

# Reconstructing 3D Structure from Amazon Product Images

Jason van der Merwe  
jasonvdm@stanford.edu

Andrew Giel  
agiel@stanford.edu

Bridge Eimon  
beimon@stanford.edu

CS 229, CS 221, Stanford University  
December 8, 2013

---

## Abstract

Constructing 3D structure from two dimensional images usually requires many images, covering a variety of angles of the object you are trying to reconstruct. However, in this case, we only had a few images, usually only two.

## Introduction

This project, reconstructing 3D models from Amazon product images, is the first step of a larger goal to build a database of 3D objects by crawling the web. This database can then be used to assist in many other object detection tasks, such as identifying objects in your home or from a cell phone camera. One important component of making such a system is to build 3D object models from multiple views of web images. SFM is generally performed using a number of images carefully taken of the same object. Can we use SFM from product images taken by different users? Amazon often has a collection of images for each object uploaded by different users. These views were all taken with different cameras in different lighting / backgrounds, which adds a challenge. Can we combine these images to create 3D object models?

## Data

To begin, we acquired a text file representation of every title in IMDBs database. Then, we queried the RottenTomatoes API with these titles (limiting our search to movies from 2000 and later due to a lack of informa-

tion available for earlier movies). Each data entry included a movie id, the director, title, genres, rating, runtime, release date, actors, studio and RottenTomatoes url. Our final dataset contained 3456 data points, ranging from movies which grossed \$181 to \$761 million. After plotting the data points box office gross, we noticed a nearly exponential scale (figure 1). So, we decided to split our data labels into 5 logarithmic scale buckets to linearize the data.

## 3.1 Technical Summary

Before we began using Amazon product images, we began with a proof of concept using photos of objects that we took ourselves. Once we took photos of these images in a controlled environment, we began implementing the technical aspects of the project. First, we manually selected correspondence pairs and then computed SFM using the Tomasi-Kanade factorization algorithm. We also computed the fundamental matrix using the normalized 8-point algorithm. After we were happy with these results, we moved to automatically retrieving correspondence pairs. We implemented SIFT, SURF and ORB and compared the results. We ran all matches through RANSAC, experimenting with dif-

ferent thresholds. Plotting these points in a 3D structure of the product.  
three dimensional space showed us the recov-