

1 Introduction

There has recently been a significant increase in the number of 2D and 3D shapes, and corresponding developments have progressed our understanding of measuring, classifying, and retrieving these shapes. The problem of shape is significant for several reasons, both practical and theoretical: because shape is a uniquely discerning attribute for visual understanding (compared to e.g. color, texture, scale), it has many potential applications from 3D digital design to robot vision. Thus, shape is a fundamental problem in computer vision and shape retrieval is an important piece of the quest to make machines intelligent.

The framework of shape retrieval is generally to take an unknown 2D and 3D shape known as the *query* and return a list of *results*, a series of known shapes (correspondingly 2D or 3D) in order of similarity. This is generally done solely with the geometry of the shape, not with textual metadata or other contextual information.

Practically, shape retrieval is generally broken up into three stages: shape representation, feature representation, and retrieval mechanics.

The shape representation can be 2D points, a 3D point cloud, or a 3D mesh. This shape representation undergoes preprocessing which extracts information and transforms the shape into a feature representation that can be subsequently used as a signature for clustering similar objects. This feature representation/signature is then operated on by retrieval mechanics, which stores the initial database of known shapes and, given an unknown query, retrieves the actual list of ordered known shapes.

We initially used the Laplace-Beltrami Operator signature and wavelet density estimation for our feature representation. This mapping (and, in general, our framework of wavelet density estimation) results in a useful interpretation as geometry on a unit hypersphere manifold. We investigated the best way to cluster on the multidimensional unit hypersphere's curved manifold, we executed an experiment with sample data constructed using the Von-Mises distribution (a generalization of the Gaussian, or normal, distribution, on a circle) on a sphere in 3D. We then verified the veracity and practicality of an implementation of spherical k -means on this sample data.

As an implementation detail, we investigated two possible implementations of the mean of points on a hypersphere. The first is a mathematically rigorous implementation of a differential geometry concept called the Karcher mean, while the second is an estimation of this that exploits the structure of the unit hypersphere. We verified that the two were effectively equal for our purposes.

After verifying that the clustering algorithm works as expected, we used it as a subroutine in our spherical hierarchical clustering investigation. The idea is to form hierarchies which capture different abstraction layers of the shapes, which may speed up the process of shape retrieval. We performed another experiment to decide between agglomerative and divisive hierarchical clustering. We tested its runtime, storage, and overall performance of both approaches. We also analyzed both algorithms using standard complexity theory techniques. Finally, we tested these on the SHREC 2011[?] and MPEG-7[?] datasets, which are standard shape datasets. This demonstrates that the hierarchical retrieval works as intended, and requires only $O(\log n)$ for a single query shape, instead of $O(n)$.

We then turn our attention to optimizing our feature representation. There was a significant amount of optimization to our wavelet density estimation which enabled a multiple-order-of-magnitude increase in speed across all dimensions and resolutions. This was possible by exploiting the grid-like structure of the wavelet density estimation process and utilizing parallelization. This enables the processing of much larger datasets, including the Princeton ModelNet datasets (the full version which contains 127,915 CAD models).

Finally, we investigated a clever optimization of the shape dissimilarity metric (the hypersphere distance metric) by first warping closer the shapes being compared. The intuition is that the “difficulty” of warping similar shapes is much smaller than warping two different shapes, and such a process will make the distance between two similar shapes smaller without affecting the distance between different shapes. This process utilizes a solution to classic combinatorial optimization problem of *linear assignment*, which proves key for this warping to be effective and tractable. We then test this modified dissimilarity metric to the original and expand on the results.