IS71021E Assignment 2

An Interactive Exploration of Camera Projection

Quince Gore-Rodney[*], Damian Lau[*], and Oliver J. Martin[†]

Department of Computing
Goldsmiths, University of London

February 6, 2026

[*]MA/MSc Virtual and Augmented Reality
[†]MSc Computer Games

Abstract

This is a technical report detailing and analysing the development of a computer graphics program created in Unity which explains camera projection. The techniques used to create the application, and the underlying mathematical concepts will be cross referenced from academic sources and informational books. Additionally, the effectiveness of the interactive applications ability to intuitively educate users on the topic will be examined and constructively critiqued. Consequently, further improvements which to improve the experience are explored.

1 Objective

The aim of this project was to explain the principle of perspective through showcasing different types of cameras used in computer graphics. The variables of the camera's projection type and transformations can be adjusted, such that the relationship between the camera and the finished rasterized image result can be understood. The purpose of this is to visually demonstrate the relationship between the two visualisations, as to nurture long-term memorisation of the conceptual model so that it can be recalled in the future (Benyon, p.469-71, 2014). This report will elaborate on the methodology for how it was be executed.

2 Introduction

2.1 History

The earliest perspective projection cameras in history are known as pinhole cameras, which take advantage of the *camera obscura* (or pinhole image) phenomenon - an innovation which the rest of perspective geometry builds on. Despite the phenomenon being recorded across many civilizations dating back to antiquity, the 11th century Arabic mathematician Ibn al-Haytham is credited with being the first to provide a clear scientific and geometric description of the effect and of the practical applications of the pinhole camera – leading to him being widely considered its 'inventor' (Camera Obscura & World Illusions, 2026).Importantly, al-Haytham showed that the inverted the image of the viewing screen proved that light travelled in straights lines, as well as noting the relationship between the focal point and the size of the pinhole (Camera Obscura & World of Illusions Edinburgh, 2026).

During the European renaissance, pinhole images were used to sketch views of natural landscapes and cities. It is suggestible that the accuracy of the pinhole image allowed for accurate drawing of nature, not directly, but by observations. Artists like Samuel van Hoogstraeten to Leonardo DaVinci used the pinhole camera to understand perspective projection (Delsaute, p.113, 1998).

By approximately the end of the 16th Century, a portable camera obscura was built equipped with converging lenses. It could correct the inverted image to produce a direct representation of the perspective from the view of the user. Mathematician Johannes

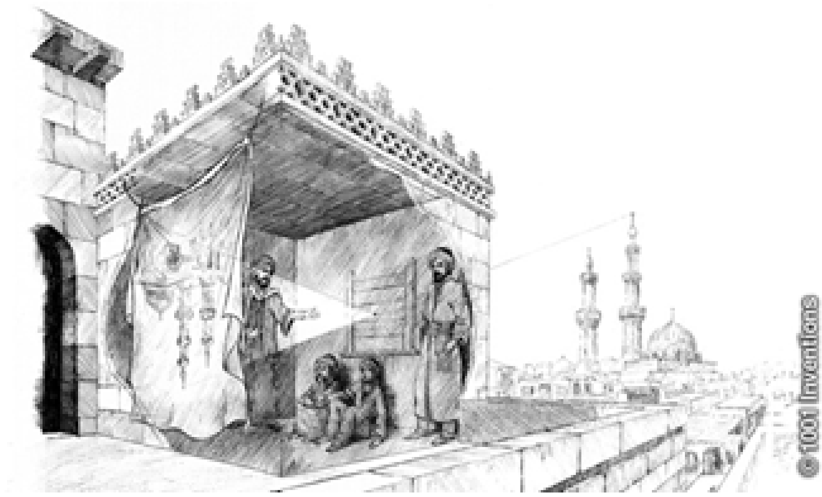

Figure 1: Visualisation of Ibn al-Haytham's camera obscura, comprising of a small hole in a dark room with a screen. (Camera Obscura & World of Illusions Edinburgh, 2026)

Kepler(1571-1630) insisted he only used it "as a mathematician, not as a painter." Similarly, through this style of visual communication, this project plans to teach users how to understand the principles of perspective with computer graphics.

2.2 Computer Graphics

The foundation of 3D computer graphics began with wireframe CAD simulations, Boeing flight simulations, and the construction of the perspective -generating algorithm by Larry G. Roberts (Manovich, L, 1993). Roberts introduced a 4x4 matrix, known as the perspective projection matrix, which serves to map points in 3D eye space to the 2D surface of the computer monitor. To do this, the viewing volume of the camera (in eye space) is transformed into the canonical view volume - a normalised 3D cube ranging from -1 to 1. For a given point in eye space $(x_{eye}, y_{eye}, z_{eye})$ and projection matrix P, this can be calculated as:

$$\begin{bmatrix} x_c \\ y_c \\ z_c \\ w_c \end{bmatrix} = P \begin{bmatrix} x_{eye} \\ y_{eye} \\ z_{eye} \\ w_{eye} = 1 \end{bmatrix} \tag{1}$$

Where (x_c, y_c, z_c, w_c) is the representation of the point in the homogenous 'clip' coordinate system. It is worth noting that the eye space point is implicitly converted into a homogenous point by assigning the w component to 1. In clip coordinates, it is efficient to determine whether a point is within the viewing volume - as a point is visible if $-w_c \leq x_c, y_c, z_c \leq w_c$. This point can then be transformed into normalised canonical

Figure 2: Samuel van Hoogstraten (Dordrecht, 1627-1678), Perspective of an Open Gallery ('The Tuscan Gallery').)

view volume coordinates through a process known as perspective division:

$$\begin{bmatrix} x_n \\ y_n \\ z_n \end{bmatrix} = P \begin{bmatrix} \frac{x_c}{w_c} \\ \frac{y_c}{w_c} \\ \frac{z_c}{w_c} \end{bmatrix} \tag{2}$$

Where (x_n, y_n, z_n) is the point in the normalised canonical view volume. x_n and y_n are the normalised screen coordinates, and can be converted to pixel coordinates by remapping them from the range (-1, 1) to the pixel resolution of the target screen (with width W and height H):

$$\begin{aligned} x_{screen} &= (x_n + 1) \cdot \frac{W}{2} \\ y_{screen} &= (y_n + 1) \cdot \frac{H}{2} \end{aligned} \tag{3}$$