Abstract

Biomolecules like proteins are the basic building blocks of living systems. It has been observed that the structure of a biomolecule plays an important role in defining its function. In this thesis, we describe novel geometric and topological techniques to understand the structure of molecules. In particular, we focus on the problems related to identification and visualization of cavities and channels in proteins. Cavities refer to empty regions within the molecule, while channels are pathways through the cavities. We pursue an integrated geometric and topological approach towards solving the problems in this domain. While topological structures provide efficient data structure representations of molecular space, geometric techniques allow accurate computation of various geometric measures having biological significance. In the first part of the thesis, we describe two methods: one for extraction and visualization of biomolecular channels, and the other for extraction of cavities in uncertain data. We also describe the two software tools based on the proposed methods targeted at the end-user, the biologists. These two web server tools publicly available for use are called ChExVis and Robust Cavities. The first method uses an alpha complex-based framework for extraction and visualization of geometrically feasible channels in biomolecules. We show that our proposed method has several advantages in terms of representation power over existing channel finding algorithms. In addition, we present novel ways of visualizing the amino-acids lining the channels together with their physicochemical properties. The second method addresses the problem of cavity extraction in biomolecules while taking into account uncertainties associated with empirically determined atomic positions and radii. We propose an approach that connects user-specified cavities by computing an optimal conduit within the region occupied by the molecule. The conduit is computed using a topological representation of the occupied and empty regions and is guaranteed to satisfy well defined geometric optimality criteria. We also describe a user interface with multiple linked views for interactive extraction and exploration of stable cavities. We demonstrate the utility of both the proposed methods using multiple case studies. In the second part of the thesis, we describe efficient parallel algorithms for two geometric structures widely used in the study of biomolecules. One of the structures we discuss is discrete Voronoi diagram which finds applications in channel visualization, while the other structure is alpha complex which is extremely useful in studying geometric and topological properties of biomolecules. We introduce a variant of the jump flooding algorithm to compute the discrete Voronoi diagram called Facet-JFA. The algorithm optimizes the number of pixels processed by computing only the faces of the Voronoi tessellation. We observed speed-up of up to 10x over JFA. As an application of the proposed algorithm, we present a GPU based method for extraction of channel centrelines in biomolecules. Secondly, we propose a GPU based parallel algorithm for the computation of the alpha complex, a subcomplex of the Delaunay triangulation that is widely used to represent biomolecules. The algorithm exploits the knowledge of typical distribution and sizes of atoms in biomolecules. Practically, we observed speed-up of up to 22x over the state-of-the-art algorithm using our implementation.