Alex Wang

CS283 Computer Vision

Final Project Proposal

I am interested in iterating on the paper “Rendering Synthetic Objects into Legacy Photographs” (Karsh et al., 2011). In this paper, Karsh describes a process of inserting a 3D object into a 2D image by estimating the geometry (camera, surfaces, etc.) and lighting (internal and external) of the scene in order to create realistic composite images. The process described requires some user annotations of the photo provided, but does not require access to the scene itself. In later a later paper, “Automatic Scene Inference for 3D Object Compositing” (Karsh et al., 2014), Karsh enhances this pipeline to automatically insert the object at a specified location, without any additional user input.

A limitation, though, of both of Karsh’s papers is that they only look at single images. A natural extension, and what I am proposing is to apply this process to video. The added benefit of inserting into video is that successive frames can be used to more robustly estimate the geometry and lighting of the scene. My project then would involve implementing Karsh’s algorithm for a single scene, sampling frames from a video, estimating the geometry and lighting for each frame from successive frames, and then combining the frames in order to create a composite video with the inserted object.

Note: I am confident that this falls under the purview of the class, however I am not entirely sure how reasonable this proposal is for a final project. Therefore, any feedback regarding ways to break up the project into reasonable benchmarks to keep the project manageable would be greatly appreciated.

References

Karsch, Kevin, Kalyan Sunkavalli, Sunil Hadap, Nathan Carr, Hailin Jin, Rafael Fonte, Michael Sittig, and David Forsyth. "Automatic Scene Inference for 3D Object Compositing." *TOG ACM Trans. Graph. ACM Transactions on Graphics* 33.3 (2014): 1-15. Web.

Karsch, Kevin, Varsha Hedau, David Forsyth, and Derek Hoiem. "Rendering Synthetic Objects into Legacy Photographs." *TOG ACM Trans. Graph. ACM Transactions on Graphics* 30.6 (2011): 1. Web.