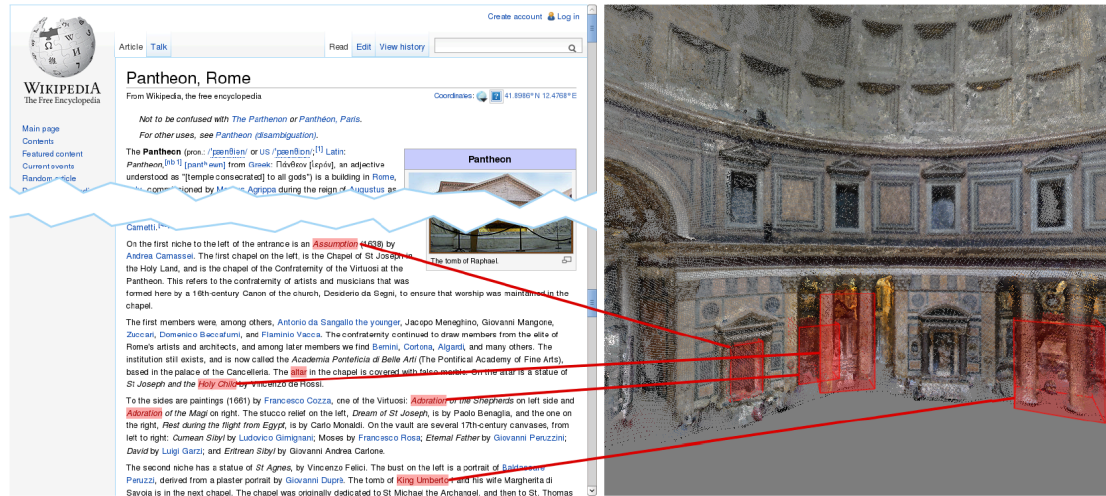


# 3D Wikipedia: Using online text to automatically label and navigate reconstructed geometry

Bryan C. Russell    Ricardo Martin-Brualla    Daniel J. Butler    Steven M. Seitz  
Luke Zettlemoyer



Several websites like Wikipedia provide users with historical facts and descriptions of popular sites. However, they rely on static images embedded in the text. The purpose of the presented system in this paper is to provide an interactive photorealistic 3D visualization of popular sites where the user can visualize the contents of the scene and their spatial relationships in 3D. The goal is to create such visualization completely automatically by analyzing the Wikipedia page itself together with many photos of the sites that are available on the Internet.

Automatically creating such a visualization is a challenging process. First, the system should be able to automatically extract the names of objects included in the scene from the Wikipedia page. The second challenge is to exactly determine the 3D location of each object within the scene. By addressing these two challenges the system would be able to link each object in the scene with its corresponding description in the reference text.

First the system downloads images of the targeted site from Flickr to automatically reconstruct a 3D model of the site. The idea of the proposed approach is to mine the co-occurrences of the photo and the text across all of the Internet. So the solution to the challenge of finding object names is to extract every noun phrase from the reference text because primarily noun phrases are used to name physical objects in English.

After obtaining the list of candidate named objects, the system search for and download images for each candidate object. Image search engine can perfectly yield cropped images of the desired object. If any of the returned photos matches the 3D model, the corresponding region is linked to the named object.

However, image searching can be noisy and multiple matching images can be produced in which they don't describe the actual object. Therefore, a linear classifier trained over the text and the visual features to filter false positive matches.

Generally, I think this system provided some impressive results despite of different sources of errors that lead to inaccurate labels of objects in some cases.

### **Reference**

Russell, Bryan C., et al. "3D Wikipedia: using online text to automatically label and navigate reconstructed geometry." *ACM Trans. Graph.* 32.6 (2013): 193.