# DGTk: A Data Generation Tool for CV and IBR

V.Vamsi Krishna

August 14, 2006

Computer Vision (CV) and Image Based Rendering (IBR) are the fields which have emerged in search of a means to make the computers understand the images like humans and the never ending pursuit of the Computer Graphics community to achieve photo realistic rendering. Though each of these fields deal with a completely different problems, both CV and IBR algorithms require high quality ground-truth information about the scenes they are applied on.

Traditionally research groups have spent large amounts of resources on creating data using high-resolution equipment for qualitative analysis of CV and IBR algorithms. Such high quality data provided a platform for comparison of CV and IBR algorithms. Though these datasets have enabled comparison of algorithms, during the past decade, the development in the fields of CV and IBR have outpaced the ability of such standard datasets to differentiate among the best performing algorithms. All the resources invested for generating these datasets become wasted. To overcome this problem, researchers have resorted to creating synthetic datasets by extending existing 3D authoring tools, developing stand alone tools for generating synthetic data and developing novel methods of data acquisition for acquiring high quality real world data. The disadvantage of acquiring data using high resolution equipment include (1) Time required for setting up the configuration of equipment, (2) Errors in measuring devices due to physical limitations, (3) Repeatability of experiments due to un-controllable parameters like wind, fog, rain etc. Synthetic data is preferred for the early testing of algorithms, since they make qualitative and quantitative analysis possible. The performance of an algorithm on synthetic data generally provides a good indication of it's performance on the real world data.

Synthetic data can be generated using some standard 3D authoring tools which provide functionality for generating high quality, high-resolution images of the 3D scenes created using these tools. Some of these tools (3DSMax, Maya, Blender, etc) provide the user with the flexibility of creating extensions through scripting support. It may not always be possible for extending the current tools to generate novel representations of data like LDI, Alpha-map, Object-maps etc due to lack of information about the internal representation of data in such tools. Though open source tools like Blender are available, the user may have to develop very unintuitive extensions for these tools to generating the required

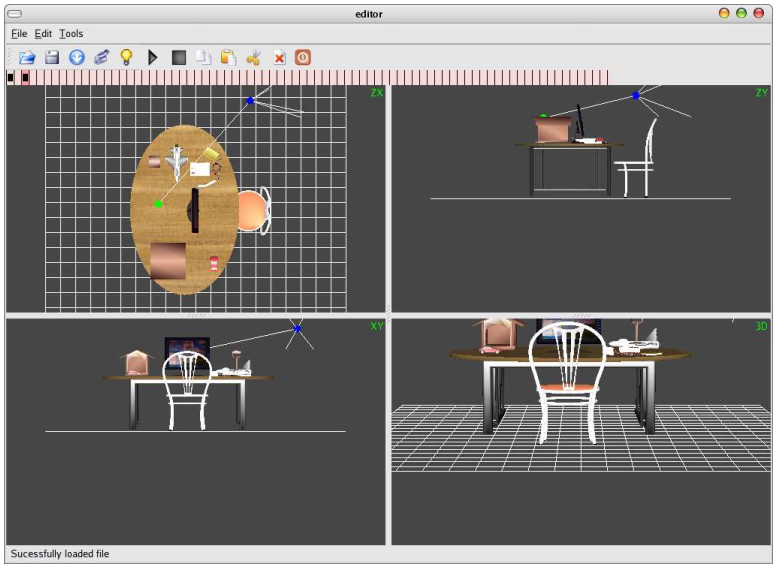


Figure 1: The screen shot of the GUI of DGTk

representation. For generating synthetic data with such novel representations, researchers generally develop their own tools. These tools are generally cannot be easily extended for supporting other representations and are difficult to use. A generic standalone open source tool which is flexible and versatile is desirable for the research community. These are the motivating factors for this thesis.

We have developed a standalone, light weight tool which enables users to generate a wide variety of representations of a synthetic 3D scene. The major contribution of our work is in the development of a unique tool which is computer vision aware. It can generate Depth maps, Layer Depth Images, Object maps, Corresponding points, Alpha maps, High-resolution images, and High-Quality images which are very important for CV and IBR algorithms. The tool not only enables the researchers to generate high quality data with ease, but also provides a method of sharing the data among researchers. The GUI of our tool (Figure 1) is similar to the that used by most of the 3D authoring tools this makes the learning curve for our tool very low. Users can import 3D models in standard formats like AC3D, 3DS, MD2 and POV-ray and place them at required positions to create complex scenes which mimic those in the real world. The user can place any number of cameras, lights and objects in a scene and generate the required representation. The tool provides a very easy to use interface for generating dynamic scenes using some information provided by the user in the form of key frames. The user can manipulate the cameras and lights just like any other objects imported into the tool. This enables the users to create dynamic scenes involving camera motion with ease. The advantage of using our

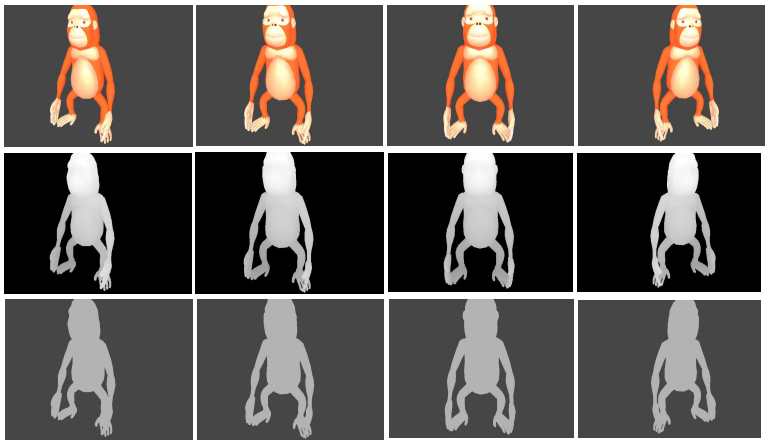

Figure 2: shows three strips each showing a different representation of a dynamic scene. The first strip is the actual scene image, the second strip of images is the depth map representation of the same corresponding scene images. The pixels close to the camera are represented using brighter colors in the depth map images. The third strip of images is the Object maps.Each object is given a unique id in terms of R,G and B values. In Object maps assign we a unique id to each pixel in the image based on the object they correspond to.

tool for generating dynamic scenes is that it can generate all the representations for each frame of the dynamic sequence (i.e, Depth maps, Alpha maps, Object maps, etc along with the actual dynamic scene (Figure 2)) based on the selection of some check boxes provided in the GUI. This would be for first time where both Object maps and Alpha maps are generated as ground truth information. These representations are very useful for the algorithms which try to estimate silhouette of objects and object matting. Our tool enables the user to generate high-quality images by generating scene description files for POV-Ray (a ray tracing software). The ASCII scene representation used by our tool helps the researchers to share the camera and light configurations used for generating a data which was not possible before. This feature enables sharing of not only the data but the means of generating the data.