

Geometrical Dimensional Analysis of Cuboidal Objects using Monocular Vision

Priyanka Sherkhane¹, Archit Konde², Saakshi Deokar³, Priyanka Gate⁴, Rugved Shinde⁵

¹Assistant Professor, Department of Computer Engineering, University of Mumbai Affiliated Institute Terna Engineering College, Mumbai, Maharashtra, India

²⁻⁵Undergraduate Students, Department of Computer Engineering, University of Mumbai Affiliated Institute Terna Engineering College, Mumbai, Maharashtra, India

Abstract — We explore methods to calculate the dimensions of a cuboid from a 2D image. Understanding the 3D representation of the world from a single monocular image is an important problem in computer vision which is why our project Geometrical Dimensional Analysis of Cuboidal Objects plays a crucial role today in a lot of vision applications. The project aims at describing a method to get accurate measurements of the dimensions of a cuboid and thus its volume from a single image.

Keywords — Monocular vision, Dimensional Analysis, Cuboids, Single View Metrology, 3D measurements, Computer Vision, Homography,

I. INTRODUCTION

While our human vision system is quite capable, it tends to be more qualitative than quantitative. It is not particularly good at making precise measurements of things in the physical world. Hence, vision systems can be designed to surpass the capability of human vision and extract information about the world that we humans simply cannot perceive.

Nowadays, smartphones come with superior-quality cameras. These cameras can be utilized for solving a lot of complex vision problems and simplifying a lot of tasks, reducing turnaround time and minimizing human errors. Consider clicking a photo of e.g., a box and getting an estimation of its size almost instantly. There is a whole field in science dedicated to the art of making measurements from images called photogrammetry. The field of photogrammetry has been extensively studied throughout the years and is commonly used for 3D modelling [1], [2] and improvements and new applications are continuously created. As computers and digital cameras have become ubiquitous and cheaper photogrammetry has become an economically sustainable and accurate measurement tool, which has led to many industrial applications [3], [4] as well as more specific applications.

To be able to perform measurements of objects in images we need to understand how the scene is reproduced by the camera. This depends on the traits of the camera, called the intrinsic parameters. These parameters are unknown unless a camera calibration is performed, which is a non-trivial task. [5]

Geometric camera calibration is used to estimate the parameters of a lens and image sensor of an image or video camera. One can use these parameters to correct measure the size of an object in world units. These tasks are used in various machine vision applications to detect and measure objects. Some of the prominent applications include robotics, for navigation systems, and 3-D scene reconstruction.

Another important factor that we need to take into consideration is a reference object, even humans need

references in images to be able to determine size. Giving the viewer something of known, constant size puts the object of interest in some context.



Figure 1. Banana for scale

The objective of the project will be to create a reliable user-friendly model which helps to calculate the length, width, height and thus the volume of the cuboidal objects.

II. LITERATURE SURVEY

TABLE I. Existing Systems

Author's name	Title and year of publication	Findings
Louise Lennartsson	Photogrammetric methods for calculating the dimensions of cuboids from images, 2015	This project uses a credit card as a reference object, which is placed on top of the cuboid of interest. The corners of the reference object are used to determine the dimensions of the cuboid.
Rahul Swaminathan, Robert Schleicher, Simon Burkard, Renato Agurto, Steven Koleczko	Happy Measure: Augmented Reality for Mobile Virtual Furnishing	The authors present a vision based augmented reality system called Happy Measure to facilitate the measurement, 3D modeling, and visualization of furniture and other objects using a smartphone or mobile device equipped with a camera.

III. METHODOLOGY

In this section we will discuss all the necessary components for understanding the relation between the real-world coordinates and image coordinates.

A. Homography

The relation between two images of the same planar surface is known as homography.

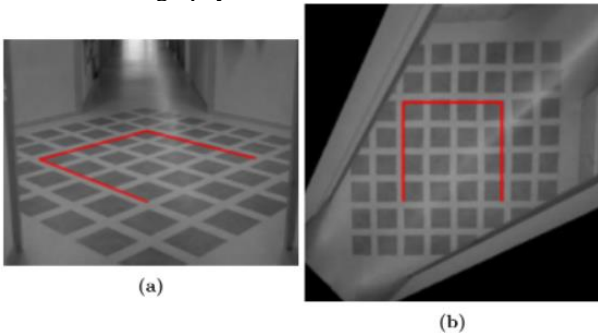


Figure 2. (a) Image of a checkered floor seen from an oblique angle. (b) The same image rectified using homography transformation. Images taken from ref. [5]

B. From 3D to 2D

Without lighting, there is no vision. And the light from various light sources falls on the scene and is scattered or reflected by the scene in many different directions. And a small fraction of this light is scattered in the direction of the camera, which plays the role of the human eye. The camera takes light from the three-dimensional scene to produce a two-dimensional image. The intrinsic and extrinsic camera matrices play a very vital role in the mapping of 3D world points to 2D image coordinates.

Mapping of world coordinates to image coordinates is given by,

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K[R|t] \begin{bmatrix} xw \\ yw \\ zw \\ 1 \end{bmatrix}$$

s is a scale factor, u and v are the image coordinates, R is the rotation matrix; t is the translation vector (R and t are extrinsic parameters of the camera) and xw, yw and zw are the points in the world coordinate frame.

C. Finding world coordinates from image coordinates

Expanding the matrix gives us the following equation

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} fx & 0 & cx \\ 0 & fy & cy \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r11 & r12 & r13 & t1 \\ r21 & r22 & r23 & t2 \\ r31 & r32 & r33 & t3 \end{bmatrix} \begin{bmatrix} xw \\ yw \\ zw \\ 1 \end{bmatrix}$$

Performing the matrix multiplications gives a system of three linear equations.



Figure 3. A cuboid with a reference object.

Now to find the dimensions of the cuboid we consider a reference object in the xy plane with the origin of the world coordinate system placed on one of the corners of the reference. Now we have two known image coordinates, u and v, and the four unknown variables s, xw, yw and zw. But as our points of interest are on the same plane we can substitute zw = 0 and now it is possible to find a unique solution for xw and yw. And thus calculating the value of zw.

Now having solved for values of the world coordinates we have the dimensions of the cuboid and consequently its volume.

IV. FUTURE SCOPE

We believe that this project can be used for solving an important vision problem and replace human efforts. Further work in this project can be a mobile app to let the user get the dimensions and volume of a cuboidal object seamlessly and instantly.

V. CONCLUSION

We have explored how the affine measurements of 3D objects may be recovered from 2D images with the help of a reference object and the intrinsic and extrinsic parameters of the camera.

VI. REFERENCES

- [1] H. Murat Yilmaz, Murat Yakar, and Ferruh Yildiz. Digital photogrammetry in obtaining of 3d model data of irregular small objects. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 37:125–130, 2008.
- [2] Surendra Pal Singh, Kamal Jain, and V.Ravibabu Mandla. A new approach towards image based virtual 3d city modeling by using close range photogrammetry. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, 1:329–337, 2014.
- [3] Thomas Luhmann. Close range photogrammetry for industrial applications. ISPRS Journal of Photogrammetry and Remote Sensing, 65(6):558–569, 2010.
- [4] Clive S Fraser. A resume of some industrial applications of photogrammetry. ISPRS Journal of Photogrammetry and Remote Sensing, 48(3):12–23, 1993.
- [5] Louise Lennartsson. Photogrammetric methods for calculating the dimensions of cuboids from images, 2015