

3D Model Retrieval Using Constructive Solid Geometry in Virtual Reality

Bachelor thesis

Natural Science Faculty of the University of Basel
Department of Mathematics and Computer Science
Database and Information Systems Research Group
<https://dbis.dmi.unibas.ch/>

Examiner: Prof. Dr. Heiko Schuldt
Supervisor: Ralph Gasser, MSc.

Samuel Börlin
samuel.boerlin@stud.unibas.ch
16-051-716

June 29th 2020

Acknowledgments

First and foremost I want to thank Prof. Dr. Heiko Schuldt for providing me the chance to realize this bachelor's thesis, as well as Ralph Gasser for suggesting the initial idea and for his continuous support as my supervisor.

I also want to thank Carlos Tejera and Kalthoum Nemmour for implementing the prototype precursor of this thesis' application with me during the semester of Fall 2019.

Furthermore, I extend my thanks to all the participants of the user study for taking their time to test and rate my application and for giving insightful feedback.

Last but not least, I'd like to thank my friends for helping me by proofreading the thesis.

Abstract

The Cineast multimedia retrieval system aims to provide a versatile system enabling users to browse and run targeted searches for various types of media. One such type of media supported by Cineast are 3D models. Cineast provides a query-by-example and query-by-sketch functionality where the user can provide an example 3D model, respectively draw a 2D sketch thereof, which it then uses to search its database, Cottontail DB, for similar models. However, this functionality of Cineast currently does not take the color or texture of the 3D models into consideration.

In this bachelor's thesis we present a novel 3D model retrieval system that makes use of virtual reality to enable the user to sculpt models in 3D using constructive solid geometry. These sculptures can then directly be used as a 3D sketch input for Cineast's 3D model similarity search, whose results are then shown in 3D in an interactive manner, all within the same virtual reality application. Furthermore, we add an extension to Cineast that improves the accuracy of its similarity search functionality for colored and textured models.

Table of Contents

Acknowledgments	ii
Abstract	iii
1 Introduction	1
1.1 Motivation	1
1.2 Approach	1
2 Concepts and Architecture	2
2.1 Multimedia Retrieval	2
2.2 REST API	3
2.3 3D Models	3
2.4 Voxels	4
2.4.1 Voxelize	4
2.4.2 Isosurface Polygonization	5
2.5 CSG	6
2.5.1 Signed Distance Functions	7
2.6 Virtual Reality	7
2.7 System Architecture	8
3 Related Work	10
3.1 3D Model Retrieval	10
3.2 Virtual Reality Sculpting	10
4 Implementation	12
4.1 Cineast	12
4.1.1 Core Changes	12
4.1.2 Model Formats	13
4.1.3 ClusterD2+Color Feature Extraction	13
4.1.4 Comparing Features For Similarity	16
4.1.5 C# Cineast API	17
4.2 Hermite Data Voxels	19
4.2.1 Voxel Storage	19
4.2.2 Cubical Marching Squares	21
4.2.2.1 Multi-material Extension for CMS	23
4.2.2.2 Lookup Table Based Implementation	25

Table of Contents	v
4.2.3 CSG Operations on a Voxel Grid	27
4.2.4 Rendering	29
4.2.5 Voxelizer	30
4.3 VR Sculpting	32
4.3.1 SteamVR Plugin	32
4.3.2 Sculpting Functionality	32
4.3.3 Sculpting Tools	34
4.3.4 VR User Interface System	37
4.3.5 Similarity Query	38
5 Evaluation	40
5.1 Technical Evaluation	40
5.1.1 ClusterD2+Color Performance	40
5.1.2 Voxel Polygonizer Performance	41
5.1.3 Voxelizer Performance	41
5.2 User Evaluation	41
5.2.1 Structure	41
5.2.2 Sample Size	42
5.2.3 Results Overview	42
5.2.4 Feedback	43
6 Discussion	48
6.1 Conclusion	48
6.2 Outlook	48
6.2.1 Performance	48
6.2.2 Usability	49
Bibliography	51
Appendix A Appendix	54
A.1 Example Images	54
A.2 User Evaluation Questionnaire	55
Declaration on Scientific Integrity	59

1

Introduction

Internet search engines like Google are part of our daily lives, so much so that the verb "to google" was added to the Oxford English Dictionary in 2006 and has been used commonly ever since. However, in the recent years interest in both on- and offline search engines for media other than text documents has increased a lot, and one such type of media are 3D models. Looking for a specific 3D model in a large collection is a laborious task that can be solved by multimedia retrieval applications capable of searching for 3D models not only by labels or title, but also by shape and looks. In this bachelor's thesis we present a novel way of searching databases for certain 3D models by making use of virtual reality to enable the user to sculpt 3D sketches that the Cineast multimedia retrieval system can then use as input.

1.1 Motivation

Multimedia retrieval applications presented in recent research use a drawing area where the user can draw a 2D sketch of the model they would like the search. This may work well for simple shapes but it does not allow the user to make full use of the 3D space. Furthermore, sculpting virtual models with keyboard and mouse is not intuitive due to the mismatch between the 3D nature of sculpting and the 2D movement of the mouse. Therefore, our goal is to build a novel system in virtual reality making use of its 3D motion tracking controllers that enable the user to sculpt 3D shapes or figures and then use them as a sketch to automatically search a database for similar 3D models.

1.2 Approach

During the semester of Fall 2019 C. Tejera, K. Nemmour and S. Börlin built a prototype of such an application in the Unity game engine for the Modern Human-Machine Interaction seminar at the University of Basel. We build upon this prototype by extending its functionality, improving its user interfaces and implementing a new feature module in Cineast in order to increase the accuracy of the database search for colored or textured 3D models.

2

Concepts and Architecture

In this chapter we discuss the relevant concepts used throughout this thesis and describe how they fit together into the high level system architecture of our application. First, it describes how our application makes use of a multimedia retrieval systems in order to manage a multimedia database and to be able to search such a database for certain 3D models. Then, it explains two representations of 3D models and why they are relevant to this thesis. Lastly, it describes how our application makes use of virtual reality and how our virtual reality sculpting and 3D model retrieval system combines all these topics.

2.1 Multimedia Retrieval

Multimedia retrieval has become an ever more important topic over the last decades. As the name implies, multimedia retrieval entails the retrieval of various different media such as text, images, video and audio but also 3D models. The goal is to allow the user to find such media based on an example input or a subset of the desired result, i.e. retrieve them from some multimedia database.

The DBIS research group [6] is working on such a system called Cineast [4]. Cineast provides the backend functionality for multimedia retrieval queries and Vitrivr NG provides a web based frontend to enable users to build queries. Currently it supports queries for text, images, video, audio and 3D models.

In order to run multimedia queries a multimedia database is required to store the data to compare media. Cineast currently supports the Cottontail DB [5] column store database for multimedia retrieval which is also being developed by the DBIS research group.

The focus of this thesis is on 3D model retrieval which is also part of multimedia retrieval. Research in this area has so far focused on building 3D model queries from example images, sketches or common 3D model file formats. In order to create one of these 3D model files the user requires a separate sculpting program and needs to learn how to use it.

Multimedia retrieval systems commonly rely on the Vector Space Model [28] in order to compare media for similarity. The goal is to find an algorithm to convert the media into a high dimensional feature vector $D_i = (d_{i1}, d_{i2}, \dots, d_{in})$, such that the feature vectors of similar media are as close to each other as possible and those of dissimilar media respectively as far away from each other as possible. How close two feature vectors are to each other is determined by a distance function $s(D_i, D_j)$. Retrieving a list of media similar to some input is then a matter of using a k -nearest

neighbors search [14] algorithm to find the k closest feature vectors to the input media's feature vector.

For example to compare a 3D model for similarity one may use an algorithm that converts the 3D model into an n-dimensional real-valued feature vector D_i , such that the feature vectors of 3D models with a similar geometry or shape are close to each other. Then use $s(D_i, D_j) = \|D_i - D_j\|$, which corresponds to the L2 distance between D_i and D_j , to obtain a metric for the similarity of two feature vectors, respectively their 3D models.

2.2 REST API

In order for a multimedia retrieval system to become useful it needs to provide some interface through which other applications can access its functionality. This is what's commonly known as an application programming interface (API). One such type of APIs are REST APIs, an acronym for Representational State Transfer. These REST APIs aim to communicate between two programs via requests and responses over the HTTP web protocol. The following incomplete list should give an insight into how REST APIs operate - according to R. Fielding [8] REST APIs follow these and some more principles:

- Client-Server: User interface concerns should be separated from data storage concerns.
- Stateless: Communication must be stateless. All information required for a request to be understood must be included in said request.
- Cache: Responses' data must be labeled whether the data can be cached on the client side or not. If it is cacheable the client may reuse the same data for future equivalent requests.
- Uniform Interface: The REST API should form a uniform interface. Said interface should remain the same independent of the back-end implementation.
- And several other principles.

Usually REST requests are exclusively executed by sending an HTTP *GET*, *PUT*, *POST* or *DELETE* request through certain server provided URLs.

2.3 3D Models

Usually virtual 3D model consist of one or multiple polygon meshes. Formally, a polygon mesh consist of a list of vertices $\mathcal{V} = \{\mathcal{P}, \dots\}$. Each vertex has at least a position $\mathcal{P} = (x, y, z)$, but can also contain other attributes such as a color. This list of vertices is then used to define the surface of the mesh. The polygons of choice to represent the mesh's surface are usually triangles, where each triplet of consecutive vertices of the mesh forms one triangle. Triangles are the primary choice for representing 3D model surfaces in computer graphics because they have some desirable properties like always being planar. In addition to that modern hardware is built mostly around rendering triangles and is thus very fast at it.

However, such 3D models that consist of polygon meshes are restricted to representing surfaces. They cannot represent a volume, since there is no way to specify what's inside a mesh. Another way to represent 3D models, but in a volumetric way, is discussed in Section 2.4.

2.4 Voxels

Our application provides a 3D sculpting module and thus needs a suitable virtual representation of the sculptures. Usually 3D model formats only store the surface of a sculpture, but our application requires the sculptures to be truly volumetric and not just an empty shell, so that when the user removes some material from the sculpture they should be able to see the material filling the inside of the sculpture.

The foundation of the sculpting module is thus based on voxels. In the most basic sense a voxel represents a value on a regular grid in 3D space [35]. The value contained in a voxel usually describes which material it is made of, but can also contain additional data, e.g. to describe the shape of the voxel.

These voxels are very useful to represent volumetric structures, such as sculptures or terrain, since they fill the 3D space and each voxel represents one volume element. One can imagine voxels being like LEGO bricks with which the sculptures are built, or the analog to atoms on a much larger scale.

Compared to polygon mesh based 3D models these voxels have the advantage of being able to define the inside of a 3D model. This however comes at a cost of higher memory usage, since not only the surface must be defined but also what's inside. Furthermore, current GPU hardware is lacking extensive hardware support for voxels and thus they are not the primary choice to represent 3D models. In the recent years however voxels have gained a lot of popularity in other areas of computer graphics, especially for lighting calculations such as global illumination [33] or ambient occlusion [31].

2.4.1 Voxelization

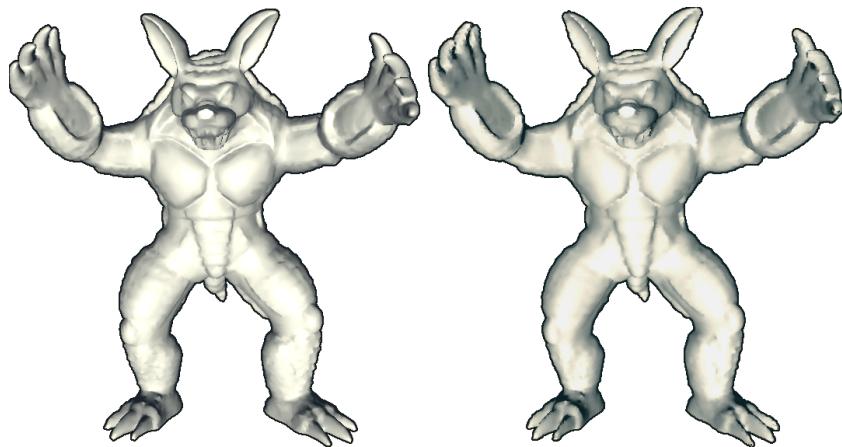


Figure 2.1: The Stanford Armadillo mesh [30] (left) and its voxel representation at a resolution of 128x128x128 voxels (right).

An important part of multimedial retrieval is the ability to take query results and reuse them to formulate a new, refined, query. In the context of this thesis multimedia retrieval is only concerned with polygonal meshes. To facilitate the modification of such polygonal meshes our application includes a voxelizer module. The voxelizer takes a polygonal mesh as input and converts it into a suitable voxel representation to enable the user to edit the model volumetrically. Since our voxels contain not only the voxel material, but also the surface normals, we can

reconstruct the mesh from our voxels while keeping most of the mesh's features intact with high accuracy, given a large enough voxel grid with sufficient resolution. In practice a resolution of 128x128x128 seems to be sufficient for most models that aren't highly detailed or noisy. Fig. 2.1 depicts an example of a mesh that has been converted to a voxel grid using an algorithm described in Section 4.2.5.

2.4.2 Isosurface Polygonization

Since the voxels are just an internal volumetric representation an algorithm is required to convert them into a representation that a computer can display on the monitor. Usually the representation of choice are meshes, as described in 2.3, since that's what modern GPU's are specialized in. In general these kinds of algorithms are called isosurface polygonization algorithms, since originally they were conceived to convert implicit functions, or e.g. computed tomography (CT) density fields, into a polygonal surface. The prefix "iso-" means equal, hence the word "isosurface" comes from the fact that the implicit function value or density is the same everywhere on the polygonal surface. There are some variants that do not directly work off of density fields but instead some intermediary representation like Hermite data which will be discussed in detail later on.

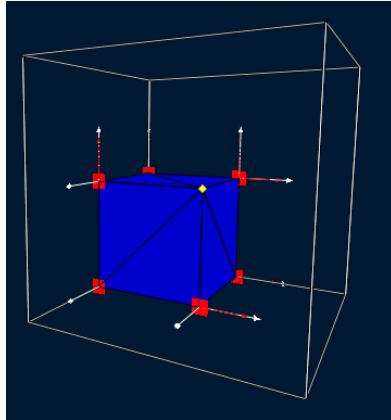


Figure 2.2: A polygonized voxel cell (white) containing the polygon surface of a cube's corner (blue) with a sharp feature (yellow).

The two main groups of isosurface polygonization algorithms are the primal, e.g. [21, 24], and dual, e.g. [10, 17], contouring algorithms. Most of these algorithms cannot polygonize a single voxel by itself, but instead require a cell consisting of eight voxels, usually arranged in the form of a cube. Primal and dual contouring algorithms differ in that the primal variants place polygon vertices somewhere on the cell's edges and the dual variants place them not on the edges but somewhere within the cell's volume. Therefore dual contouring algorithms have the advantage that they can more reliably polygonize surfaces with sharp features, such as e.g. the corners of a cube. Primal contouring algorithms can also polygonize sharp features in certain cases, however only if said sharp feature lies exactly on a cell's edge. Fig. 2.2 depicts a polygonized voxel cell where the sharp feature lies within the voxel cell.

Marching Cubes [24] (MC), a primal contouring algorithm, checks whether the material at each corner of the voxel cell is filled or empty and then packs all 8 resulting booleans into a single 8 bit integer. That integer is then used to index into a lookup table that contains which cell edges

are to be connected to each other to form polygons. Vertex positions can either lie on the middle of their respective cell edge or if the voxel contains a density value the vertex position can be smoothly linearly interpolated.

Dual Contouring [17] (DC) uses another approach. A valid voxel cell always contains either no edge with a material, or at least three edges with a material change. If a voxel cell does contain edges with a material change then Dual Contouring finds the sharp feature that minimizes the squared distance between the sharp feature and all planes spanned by the normals at the edges with the material changes. In a second pass the generated vertices are then connected between neighboring voxel cells to form polygons.

Cubical Marching Squares [13] (CMS) makes use of both primal and dual contouring features. Much like MC it places vertices on voxel cell edges. However it can also place vertices on voxel cell faces and within the voxel cell volume, like dual contouring algorithms do. This gives it the advantage that each cell can be polygonized individually while still being able to reliably polygonize surfaces with sharp features. This algorithm is described in more detail in Section 4.2.2.

2.5 CSG

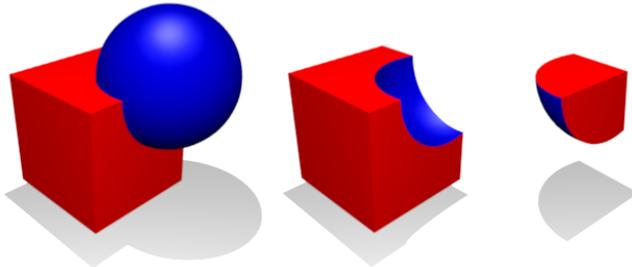


Figure 2.3: Boolean union (left), difference (middle) and intersection (right) of a cube and sphere primitive. Adapted from https://en.wikipedia.org/w/index.php?title=Constructive_solid_geometry&oldid=962220829

Constructive Solid Geometry (CSG) is a method to create geometric shapes by constructing them from primitives and a few operations. The three most common operations are the boolean union (\cup), difference ($-$) and intersection (\cap) operations are illustrated in Fig. 2.3. With just those three simple tools and a couple of primitives, e.g. spheres, cubes, cylinders, etc., the user can easily create arbitrary shapes or sculptures. Additionally CSG need not be restricted to only boolean operations, but can also make use of smooth operations that blend two shapes together in a continuous manner. In the context of signed distance functions, which will be discussed in Section 2.5.1, these smooth operations correspond to the smooth minimum and maximum operations.

The application developed for this thesis makes extensive use of CSG to enable the users to edit and shape their voxel based sculptures. The union ("add") and difference ("remove") operations are intuitive and easy to use. Additionally it also provides a "replace" operation that allows the user to replace solid materials inside a primitive with another material. Logically it is equivalent to $(A - B) \cup (B \cap A)$ where A is the existing sculpture and B is the primitive or bounds within which solid materials should be replaced with B's material.

2.5.1 Signed Distance Functions

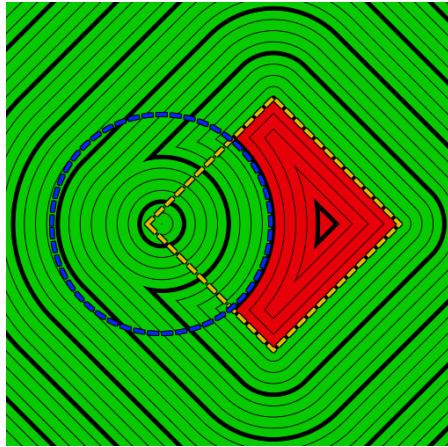


Figure 2.4: The resulting SDF from subtracting a circle SDF (blue) from a square SDF (yellow). The background is green where the resulting SDF is positive and red otherwise. The black lines represent the isolines of the resulting SDF, i.e. the value is the same everywhere on one of those lines. The subtraction operation results in an inexact SDF as illustrated by the isolines on the left half that are evidently not equidistant from the red area. Adapted from <https://sudonull.com/post/1035-Combining-Signed-Distance-Fields-in-2D>

Signed distance functions (SDF) are mathematical formulas that describe the signed distance between a point and the surface of a certain shape. For points inside the shape the signed distance is negative. For example it is trivial to derive the signed distance function $f(x, y, z) = \sqrt{x^2 + y^2 + z^2} - r$ of a sphere at $(0, 0, 0)$ from the implicit formula of the sphere's surface, $x^2 + y^2 + z^2 - r^2 = 0$. There exist readily available formulas for SDFs of various shapes, e.g. Inigo Quilez' collection of such functions [26].

SDFs are well suited as a representation of CSG primitives because one can approximate their union and difference by taking the minimum, respectively maximum, of two SDFs. This approach generally works well and can produce the exact union or difference, however in certain cases it can result in an incorrect approximation, especially for the difference and intersection operations as illustrated in Fig. 2.4 and for the inside of union'd SDF's. For the purposes of this application the SDFs need not be exact, hence we have decided to use SDFs as the representation of our CSG primitives. Additionally since the domain and range are the same for all SDFs, namely $f: \mathbb{R}^3 \mapsto \mathbb{R}$, they can all be treated exactly the same besides their formulas and can be encapsulated as a simple method in code.

2.6 Virtual Reality

Virtual reality (VR) applications and hardware aim to immerse the user into a virtual scene or application as much as possible. This is achieved by the use of a headset with an integrated display that the user wears, whereby all the user can see is the content shown on the integrated display. Furthermore, these headsets track the user's head position and rotation so that it can be precisely reflected in the virtual scene. Another large portion of the immersion also comes from the fact that the integrated display is stereoscopic, i.e. it can trick the user's eyes into seeing the depth of the virtual scene by showing a separate and slightly offset image of the scene

to each eye. In addition to the headset there are also VR controllers that are, similarly to the headset, tracked in 3D space and reflected in the virtual scene. The user can hold these and then precisely control the position of their virtual hands and even use them to manipulate objects in the virtual scene.

Current 3D sculpting applications usually have a steep learning curve. This comes from the fact that keyboards and computer mice, which serve as the primary human to machine interface, are not intuitive for sculpting. Modifying parts of a physical clay sculpture by hand is much more intuitive for a human than doing it virtually with a keyboard and mouse, because there is no direct mapping between these actions and the movement of a mouse or the buttons of a keyboard. There are companies that produce hardware specifically made to alleviate some of these issues, e.g. 3Dconnexion¹, but the issue of not having a direct mapping mostly persists.

This is where virtual reality excels and why it is a crucial part of our application. The VR controllers create a direct mapping between the position and movement of the hands and the ingame sculpting brush. In addition to that the VR headset also provides a true sense of depth with its stereoscopic display and makes it easier to view the virtual sculpture from all sides because head movement is tracked accurately. When combined these advantages provide a more immersive experience and hopefully also a more intuitive human machine interface.

2.7 System Architecture

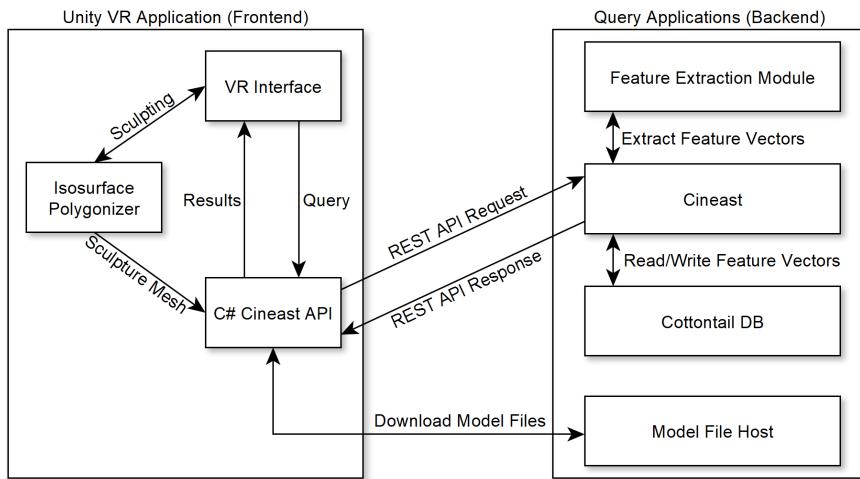


Figure 2.5: The high level architecture of our system.

To tie this all together Fig. 2.5 illustrates the high level architecture of our system. The frontend, i.e. our VR application running in the Unity engine, contains three main components relevant for the querying functionality: the VR interface, isosurface polygonizer and generated C# Cineast API module.

Through the VR interface the user can sculpt using voxels and CSG operations and start queries or view query results. On starting a query from the VR interface the entire sculpture is converted into a single mesh which is then sent to the multimedia retrieval system Cineast as a REST API request via the C# Cineast API. Cineast provides the functionality required to search a database

¹ <https://www.3dconnexion.de/>

for similar media, such as images, videos and 3D models.

Cineast then tasks the feature extraction module to convert the mesh into a feature vector. This feature vector is then sent over to the Cottontail DB database to be compared with the feature vectors of other models. Cottontail DB is Cineast's database backend and is primarily responsible for feature vector storage and running k -nearest neighbors searches with different kinds of distance functions. All models in that multimedia database that have a similar feature vector are then scored, ranked and their IDs are sent back to be received by the C# Cineast API. The C# Cineast API then uses the received IDs to download the respective model files and hands them over to the VR interface to be displayed.

3

Related Work

3.1 3D Model Retrieval

An important part of 3D model retrieval are the feature extraction algorithms. These have been, and still are, an active research topic and thus there are many algorithms available. R. Osada et al. describe the D2 shape distribution [25] which results in a histogram of distances between randomly sampled pairs on the 3D model's surface. Y.-J. Liu et al. then build upon this and propose ClusterD2+Color and ClusterAngle+Color [23], which work in a similar way and also take color of the 3D model into consideration. Kazhdan et al. describe a rotation invariant spherical harmonics representation of 3D models [18] making use of voxel grids and spherical harmonics functions to compute a descriptor. A different approach is described by D.-Y. Chen et al. [3], where the algorithm takes pictures of the 3D model from different angles, finds their contours and then uses them to compute a feature vector. Yet another method, Spin Images [16], performs well at matching partial surfaces of 3D models.

However, these algorithms are only useful as part of a retrieval system that can use them to search through a database. Cineast was initially conceived by L. Rossetto [27] as a retrieval system for videos. R. Gasser then extends Cineast to a multimedia retrieval system [11] supporting various other types of media, such as text or 3D models. Flickner et al. [9] first described the QBIC System, where the user can specify or sketch an example image that the retrieval system then uses to find similar looking content. The retrieval system user interfaces demonstrated in [11, 23, 34] all make the use of such a system.

3.2 Virtual Reality Sculpting

While virtual reality sculpting has not yet attracted the attention of the scientific research community, it has gained a lot of popularity in the games industry in recent years. Medium² offers a vast and mature toolbox to enable users to sculpt professional grade character models using the Oculus VR hardware. Tilt Brush³ takes a more painterly approach and mostly uses flat brush strokes to create scenes with a stylized hand drawn look. However, it also has some brush strokes that are 3D or can be animated. Quill⁴ uses a similar style and also offers extensive tools to let

² <https://www.oculus.com/medium>

³ <https://www.tiltbrush.com/>

⁴ <https://quill.fb.com/>

the user animate scenes smoothly. SculptVR⁵ takes a voxel based approach similar to some of the methods discussed in this thesis.

⁵ <http://www.sculptrvr.com/>

4

Implementation

This chapter goes more into detail about the concepts and architecture mentioned in Chapter 2 and how they are implemented in our application. Text IN THIS FONT refers to code classes, methods or fields. Method parameter types are specified in the parentheses following the method name.

The chapter starts out by elaborating the implementation details about the multimedia retrieval application Cineast and how it was extended to support model similarity queries that take 3D model color and texture into consideration. Furthermore, it explains how our VR sculpting application communicates with Cineast.

Following that, it explains how voxels were used to implement the sculpting functionality, sculpture rendering and how the CMS algorithm was extended to support multi-material voxel volumes. In addition to that it presents an algorithm to convert 3D model meshes into a voxel representation and we propose a new lookup table based implementation of the CMS algorithm. Finally, it ties all these features together and describes how they were combined to form our VR sculpting application.

4.1 Cineast

Cineast is the multimedia retrieval backend of choice for our application. In order to implement the requirements of our application several changes had to be made to Cineast. However due to Cineast's modular structure most of the changes could be applied without having to modify the core code.

4.1.1 Core Changes

Cineast's mesh classes MESH, READABLEMESH and WRITABLEMESH did not support model textures and thus had to be extended for our feature module. Methods to get/set a BUFFEREDIMAGE texture and a STRING texture path, for later retrieving the model texture through the API, were added to the respective classes. The mesh's VERTEX class was also extended to support storing VECTOR2F texture UV attributes.

The OBJMESHDECODER, which is responsible for reading an OBJ-model file into a MESH, was extended to also read texture UV coordinates. Moreover, it looks for a texture file with the same name as the OBJ-model file. If such a texture file is found it is added to the MESH and

the texture file's path is persisted in the database as metadatum with the domain "TEXTURE" and key "texture_path". To do so the SEGMENTER class received an additional method `PERSIST(SEGMENTCONTAINER, MEDIASEGMENTDESCRIPTOR)` which is called when a SEGMENTCONTAINER is persisted by `GENERICEXTRACTIONITEMHANDLER`. MODELSEGMENTER then uses that method to persist the mesh's texture file path metadatum. This metadatum can later be used by clients to download the model's texture.

A similar change was required in the MODELQUERYCONTAINER. Model data is being sent to Cineast over its REST API, described in Section 4.1.5, in the form of a custom string based model format. This format however does not support vertex colors which are required for color sensitive queries. For this reason a new COLORMESHPARSER class was added that uses a similar format but also supports vertex colors. Based on the prepended MIME type, an identifier used to describe the format of data, the MODELQUERYCONTAINER can decide which parser to use. Like this the API remains backwards compatible with the model format that does not contain vertex color data.

Throughout the course of development of this thesis Cineast's REST API was partially rewritten to use the Javalin⁶ web framework instead of Spark⁷ and SparkSwagger⁸. Javalin makes great use of Cineast's already built-in JSON serialization structure using Jackson⁹. This enables Javalin, in combination with its OpenAPI plugin, to generate an OpenAPI specification [15] that describes Cineast's REST API in a standardized way. An excerpt of this OpenAPI specification is shown in Listing 4.1 of Section 4.1.5. Since Javalin's OpenAPI plugin makes use of Jackson, it is capable of more reliably creating the specifications for JSON data structures than SparkSwagger, especially for subclasses of superclasses with generic types.

4.1.2 Model Formats

The current implementation supports the Wavefront OBJ format [36] as it is comparatively simple to parse. However, the feature extraction algorithm is agnostic to the original 3D model format and only requires a 3D mesh consisting of triangles and vertices that contain the following attributes: position and texture coordinates or vertex colors. Because of Cineast's modular nature it is possible to add parsers for other 3D model formats, or to extend the currently existing parsers to also read the aforementioned vertex attributes, without having to alter the feature extraction algorithm.

4.1.3 ClusterD2+Color Feature Extraction

Since our application allows the user to color their sculptures it makes sense that our similarity queries for 3D models also take color into account. Hence part of the color information must be contained in the feature vector, which will then be used to compare 3D models for similarity as described in Section 2.1. Currently, none of Cineast's feature extraction modules take the color of 3D models into consideration.

The ClusterD2+Color algorithm [23], which was used in our application, achieves this by creating random sample points on the 3D model's surface according to certain criteria and then use

⁶ <https://javalin.io/>

⁷ <http://sparkjava.com/>

⁸ <https://github.com/manusant/spark-swagger>

⁹ <https://github.com/FasterXML/jackson>

those sample points to compute a feature vector.

Our feature module, CLUSTERD2ANDCOLORHISTOGRAM, extends Cineast's STAGEDFEATUREMODULE class which takes care of the querying base functionality, such as searching for similar feature vectors and sorting results by score or persisting an extracted feature vector. The CLUSTERD2ANDCOLORHISTOGRAM class mostly takes care of extracting the ClusterD2+Color feature vector from a given input model.

In order to obtain the data required to construct a feature vector the algorithm runs through the following stages:

Sample Generation Many isosurface polygonization algorithms, such as the non-adaptive CMS implementation used in this application, can produce highly tessellated meshes. On the other hand meshes exported by 3D modeling software are often optimized to have large triangles where the surface has little detail, and small triangles where the surface is more detailed. Hence it is important that the sample point generation is mostly independent of the triangle sizes of the mesh, so that meshes of varying levels of tessellation can be compared for similarity effectively. The ClusterD2+Color algorithm achieves this by making the sampling directly proportional to the surface area of the triangles, hence a sample point has a higher chance to be positioned on a large triangle, and a smaller chance to be positioned on a small triangle. Before any sample points are generated the model is first scaled such that its total surface area equals 100. After generating a certain number of samples, the result is a collection of samples that are uniformly distributed on the mesh's surface. The number of samples is determined by a function of the integral of the absolute Gaussian curvature over the model's surface as described in [23].

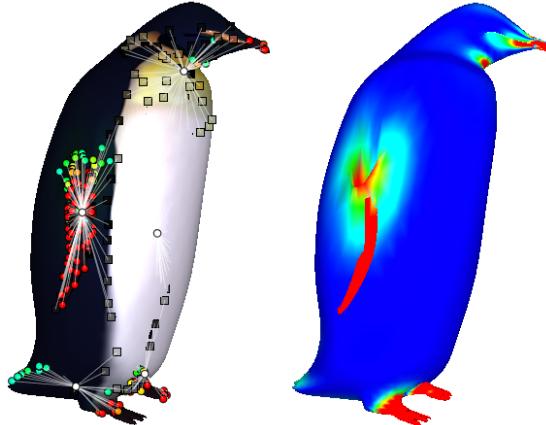


Figure 4.1: Visualization of the sample selection and clustering. On the left is shown the textured input model and the resulting shape (circles) and color (squares) samples and their cluster assignments (white circles and lines). The picture on the right shows the computed difference-of-Laplacian D of the model (red: high, blue: low).

Shape Sample Selection These samples should convey relevant information about the geometric shape. Previous work done by C. H. Lee et al. [20] has shown that mean curvature is a good indication for visual saliency and thus for comparing geometric shapes for similarity. The mean curvature of the mesh vertices is computed using Taubin's method [32]. Our implementation of the ClusterD2+Color algorithm computes the mean curvature of each vertex and stores

it as a vertex attribute. This mean curvature attribute is then smoothed over the mesh's surface by a Laplacian smoothing approximation described in [7], resulting in a curvature map. In particular, our implementation sets $\lambda = 8$ and, to satisfy the stability criterion $\lambda dt \leq 1$, $dt = \frac{1}{8}$, i.e. equivalently it applies small smoothings with $\lambda = 1$, $dt = 1$ a total of eight times for one full smoothing operation. This is repeated multiple times to create three smoothed curvature maps C^1 through C^3 . Each curvature map C^i has been smoothed i times.

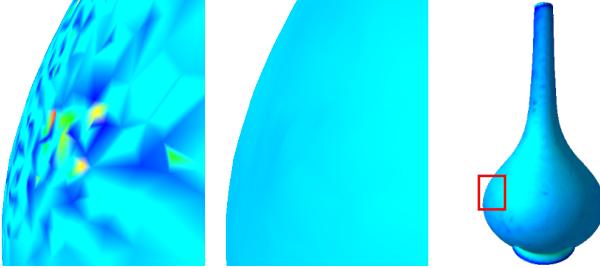


Figure 4.2: C^0 (left) exhibits significant artefacts on a non-manifold vase mesh (right) which has many invisible touching edges. C^1 (middle) is a more sensible curvature map as baseline for the shown model.

As opposed to [23] we use C^1 as the baseline curvature map instead of C^0 , because certain non-manifold models exhibit significant artefacts in C^0 as illustrated in Fig. 4.2. A mesh is considered non-manifold when it has unconnected vertices or edges that are touching. Lastly each shape sample calculates its $D = D^2 \oplus D^3$ value, where $D^i = C^i - C^{i-1}$ is the difference-of-Laplacian space, as depicted in Fig. 4.1, and only the samples with the top 10% highest D values are kept. Indeed the remaining top 10% D value shape samples in Fig. 4.1 are concentrated to the relevant features of the penguin model, namely its arms, feet and beak.

Color Sample Selection The color samples' purpose is to allow the algorithm to distinguish between models that have the same geometric shape but differ in color or texture. This is achieved by only selecting those samples that lie on significant color transitions. As described by Y.-J. Liu et al. [23], all generated samples extract the mesh's color at their position and then transform the color to the CIE-LAB color space [1], which aims to provide a perceptually uniform color space, i.e. same distances in CIE-LAB color space correspond to the same amount of perceptual difference. The resulting three CIE-LAB coordinates, L^* , a^* and b^* are then quantized into 6 (L^*), respectively 4 (a^*) and 4 (b^*) integer values. These three, now quantized, coordinates are combined into a single color code integer, such that each of those color codes corresponds to a unique combination of the quantized coordinates. Perceptually similar colors are thus mapped to the same color code and dissimilar colors are mapped to different color codes.

Lastly the algorithm only selects those samples as color samples of which less than 80% of the neighboring samples' color codes are the same. In our implementation the neighborhood range is defined as a function of the model's mean edge length, or if less than 5 samples are found within said range then it chooses the 5 nearest samples as the neighborhood. In Fig. 4.1 the color samples are selected mostly along the transition between the white, black and beige hair of the penguin

4.1.4 Comparing Features For Similarity

The selected shape and color samples now need to be reduced into an N-dimensional feature vector. ClusterD2+Color algorithm achieves this by first clustering nearby samples together and then creating a histogram from the euclidean distances, i.e. L2 norm, between all samples that are not in the same cluster.

More specifically, the samples are clustered together using a modified ISODATA algorithm described in [23]. In our implementation we set $P_n = 5\%$ of the total number of sample points, $P_s = 2 * E$, $P_c = 10 * E$, $I = 100$ with $E = \text{mean edge length of the model}$. Additionally we introduce two new parameters: $I_{min} = 10$, the minimum number of iterations, and $I_{nc} = 200$, the total maximum number of iterations after which the algorithm is stopped if no convergence is happening. Without the I_{min} parameter we have observed that the algorithm would commonly stop too early with inadequate cluster assignments. Then, once the samples are clustered together, the euclidian distances between all samples in different clusters are computed and stored in an array. The values in this array are then normalized such that the lowest value in the array is mapped to zero and the highest value is mapped to one. Finally this array is then converted into a histogram with 20 bins.

Y.-J. Liu et al. [23] have also proposed a second method called ClusterAngle+Color which calculates the values for the histogram from the angle between triplets of samples that are each in different clusters. According to the authors this method results in a better similarity measure, however due to the $O(N^3)$ complexity we have unfortunately found it to be infeasible with the large number of samples generated from the user sculpted models in our application.

In either case the result is a histogram which serves as 20-dimensional feature vector and in fact, since they are a geometric signature of the model, they can be used to compare models for similarity, e.g. by using the χ^2 distance metric, a commonly used distance metric to compare two histograms for similarity. The lower the χ^2 distance the more similar the models. Conveniently Cineast and its multimedia retrieval database Cottontail DB offer an efficient k-nearest neighbors query for the χ^2 distance metric.

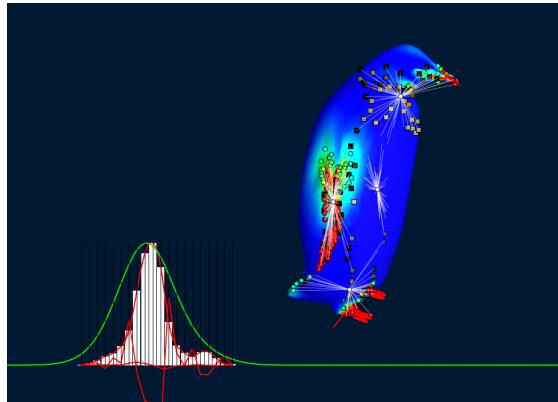


Figure 4.3: The feature visualization tool that was developed during the feature extraction implementation. White: histogram feature vector. Green: probability density function of the histogram.

Along with the χ^2 distance metric the Jensen-Shannon divergence (D_{JS}) was implemented. This is the metric Y.-J. Liu et al. [23] chose for their implementation. In short, the D_{JS} metric

treats the histogram feature vector as an independent and identically-distributed sample of a random variable and thus first converts the histogram into a probability density function by using kernel density approximation with a gaussian kernel. The Jensen-Shannon divergence of the two probability density functions to compare is then computed and serves as a probabilistic measure ranging from 0 to 1 for how similar they are, so in a sense how likely it is that the two histograms were generated from the same model. Fig. 4.3 depicts the histogram of a penguin model and its probability density function, along with some debug information of curves required for the D_{JS} calculation.

In our ClusterD2+Color implementation in Cineast D_{JS} is currently significantly slower than χ^2 because Cottontail DB does not natively support k-nearest neighbors queries for D_{JS} . Therefore Cineast must retrieve the feature vectors of all currently existing models in the database and compare all of them one by one to the input feature vector using D_{JS} .

The exact reason why Y.-J. Liu et al. used the Jensen-Shannon divergence was not mentioned, other than the fact that for a given 3D model they "[...]" regard its shape histogram as an independent and identically-distributed sample of a random variable [...] and thus convert it to a probability density function which can be compared using D_{JS} . We speculate that D_{JS} may be more robust to noise from the random sample generation than the χ^2 distance metric due to the conversion to a smooth probability density function. However this may incur a slight general loss of accuracy due to the smoothing.

Finally Cineast can run multiple features or distance metrics for a single query and then combine their scores by weights specified in Cineast's configuration. This can be useful to capture different kinds of feature types in a single similarity score.

4.1.5 C# Cineast API

Cineast provides two main APIs, a WebSocket API and more importantly a REST API which was used for the purposes of this thesis. The operating principles of a REST API are described in Section 2.2.

For example Cineast's similarity query endpoint is executed by sending a *POST* request to `http://cineast.domain/find/segments/similar`. Additional data, such as the model to run the similarity query for, is included in the HTTP request body in the form of JSON data as specified in "requestBody" of Listing 4.1. The structures of the used schemas are not shown in the listing but they are specified in the same file.

Listing 4.1: Part of the Cineast OpenAPI specification responsible for the similarity query API endpoint.

```

1  "/find/segments/similar": {
2    "post": {
3      "tags": ["/api/v1"],
4      "summary": "Finds similar segments based on the given query",
5      "description": "Finds similar segments based on the given query",
6      "operationId": "FindSegmentSimilarActionHandler#POST",
7      "requestBody": {
8        "content": {
9          "application/json": {
10            "schema": {
11              "$ref": "#/components/schemas/SimilarityQuery"

```

```

12      }
13      }
14    },
15  },
16  "responses": {
17    "200": {
18      "description": "OK",
19      "content": {
20        "application/json": {
21          "schema": {
22            "$ref": "#/components/schemas/SimilarityQueryResultBatch"
23          }
24        }
25      }
26    }
27  }
28}
29

```

The OpenAPI specification mentioned in Section 4.1.1 allows programs like Swagger Codegen¹⁰ to generate fully functional code libraries in various programming languages capable of interfacing with the REST API defined by the specification. The C# library that was used to access Cineast's API endpoint from Unity was generated by Swagger Codegen from Cineast's OpenAPI specification.

Listing 4.2: Calling the generated C# code of the similarity query API endpoint.

```

1 var terms = new List<QueryTerm> {
2   new QueryTerm {
3     Categories = categories,
4     Type = QueryTerm.TypeEnum.MODEL3D,
5     Data = "data:application/3d-color-json," + base64ModelData
6   }
7 };
8 var components = new List<QueryComponent> {
9   new QueryComponent {
10   Terms = terms
11 }
12 };
13 var conf = new QueryConfig();
14 var query = new SimilarityQuery {
15   Components = components,
16   Config = conf
17 };
18 var response = await Api.FindSegmentSimilarActionHandlerPOSTAsync(query);

```

Listing 4.2 shows an excerpt from our code that sends Cineast a similarity query REST request by calling the Swagger Codegen generated C# method FINDSEGMENTSIMILARACTIONHANDLER-POSTASYNC and using the generated classes QUERYTERM, QUERYCOMPONENT, QUERYCONFIG and SIMILARITYQUERY.

¹⁰ <https://swagger.io/tools/swagger-codegen/>

4.2 Hermite Data Voxels

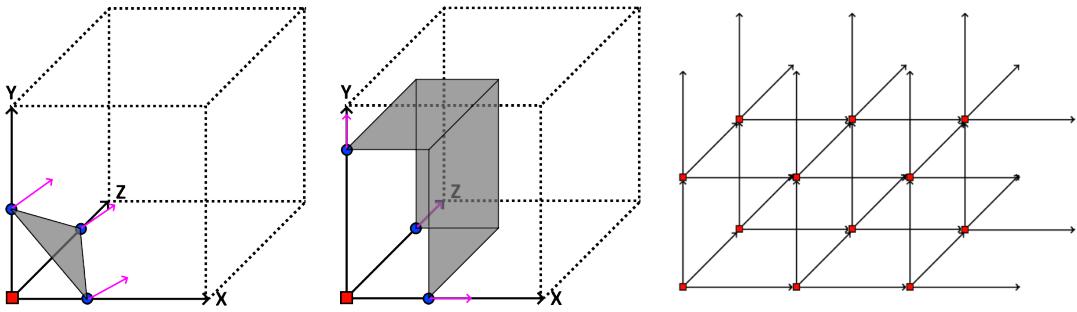


Figure 4.4: Hermite data voxels. One Hermite data voxel contains 1 material ID (red squares), up to three intersection points (blue circles) and their respective normals (pink arrows). Depending on the configuration of the intersection points and normals one voxel can take various shapes (gray areas), e.g. a single triangle (left) or a sharp corner (middle). The Hermite data voxels can be arranged in a grid (right) to form closed surfaces.

Since voxels are an abstract representation of volumes they can be visualized in various ways depending on the data the voxel contains. Originally a voxel consisted of a single attribute, a color or binary solid/non-solid, in which case the voxel can be visualized for example by cubes, spheres or splats. More sophisticated voxels can contain additional attributes such as normals or texture coordinates and thus can have a different look based on these attributes.

In particular, our application uses voxels that consist of a 32 bit material ID plus Hermite data [17]. Hermite data, first used for voxels in [17], consists of the following attributes: exact intersection points and normals. This allows the voxel to take on various shapes much like MC, but with more precision since the exact intersection points are stored explicitly and sharp features can be constructed from the stored normals. Fig. 4.4 illustrates the Hermite data and how it can represent diverse shapes. When arranged in a grid these Hermite data voxels form closed surfaces like in Fig. 4.5.

4.2.1 Voxel Storage

Data Structure There are many different data structures to store voxels, the simplest of which is simply storing the voxels' values in large 3D arrays. Other data structures however need not necessarily store all voxel values on a regular grid. Often there are regions with many voxels of the same value, and in that case it is beneficial to use a more advanced data structure that can store voxel values more memory efficiently.

The most common method to store voxels in memory is to split the voxels into fixed size blocks or chunks. Within each chunk the voxels are stored in an array. The chunks are stored separately in a sparse data structure, e.g. a hash map, where they can be accessed by their position. This gives a good trade-off between memory usage and access speed, because often voxel accesses happen within the same chunks and hence the hash map lookups can usually be cached. Instead of multi-dimensional arrays for the chunk voxel storage we use one contiguous linear array per chunk to avoid having to dereference multiple pointers for a single access. The index into this linear array can be directly computed from a given X, Y and Z position within the chunk by the following formula: $\text{index}(x, y, z) = x * N * N + y * N + z$, where N is the chunk size. Our

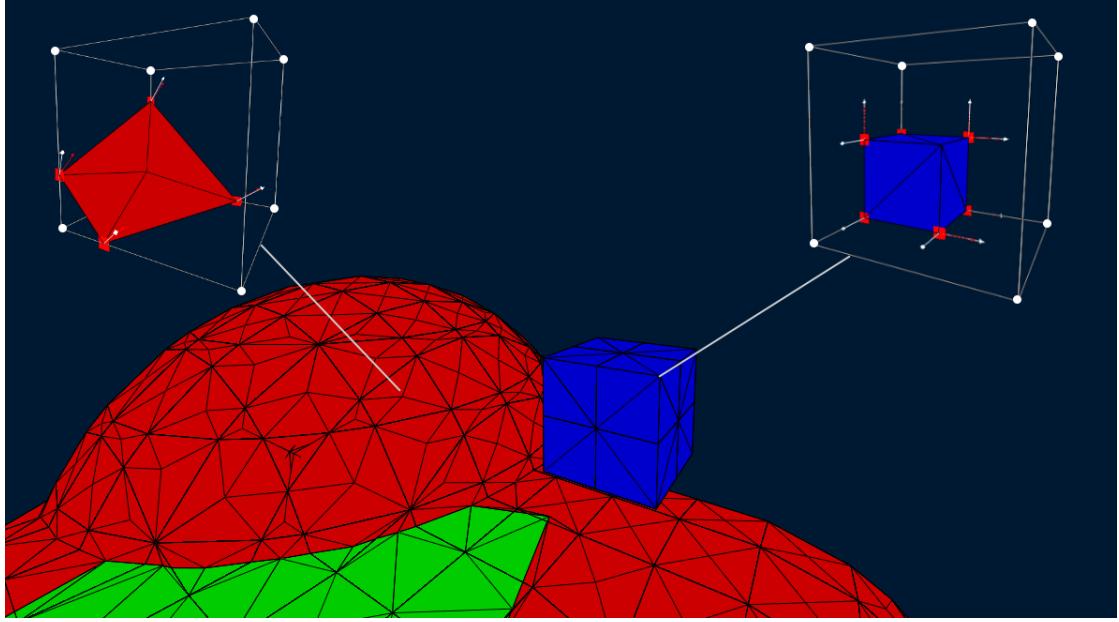


Figure 4.5: A voxel scene showing the mesh resulting from a grid of Hermite data voxels. Two voxel cells (top left and right) consisting of eight voxels each are shown separately to highlight how the mesh is constructed by combining the partial surfaces of all voxel cells. Colors signify different materials.

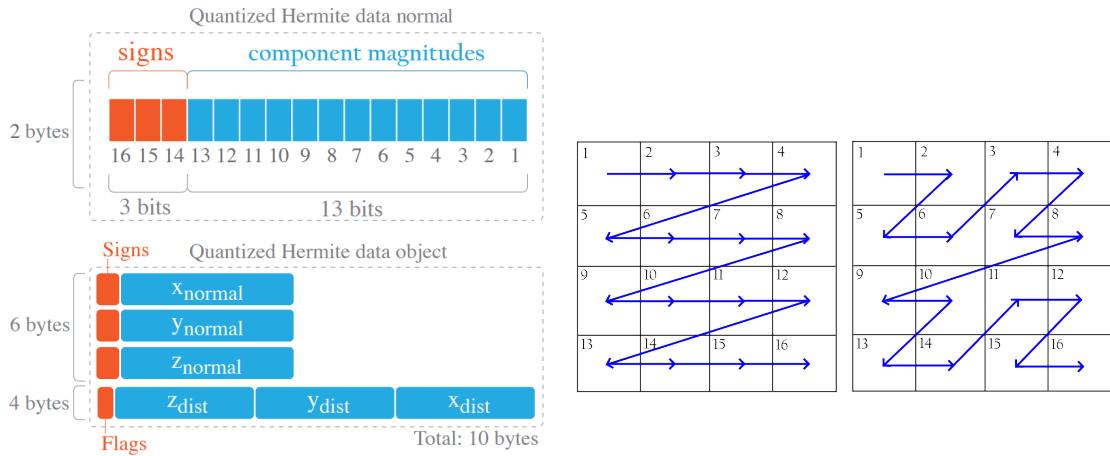


Figure 4.6: Quantized Hermite data structure. Top: a normal's component magnitudes are packed into 13 bits plus 3 bits for the signs. One normal thus fits into two bytes. Bottom: the entire quantized Hermite data object. x, y, z_{dist} refer to the intersection points. Adapted from [2], pages 18 and 20.

Figure 4.7: Simple linear indexing (left), Z-Order indexing (right). Each square represents a voxel in a 2D array. The blue line represents the order in which they are stored in main memory.

implementation uses a chunk size of $N = 32$. Voxels on the faces towards the positive X, Y and Z axes are padding voxels that always mirror the state of the neighbor voxel in the respective neighbor chunk. These padding voxels allow the voxel polygonizer to run faster since it entirely removes the need to query neighbor chunks during meshing. This method is described by [12] in more detail.

Memory Consumption The voxels in this application consist of a 32 bit material ID plus Hermite data. Therefore the size of a single voxel in memory is: 4 (material) + $3 * 3 * 4$ (three normals, each with three 32 bit float components) + $3 * 4$ (three intersection points) = 52 bytes in total. Considering that a single chunk contains $32 * 32 * 32 = 32768$ voxels, hence consumes roughly 1.7MB, and one scene can contain many of these chunks, the main memory will be exhausted rather quickly. To remedy this issue the Hermite data compression scheme described in [2] was used. This quantized Hermite data compression is lossy since it relies on quantizing the floating point values into small integers that are then packed together using bitwise operations. The errors introduced by the compression are negligible: the exact intersection points remain accurate to a $\frac{1}{1024}$ th, respectively less than 0.1%, of the voxel's edge length. Similarly the maximum error in a normalized normal's component is $\frac{1}{1024}$. The effective size of the quantized Hermite data is 12 bytes, two of which are struct alignment padding, i.e. only 23% of the uncompressed size. Fig. 4.6 illustrates the quantized Hermite data structure in more detail.

Access Optimization When a program accesses a piece of memory from main memory the CPU loads the surrounding address space into one or multiple fixed size cache lines, contiguous pieces of the heap memory around the access, into the cache. If the program then accesses a nearby address again the CPU can read the value from the cache which is a lot faster than reading it from main memory. This is called a cache hit. On the other hand if the program accesses an address too far away from the previous access or not cached then the CPU will have to read the value from main memory instead, ergo a cache miss. Hence by optimizing for cache hits a program's speed can be improved if it is memory bound. The most common access pattern for our voxel storage is to read eight neighbor voxels as a voxel cell for polygonization. For simplicity's sake we will consider a 2D instead of 3D array: if we use simple linear indexing as illustrated in Fig. 4.7 to read the voxel cell (1, 2, 5, 6) there will be a large address difference between the voxels on the first and second row. With large arrays this will often lead to cache misses. By instead using Z-Order indexing, also known as Morton indexing, spacially nearby voxels are much more likely to be nearby to each other in the address space, increasing the chance for cache hits. However, such indexing schemes usually have a larger overhead due to the calculation of the index. To compute the Z-Order indices parts of the Libmorton library¹¹ were used and ported to C#. Unfortunately in our case using Z-Order indexing has made no significant difference.

4.2.2 Cubical Marching Squares

For our application we have decided to use the Cubical Marching Squares algorithm [13] to polygonize our voxels. The reason for this choice was that CMS provides both primal and dual features. As such each cell can be polygonized independently but it can still also preserve sharp features well, e.g. the top of a voxel pyramid.

Our implementation of the CMS algorithm runs only on regular voxel grids and there is no adaptivity or level of detail. Therefore all future explanations on CMS are based on these two limitations.

¹¹ <https://github.com/Forceflow/libmorton>

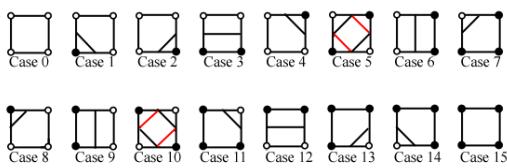


Figure 4.8: All marching squares cases. Red lines indicate an ambiguous case, i.e. either pair of lines is valid if only the corner values of the cell are available. Source: [22]

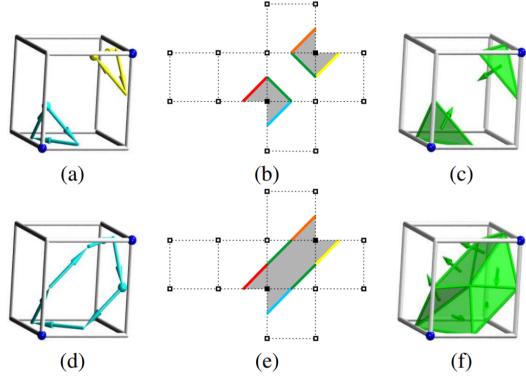


Figure 4.9: Cubical Marching Squares voxel cell components (a, d), their Marching Squares cell faces (b, e) and their triangulations (c, f). Source: [13], page 3

The first step of the CMS algorithm is to run the Marching Squares (MS) algorithm, the 2D counterpart of the MC algorithm, on each face of a voxel cell. Fig. 4.8 gives an overview over the total 16 different MS cases for a single voxel cell face. If the normals of the Hermite data span an angle greater than the 2D sharp feature angle threshold then the straight line segments are replaced with two lines that connect the two edges and the intersection point of the normals' tangents. This allows CMS to preserve sharp features on the voxel cell faces. Furthermore this sharp feature helps CMS to resolve the ambiguous MS cases 5 and 10, because the two line segments must not overlap. With the two sharp features in cases 5 and 10 it is simple to check whether the two sharp segments overlap or not, and then use the non-overlapping configuration. Once all six voxel cell faces have been processed by MS the CMS algorithm connects or links all touching line segments together. The result of this is a collection of components, i.e. groups of line segments that form a loop.

Then, for each component in a voxel cell, and all their line segments, CMS outputs vertices. Additionally for each component in a voxel cell, if the component's normals of a component span an angle greater than the 3D sharp feature angle threshold the CMS algorithm finds the component's 3D sharp feature that minimizes the sum of squared distances between the sharp feature and the planes spanned by the component's normals. Then it places a vertex at said sharp feature. If the angle threshold is not met then CMS instead places a vertex at the mean position of the component's vertices. This results in a list of triangles for each component, like illustrated in Fig. 4.9, that are added to the output mesh.

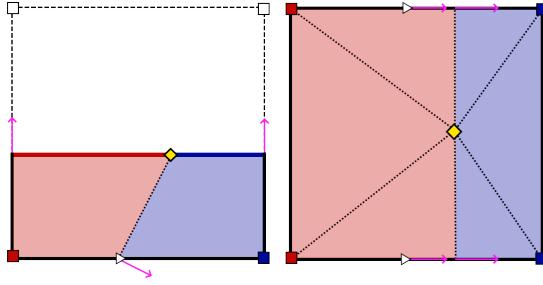


Figure 4.10: Multi-material CMS. The side view (left) shows how the material transitions (white triangles) are projected onto the segment, causing a split where one would expect the material transition to occur on the surface. The top view (right) illustrates how those splits are included in the sharp feature (yellow diamond) solution. The right side also shows the resulting triangulation (dotted lines).

4.2.2.1 Multi-material Extension for CMS

Algorithm 1 **GenerateSegment.** *This procedure remains mostly the same as described in the CMS paper. The only difference is the last line 10.*

```

1: procedure GENERATESegment(Face  $f$ )
2:   if there are two segments then
3:      $\{l_1, l_2\} \leftarrow \text{RESOLVEFACEAMBIGUITY}(f)$ 
4:      $f.list \leftarrow \{l_1, l_2\}$ 
5:     DETECTFACESHARPFEATURE( $f.list$ )
6:   else if there is one segment  $l$  then
7:      $f.list \leftarrow \{l\}$ 
8:     DETECTFACESHARPFEATURE( $f.list$ )
9:   end if
10:  DETECTMATERIALTRANSITION( $f$ )
11: end procedure

```

Algorithm 2 DetectMaterialTransition. This procedure finds and inserts material transitions into the segments of a face f . INSERTTRANSITION inserts the transition points into $f.list$ at their appropriate indices s.t. the segment's pieces remain straight and don't form folds, i.e. s.t. they subdivide the segment, and transition2 is to be inserted after transition1. The MATERIAL_TRANSITIONS table returns one or two edges connecting two neighboring face corners with solid materials given the marching squares case, since those could contain material transitions.

```

1: procedure DETECTMATERIALTRANSITION(Face  $f$ )
2:   if  $f.marchingSquaresCase \in \{3, 6, 7, 9, 11, 12, 13, 14\}$  then
3:     edges  $\leftarrow$  MATERIAL_TRANSITIONS[ $f.marchingSquaresCase$ ]
4:     for all edge  $\in$  edges do
5:       normal  $\leftarrow$  GETNORMAL( $f, edge$ )
6:       if normal  $\neq \emptyset$  then  $\triangleright$  There can only be a transition if a normal exists.
7:         intersection  $\leftarrow$  GETINTERSECTION( $f, edge$ )
8:         tangent  $\leftarrow$  CONSTRUCTTANGENT(intersection, normal)
9:         for all  $l \in f.list$  do
10:           for all  $p \in l$  do  $\triangleright$  The variable  $p$  is one piece of the segment  $l$ .
11:             if tangent intersects with  $p$  then
12:               transitionPoint  $\leftarrow$  INTERSECTIONPOINT(tangent,  $p$ )
13:               transition1  $\leftarrow$  CREATETRANSITION()
14:               transition1.position  $\leftarrow$  transitionPoint
15:               transition1.material  $\leftarrow$  edge.points[0].material
16:               transition1.normal  $\leftarrow$  normal
17:               transition2  $\leftarrow$  CREATETRANSITION()
18:               transition2.position  $\leftarrow$  transitionPoint
19:               transition2.material  $\leftarrow$  edge.points[1].material
20:               transition2.normal  $\leftarrow$  normal
21:               INSERTTRANSITION( $f.list, transition1, transition2$ )
22:             end if
23:           end for
24:         end for
25:       end if
26:     end for
27:   end if
28: end procedure

```

The regular CMS algorithm does not support multiple materials. Hence due to the need for multi material volumes in this thesis's application a multi-material CMS extension was developed. In order for this to work the Hermite data normals must additionally also be stored for each material change edge, and not only for non-solid to solid changes or vice versa. Furthermore each segment vertex must store a material. Only one change was required in the algorithms presented in the CMS paper: GENERATESEGMENT's modification is shown in Alg. 1. Since our application only uses regular voxel grids the modification was made to work on regular voxel grid cells. It can however be extended to work with the adaptive CMS version in a similar manner.

Additionally to the GENERATESEGMENT modification the DETECTSHARPFEATURE implementation also requires a change. The transitions generated in DETECTMATERIALTRANSITION, which are added to $f.list$, and their normals must be included in the linear system of equations to be solved [13] for the sharp feature detection, as suggested in [17]. This causes the sharp feature to be placed where one would expect the material transition to occur on the surface, as illustrated in Fig. 4.10. This of course means that the multi-material CMS algorithm must always generate a sharp feature when material transitions are present, even if the sharp feature

conditions/heuristics described by Kobbelt et al. [19] , which are used in CMS, are not met.

4.2.2.2 Lookup Table Based Implementation

Much of the CMS algorithm can be done and perhaps be sped up using appropriate lookup tables, especially in the case of regular grid voxel cells. We have generated such lookup tables in order to show the viability of such an approach. Due to time constraints and it breaking the multi-material extension from Section 4.2.2.1 however, it has ended up not being used in the final implementation. Nevertheless with some more work it can also be made to work with multi-material volumes. Furthermore this lookup table based approach may prove useful for a future GPU based implementation as it splits the problem into several small steps that can all be run in parallel with very little divergent behaviour.

A tool was implemented to generate the lookup tables for all $2^8 = 256$ possible cell states. There are two states per cell corner, solid vs. non-solid, hence 256 states per cell. Using rotational symmetries several lookup tables have been reduced in size - the initial lookup table maps one of the 256 cell states to one of 23 base configurations by rotating it appropriately. Each of the in total 24 unique such rotations, consisting of one or multiple 90° rotations around the X, Y and Z axes, gets its own index. All other lookup tables are based on these base configurations, their CMS components and the rotation index. From all the possible components that CMS can generate in a regular grid voxel cell, the lookup table tool has identified a total of 86 unique components for all base configurations. These components have then been indexed and their size and edges were stored in a separate lookup table.

Algorithm 3 DisambiguationBit. *Determines the disambiguation bit for the specified ambiguous face. In our implementation we use the CASE_AND_AMBIGUITY_NR_TO_EDGES lookup table described further below to obtain the face edges.*

```

1: procedure DISAMBIGUATIONBIT(Face face)
2:   edges  $\leftarrow$  face.edges
3:   segment1  $\leftarrow$  CONSTRUCTSEGMENT(edges[0], edges[1])  $\triangleright$  Constructs a segment from the
   intersections on edge[0] and edge[1]
4:   DETECTFACESHARPFEATURE(segment1)
5:   segment2  $\leftarrow$  CONSTRUCTSEGMENT(edges[2], edges[3])
6:   DETECTFACESHARPFEATURE(segment2)
7:   if segment1 intersects with segment2 then
8:     return 1
9:   else
10:    return 0
11:   end if
12: end procedure

```

One complication of the lookup table based approach are the potential 2D cell face ambiguities. All tables thus require as index not only the base configuration and rotation index, but also a 6 bit disambiguation index, also called "ambiguityRes", where each bit corresponds to one of the cell faces. For a given base configuration case and an ambiguity number, i.e. the index of the ambiguity ranging from 0 to 5 to check, Alg. 3 calculates a disambiguation bit. These disambiguation bits are then bit packed into a single integer where the bit with *ambiguityNr* = 0 is the most right bit and then from right to left according to *ambiguityNr*, resulting in the 6 bit disambiguation index. The 3D ambiguities as mentioned in the CMS paper are not being

handled by lookup tables, those still need to be resolved dynamically.

With the base configuration, rotation index and now also the disambiguation index a list of components can be obtained from yet another lookup table. The proposed lookup table based CMS algorithm then creates the segments for each component, finds their 2D and 3D sharp features similarly to the regular CMS algorithm, and finally triangulates them.

In total the proposed lookup table based CMS algorithm primarily requires the following lookup tables which have been generated by our lookup table tool:

- **RAW_CASE_TO_CASE_ROTATION_AND_AMBIGUITY_COUNT**[$caseIndex * 3 + value$]
Returns either the base configuration case, rotation index or ambiguity count
 - $caseIndex$: One of the 256 possible cell cases
 - $value$: 0 for base configuration case, 1 for rotation index, 2 for ambiguity count
- **CASE_AND_AMBIGUITY_NR_TO_FACE**[$baseCase * 6 + ambiguityNr$]
Returns the base configuration face index (i.e. after reducing the rotational symmetries) for the given $ambiguityNr$
 - $baseCase$: One of the 23 base configuration cases
 - $ambiguityNr$: Index of the ambiguity, as discussed above
- **ROTATION_TO_RAW_FACES**[$rotationIndex * 6 + baseFaceIndex$]
Returns the original face index (i.e. before reducing the rotation symmetries)
 - $rotationIndex$: One of the 24 unique rotations
 - $baseFaceIndex$: Base configuration face index
- **CASE_AND_AMBIGUITY_NR_TO_EDGES**[$baseCase * 6 * 4 + ambiguityNr * 4 + entry$]
Returns the base configuration edge index (see Alg. 3)
 - $baseCase$: One of the 23 base configuration cases
 - $ambiguityNr$: Index of the ambiguity, as discussed above
 - $entry$: Which edge index to return, i.e. 0 - 3 (four edges per face)
- **ROTATION_TO_RAW_SIGNED_EDGES**[$rotationIndex * 12 + baseEdgeIndex$] & 0b01111
Returns original edge index (i.e. before reducing the rotation symmetries)
 - $rotationIndex$: One of the 24 unique rotations
 - $baseEdgeIndex$: Base configuration edge index
- **CASE_AND_AMBIGUITY_RES_TO_SIZE_AND_COMPONENTS**[$baseCase * 64 * 5 + ambiguityRes * 5 + componentNr$]
Returns the base configuration component index
 - $baseCase$: One of the 23 base configuration cases
 - $ambiguityRes$: Disambiguation index, as discussed above
 - $componentNr$: Which component to return, i.e. 1 - 4 (at most 4 configurations per voxel cell). If 0, the number of components in the voxel cell is returned.

- COMPONENT_TO_SIZE_AND_EDGES[*componentIndex* * 13 + *edgeNr*]
Returns the base configuration edge index
 - *componentIndex*: One of the 86 unique base configuration components
 - *edgeNr*: Which component to return, i.e. 1 - 12 (at most 12 edges per component). If 0, the number of edges in the component is returned.

For further details about the proposed lookup table CMS algorithm please refer to our lookup table tool and the sample implementation, both in the same repository, written in Java. The sample implementation demonstrates how the lookup tables may be used.

4.2.3 CSG Operations on a Voxel Grid

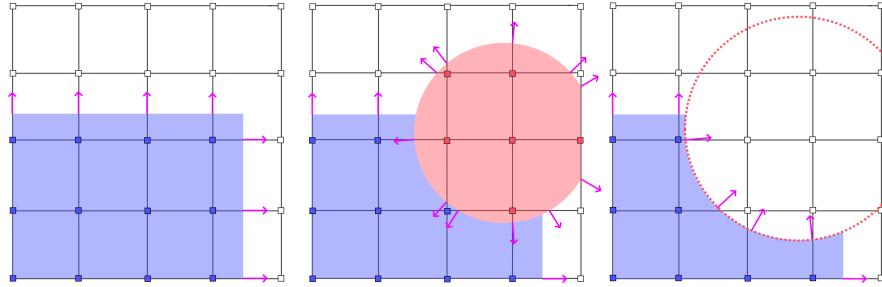


Figure 4.11: CSG operations on a 2D Hermite data voxel grid. From left to right: initial grid, union and difference. Each square and its line segments to the right and above, including any normals (pink) on them, represent one Hermite data voxel.

The CSG operations that our application uses for the sculpting brushes can be applied using variants of Alg. 4.

Algorithm 4 Union. Applies a CSG union to a voxel grid. FINDINTERSECTION finds the intersection of a primitive with the specified edge and returns an intersection object. intersection.distance is the distance between the intersection and a voxel or edge. voxels[0], and intersection.normal is the surface normal vector at the intersection.

```

1: procedure UNION(VoxelGrid grid, CSGPrimitive primitive)
2:   for all  $v \in \text{grid}.\text{voxels}$  do
3:     if  $v$  is inside primitive then
4:        $v.\text{inside} \leftarrow \text{true}$ 
5:        $v.\text{material} \leftarrow \text{primitive}.\text{material}$ 
6:     else
7:        $v.\text{inside} \leftarrow \text{false}$ 
8:     end if
9:   end for
10:  for all  $e \in \text{grid}.\text{edges}$  do
11:    if  $e.\text{voxels}[0].\text{inside} \neq e.\text{voxels}[1].\text{inside}$  then
12:      intersection  $\leftarrow \text{FINDINTERSECTION}(e, \text{primitive})$ 
13:      if  $e.\text{voxels}[0].\text{inside} \wedge \text{intersection}.\text{distance} \geq e.\text{voxels}[0].\text{intersections}[e.\text{axis}]$  then
14:         $e.\text{voxels}[0].\text{intersections}[e.\text{axis}] \leftarrow \text{intersection}.\text{distance}$ 
15:         $e.\text{voxels}[0].\text{normals}[e.\text{axis}] \leftarrow \text{intersection}.\text{normal}$ 
16:      else if  $e.\text{voxels}[1].\text{inside} \wedge \text{intersection}.\text{distance} \leq e.\text{voxels}[0].\text{intersections}[e.\text{axis}]$  then
17:         $e.\text{voxels}[0].\text{intersections}[e.\text{axis}] \leftarrow \text{intersection}.\text{distance}$ 
18:         $e.\text{voxels}[0].\text{normals}[e.\text{axis}] \leftarrow \text{intersection}.\text{normal}$ 
19:      end if
20:    end if
21:  end for
22: end procedure

```

First, the algorithm for a CSG union operation tags all voxels of the voxel grid that are inside the CSG primitive. Then it iterates over all edges of the voxel grid and checks if the two voxels that form the edge have a different *inside* state. If that is the case then the edge must intersect with the primitive's surface and the exact intersection is found using FINDINTERSECTION. The algorithm then checks if any already existing intersections can and must be replaced by comparing their values. For multi-material voxel volumes some additional checks are required, mostly so that intersections are placed correctly at solid to solid material transitions. In addition to that, our implementation also gives each CSG primitive an axis aligned bounding box, such that only those voxels and edges inside the bounding box need to be checked.

With some minor adjustments, namely swapping the \geq and \leq , inverting the normals and changing the material assignment at the top to set to empty voxels, the algorithm for the CSG difference operation is obtained.

In Section 2.5.1 it was already mentioned that SDF's are well suited for CSG primitives. The reason for this is that they are versatile and checking whether a voxel lies inside the SDF or finding line to surface intersection points is simple. Checking whether a voxel is inside the SDF is as simple as checking whether the value of the SDF evaluated at said voxel is negative. Finding intersection points (FINDINTERSECTION) on an edge is done by running a binary search on the edge, starting with the edge start- and endpoints and repeatedly moving them closer together. The binary search runs as long as the sign at the start- and endpoints are different, or until the value returned by the SDF for either start- or endpoint is within a certain threshold from zero. Both the start- and endpoints will thus converge towards the position where the SDF evaluates to zero, i.e. the SDF's surface.

This binary search can be optimized by using the values returned by the SDF for each iteration.

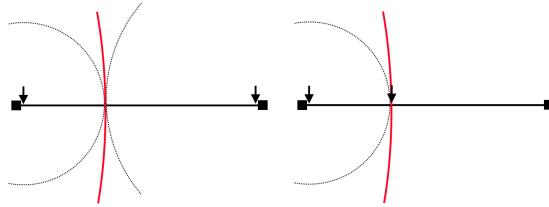


Figure 4.12: Adaptive binary search. Red circle: SDF’s surface, Arrows: start- & endpoint, Black circles: circles with radius of SDF’s value at start- or endpoint. The adaptive binary search finds the zero point from the initial state (left) after just one iteration (right).

Those values are a lower bound of the distance the start- or endpoint can be moved without accidentally crossing the SDF’s surface. Using this fact the adaptive binary search can often make larger and more precise steps than a blind binary search. Consider a nearly linear surface perpendicular to the edge, like in Fig. 4.12: the adaptive binary search will find the exact zero point after just one iteration because the two values returned by the SDF are the exact distances between the start- and endpoints and the surface. This may not always work for inexact SDF approximations since the returned values may not be actual lower bounds, hence the algorithm should still be able to use the regular binary search method as fallback for robustness.

4.2.4 Rendering

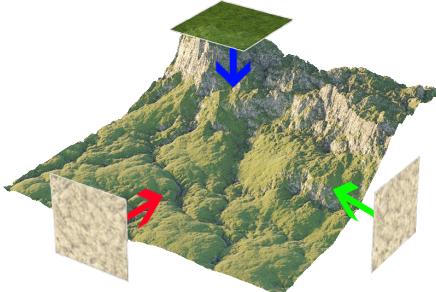


Figure 4.13: Tri-planar texture mapping of a mesh. Note how the hill-sides of the terrain are rocky and the top is grassy. Adapted from https://corona-renderer.com/blog/wp-content/uploads/2017/04/Triplanar_01.jpg

The voxel polygonizer already does most of the heavy lifting for the voxel rendering. Once the voxel polygonizer has converted the voxels into a triangle mesh the rendering is done using the typical mesh rasterization techniques, provided by Unity, and shaders. Shaders are GPU programs that determine how objects in the scene are rendered to the screen, e.g. by computing lighting or adding special effects. The voxel materials and their colors are packed into the 32 bit RGBA color attachment of the mesh vertices. The first three components, RGB, are used to store the voxel color, the fourth component A for the voxel material ID.

Since our voxel meshes are not UV mapped, i.e. they do not contain information about where the textures should be sampled, we have resorted to a tri-planar texture mapping shader to texture them. These kinds of shaders use the position of a mesh’s pixels to determine where to sample a texture, without the need of any explicit UV mapping. In order to achieve a realistic looking texture mapping the tri-planar texture mapping shader projects an infinitely repeating texture from the X, Y and Z axes onto the mesh like depicted in Fig. 4.13. The weight, i.e. how visible

it is, of each texture projection is determined by the mesh's normals. For example the more a normal points along the X axis, the stronger the X axis texture projection is weighted, and the less the other two. The result of this is a mesh that is usually textured nicely on all sides. In certain situations, especially when the weights of two texture projections are similar, e.g. for 45° diagonal parts of the mesh with respect to axes, there can be unnatural overlapping of two texture projections. This however is only a problem for textures with obvious structures, e.g. bricks, as opposed to natural or noisy textures like dirt or grass. To support multiple textures we store all voxel material textures in a single texture array on the GPU. The tri-planar texture mapping shader can then select the appropriate texture to sample from for each vertex by using the vertex' voxel material ID, stored in the alpha component, to index into the texture array.

4.2.5 Voxelizer

The precursor of our application already had an implementation of a voxelizer, however throughout the development of this thesis the voxelizer was rewritten to scale better with high triangle counts.

The implementation of a basic voxelizer is described in Alg. 5. This algorithm however will only work properly for meshes without holes.

Algorithm 5 Voxelize. *Voxelizes the given mesh into the voxel grid. FILLSECTION fills the voxel grid along and inside the given section with the given material and also sets the voxel Hermite data at the section's start and end points.*

```

1: procedure VOXELIZE(VoxelGrid grid, Mesh mesh, Material material)
2:   for all axis  $\in \{X, Y, Z\}$  do
3:     for all voxel  $\in \text{grid.sideVoxels}[\text{axis}]$  do
4:       for all section  $\in \text{FINDSECTIONS}(\text{axis}, \text{voxel.position}, \text{mesh})$  do
5:         FillSection(grid, section, material)
6:       end for
7:     end for
8:   end for
9: end procedure

```

Algorithm 6 `FindSections`. *Finds all intersections of a line with the mesh and then returns all sections that are between an ingoing and outgoing intersection, i.e. inside the mesh.*

```

1: procedure FINDSECTIONS(Axis axis, Position position, Mesh mesh)
2:   line  $\leftarrow$  CREATELINE(position, position + axis * gridSize)
3:   intersections  $\leftarrow \emptyset$ 
4:   for all triangle  $\in$  mesh do
5:     intersections  $\leftarrow$  intersections  $\cup$  FINDINTERSECTION(line, triangle)
6:   end for
7:   SORT(intersections)
8:   sections  $\leftarrow \emptyset$ 
9:   previous  $\leftarrow \emptyset$ 
10:  inside  $\leftarrow \text{false}$ 
11:  for all intersection  $\in$  intersections do
12:    if  $\neg$ inside  $\wedge$  previous  $\neq \emptyset$  then
13:      sections  $\leftarrow$  sections  $\cup$  CREATSECTION(previous, intersection)
14:    end if
15:    previous  $\leftarrow$  intersection
16:    inside  $\leftarrow \neg$ inside
17:  end for
18:  return sections
19: end procedure

```

Essentially the algorithm checks the grid lines perpendicular to the +X, +Y and +Z faces of the voxel grid for intersections with the mesh. Doing this for all the three grid faces will ensure that no triangles or voxels to be filled are missed. The found intersections are then sorted from minimum distance from grid face to maximum distance. When iterating over these sorted intersections the algorithm alternates between inside the mesh and outside the mesh, since for each incoming intersection the next intersection must be an outgoing intersection. All voxels inside such a pair of ingoing and outgoing intersections are then filled with a chosen material. For the first and last voxel in such a section the algorithm also needs to set the Hermite data, i.e. the voxel intersection values and normals.

We have found that in practice some triangles seem to be missed during the intersection checks due to floating point precision errors. To address this problem our implementation performs a post-processing hole patching algorithm that checks for missing normals in the voxel Hermite data and then obtains them by sampling from the closest triangle to the hole's position.

Note that the complexity of the algorithm, disregarding FILLSECTION, is $O(N^2 * K)$ where N is the voxel grid size and K the triangle count of the mesh, since for each side voxel touching one of the grid's +X, +Y or +Z faces, of which there are $3 * N^2$ in total, the algorithm needs to check all triangles. For large N and K this poses a problem. Hence as improvement our new algorithm first runs Alg. 7.

Algorithm 7 AssignTriangles. Projects the AABB’s of all triangles onto all three voxel grid faces and then assigns the triangles to the bin of the according face and position.

```

1: procedure ASSIGNTRIANGLES(Bins bins, Mesh mesh)
2:   for all axis  $\in \{X, Y, Z\} do
3:     for all triangle  $\in$  mesh do
4:       box  $\leftarrow$  CREATEAABB(triangle)
5:       area  $\leftarrow$  PROJECTONGRIDFACE(axis, box)
6:       for all position  $\in$  area do
7:         bins[axis][position]  $=$  bins[axis][position]  $\cup$  triangle
8:       end for
9:     end for
10:   end for
11: end procedure$ 
```

Following that Alg. 6 is then modified to only iterate over the triangles in *bins[axis][position]*. In practice this provides a good speed improvement, although at the cost of more memory, because Alg. 6 will only have to check a tiny fraction of all the triangles for each position.

In our implementation FINDSECTIONS runs in parallel for all the voxels touching one grid face since all the voxels along one of the face’s perpendicular grid lines can be written to independently in parallel. This is done three times, once per grid face, because otherwise the algorithm could potentially write to the same voxel multiple times at the same time, causing a write conflict.

4.3 VR Sculpting

The feature that ties all these technologies together is the VR sculpting. It allows the user to freely sculpt arbitrary shapes or figures by using voxels, CSG operations and various brush shapes.

4.3.1 SteamVR Plugin

The implementation of our VR application was realized through the Unity SteamVR Plugin¹². This Unity plugin offers most of the base functionality required for VR applications, such as setting up the headset and controller tracking and handling controller input. Furthermore it provides some additional scripts that can be added to objects in the scene, for example to allow the player to teleport or to grab and hold objects.

Unfortunately certain features of the SteamVR Plugin are currently not yet compatible with Unity’s new Universal Render Pipeline¹³ (URP) that handles the rendering of scenes. With the introduction of the URP the format of Unity’s shaders has changed. Some of these shaders, most notably the highlight when grabbing an object, were required for our application and thus had to be partially rewritten to work in the new URP.

4.3.2 Sculpting Functionality

The sculpting functionality is split into several scripts and components. The following list goes through the most important components for the sculpting functionality. The term ”script” refers

¹² <https://assetstore.unity.com/packages/tools/integration/steamvr-plugin-32647>

¹³ <https://docs.unity3d.com/Packages/com.unity.render-pipelines.universal@8.1/manual/index.html>

to a C# class extending MONOBEHAVIOUR in the Unity project and "component" refers to a script that has been instantiated and added to an object. MONOBEHAVIOUR's are scripts that can be attached to objects in the scene.

Voxel Sculpture The voxel sculpture/terrain functionality is provided by the DEFAULTVOXELWORLDCONTAINER script and can be added as a component to an empty object. This object is then capable of storing and rendering voxels and can be modified through the APPLYSDF method in DEFAULTVOXELWORLDCONTAINER.INSTANCE.

Texture Array As mentioned in Section 4.2.4 our application uses a texture array to store all the voxel material textures on the GPU for rendering. This texture array is created by the VOXELTERRAINTEXTURES script which is to be added to the same object that contains the DEFAULTVOXELWORLDCONTAINER component. On start of the application it reads all the textures that have been added to the VOXELTERRAINTEXTURES component's textures array into a texture array on the GPU and then assigns it to the ".TextureArray" attribute of the object's material such that the tri-planar texturing shader can access it.

Brush Type Renderers All regular brush types need some way to be rendered as a preview. For each brush type a unique ISDFRENDERER instance is responsible for rendering. The regular way to obtain such a renderer is to create a MESHSDFRENDERER object which extends ISDFRENDERER and handles all the rendering when given a mesh. Such an object is created in Unity's context menu under "Create → ScriptableObject → MeshSdfRenderer". Once created a mesh and material must be assigned to the object. Optionally an offset and scale can be specified. For advanced brushes one may create a custom renderer by extending ISDFRENDERER.

These renderers by themselves aren't useful yet. They need to be registered to the "Static Renderers" menu of the SDFSHAPEREHANDLER component on the player object.

Custom Brush The main functionality of the custom brushes is contained in the DEFAULTCUSTOMBRUSHCONTAINER script which is found on a separate object. It mostly handles creating the SDF from a list of custom brush primitives. In order to be able to render the custom brush is in the DEFAULTCUSTOMBRUSHSDFRENDERER script which is usually on the same object as the DEFAULTCUSTOMBRUSHCONTAINER. The custom brush can be edited by changing the DEFAULTCUSTOMBRUSHCONTAINER.INSTANCE.PRIMITIVES list which contains the current custom brush primitives.

Undo/Redo The undo/redo functionality is handled by the DEFAULTVOXELEDITMANAGERCONTAINER script. It exposes an UNDO and REDO method in DEFAULTVOXELEDITMANAGERCONTAINER.INSTANCE. In order to store edits DEFAULTVOXELEDITMANAGERCONTAINER.INSTANCE.CONSUMER() must be passed to APPLYSDF. Brush strokes are counted as a single edit. Each edit stores a snapshot of the previous state of the chunks it affects so the edit can be undone by replacing the chunk data with the snapshots.

VR Interaction The core of the application that ties everything together is the VRSCULPTING script found on the player object. It handles all the interactions between the sculpture, voxels, brushes, queries, UI's and so on.

Cineast API Communication with Cineast is handled through the `UNITYCINEASTAPI` script. A query is started by calling `STARTQUERY` with the model data and once the results are back the `UNITYCINEASTAPI` relays them to the `QUERYRESULTSPAWNER` which handles displaying the results.

4.3.3 Sculpting Tools

Our VR sculpting application provides several tools to enable the user to sculpt.

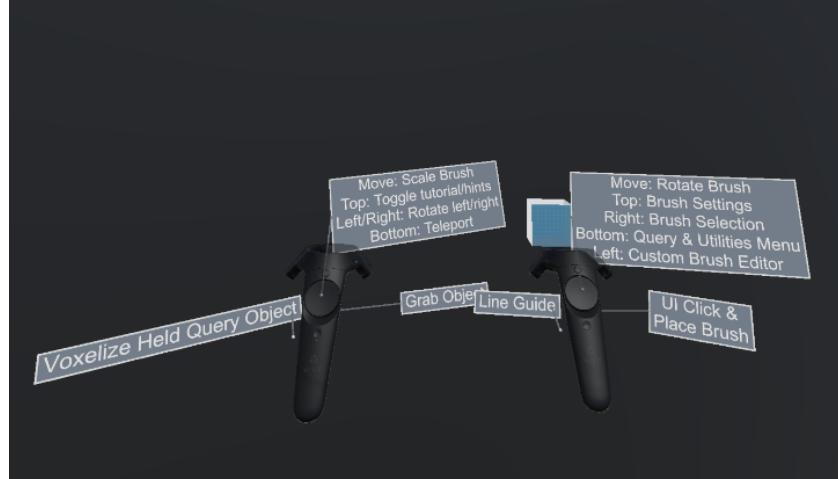


Figure 4.14: Controller hints

Controller Hints The controller hints portrayed in Fig. 4.14 help the user become familiar with the VR sculpting application. They are always shown when the application is started but can then be toggled on or off by pressing a button on the controller.

This controller hints functionality is provided as a script in the SteamVR plugin package, however we modified it to make it possible to specify on which side to place the hints. Before this modification all hints were always placed on the right side of the controller, causing some of them to overlap.

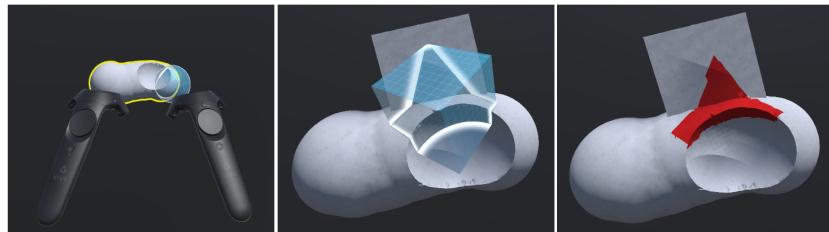


Figure 4.15: Illustration of the sculpting process. Left: modifying a sculpture. Middle: brush preview with intersection highlight. Right: after applying brush in replace mode with a red material.

Brush The most important tool available to the user is the brush. It is attached to the right controller as depicted in Fig. 4.15 and displayed as a translucent preview of the brush's shape. Using a custom shader this brush preview shows the intersection between the brush and

the sculpture. This feature has turned out to be an important addition, because without it it's quite difficult to tell where the brush intersects with the sculpture despite the stereoscopic display of the VR headset.

By holding the brush button, which is by default the index finger trigger on the right controller, the user can drag the brush and continuously apply a brush stroke so long performance allows. The brush size can be changed by swiping up or down on the trackpad of the left controller. Similarly the brush can be rotated by swiping any direction on the trackpad of the right controller. Rotating the brush however takes some practice because the trackpad only provides two degrees of freedom whereas there are three rotation axes for the brush.

Currently our application provides five brush types: cube, sphere, pyramid, cylinder and custom. These brush types are defined from SDF's, thus it is easy to add new brush types. Our VR application provides some Unity editor interfaces to enable a developer to easily add such new brushes to the application.

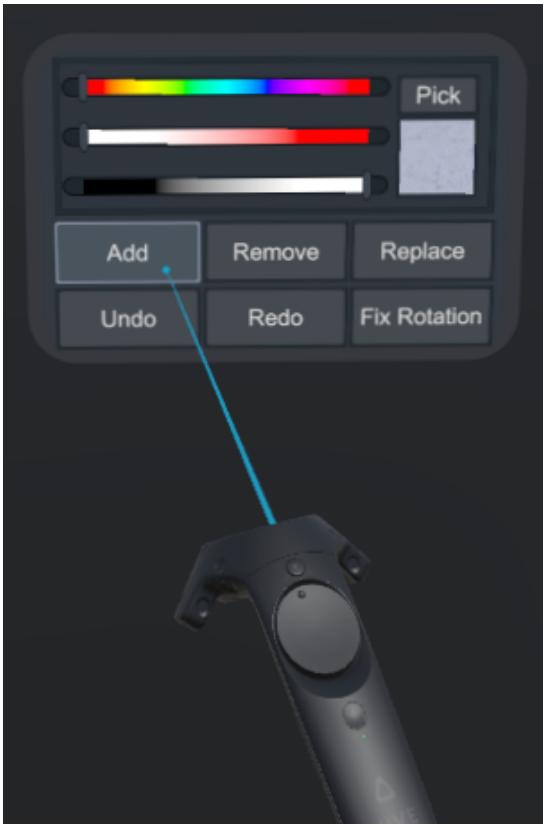


Figure 4.16: Brush properties menu

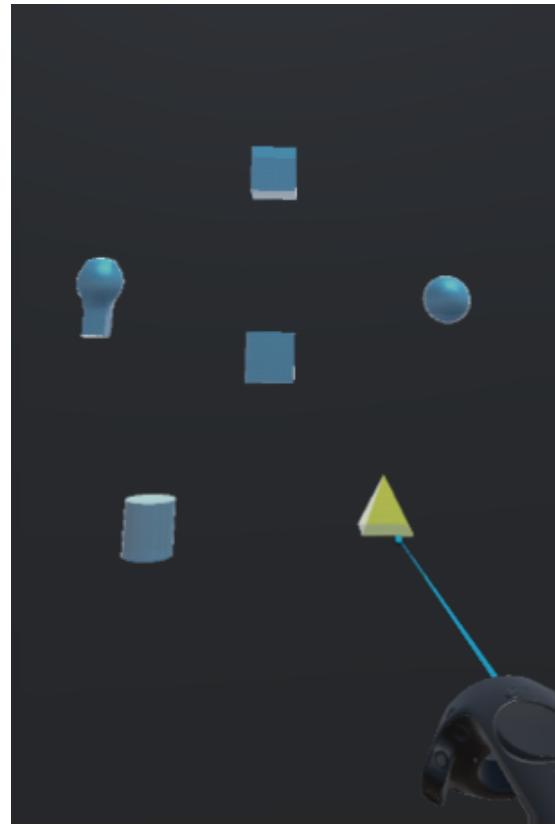


Figure 4.17: Brush selection menu

Brush Settings These brushes are only useful if the user can adjust the brush and its behaviour to their liking. Fig. 4.16 shows the brush properties menu which contains all the brush settings that can be changed.

The user can change the CSG operation using the "Add", "Remove" and "Replace" buttons. During development we decided to replace the CSG terminology "Union" and "Difference" with the more common words "Add" and "Remove" to make it more intuitive for users that do not know anything about CSG.

The three sliders are used to change the color of the brush material. We decided to use this

three slider style Hue/Saturation/Value (HSV) color selection because it immediately presents the user with a large spectrum of colors that can be finely adjusted with the saturation and value, i.e. brightness, sliders. The color preview box on the right shows how the brush material will look with the color applied. Once the sculpture has been colored with a certain color the user can always reselect that color by pressing the "Pick" button and then clicking on the spot of the sculpture with the desired color.

The brush material texture is changed by clicking on the color preview box which opens a submenu and then clicking on another texture in that submenu.

By default, the brush is fully attached to the right controller which means that it also rotates with the controller. If that is not desired then the user can toggle the "Fix Rotation" button. This disables the automatic rotation of the brush with the controller making it easier to for example to create straight lines with the brush.

If the user has made a mistake it can be undone by clicking the "Undo" button. Currently up to ten edits can be undone. Edits that have been undone can be redone by clicking the "Redo" button, however only if the user has not edited the sculpture again since the undo.

The brush selection menu in Fig. 4.17 lets the user select different brushes. It is arranged as a radial menu centered in front of the user's right controller. Like this the user quickly and conveniently select other brush types without having to move the controller much.

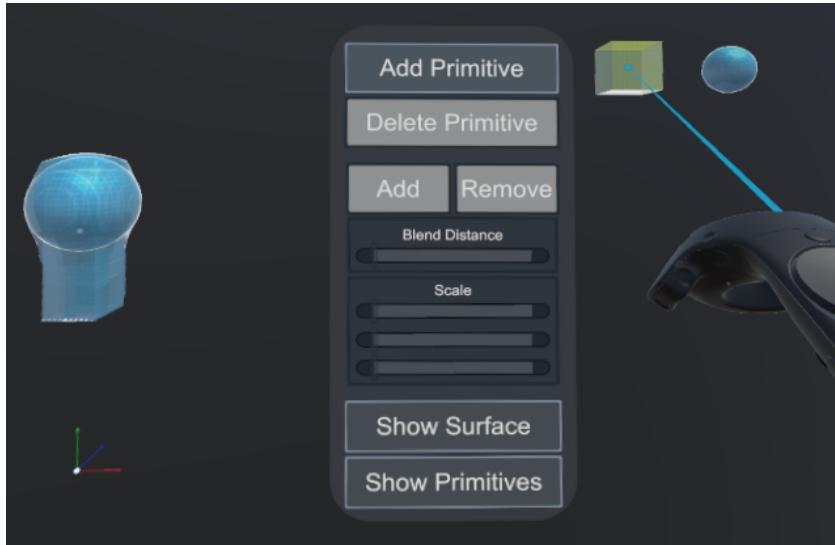


Figure 4.18: The custom brush editing menu used to create custom brush shapes. A custom brush shape composed of a sphere and scaled cube is shown on the left side. The shapes are blended together smoothly resulting in a brush shape that looks like a cuboid that bulges out at the top.

Custom Brushes Our application currently only supports hard boolean CSG, meaning that union and difference operations often lead to sharp edges and corners. This is not always desirable, especially when the user wants to sculpt organic and smooth shapes. One solution to address this would be to implement a smoothing brush that smooths out sharp edges, e.g. by averaging the positions of the surface vertices and adjusting the normals. However doing this is not trivial because the surface vertices are constructed from the underlying voxels, hence the smoothing would instead have to happen in the voxel data instead. This also leads to several edge

cases like when a voxel Hermite data intersection is moved from one voxel to another care must be taken to also adjust the intersections on the other voxel edges such that it doesn't produce a hole in the mesh. The multi-material voxels also pose a problem because the smoothing brush should not disturb or even destroy the material transitions.

To avoid all these troubles we have conceived a proof of concept custom brush that can be shaped by the application's user. As discussed previously SDF's can be smoothly blended together using the smooth minimum and maximum operators. Our custom brushes take advantage of that and enable the user to smoothly blend two or more SDF primitives together to create a brush shape. Currently our application supports combining cube and sphere primitives together, but new primitives can be added easily in almost the same way as it is done for the regular brushes. In addition to that each SDF primitive can be scaled on all three axes independently.

These custom brushes have several use cases: they can be used to create a single primitive shape that can be scaled, e.g. creating an ellipsoid from a sphere primitive, or it can be used to create more complex shapes, such as e.g. a rounded disc by combining one sphere primitive and smoothly subtracting two cube primitives from the top and bottom of the sphere.

4.3.4 VR User Interface System

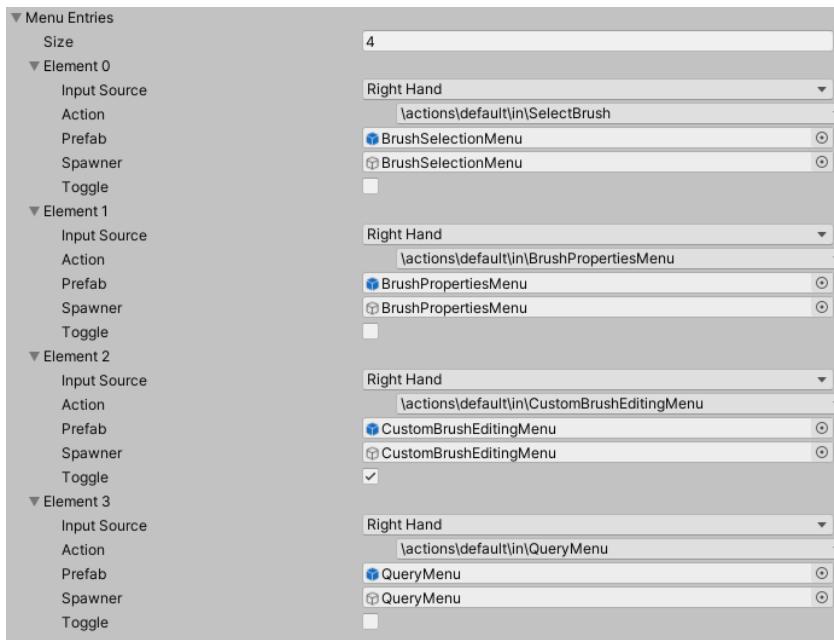


Figure 4.19: Menu entries section in the VRSCULPTING component where developers can register new UI's.

We have implemented a flexible system for VR user interfaces (UI) in Unity. Most of our menus are similar to what you would expect in a regular non-VR application. The reason for this is that most users will already be familiar with the typical window based UI's found in modern operating systems. Furthermore, such window based UI's can densely pack information and functionality while remaining organized and hierarchical, e.g., by using grouping background elements such as the color settings group in Fig. 4.16 to group contextually similar components together.

The UI's are controlled by the VR controller, which has a pointer and cursor attached to it. Since our application uses the SteamVR Plugin, we have designed our pointer and cursor to be

similar to that of the SteamVR menu to maintain consistency. UI components can be interacted with by pressing the interaction button on the VR controllers which by default is bound to the index finger trigger on the right controller.

New UI's can easily be created by creating a new empty object and then adding Unity's CANVAS component. This turns the object into a UI canvas and the developer can add UI elements to the viewport specified in the CANVAS component. Any UI element already available in Unity should work in VR out of the box. This is made possible by our VRPOINTERINPUTMODULE script, which extends Unity's BASEINPUTMODULE and thus replaces Unity's default keyboard and mouse based input controller with our VR pointer input controller. Once a new UI is finished it is converted into a Unity prefab and then registered in order to become usable by the user. This is done in the Unity editor in the VRSCLPTING component on the player object under the "Menu Entries" section. Within this section, depicted in Fig. 4.19, the developer can add a new entry, then set the prefab field to point to the UI prefab, set which controller button should open the UI, whether it is toggled and where it should appear when opened.

4.3.5 Similarity Query

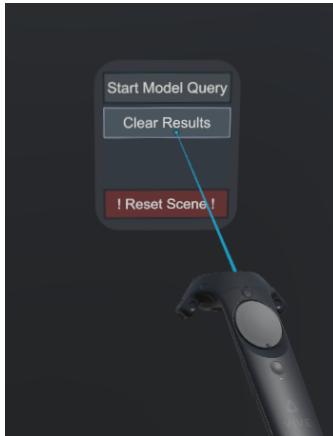


Figure 4.20: Query menu with buttons to start a query, clear query results and to reset the entire scene.

Similarity queries are executed by clicking the "Start Model Query" button in the query menu shown in Fig. 4.20. The application then converts the current sculpture into a single mesh and sends it to Cineast to be processed.

Once the results have arrived back at the VR sculpting application they are displayed in a curved UI centered on the player's head. The intention is to be able to use the entire 360° available in VR to display as many results as possible. Unfortunately due to performance problems while loading the results in and a networking problem in the API communication this is not yet possible, but certainly achievable in the future.

The user can point at query results which makes a 3D model object appear and spin around their Y axis. Additionally this shows their name and the similarity score value on a label above the model. These model objects can be grabbed with the left controller for closer inspection.

Fig. 4.21 portrays the curved query results UI with the user pointing at a result object. When held in the left controller the 3D model objects can be voxelized, i.e. converted into a voxel sculpture like in Fig. 4.22. This sculpture is then fully modifiable using the sculpting tools and

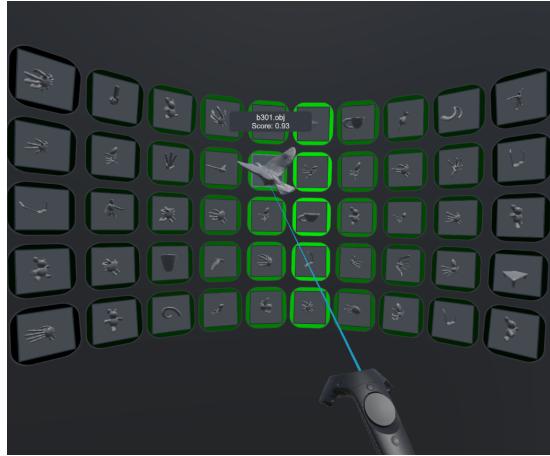


Figure 4.21: Query results UI showing the results of a similarity query.

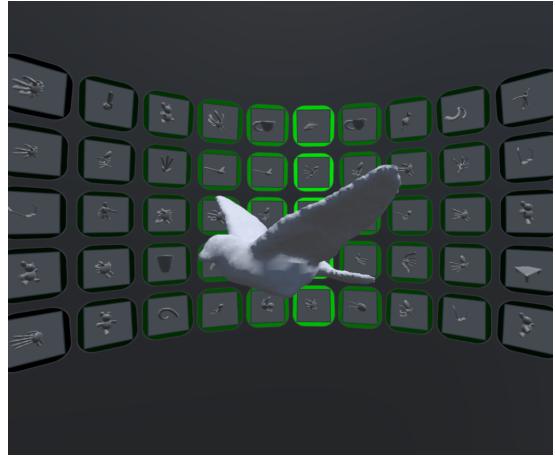


Figure 4.22: A voxelized and fully editable model from the query results UI.

can be used to create a new query. This allows the user to start with a rough 3D sketch to perform an initial query and then successively refine the query until the user finds the desired object.

5

Evaluation

5.1 Technical Evaluation

This section is concerned with the technical performance of our application. These results should provide insight into how the application performs to related work and reveal potential problems that may need to be improved upon in future work.

5.1.1 ClusterD2+Color Performance

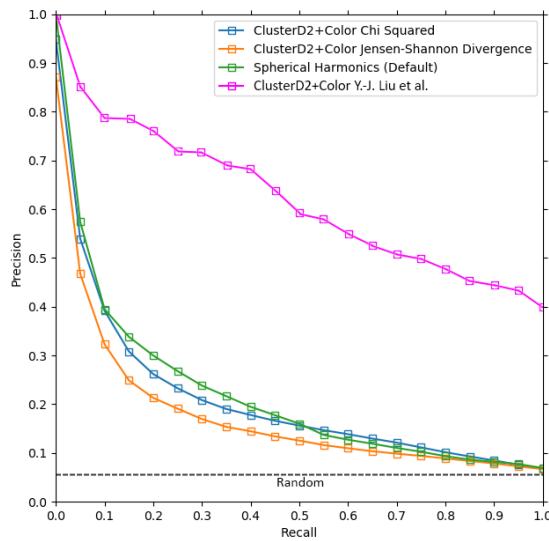


Figure 5.1: Precision/recall plots comparing the performance of Y.-J. Liu et al.’s (pink) and our implementation (blue, orange) of the ClusterD2+Color algorithm using the McGill 3D shape benchmark [29] and Cineast’s ”Spherical Harmonics (Default)” feature module (green). The random scoring baseline is calculated as the average of $\frac{P}{P+N}$ over all classes where P is the number of positives and N the number of negatives.

There are several ways to compare the performance and retrieval quality of different algorithms. One way to do so are precision/recall plots as shown in Fig. 5.1. The larger the area under the curve the better the algorithm performs. Our precision/recall plots have been computed in the same way as described by Y.-J. Liu et al. [23] and also using the same model benchmark dataset. As can be seen in Fig. 5.1 our implementation of the ClusterD2+Color algorithm with both the

Jensen-Shannon divergence distance metric, which Y.-J. Liu et al. used, and the χ^2 distance metric perform significantly worse than that of the original authors of the algorithm. This begs the question whether our implementation is faulty. Visual inspection, using our feature extraction visualization tool, however shows no obvious errors. Both the shape and color feature samples are concentrated where one would expect and the cluster assignments also look sensible. The code that generates the histogram feature vector from the clustered samples is simple and short so we deem it unlikely that it contains any errors.

One possible explanation could be the many parameters of the algorithm that can be tweaked. Several of these parameters were not specified by Y.-J. Liu et al. and thus had to be chosen and adjusted by us empirically and may not be optimal.

On the other hand our ClusterD2+Color implementation performs similarly to Cineast's "Spherical Harmonics (Default)" feature module, whereas Y.-J. Liu et al.'s is surprisingly vastly better. So perhaps the discrepancy between our and Y.-J. Liu et al.'s results could be caused by a difference in the McGill 3D shape benchmark model dataset [29] or in the way the precision/recall plots are computed. Finally, there's also the possibility that the result in [23] is incorrect.

Missing: Query time

5.1.2 Voxel Polygonizer Performance

Time for an average chunk? Chunk size vs. time?

5.1.3 Voxelizer Performance

Before vs. after optimization. Plot tri count vs. time taken.

5.2 User Evaluation

5.2.1 Structure

The user evaluation consists of 11 tasks, grouped into three categories, that the participant should complete within a certain time limit. The first category with two tasks is about the user's background with virtual reality and 3D sculpting applications. The next category with six tasks concerns the usability of the sculpting feature. The remaining three tasks are about the similarity search feature.

Before any of the tasks the user was introduced to the topic and was handed the questionnaire to gain an understanding of the user evaluation's structure.

The tasks are intended to start out simple and then increase in complexity.

The first hands-on task is for the user to read the controller hints, as depicted in Fig. 4.14, to get familiar with the control scheme. During the first few runs it was decided that the user may also try out the controls during this first task.

Instead of having to take off the VR headset after every task, the tasks were read to the participant by the user evaluation supervisor and the questionnaire was filled out after all tasks had been completed.

Our questionnaire uses a five-point Likert scale, which is the most common approach to capturing data for surveys. These scales offer two or more choices, where each end of the scale corresponds to opposing sentiments or ratings.

The complete questionnaire can be found in the Appendix.

5.2.2 Sample Size

In total our user evaluation had 14 participants. Nine of which claim to be moderately or better experienced with virtual reality. Three participants claim to be moderately or better experienced with 3D sculpting applications.

This sample size is not large or diverse enough to produce a statistically significant result, however it still allowed us to collect valuable feedback and identify shortcomings in the user interface design or tools of our VR sculpting application.

5.2.3 Results Overview



Figure 5.2: The distribution of ratings for each task of the user evaluation.

Deducing from the ratings listed in Fig. 5.2 and the feedback in Section 5.2.4 the application was perceived as positive and easy to use for the most part. However, it needs to be noted that many participants had stated, while trying out the application, that it became intuitive once they had figured out the controls to use the tools. This means that it may be a bit less intuitive than the final ratings portray.

Table 5.1: Average rating of each task rounded to three significant digits.
Same scales apply as in Fig. 5.2.

Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	Task 7	Task 8	Task 9	Task 10	Task 11
2.86	1.79	1.15	1.36	1.5	1.36	2	2.43	1.36	2.07	1.36

According to Table 5.1 none of the practical tasks, 3 through 11, were rated 3 or higher on average. This means that, on average, they were all considered very easy, easy or somewhere between easy and neutral. Task 8, which had users create a custom brush shape using the custom brush editor, was clearly considered the most difficult, which can be attributed to its partially unintuitive user interface and performance issues. Task 10, which had the users voxelize a query result, was also considered a difficult task because the grip buttons of the VR controllers are difficult to find and

uncomfortable use. Lastly task 7, which was about changing the texture/material of the brush, was also rated comparatively difficult because the participants had difficulty finding the correct button on the user interface.

5.2.4 Feedback

Tasks one and two, which ask how much experience the participants has with virtual reality and 3D sculpting applications, will be skipped here since there is no feedback from those.

Some of the comments have been edited in square brackets by us to clarify the intent of the comment to the reader.

P1 through P14 refer to the anonymized user evaluation participants.

Task 3

Description: "Read the controller hints to become familiar with VR and the control scheme. Disable the controller hints once you're ready. Please note that you can activate these hints again at any time if necessary."

Feedback:

- P1: Put the arrows for the side buttons [of the trackpad] at the exact position. Picture for direction field [trackpad] (for easier left/right/top/bottom) if possible
- P2: The hints on the buttons were tiny bit large where required me to take them away to read them, maybe smaller font size?
- P4: For the buttons at the side [grip buttons], it is not immediately clear, that they belong to the voxelize interaction
- P5: Appears very cluttered at first
- P13: It was sometimes difficult to get the VPad [trackpad] to accept the direction I wanted (e.g. it showed query settings [menu] instead of brush [menu])

One participant did not rate this task because they found it unclear what to rate.

As depicted in Fig. 4.14 the controller hints for the trackpad all point to the center of the trackpad. At this point in time the Unity SteamVR plugin does not provide a way to get them to point to the sides of the trackpad, but this problem could be fixed in the future since the source code of the scripts responsible for said feature can be fully edited.

Task 4

Description: "Select the sphere brush and place it the world to create a shape or simple sculpture using the 'Add' mode."

Feedback:

- P2: Once realizing the selection method, it's easy to continue & apply
- P9: Rethink menu titles/labels
- P11: It's confusing the first time and gets very intuitive after the 5th time using
- P13: Well done! Feature request: point to point fill (e.g. for straight legs)

Task 5

Description: "Select a brush and create a shape, then select another brush and remove a piece of your sculpture with it using the 'Remove' mode."

Feedback:

- P2: The only difficulty of this task is the trackpad sensitivity which activated different menus [than the one intended].
- P9: Rethink menu titles/labels
- P13: Well done, this is very intuitive. Request: "Smooth surface" [tool or brush]

Task 6

Description: "Select a brush and create a shape, then pick another colour and colour a piece of your sculpture using the 'Replace' mode."

Feedback:

- P2: The saturation selection is not the most intuitive color selection. It's better you have saturation on a value before [the participant probably meant that the default value should be 1 instead of 0]
- P9: Rethink menu titles/labels
- P12: Replace makes sense when considering voxels, but some people may not realize this can be used to color
- P13: Request: Define groups [on the sculpture] and then a "colour entire group" option (e.g. all four legs of a stool)
- P14: Use other word than replace

It was observed that a few users had trouble selecting a color because the saturation value is by default at 0, hence without changing that value the color of the preview square does not change at all. An improvement for this was implemented after the use evaluation. Now, when the saturation value is close to 0 and the user changes the hue, the saturation value is automatically set to 1.

Task 7

Description: "Select a brush and create a shape, then pick another material (i.e. texture) and change the material of a piece of your sculpture using the 'Replace' mode."

Feedback:

- P1: I thought I have to click the button above/below[?] the texture field [the participant probably means the "Pick" button]
- P2: I would suggest there is a different square for color & texture. Took me a few seconds to realize I need to click on the color square.
- P5: Easy if you know how. However, the texture selection is somehow hidden.
- P9: Rethink menu titles/labels
- P10: Could use a label for texture
- P11: Textures are too hidden
- P13: Bug/inconvenient: it remembers the color selected → unexpected color of material + unclear where to select material.
- P14: Difficult to find

It is evident from the comments and ratings that the texture/material selection button was difficult to find for many users. One reason for this is that the default texture is a marble texture that is almost entirely white and thus it is easily missed that the color/texture preview square also shows a texture. This problem was addressed by labeling the color/texture preview square like a regular button in a revision after the user evaluation took place.

Task 8

Description: "Create your own brush (with at least two primitives) using the custom brush editor and then use your own brush to create a shape."

Feedback:

- P1: More difficult to handle than the previous tasks
- P2: The middle point (white dot) [center of custom brush] is not very visible. Maybe limiting the space the brush needs to be made in [that was already the case, but probably too large since the participant did not notice], or centering [centering] the shape afterwards would be more intuitive.
- P4: It is not very clear how you switch between different components of the texture [participant probably meant brush instead of texture]
- P5: That part of the UI is not very intuitive and involves a lot of buttons. Maybe implement as a dedicated "mode".
- P11: Kind of hidden
- P13: Not that intuitive. Improvements in: snapping to center, rotate [to] fixed value etc...
→ tried to get a triangle shape and that's hard.

The ratings of this task were quite scattered. This is probably the case because some participants tried to make a simple shape and succeeded, whereas others had a more complex shape in mind and then failed due to the unintuitive or lacking tools of the custom brush editor. These results show that there is still a lot of work required in order to turn this feature into something usable. Another major issue was also the performance drops while using the custom brush editor.

Task 9

Description: "Select a brush and create a shape, then use the query menu to run a similarity search."

Feedback:

- P2: Except the framerate drop [when loading in the query results to be displayed] which feels like a VR earthquake it's very easy.
- P5: Querying takes a long time.

This task was generally perceived as simple. Several participants noted that the query seems to take a long time. Furthermore for some participants there was a significant framerate drop while the query results were being loaded in. This is definitely something to be improved in the future.

Task 10

Description: "Select a brush and create a shape, then use the query menu to run a similarity search. After that, pick one of the results and voxelize it into the world."

Feedback:

- P2: Button nearly impossible to find. I recommend against using the side buttons for such important interactions
- P4: Buttons are chosen suboptimally
- P5: Grip feels very uncomfortable to me.
- P11: It's hard to figure out that you have to grab the item from very near, when sculptures can be grabbed from farther away

Nearly all participants had trouble finding the correct button on the left VR controller. These grip buttons are intended to be pressed by squeezing ones hand or by pressing them with the pinky finger. However, unless one already knows these buttons they're very difficult to find since they don't feel or look like buttons. A solution for this is proposed in Section 6.2.

Task 11

Description: "Select a brush and create a shape, then use the query menu to run a similarity search. After that, pick one of the results and voxelize it. Using a brush, remove a piece of it and then run a similarity search for the modified sculpture."

Feedback:

- P1: Only voxelize was difficult

General Feedback

Description: "If you have any additional remarks or suggestions for improvements please write them down here."

Feedback:

- P1: I needed some time to understand the program but then it was relatively easy to use. A visual info about the current mode (add/remove/etc.) would be helpful (adapt color of brush?)
- P2: It's a very well done demo, quite entertaining & interesting. The only drawback of using complex & memory consuming applications in VR is the very visible & felt framerate drop which affect the experience. Other than that, the implementation is very smooth. Well done!
- P4: Sculpting in general felt easy and intuitive. The evaluation might have benefited from specifying which object the user should sculpt. Good job, impressive result!
- P5: Very cool tool! More results would be nice. Sometimes performance drops considerably.
- P6: It's a very good project. It would be nice to add some physics and [user created] animations to the model.
- P8: The assignment of the functions to the buttons can be improved. Allow the modelling of a second sculpture while the first one is queried. Add more functions which help modelling, similar to the line functionality. (that could also draw for you once you have aligned the line).
- P9: Change brush width, height, depth [of regular brushes, not custom brushes]. Works well and is quite easy to understand. Fun to use!
- P10: Thought it was excellent. Easy to use, the [controller] hints were incredibly useful and can definitely see the potential.
- P11: The whole experience felt surprisingly real and immersive. "Mouse clicks" [from the VR controller pointer] are a thing I had to get used to.
- P13: Great demo/prototype! Apparently, there are some minor differences during feature extraction, as the stool was not found using it [voxelized stool] as query, but returned almost always as if something other was queried for. Request: Add larger "brushes" for weights → prioritise the shape of a certain part... [the participant probably means that there should be a tool to select certain parts of the sculpture such that brushes only (or mostly) affect those selected parts]
- P14: Better manual to know where the textures are. Voxelizing is a little tricky.

The suggestion of P1, changing the color of the brush depending on the brush mode, was implemented after the user evaluation. In the union mode the brushes remain the default blue color, in remove mode they become red and in replace mode green. Overall a lot of useful feedback was provided. Several interesting ideas have been brought up and are mentioned later in Section 6.2.

6

Discussion

6.1 Conclusion

The goal of this bachelor thesis was to implement a VR sculpting application enabling users to sculpt their own figures which can then, in conjunction with a multimedia retrieval application, be used to search a database for similar models.

Our VR sculpting application provides a brush tool with which the user can add or remove material from a fully volumetric interactable sculpture. Additionally the brush can take different shapes, including custom shapes made by the user, and can be used to color or texture the sculpture. The undo/redo functionality allows the user to undo mistakes and keeps them from accidentally damaging their sculpture.

User made sculptures can be used to execute a similar search in a database connected to the Cineast application. The search results are presented to the user in 3D in the virtual scene. These result 3D models can be grabbed by the user for close up inspection or to convert them into a voxel representation to refine their similarity query. The new ClusterD2+Color feature module, based on the work of Y.-J. Liu et al. [23], enables Cineast to execute similarity queries that take the color and texture of 3D models, and thus also our sculptures, into consideration, which was previously not the case.

Our main goal was achieved and the result is a functional VR sculpting application with multimedia retrieval capabilities and more mature user interfaces and tools compared to the precursor prototype mentioned in Section 1.2. As is often the case there are several things left that can be improved, which will be discussed in Section 6.2.

6.2 Outlook

While our VR sculpting application already provides a solid base to work off of there are still many improvements required for it to become a mature sculpting application.

6.2.1 Performance

The main problem currently is the performance of the sculpting and rendering. The sculpture mesh does not use any level of detail (LOD) system or adaptive data structure and as such the resulting meshes have very high triangle counts. Eventually when a sculpture becomes large enough the framerate will drop below the desired 120Hz for VR. This problem could be addressed

by using an adaptive data structure like Octrees to reduce the complexity of the meshes and only keep the details where necessary. This approach has already been tried and discussed in [17]. In addition to improving rendering performance this would also greatly reduce the duration of queries, because it requires less data to be sent over the network and to be processed.

Closely related is also the memory usage of sculptures which could be reduced by using adaptive data structures and compression schemes. A commonly used compression scheme for voxels is Run Length Encoding (RLE), where instead of storing each voxel in an array the program stores runs of voxels. Each run is described by the length of the run and the single voxel data. In most cases this leads to a good compression because large parts of sculptures are uniform, i.e. contain the same voxel data.

6.2.2 Usability

Sculpting The sculpting module currently misses some features that would increase usability of our application by a lot.

Our brushes are based entirely on SDF's and while those are quite flexible a user may instead want to use a mesh as brush shape. The voxelizer module could be used to compute the voxel representation of such a mesh, store it in a separate voxel grid, and then copy that voxel grid to the sculpture when the brush is being applied.

Currently only one sculpture can exist at once. Technically multiple sculptures can exist at once, but our application does not yet offer the interaction functionality yet to select which sculpture the user wants to edit. Furthermore when adding multiple sculptures it would make sense to implement an algorithm that is able to detect when a sculpture has multiple disjoint parts so that they can each be converted into separate sculptures.

One participant of the user evaluation has suggested to add a brush mode where the user selects two or more points to form a line or polygon and the application then automatically fills in that shape. This feature would be very useful to construct more precise shapes.

Another suggestion was to implement a form of weight painting. The weight painting tool could be used to give each voxel of the sculpture a certain editing weight. Regular brushes would then only affect voxels directly proportional to their editing weight. Like this the user could selectively only edit certain parts and influence the area of effect of the brush.

Custom Brush Editor While the custom brush editor is functional it is not yet useful to create brush shapes with more than three or four primitives. The performance decreases with every primitive added to the brush because whenever a primitive is changed the entire custom brush shape needs to be recomputed. Secondly the editor interface is not yet intuitive and users have found it difficult to edit their custom brushes. One reason for this was the confusion between the "Add/Remove" buttons, which changes the selected primitive's CSG operation, and the "Add/Delete Primitive" buttons, which add or remove a primitive from the custom brush. This could be fixed by using better wording and appropriate icons. Selecting a primitive was also perceived as difficult because it requires the user to point at the center of the primitive with the right controller and then click the trigger, in the same way one would click a button. It would make more sense to change selecting a primitive to be done with the left controller similarly to grabbing an object.

Saving/Loading There is currently no way to persistently save a sculpture or custom brush. It would make sense for the application to be able to export sculptures or custom brushes into a file on disk so the user can later load them again.

Queries Due to some performance issues while loading results and a networking problem the number of results shown in the query results UI is currently limited to roughly 50 results. The networking problem is caused by the C# Cineast API wrapper sending too many REST requests and the RestSharp¹⁴ library responsible for sending the requests retains the socket connections open for a while instead of immediately closing them¹⁵. This causes the machine to run out of sockets and then refuses to send any further REST requests. The solution would be to use a socket pool and to specify a maximum number of requests that can be in process at once.

Controls It was mentioned several times during the user evaluation that the current control scheme is not optimal, especially the usage of the grip buttons which are pressed by squeezing ones hand or pressing them with the pinky finger. As depicted in Fig. 4.14 nearly all the controller buttons have been assigned a function, including the difficult to use side grip buttons of the controller. Many participants had trouble finding these side grip buttons and then use them as intended. These problems could be resolved by introducing multiple control schemes or modes that can be switched between while using the application. There could be a mode for just editing the sculpture, including brush controls and the two brush UI's, a mode for just the queries, and so on. This would reduce the amount of buttons required at any one point because the controls are distributed over the different modes.

¹⁴ <https://github.com/restsharp/RestSharp>

¹⁵ <https://github.com/restsharp/RestSharp/issues/1322>

Bibliography

- [1] Cie recommendations on uniform color spaces, color-difference equations, and metric color terms. *Color Research & Application*, 2(1):5–6, 1977.
- [2] Mihai Aldén. Fast and adaptive polygon conversion by means of sparse volumes. 2011.
- [3] Ding-Yun Chen, Xiao-Pei Tian, Yu-Te Shen, and Ming Ouhyoung. On visual similarity based 3d model retrieval. *Computer Graphics Forum*, 22(3):223–232, 2003.
- [4] DBIS Group of the University of Basel. Cineast, 2020. URL <https://github.com/vitrivr/cineast>.
- [5] DBIS Group of the University of Basel. Cottontail db, 2020. URL <https://github.com/vitrivr/cottontaildb>.
- [6] DBIS Group of the University of Basel. Databases and information systems group, 2020. URL <https://dbis.dmi.unibas.ch/>.
- [7] Mathieu Desbrun, Mark Meyer, Peter Schröder, and Alan H. Barr. Implicit fairing of irregular meshes using diffusion and curvature flow. page 317–324, 1999.
- [8] R. T. Fielding. Architectural styles and the design of network-based software architectures, 2000. URL <https://www.ics.uci.edu/~fielding/pubs/dissertation/top.htm>.
- [9] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, B. Qian Huang, Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker. Query by image and video content: the qbic system. *Computer*, 28(9):23–32, 1995.
- [10] Sarah Frisen. Constrained elastic surface nets: Generating smooth surfaces from binary segmented data. 07 2001.
- [11] R. Gasser. Towards an all-purpose, content-based multimedia information retrieval system, 2017. URL <https://dbis.dmi.unibas.ch/teaching/studentprojects/towards-an-all-purpose,-content-based-multimedia-information-retrieval-system/>.
- [12] Sören Grimm, Stefan Bruckner, Armin Kanitsar, and Eduard Gröller. A refined data addressing and processing scheme to accelerate volume raycasting. *Computers and Graphics*, 28(5):719 – 729, 2004. ISSN 0097-8493.
- [13] Chien-Chang Ho, Fu-Che Wu, Bing-Yu Chen, Yung-Yu Chuang, and Ming Ouhyoung. Cubical marching squares: Adaptive feature preserving surface extraction from volume data. *Comput. Graph. Forum*, 24:537–545, 09 2005.

- [14] Piotr Indyk and Rajeev Motwani. Approximate nearest neighbors: Towards removing the curse of dimensionality. *Conference Proceedings of the Annual ACM Symposium on Theory of Computing*, 604-613, 10 2000.
- [15] OpenAPI Initiative. The openapi specification: a broadly adopted industry standard for describing modern apis., 2020. URL <https://www.openapis.org/>.
- [16] Andrew E. Johnson and Martial Hebert. Using spin images for efficient object recognition in cluttered 3d scenes. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 21(5):433–449, 1999.
- [17] Tao Ju, Frank Losasso, Scott Schaefer, and Joe Warren. Dual contouring of hermite data. *ACM Trans. Graph.*, 21(3):339–346, July 2002. ISSN 0730-0301.
- [18] Leif Kobbelt, P. Schrder, Michael Kazhdan, Thomas Funkhouser, and Szymon Rusinkiewicz. Rotation invariant spherical harmonic representation of 3d shape descriptors. *Proc 2003 Eurographics*, vol. 43, 07 2003.
- [19] Leif P. Kobbelt, Mario Botsch, Ulrich Schwanecke, and Hans-Peter Seidel. Feature sensitive surface extraction from volume data. *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques*, page 57–66, 2001.
- [20] Chang Ha Lee, Amitabh Varshney, and David W. Jacobs. Mesh saliency. *ACM Trans. Graph.*, 24(3):659–666, July 2005. ISSN 0730-0301.
- [21] E. Lengyel. The transvoxel™ algorithm, 2009. URL <https://transvoxel.org/>.
- [22] D. Lingrand, A. Charnoz, R. Gervaise, and K. Richard. The marching cubes, 2002. URL <http://users.polytech.unice.fr/~lingrand/MarchingCubes/algo.html>.
- [23] Yong-Jin Liu, Yi-Fu Zheng, Lu Lv, Yuming Xuan, and Xiaolan Fu. 3d model retrieval based on color + geometry signatures. *The Visual Computer*, 28:75–86, 01 2012.
- [24] William Lorensen and Harvey Cline. Marching cubes: A high resolution 3d surface construction algorithm. *ACM SIGGRAPH Computer Graphics*, 21:163–, 08 1987.
- [25] Robert Osada, Thomas Funkhouser, Bernard Chazelle, and David Dobkin. Shape distributions. *ACM Trans. Graph.*, 21(4):807–832, October 2002. ISSN 0730-0301.
- [26] I. Quilez. 3d sdf functions, 2009. URL <https://iquilezles.org/www/articles/distfunctions/distfunctions.htm>.
- [27] L. Rossetto. Cineast: a content-based video retrieval engine, 2014. URL <https://dbis.dmi.unibas.ch/teaching/studentprojects/cineast--a-content-based-video-retrieval-engine/>.
- [28] G. Salton, A. Wong, and C. S. Yang. A vector space model for automatic indexing. *Commun. ACM*, 18(11):613–620, November 1975.
- [29] Shape Analysis Group of the McGill University. Mcgill 3d shape benchmark, 2005. URL <http://www.cim.mcgill.ca/~shape/benchMark/index.html>.
- [30] Stanford Computer Graphics Laboratory. The stanford 3d scanning repository, 2014. URL <http://graphics.stanford.edu/data/3Dscanrep/>.

-
- [31] A. Tatarinov and A. Pantelev. Advanced ambient occlusion methods for modern games, 2016. URL https://developer.nvidia.com/sites/default/files/akamai/gameworks/downloads/papers/vxao/atatarinov_alpantelev_advanced_ao.pdf.
 - [32] G. Taubin. Estimating the tensor of curvature of a surface from a polyhedral approximation. *Proceedings of IEEE International Conference on Computer Vision*, pages 902–907, 1995.
 - [33] Sinje Thiedemann, Niklas Henrich, Thorsten Grosch, and Stefan Müller. Voxel-based global illumination. In *Symposium on Interactive 3D Graphics and Games*, I3D ’11, page 103–110, New York, NY, USA, 2011. Association for Computing Machinery. ISBN 9781450305655.
 - [34] Dejan V. Vranic. 3d model retrieval. 2001.
 - [35] Wikipedia contributors. Voxel — Wikipedia, the free encyclopedia, 2020. URL <https://en.wikipedia.org/w/index.php?title=Voxel&oldid=964604373>.
 - [36] Wikipedia contributors. Wavefront .obj file — Wikipedia, the free encyclopedia, 2020. URL https://en.wikipedia.org/w/index.php?title=Wavefront_.obj_file&oldid=962279987.

A

Appendix

A.1 Example Images

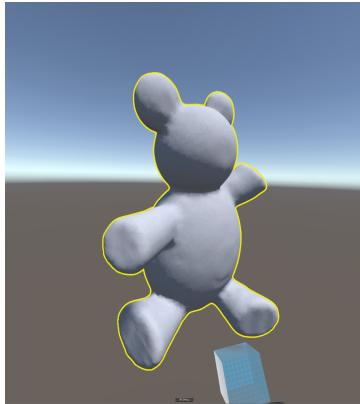


Figure A.1: A query result 3D model that has been voxelized into a sculpture.



Figure A.2: The same sculpture from A.1 but after several edits.

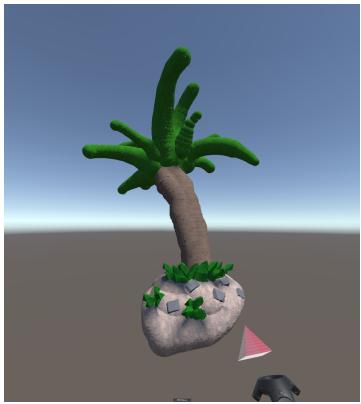


Figure A.3: A tree sculpture made in our VR sculpting application.

A.2 User Evaluation Questionnaire

The purpose of this user evaluation is to gather insight and feedback about the current implementation of the sculpting and querying functionality and the impression it has on users. The results will be helpful in identifying shortcomings and to improve the user experience.

If at any point you feel unwell or nauseous while using the VR headset please let me know. That can be a common reaction when not used to VR or when the program is not responsive enough. Each time before moving on a next task please press the ”! Reset Scene !” button in the ingame ”Query & Utilities Menu” to reset the scene.

Background

1: not at all, 2: slightly, 3: moderately, 4: very, 5: extremely

1. How experienced are you with Virtual Reality?	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
2. How experienced are you with 3D sculpting applications?	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5

Sculpting

1: very easy, 2: easy, 3: neutral, 4: difficult, 5: very difficult

- 3.** Read the controller hints to become familiar with VR and the control scheme. Disable the controller hints once you're ready. Please note that you can activate these hints again at any time if necessary.

Time limit: 5min.

1	2	3	4	5
---	---	---	---	---

Feedback:

- 4.** Select the sphere brush and place it the world to create a shape or simple sculpture using the 'Add' mode.

Time limit: 5min.

1	2	3	4	5
---	---	---	---	---

Feedback:

- 5.** Select a brush and create a shape, then select another brush and remove a piece of your sculpture with it using the 'Remove' mode.

Time limit: 5min.

1	2	3	4	5
---	---	---	---	---

Feedback:

- 6.** Select a brush and create a shape, then pick another colour and colour a piece of your sculpture using the 'Replace' mode.

Time limit: 5min.

1	2	3	4	5
---	---	---	---	---

Feedback:

- 7.** Select a brush and create a shape, then pick another material (i.e. texture) and change the material of a piece of your sculpture using the 'Replace' mode.

Time limit: 5min.

1	2	3	4	5
---	---	---	---	---

Feedback:

- 8.** Create your own brush (with at least two primitives) using the custom brush editor and then use your own brush to create a shape.

Time limit: 8min.

1	2	3	4	5
---	---	---	---	---

Feedback:

Querying

1: very easy, 2: easy, 3: neutral, 4: difficult, 5: very difficult

- 9.** Select a brush and create a shape, then use the query menu to run a similarity search.

Time limit: 6min.

1	2	3	4	5
---	---	---	---	---

Feedback:

- 10.** Select a brush and create a shape, then use the query menu to run a similarity search. After that, pick one of the results and voxelize it into the world.

Time limit: 6min.

1	2	3	4	5
---	---	---	---	---

Feedback:

- 11.** Select a brush and create a shape, then use the query menu to run a similarity search. After that, pick one of the results and voxelize it. Using a brush, remove a piece of it and then run a similarity search for the modified sculpture.

Time limit: 8min.

1	2	3	4	5
---	---	---	---	---

Feedback:

- 12.** That was all, thank you! If you feel like playing around some more with the program feel free to do so for a couple more minutes. On the next page you can give general feedback.

Time limit: 5min.

General feedback

- 13.** If you have any additional remarks or suggestions for improvements please write them down here.

Declaration on Scientific Integrity

Erklärung zur wissenschaftlichen Redlichkeit

includes Declaration on Plagiarism and Fraud
beinhaltet Erklärung zu Plagiat und Betrug

Author — Autor

Samuel Börlin

Matriculation number — Matrikelnummer

16-051-716

Title of work — Titel der Arbeit

3D Model Retrieval Using Constructive Solid Geometry in Virtual Reality

Type of work — Typ der Arbeit

Bachelor thesis

Declaration — Erklärung

I hereby declare that this submission is my own work and that I have fully acknowledged the assistance received in completing this work and that it contains no material that has not been formally acknowledged. I have mentioned all source materials used and have cited these in accordance with recognised scientific rules.

Hiermit erkläre ich, dass mir bei der Abfassung dieser Arbeit nur die darin angegebene Hilfe zuteil wurde und dass ich sie nur mit den in der Arbeit angegebenen Hilfsmitteln verfasst habe. Ich habe sämtliche verwendeten Quellen erwähnt und gemäss anerkannten wissenschaftlichen Regeln zitiert.

Basel, June 29th 2020

Signature — Unterschrift