

ENTROPY GUIDED VISUALIZATION AND ANALYSIS OF
MULTIVARIATE SPATIO-TEMPORAL DATA GENERATED BY
PHYSICALLY BASED SIMULATION

by

SELCUK SUMENGEN

Submitted to the Graduate School of Engineering and Natural Sciences
in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

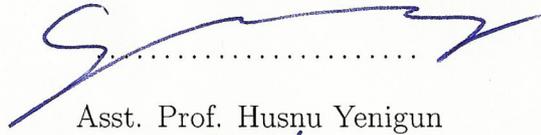
Sabanci University

August 2013

ENTROPY GUIDED VISUALIZATION AND ANALYSIS OF MULTIVARIATE
SPATIO-TEMPORAL DATA GENERATED BY PHYSICALLY BASED
SIMULATION

APPROVED BY:

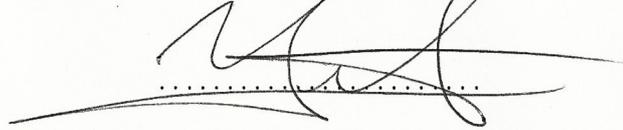
Assoc. Prof. Selim Balcisoy, (Dissertation Supervisor)



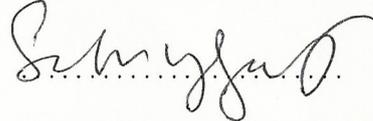
Asst. Prof. Husnu Yenigun



Assoc. Prof. Yucel Saygin



Assoc. Prof. Serhat Yesilyurt



Assoc. Prof. Marcelo Kallmann



DATE OF APPROVAL: 13.08.2013

© Selcuk Sumengen 2013

All Rights Reserved

ENTROPY GUIDED VISUALIZATION AND ANALYSIS OF
MULTIVARIATE SPATIO-TEMPORAL DATA GENERATED BY
PHYSICALLY BASED SIMULATION

Selcuk Sumengen

Computer Science and Engineering, PhD Thesis, 2013

Thesis Supervisor: Assoc. Prof. Dr. Selim Balcisoy

Keywords: Information Theory, Entropy, Vector Fields, Histogram Generation, Smoothed Particle Hydrodynamics, Multivariate Spatio-temporal Data, Visualization

Abstract

Flow fields produced by physically based simulations are subsets of multivariate spatio-temporal data, and have been in interest of many researchers for visualization, since the data complexity makes it difficult to extract representative views for the interpretation of fluid behavior. In this thesis, we utilize Information Theory to find entropy maps for vector flow fields, and use entropy maps to aid visualization and analysis of the flow fields. Our major contribution is to use Principal Component Analyses (PCA) to find a projection that has the maximal directional variation in polar coordinates for each sampling window in order to generate histograms according to the projected 3D vector field, producing results with fewer artifacts than the traditional methods.

Entropy guided visualization of different data sets are presented to evaluate proposed method for the generation of entropy maps. High entropy regions and coherent directional components of the flow fields are visible without cluttering to reveal fluid behavior in rendered images. In addition to using data sets those are available for research purposes, we have developed a fluid simulation framework using Smoothed Particle Hydrodynamics (SPH) to produce flow fields. SPH is a widely used method for fluid simulations, and used to generate data sets that are difficult to interpret with direct visualization techniques. A moderate improvement for the performance and stability of SPH implementations is also proposed with the use of fractional derivatives, which are known to be useful for approximating particle behavior immersed in fluids.

FİZİK TABANLI SİMÜLASYONLARDAN ELDE EDİLMİŞ
ÇOK-DEĞİŞKENLİ UZAM-ZAMANSAL VERİLERİN
ENTROPİ REHBERLİĞİNDE ANALİZİ VE GÖRSELLEŞTİRMESİ

Selçuk Sümengen

Bilgisayar Bilimleri ve Mühendisliği, Doktora Tezi, 2013

Tez Danışmanı: Doç. Dr. Selim Balcısoy

Anahtar Kelimeler: Bilgi Kuramı, Entropi, Vektör Alanları, Histogram Oluşturma,
Yumuşatılmış Parçacık Hidrodinamiği, Çok-değişkenli Uzam-zamansal Veri
Görselleştirme

Özet

Fizik tabanlı simülasyonlar ile üretilen akış alanları, çok-değişkenli uzam-zamansal verilerin alt kümesi olup, bu verilerden sıvı davranışlarını yorumlamayı sağlayan görsellerin çıkarılması veri karmaşıklığından dolayı zordur. Bu tez içerisinde, vektör akış alanlarının görselleştirme ve analizine yardımcı olmak üzere, Bilgi Kuramı'ndan faydalanılarak entropi haritaları çıkarılmaktadır. Ana katkı olarak, her örneklem penceresi için Temel Bileşen Analizi ile bulunan, polar koordinat düzleminde en yüksek yönsel varyasyonu veren projeksiyon kullanılarak yansıtılmış 3 boyutlu vektör alanlarının histogramları hesaplanmış ve geleneksel metotlardan daha az hatayla sonuçlar elde edilmiştir.

Entropi haritaları üretilmesi için önerilen metodun değerlendirilmesi için, entropi rehberliğinde farklı veri setlerinin görselleştirilmesi sunulmuştur. Oluşturulan imgelerde yüksek entropili alanlar ve uyumlu yönsel bileşenler karışıklığa yol açmadan görünür haldedir. Araştırma amaçlı hazır veri setlerine ek olarak, geliştirilen Yumuşatılmış Parçacık Hidrodinamiği (YPH) simülasyon altyapısı ile üretilmiş akış alanları da kullanılmıştır. YPH akışkan simülasyonları için yaygın olarak kullanılan bir metot olup, doğrudan görselleştirme teknikleri ile yorumlanması zor veri setleri oluşturmaktadır. Sıvı içerisinde batmakta olan parçacık davranışına yakınsama hesabında faydalı olduğu bilinen kesirli türevler kullanılarak, YBH uygulamasının performans ve kararlılığını artıran iyileştirme de sunulmaktadır.

To my family

Acknowledgements

I would like to express my deepest gratitude and appreciation to my advisor, Selim Balcisoy for his continuous support and excellent guidance. I am also thankful for his trust in my abilities and for his efforts in motivating me to pursue graduate study.

I have been honored to have Husnu Yenigun, Yucel Saygin, Serhat Yesilyurt, and Marcelo Kallmann as members of my thesis committee. I am grateful for their valuable review and comments on the thesis.

I would like to thank my collaborators Oktar Ozgen and Ekrem Serin for leading with the research that lies behind my thesis.

Thanks to all my friends and colleagues for their support. I am very grateful for the time spent with the friends and memories.

Finally, I would like to thank my parents for their love and support.

TABLE OF CONTENTS

1	INTRODUCTION	1
1.1	Motivation	1
1.2	Problem Statement	2
1.3	Outline	2
2	PREVIOUS WORK	4
2.1	Non-Information Theoretic or General Approaches for Multivariate Data Visualization	4
2.1.1	Topology Based Flow Visualization	5
2.1.2	Scientific Visualization	6
2.2	Information Theoretic Approaches for Multivariate Data Visualization	7
2.2.1	Information Theory Assisted Topology Based Flow Visualization	7
2.2.2	Scientific Visualization	8
3	PRELIMINARIES	12
3.1	Physically Based Simulation	12
3.1.1	Navier-Stokes Equations and Smoothed Particle Hydrodynamics	12
3.2	Information Theory	14
3.2.1	Entropy	14
3.2.2	Viewpoint Entropy	14
3.2.3	Viewpoint Kullback-Leibler	15
4	SMOOTHED PARTICLE HYDRODYNAMICS	17
4.1	Standard SPH	18
4.2	Fractional SPH	20
4.2.1	Computing the Half Derivative Terms	21
4.2.2	Comparison of SPH and FDSHP	22

4.2.3	Discussion	26
5	VISUALIZATION & ANALYSIS OF MULTIVARIATE SPATIO-TEMPORAL DATA	28
5.1	Visualizing Flow Fields	29
5.1.1	Direct Methods for Visualization	29
5.1.2	Integration Based Geometry Extraction Methods	33
5.2	Use of Information Theory in Visualization	35
5.2.1	Case Study and Inspiration: Viewpoint Selection for 3D Poly- onal Meshes	36
5.2.2	Entropy Guided Visualization of Multivariate Spatio-Temporal Data	39
6	CONCLUSION	51
6.1	Future Work	53
	Bibliography	65

List of Tables

2.1	Table of reviewed previous work is given, and categorized by the use of Information Theory.	4
-----	---	---

List of Figures

4.1	Example of a typical SPH simulation scenario. As demonstrated in several evaluations, our fractional SPH model will improve the realism of the simulation in a chosen resolution. The colors represent velocity magnitudes in a scale ranging from red (high), to green (medium), and to blue (low).	18
4.2	The figure shows the results of the Shear Driven Cavity Test: Low resolution standard SPH is on the left, low resolution Fractional SPH is on the middle, and high resolution standard SPH is on the right.	23
4.3	Lid-driven cavity test comparing OpenFOAM’s grid-based Navier–Stokes solution (black curve), standard SPH (blue curve), and fractional SPH (dashed red curve) with 40k particles. The velocities along the vertical line $x = D/2$ passing by the center of the box are demonstrated at $t = D/5$ s when the simulations are in steady state. The horizontal axis in the graph represents the vertical coordinates along the line $x = D/2$. The similarity of the curves validate the viscosity behavior of the fractional SPH simulation in a steady flow scenario.	24
4.4	Lid-driven cavity test comparing the solution of a high resolution regular SPH simulation with low resolution regular SPH and Fractional SPH simulations.	25
4.5	Lid-driven cavity test for comparing the stability of regular SPH and Fractional SPH simulations.	26
4.6	The figure shows the average velocities and the velocity directions of particles in the context of Shear Driven Cavity Test. The figures correspond to 21k standard SPH, 6k Fractional SPH and 6k standard SPH, respectively. Colors red, green and blue represent high, medium and low velocities, respectively. The color distribution and regional velocity directions of Fractional SPH simulation are similar to the ones of the high resolution reference simulation.	26
5.1	Direct visualization of 2D Lid-driven cavity test using color codes for velocity magnitude, and directional components separately.	30
5.2	Direct visualization of 3D fluid flow generated by SPH simulations, colors are representing velocity magnitudes.	30
5.3	A 2D slice from the simulation data of Hurricane Isabel, rendered using color coding representing velocity magnitudes.	31

5.4	Arrow glyph visualization of a 2D slice from the simulation data of Hurricane Isabel on the left, and the use of arrow glyph technique in 3D for the whole dataset.	32
5.5	Visualization of 3D fluid flow generated by SPH simulations using arrow glyph method.	32
5.6	A 2D slice from the simulation data of Hurricane Isabel used to generate iso-lines representing vector magnitudes is shown on the left, and iso-surfaces representing the same dataset are given on the right.	33
5.7	Hurricane Isabel data volume rendered in 3D, a linear transfer function scaled to vector magnitudes is used for transparency and color.	34
5.8	Streamline generation for Hurricane Isabel data set along two different lines used for seeding.	34
5.9	Streamline generation for SPH data set along two different lines used for seeding.	35
5.10	Mesh saliency for a hand model shown in (a). HSV color model shown in (b) is used to mark the saliency of the vertices. Hot colors(red) $Hue=0$ shows the highest saliency, and $Hue=240$ for the lowest. <i>Saturation</i> and <i>Value</i> are kept fixed in distribution.[50]	38
5.11	Stanford Bunny is displayed with five viewpoints using the approach from [48] and [60] compared to greedy integrated $vSKL$ method.	39
5.12	Sample vector field with varying directions and same magnitude is given on the left, and vector fields with the same directional component, but varying magnitudes are given on the middle and the right.	42
5.13	Blue vector is assigned to a point on the outer sphere, while red one is assigned to a point on the inner sphere.	43
5.14	Entropy is calculated on Hurricane Isabel data using buckets of varying magnitude is on the left. Angular entropy field calculated with the regular approach is on the middle, and entropy field calculated from magnitudes is on the right.	43
5.15	Entropy is calculated on Hurricane Isabel data using different window sizes. Each dimension of the cubic windows are 3,5,7,9 samples wide.	44

5.16	Entropy is calculated on Hurricane Isabel data using different histograms with varying bin sizes. We have used spheres with 180, 360 and 720 patches from left to right.	44
5.17	Sample set of vectors, and the projection plane found by principal component analysis.	45
5.18	Entropy calculated with our method on Hurricane Isabel dataset; entropy calculated with angle of direction on the projected plane is given on the left, entropy calculated with the magnitudes, and z coordinate after projection are given on the middle and left.	46
5.19	Directional entropy calculated after utilizing PCA is given on the left, and directional entropy calculated using unit sphere for discretization is given on the right.	47
5.20	SPH simulation data is rendered using color coding for the entropy values and arrow glyph for the velocity vector where the entropy value is below the threshold to reveal fluid behavior.	48
5.21	Simulation data of Hurricane Isabel is rendered using color coding for the entropy values and arrow glyph for the velocity vector where the entropy value is below the threshold to reveal fluid behavior.	49

1 INTRODUCTION

There are many applications that produce multivariate spatio-temporal data sets using physically based simulations, and those data sets share certain characteristics. It's a challenging problem to visualize varying large data sets having vector attributes defined on a grid covering 3D domain in order to reveal the underlying behavior. In this work, we experiment on every stage beginning from the simulation to the visualization, and we introduce improvements on certain tasks until proposing a novel method for histogram generation to calculate entropy and aid visualization.

1.1 Motivation

Time-varying multivariate spatio-temporal data sets are produced by physically based simulations of many natural phenomenon, however fluid simulations producing flow fields are the most common ones that are frequently used for practical applications. From weather forecasts to water flow analysis in turbines of power plants, many simulations are performed at several scales producing flow fields at different complexities and characteristics in daily life. Analyses of those large data sets is difficult without assisting visualization techniques, and revealing the fluid behavior under flow field is difficult with simple visualization methods.

Information Theory is a promising field of research, already applied to many areas in Computer Graphics including scientific visualization for many years. However, there are still many problems unexplored in the field, due to the variety of spatio-temporal multivariate data characteristics, and the broad perspective of Information Theoretical approaches bringing many opportunities to evaluate the information content in several ways.

1.2 Problem Statement

In this thesis, a fluid simulator using Smoothed Particle Hydrodynamics is developed in order to experiment analysis and visualization techniques of flow fields, while fractional derivatives are also explored for the possibility of enhancing simulation performance and stability. Ultimate goal of this thesis is finding methods and techniques to assist analysis and visualization of multivariate data sets using Information Theoretical approaches.

Our main contribution is;

- The utilization of PCA to generate histograms of 3D vector fields by polar coordinate transformation.

We also accomplished to have additional contributions during our work including;

- Proposing a histogram generation method for 3D vector fields taking magnitudes and directions into account.
- Introducing vSKL distance for generating representational images of 3D polygonal meshes[50] in collaboration with Ekrem Serin,
- The development of an SPH framework utilizing fractional derivatives[38] in collaboration with Oktar Ozgen,

1.3 Outline

Previous work related to multivariate spatio-temporal data visualization is reviewed in Chapter 2. Methods for the visualization of flow fields are mentioned as well as Information Theory related approaches.

Preliminaries for Physically Based Simulation, and Information Theory are briefly summarized in Chapter 3.

In chapter 4, Smoothed Particle Hydrodynamics (SPH) is introduced for generating flow fields using physically based simulations. Standard SPH model is defined with

the governing equations, and our model based on fractional derivatives is explained. Fractional Derivatives are used for particle motion in viscous fluids. Fractional SPH is compared with standard SPH in terms of validity, performance and stability. The necessity of using special visualization techniques for flow rendering is discussed at the end.

Visualization methods existing in literature for flow fields are experimented using our simulation results as well as publicly available data sets for research purposes in Chapter 5. We introduce Information Theory to generate representative images for 3D polygonal meshes while preserving salient features, and propose a new method for calculating entropy to aid visualization of flow fields.

In Chapter 6, our contributions are revisited and summarized for a conclusion.

2 PREVIOUS WORK

In this chapter, previous work on multivariate spatio-temporal data visualization is reviewed as well as relevant approaches for flow visualization and Information Theory related scientific visualization methods. In order to put the relevant work together and have a structural organization, we group the approaches according to their main area of interest. In each category, reviewed publications are in chronological order.

	Topology Based Flow Visualization	Scientific Visualization
Non-Information Theoretic or General Approaches	Pobitzer et al. [40], McLoughlin et al. [30].	Kehrer et al. [24], Burger et al. [6], Tong et al. [58].
Information Theoretic Approaches	Tao et al. [55], Ma et al. [28], Chen et al. [8], Xu et al. [65], Marchesin et al. [29].	Sbert et al. [47], Wang et al. [64], Chen et al. [9], Ruiz et al. [46], Guoqing et al. [16], Bramon et al. [5], Tao et al. [56], Chaudhuri et al. [7].

Table 2.1: Table of reviewed previous work is given, and categorized by the use of Information Theory.

2.1 Non-Information Theoretic or General Approaches for Multivariate Data Visualization

General approaches for multivariate data visualization methods that are not involving Information Theory are briefly reviewed. Flow fields are a subset of multivariate time-

varying data, and most the techniques are referred in survey papers for topology based visualization. Rest of the work dealing with multivariate data visualization without involving Information Theory are reviewed as scientific visualization techniques.

2.1.1 Topology Based Flow Visualization

A state of art report in topology based unsteady flow visualization is published by Pobitzer et al. [40]. In this report, topology based and topology inspired visualization methods for unsteady flow fields are grouped as Lagrangian methods, space-time domain approaches, local methods, stochastic and multifield approaches. The goal of classical vector field topology is defined as segmenting the flow domain into regions where the trajectories show the same behavior for steady flows, since flow behavior can be determined at an arbitrary instance of time. Extending this approach, and keeping track of topology in time applying classical vector field topology for each time frame is classified under the category named tracking of topology. The shortcoming of approaches in this category is given as the difficulty of finding nearly stable velocity fields for unsteady flow fields. Feature extraction methods that use trajectories of particles in fluid are defined as Lagrangian based methods. The finite-time Lyapunov exponent feature detectors are in this category, which are measuring the stretching of an infinitesimal neighborhood along a finite segment of flow trajectory such as separation and repulsion boundaries. Streamlines and pathlines are categorized in space-time domain approaches, taking time as another dimension and applying steady case for each time frame. Also the feature flow fields, which are capturing topological information in 4D space-time domain are in this category. Methods that are using only point-wise information are categorized as local methods such as extracting edges or ridges of images. Stochastic and multifield approaches are looking at multiple features or multiple definition of same feature to get an understanding of the underlying field. Interactive visual analysis and fuzzy feature detectors are under this category. Note that those categories specified by Pobitzer et al [40]. are not mutually exclusive.

McLoughlin et al. [30] published a survey on integration based geometric flow visualization techniques to review and classify geometric flow visualization literature. They classify vector field visualization approaches into four categories as direct, dense texture-based, geometric, and feature based approaches. Then they focus on Integration-based, geometric flow visualization and review them under a classification based on dimensions such as integral curve objects in 2D, surface-based integral objects, and volume integral objects. So that their classification allow streamlines and their variations like streamsurfaces to be named according to domain and dimension of associated geometries, and fall into different categories.

2.1.2 Scientific Visualization

In a recent survey by Kehrer et al. [24], multivariate spatio-temporal data evaluated as multifaceted in terms of having many data models and sources from different scenarios. Thus many techniques for multifaceted data is categorized according to data model as well as the analysis approach from visual mapping to computational analysis.

Burger et al. [6] categorizes visualization techniques for multivariate scientific data, according to the specific states at the visualization pipeline, and separates data type as scalars, vectors, and tensors. Processing, filtering and visualization mapping is categorized as one visualization pipeline stage, while rendering and image stages are considered separately.

Salient Time-steps

Tong et al. [58] proposes an approach by minimizing the information loss for selecting arbitrary number of time frames from time-varying data sets. They apply dynamic programming, and define dissimilarity matrix before selecting salient time steps. Although they claim to minimize the information loss, this approach is not classified as an Information Theory based approach.

2.2 Information Theoretic Approaches for Multivariate Data Visualization

Wang et al. [61] presents a survey on information and knowledge assisted visualization, and creates a taxonomy by grouping approaches in categories named information assisted visualization, knowledge assisted visualization, intelligent visualization and visualization interface. Under information assisted visualization, subcategories are visualization enabled by statistical information, geometric information, topological information and semantic information. After briefly introducing approaches in each category, concludes that the future of visualization lies in development of information and knowledge assisted solutions.

2.2.1 Information Theory Assisted Topology Based Flow Visualization

Tao et al. [55] performs Information Theory guided streamline selection, and additionally do best viewpoint selection in a similar manner. They propose solutions to streamline clustering and viewpoint partitioning as well.

Ma et al. [28] present an importance driven and a view-dependent streamline selection methods using Information Theory considering amount of information shared by 3D streamline and its 2D projection. A large number of seeds are used to generate pool of streamlines, then their streamline selection methods eliminate excessive amount of streamlines by using view-dependent or view-independent importance measures to avoid cluttering. Coherent update of selected streamlines is also maintained while changing the viewpoint. Their shortcomings are the need of generating many streamlines than the flow-guided streamline generation methods, and selecting relatively more streamlines in comparison with the feature driven approaches. Also their entropy measure is not sensitive to small-scale features.

In article named illustrative framework for 3D vector fields, Chen et al. [8] intro-

duces streamribbons or streamtapes, in which twist and width are determined according to local flow torsion. They also apply entropy based seeding, and perform streamline clustering with k-means algorithm before generating streamtapes for illustrative rendering. They follow the same approach with Xu et al. [65] for histogram generation and entropy calculation. While clustering they don't take viewpoint into consideration, and they experimented only with steady flows.

Xu et al. [65] utilizes Information Theory for streamline placement to visualize 2D and 3D flow data. An entropy field generated to locate seeds and generate streamlines in regions with high information content. They use spherical partitioning to discretize 3D vectors and generate histograms for entropy calculation only considering directional components.

Marchesin et al. [29] proposes a method for view dependent streamline selection using occupancy buffers to minimize occlusion and cluttering. Although they have fast GPU implementation, their method is not interactive and considers single time frame for steady vector flows.

2.2.2 Scientific Visualization

In course notes prepared by Sbert et al. [47], Information Theoretical methods and their applications for computer graphics and visualization are summarized. Information-theoretic measures such as Shannon entropy, Kullback-Leibler distance, Jensen-Shannon inequality as well as divergence measures are reviewed. A framework for polygonal models with viewpoint selection and mesh saliency is introduced. In addition, applications to global illumination, shape recognition and image processing in computer graphics are exemplified. Methods utilizing Information Theory in Scientific Visualization are briefly reviewed with a focus on volume visualization.

Wang et al. [62] reviews the use of Information Theory in scientific visualization. Concepts of Information Theory are explained from entropy to distance measures and mutual information. View selection for volumetric data, streamline seeding and selec-

tion, designing transfer functions for multimodal data, selection of representative iso surfaces, multi-resolution volume visualization, and time-varying multivariate data analysis are the application areas mentioned in Scientific Visualization as well as the applications of Information Theory in Imaging and Graphics.

Wang et al. [63] proposes to utilize transfer entropy for analyzing causal connection between variables in time-varying multivariate data sets. They use information and scientific visualization techniques to display information transfer, and define a new approach volumetric and particle data sets using time plot and circular graph. They also define relative transfer entropy to generalize pair-wise transfer entropy to simultaneously handle multiple variables. Their limitation is being able to use transfer entropy only on two scalar variables, and extending their approach to work on multiple variables simultaneously should be further studied.

Chen et al. [9] presents an information theoretic view of visualization pipeline, and evaluates usability of concepts in Information Theory for visualization. They conclude that several aspects of Information Theory can be utilized for visualization.

Ruiz et al. [46] uses viewpoint information channel for illustrative rendering of volumetric data sets. An information channel is constructed between the volumetric data set and a set of viewpoints. An ambient occlusion value for each voxel is derived from the information associated, and combined with assigned color for each viewpoint and non-photorealistic effects, illustrative are obtained. Transfer function is also modulated with voxel information for the transparency.

Wang et al. [64] presents a compression scheme for visualizing large-scale time-varying data sets. They only perform scalar quantization, and consider scalar variables of multivariate data sets.

Salient Time-steps

Guoqing et al. [16] presents an Information Theory assisted method to locate spatial and temporal salient features for the visualization of large scale time varying data sets.

They use Kullback-Leibler distance for measuring dissimilarity of different time steps, and off-line marginal utility for surprising information at each newly added time step. Spatial salient features are detected by entropy for scalar data sets.

Multimodal Data Sets

Multimodal visualization aims to combine different volumetric data sets into one. Brannon et al. [5] present a framework for volume visualization that exploits Information Theory to define a transfer function for multimodal data sets. First, they generate information maps between input data sets and compute fused colors. Then, they calculate informativeness using two different information measures, global informational divergence and viewpoint informational divergence to compute opacity values by minimizing informational divergence. This approach is limited with two data sets both of which have one scalar variable.

Viewpoint Selection

Tao et al. [56] introduces structure aware viewpoint selection for volume visualization by defining shape view descriptor and detail view descriptor. Shannon's entropy is used to define shape view descriptor to measure the distribution of the relative view angle between the gradient direction and viewing direction. The detail view descriptor measures the visible detail in terms of variances between the shape volume and original volume. Their limitation is the working only with volumetric data sets having scalar variables, and they are not taking time-varying data into account to measure structure difference between consecutive time steps.

Histogram Generation

Chaudhuri et al. [7] proposes a histogram generation approach for large scale data sets. Their method is suitable for distributed computation, and they're able to produce multi-level histograms efficiently. They use geodesic grid instead of a sphere like [65],

since their multi-resolution approach requires patches at different resolutions to have parent-child relation, and they don't take vector magnitudes into account. They propose weighted vertex method, which is faster than sampling to increase data resolution before histogram generation. A histogram estimator proposed by Rudemo [45] is used for determining number of bins for the histogram with a fixed bin size.

3 PRELIMINARIES

In this chapter, preliminary definitions as well as simple derivations used in following chapters are introduced.

3.1 Physically Based Simulation

Derivation of governing equations in Smoothed Particle Hydrodynamics for fluid simulations are briefly given in the following section.

3.1.1 Navier-Stokes Equations and Smoothed Particle Hydrodynamics

Governing equations that are supposed to hold for fluid simulations are incompressible Navier-Stokes equations. Equation 3.1 is called momentum equation, while Equation 3.2 is called incompressibility condition.

$$\frac{\partial \vec{u}}{\partial t} + \vec{u} \cdot \nabla \vec{u} + \frac{1}{\rho} \nabla p = \vec{g} + \nu \nabla \cdot \nabla \vec{u} \quad (3.1)$$

$$\nabla \cdot \vec{u} = 0 \quad (3.2)$$

For the symbology, \vec{u} is used for the velocity of the fluid. ρ stands for density, and p stands for pressure. \vec{g} is used for body forces, including acceleration due to gravity. ν is used to represent kinematic viscosity of the fluid.

There are two approaches for simulating fluids, using the Eulerian viewpoint governing equations are supposed to hold on regular grid points fixed on the domain. On

the hand, the Lagrangian viewpoint treats fluid as a free mesh moving in the domain. It's possible to use particles to keep track of fluid in Lagrangian approaches. Smoothed Particle Hydrodynamics (SPH) is a method that allow to represent the attributes of fluid in the continuum using smoothing kernels which are used in Chapter 4. In order to use smoothing kernels, governing equation for SPH is derived from conventional Navier-Stokes equations in Equation 3.3 after diving Equation 3.3 by ρ . Kinematic viscosity coefficient is also replaced by dynamic viscosity coefficient using $\nu = \frac{\mu}{\rho}$.

$$\rho \left(\frac{\partial \vec{u}}{\partial t} + \vec{u} \cdot \nabla \vec{u} \right) = -\nabla p + \rho \vec{g} + \mu \nabla^2 \vec{u} \quad (3.3)$$

By the definition of material derivative in Equation 3.4, the left hand side of Equation 3.3 can be replaced by substantial derivative.

$$\vec{a} = \frac{D\vec{u}}{Dt} = \frac{\partial \vec{u}}{\partial t} + \vec{u} \cdot \nabla \vec{u} \quad (3.4)$$

Since for the Lagrangian viewpoint, substantial derivative of the velocity field is equal to the time derivative, there's no need to have a convective term in particle systems.

$$\rho \vec{a} = -\nabla p + \rho \vec{g} + \mu \nabla^2 \vec{u} \quad (3.5)$$

From Equation 3.5, the terms $-\nabla p$ for $f_{pressure}$, $\mu \nabla^2 \vec{u}$ for $f_{viscosity}$, $\rho \vec{g}$ for $f_{external}$ can be derived and used in governing equations for SPH. For the conservation of mass Equation 3.6 is derived from Equation 3.2.

$$\frac{\partial \vec{u}}{\partial t} + \nabla \cdot (\rho \vec{u}) = 0 \quad (3.6)$$

In a particle system, there are constant number of particles and each particle has a

constant mass, so Equation 3.6 can be omitted.

3.2 Information Theory

In this section, Information Theoretical concepts that are used but not explained in Chapter 5 are given briefly.

3.2.1 Entropy

The entropy [51] of a discrete random variable X with values in the set $\{x_1, x_2, \dots, x_n\}$ is defined as

$$H(x) = - \sum_{i=1}^n p(x_i) \log_b p(x_i) \quad (3.7)$$

Even though the entropy is expressed as a function of the random variable X , it is actually a function of the probability distribution p of the variable X over the number of distinct symbols N . Entropy function has following two important properties [4];

1. For a given number of symbols N , the maximum entropy occurs for the distribution p_{eq} , where $\{p_0 = p_1 = \dots = p_{N-1} = 1/N\}$.
2. Entropy is a concave function, which implies that the local maximum at p_{eq} is also the global maximum. It also implies that as we move away from the equal distribution p_{eq} , along a straight line in any direction, the value of entropy decreases (or remains the same, but does not increase).

3.2.2 Viewpoint Entropy

The properties of the entropy function expressed above give us that the calculated viewpoints in extracted regions will be the global maximum points where the object surface is perceived equally.

Viewpoint entropy [59] using Shannon Entropy is defined as

$$I(S, p) = - \sum_{i=0}^{N_f} \frac{A_i}{A_t} \log_b \frac{A_i}{A_t} \quad (3.8)$$

where A_i is the projected area of face i over the sphere, A_t is the total area of the sphere and b is the base of logarithm which is taken as $b = 2$ in this case the result is bits/symbols. In other terms the formula shown above can be translated into where A_i can denote the number of pixels in the image, and A_t can represent the number of pixels that belongs to each face of the object. A_0 is a special case for the projected model or scene onto the screen. For the closed scenes A_0 is taken as 0 and for open scenes A_0 is considered as the number of pixels that belong to the background color. With the contribution of A_0 for open scenes we can have a viewpoint entropy definition that is consistent with Shannon's entropy where $\sum_{i=1}^n p_i = 1$.

3.2.3 Viewpoint Kullback-Leibler

The relative entropy or Kullback–Leibler distance is defined between two probability distributions $p = \{p(x)\}$ and $q = \{q(x)\}$. In this metric, the distance is interpreted as the divergence between *true* probability distribution p and *target* probability distribution q . Kullback–Leibler distance is defined as,

$$KL(p | q) = \sum_{x \in X} p(x) \log \frac{p(x)}{q(x)} \quad (3.9)$$

For the continuity the convention that $0 \log 0 = 0$, $p(x) \log \frac{p(x)}{0} = \infty$ if $p(x) > 0$, and $0 \log \frac{0}{0} = 0$ is used [48]. The minimum value 0 means that the true probability distribution is equal to the target probability distribution, where $KL(p | q) \geq 0$. The viewpoint Kullback–Leibler distance is defined by

$$KL_v = \sum_{i=1}^{N_f} \frac{a_i}{a_t} \log_b \frac{a_i}{A_t} \quad (3.10)$$

where a_i is the projected area of polygon i , $a_t = \sum_{i=1}^{N_f} a_i$. A_i is the actual area of polygon i and $A_t = \sum_{i=1}^{N_f} A_i$ is the total area of the 3D object. In order to select high quality views KL_v should be minimized.

4 SMOOTHED PARTICLE HYDRODYNAMICS

In order to generate flow fields, we implemented our own fluid dynamics solver using Smoothed Particle Hydrodynamics (SPH). It's mesh-less method suitable for large displacements, and we're able to produce chaotic flow fields at interactive frame rates.

SPH has a long history in physics, developed in 1977 by Gingold and Monaghan [15] to model astrophysical phenomena, and extended to solve many problems in continuum mechanics. There are many uses of particle systems in Computer Graphics, however discrete formulation of continuous fields by particles was first introduced by Desburn et al. [13] for simulating highly deformable bodies. Muller et al. [34] reached very promising results in particle-based fluid simulation for interactive applications using the SPH method. A very detailed study of SPH since its first emergence is presented by Monaghan [33].

The method of Smoothed Particle Hydrodynamics (SPH) has become a popular particle-based approach for fluid simulation because results incorporating complex interactions (e.g., splashes, coupling, etc.) can be obtained with relatively modest computational costs [13, 34, 35, 2, 52]. Key to the quality of the results obtained is the determination of an appropriate number of particles achieving sufficient volumetric density. While better results are, in principle, obtained with high concentrations of particles, the computational penalty is significant.

In recent years, new variations to the standard SPH models have also emerged. Sotherthaler [52] proposed the PCISPH method for reducing the computation time of standard SPH and increasing the incompressibility of the fluid by employing a prediction-correction scheme based on particle pressures. Raveendran [44] introduced a hybrid approach that uses a Poisson solver along with a local density correction step to increase

the stability of SPH method in higher time steps. Solenthaler [53] proposed a two-scale simulation by merging the results from low and high resolution simulations running simultaneously. Adaptive time steps are employed by Ihmsen and Adams [22, 1] in SPH methods to increase the stability of the simulations. SPH applications based on parallel computing are also proposed by various groups [22, 21, 17].

In our work led by Oktar Ozgen[38], we present a novel approach to increase the performance and stability of SPH with the introduction of *fractional derivatives* [36] [41] to the viscosity term. In this work, the goal is producing results similar to the ones obtained with high-resolution SPH simulations. In order to compare the results of the proposed method with regular SPH, we employ some of visualization techniques in addition to a direct numerical comparison and a well known standard test in fluid dynamics.

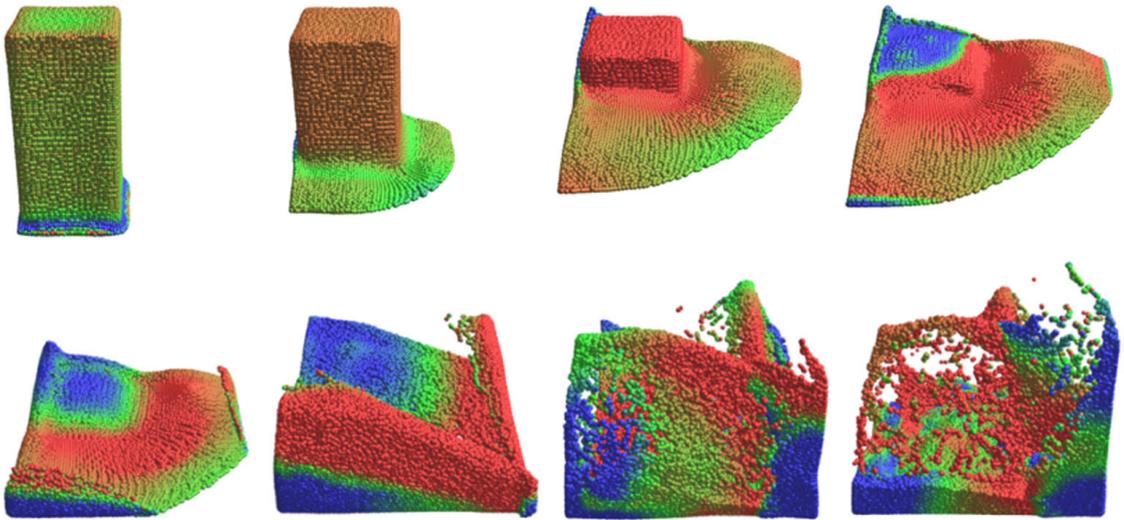


Figure 4.1: Example of a typical SPH simulation scenario. As demonstrated in several evaluations, our fractional SPH model will improve the realism of the simulation in a chosen resolution. The colors represent velocity magnitudes in a scale ranging from red (high), to green (medium), and to blue (low).

4.1 Standard SPH

The Smoothed Particle Hydrodynamics (SPH) model we employ is based on the scheme presented by Muller [34]. SPH is a Lagrangian model where the fluid is represented by

a set of particles that carry field attributes. An arbitrary attribute on a given particle's position is computed via smoothing kernels that only consider nearby particles within the core radius h . The smoothing of attributes is modeled with:

$$A_S(r) = \sum_j m_j \frac{A_j}{\rho_j} W(\mathbf{r} - \mathbf{r}_j, \mathbf{h}), \quad (4.1)$$

where m_j is the mass, r_j is the position and ρ_j is the density of a particle j within the core radius h of the smoothing kernel $W(r - r_j, h)$. A_j is the field attribute quantity at r_j .

At each timestep of the simulation, the density values of individual particles are evaluated first:

$$\rho_i = \sum_j m_j W(|r_i - \mathbf{r}_j|, \mathbf{h}), \quad (4.2)$$

then, the pressure is computed by the ideal gas state equation

$$p = k(\rho - \rho_0), \quad (4.3)$$

where k is a gas constant and ρ_0 is the rest density. Once, the density and the pressure fields are computed, the pressure and viscosity forces acting on particle pairs are computed in a symmetric manner as proposed by Muller [34]:

$$f_i^{pressure} = - \sum_j m_j \frac{p_i + p_j}{2\rho_j} \nabla W(r_i - r_j, h), \quad (4.4)$$

$$f_i^{viscosity} = \mu \sum_j m_j \frac{\dot{x}_j - \dot{x}_i}{\rho_j} \nabla^2 W(r_i - r_j, h), \quad (4.5)$$

where $\nabla W(r_i - r_j, h)$ is the gradient, $\nabla^2 W(r_i - r_j, h)$ is the Laplacian of the kernel, μ is the viscosity constant, $\dot{x}_i = v_i$ and $\dot{x}_j = v_j$, are the velocity vectors of the particles i and j , respectively.

4.2 Fractional SPH

The subject of Fractional Calculus [36], or the mathematical analysis of differentiation and integration to an arbitrary non-integer order, has recently attracted much interest especially in solid mechanics, rheology, electromagnetism, electrochemistry, and biology.

Fractional Calculus models, aside from their capability of modeling memory-intense and delay systems, have been associated with the exact description of unsteady viscous and viscoelastic phenomena. In [11, 27], definitive experimental evidence of fractional history effects in the unsteady viscous motion of small particles in suspension is presented. This formulation is exact at low particle Reynolds numbers, but can be extended to include convective effects as illustrated in [39]. Furthermore, a rich literature is available on the ability of non-integer derivatives to capture non-local behavior and to interpolate between different dynamic regimes [36, 32, 41, 20, 19, 25], including the fundamental modeling of viscoelastic behavior [42] and the unsteady drag for individual particles moving through a viscous fluid [43].

Coimbra and Rangel [12] have showed that the Basset force is mathematically equivalent to the half-derivative of the differential velocity between the particle and the far-stream flow. These results indicate that the behavior of immersed particles can be well represented with models based on fractional derivatives. The concept has been well demonstrated by Ozgen et al. [37] on the problem of simulating cloth deformations with underwater behavior.

Motivated by these fundamental results on the motion of the particles in unsteady viscous fluids, we aim to increase the physical accuracy of simulating flow collisions in low resolution simulations by utilizing a fractional derivative model. We thus propose a new SPH model with half-derivative viscosity terms to compensate the loss of information in low resolution simulations.

As discussed in [38], to demonstrate the memory-laden characteristics of the fluid body, we introduce the fractional viscosity term of order $1/2$ to the motion of particles.

We achieve so by replacing the first time derivatives of positions by the half derivatives of positions. As a result, the history-based viscosity is defined as:

$$f_i^{viscosity} = \mu \sum_j m_j \frac{D^{1/2}x_j - D^{1/2}x_i}{\rho_j} \nabla^2 W(r_i - r_j, h) \quad (4.6)$$

where $D^{1/2}x_i$ and $D^{1/2}x_j$ are the half-derivatives of the positions of particles i and j , respectively. Note that the viscosity force is now proportional to the difference of the half-derivatives, achieving the memory-laden viscosity needed to define the motion resulting from flow collisions.

The memory-laden viscosity is especially well-suited for fluid phenomena occurring in intense flow collision regions. We recognize that most of the time, a fluid simulation contains both flow collision regions and collision-free regions. Therefore, it is not necessary to apply memory-laden viscosity to particles creating steady flows.

4.2.1 Computing the Half Derivative Terms

In Coimbra [10], a first-order accurate numerical solution to the history integral of Riemann–Liouville differential operator is suggested. Following this solution, the 1/2 order derivative of x can be expressed as:

$$\begin{aligned} D^{1/2}x_n = & \frac{h}{6\sqrt{\pi}} \sum_{i=1}^{n-1} \left[\frac{\dot{x}_{i-1}}{(nh - (i-1)h)^{1/2}} \right. \\ & \left. + \frac{2(\dot{x}_{i-1} + \dot{x}_i)}{(nh - (i-1/2)h)^{1/2}} + \frac{\dot{x}_i}{(nh - ih)^{1/2}} \right] \\ & + \frac{0.15h}{\sqrt{\pi}} \left[\frac{\dot{x}_{n-1}}{h^{1/2}} + \frac{2(\dot{x}_{n-1} + \dot{x}_n)}{(0.55h)^{1/2}} + \frac{\dot{x}_n}{(0.1h)^{1/2}} \right] \\ & + \frac{0.05h}{\sqrt{\pi}} \left[\frac{8\sqrt{2}}{3} \frac{\dot{x}_n}{(0.05h)^{1/2}} - \frac{4}{3} \frac{\dot{x}_n}{(0.1h)^{1/2}} \right], \end{aligned} \quad (4.7)$$

where h is the timestep, i is the timestep index and n is the index of the most recent

computed timestep. This formulation is also used in Ozgen [37].

In Coimbra [54], a more general and second-order accurate quadrature formula derived using the product trapezoidal method is suggested for derivative orders q in the $0 < q < 1$ range. This fractional-order differential operator reads

$$D^q x_n = \frac{h^{1-q}}{\Gamma(3-q)} \sum_{i=0}^n a_{i,n} D^1 x_i, \quad (4.8)$$

$$a_{i,n} = \begin{cases} (n-1)^{2-q} - n^{1-q}(n+q-2) & \text{if } i = 0, \\ (n-i-1)^{2-q} - 2(n-i)^{2-q} + (n-i+1)^{2-q} & \text{if } 0 < i < n, \\ 1 & \text{if } i = n, \end{cases}$$

where q is the derivative order and $0 < q < 1$. n is the index of the most recently computed timestep, $a_{i,n}$ is the weight of timestep index i at timestep n and $D^1 x_i = v_i$ is the velocity of a particle at timestep i . In comparison to the method employed by Ozgen et al. [37], this formulation is relatively simpler and more accurate. In the presented simulations we have used this latter formulation with $q = 0.5$ to acquire the half derivatives.

The fact that computing the half derivative of the position of a particle makes use of all the past velocities of that particle seems to be a computational barrier at first. However, as stated in Ozgen [37], an analysis on the evolution of the weights used for the fractional derivative computation shows that the most recent states have much more influence on the final result of the equation. Thus, we only consider ten last timesteps when computing the half-derivative.

4.2.2 Comparison of SPH and FDSPH

We validate both our standard and Fractional SPH implementations with a standard test known as Shear Driven Cavity Test in Fluid Dynamics [14]. In this test, flow is generated by moving the top wall of a square box full of fluid while the other three walls are stationary. The top wall of the box moves in the x direction with a constant speed,

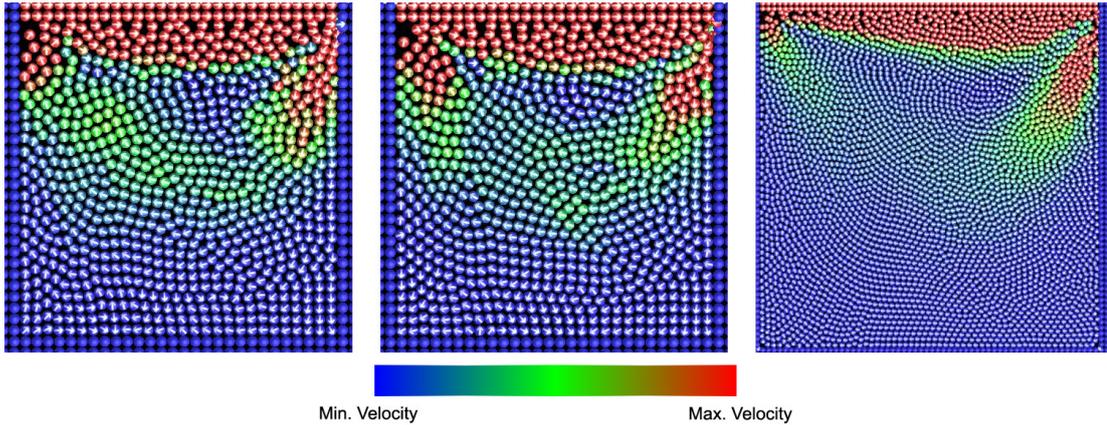


Figure 4.2: The figure shows the results of the Shear Driven Cavity Test: Low resolution standard SPH is on the left, low resolution Fractional SPH is on the middle, and high resolution standard SPH is on the right.

and the flow reaches a steady state after running the simulation for awhile. The visible flow patterns and the time required to reach a steady state vary according to the Reynolds Number, so we repeated the experiment with various viscosity parameters. We evaluated the results by mapping the magnitude of the velocity vectors to colors and rendered the velocity directions as vectors on top of the particles as shown in Figure 4.2.

All tests demonstrated that our fractional SPH model produced results closely matching the results computed by a high-precision fluid solver. We compared our results against the results generated by OpenFOAM, a grid-based solver widely employed by the Computational Fluid Dynamics community [18]. One example of the obtained results are demonstrated in Figure 4.3. As it can be seen in the figure, standard SPH and fractional SPH simulations with 40k particles follow the grid-based solution tightly, showing that the viscosity behavior of both fluids are valid and that the use of fractional derivatives in the viscosity formulation does not introduce any additional viscosity to the standard formulation.

We have also evaluated our method by comparing low resolution regular SPH and Fractional SPH simulations with a higher resolution regular SPH simulation. We have measured velocities along a vertical axis of the simulation shown in Figure 4.4. The errors introduced by using lower resolution regular SPH and Fractional SPH simulations

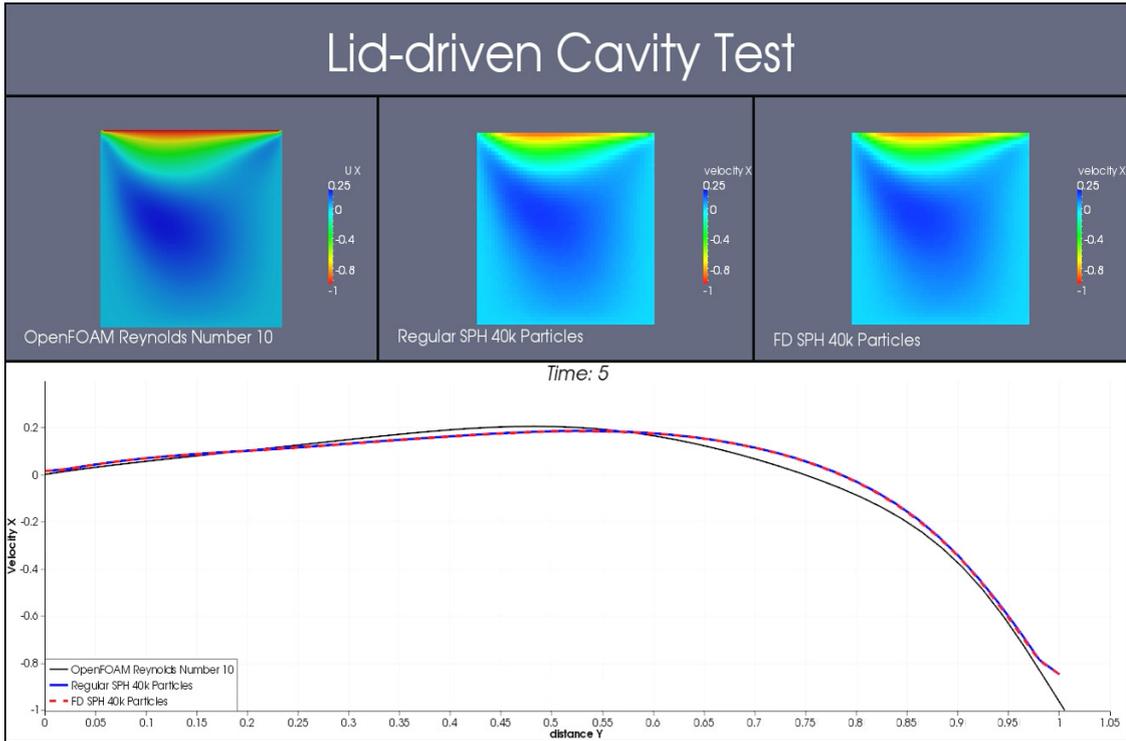


Figure 4.3: Lid-driven cavity test comparing OpenFOAM’s grid-based Navier–Stokes solution (black curve), standard SPH (blue curve), and fractional SPH (dashed red curve) with 40k particles. The velocities along the vertical line $x = 0.5$ passing by the center of the box are demonstrated at $t = 5$ s when the simulations are in steady state. The horizontal axis in the graph represents the vertical coordinates along the line $x = 0.5$. The similarity of the curves validate the viscosity behavior of the fractional SPH simulation in a steady flow scenario.

are also presented in the same figure. The absolute values of the differences in the velocities among line $x = 0$ are shown in the line chart on the bottom right for the timestep $t = 5.008$. The errors are measured and compared for each time step during the experiment in order to determine a comparison over several frames of simulation. In the end, Fractional SPH produced more precise results 59% of the time. A video is also provided showing the evolution of the errors. It can be seen that the inclusion of the fractional derivative terms does not influence the trajectories in this simulation and still the overall error showed to be favorable.

In terms of performance, our method runs real-time for simulations with up to 5k particles and runs with 4 FPS for simulations with 20k particles on a MD Athlon II X4 3.2 GHz computer. The use of half derivatives in the SPH implementation does not affect

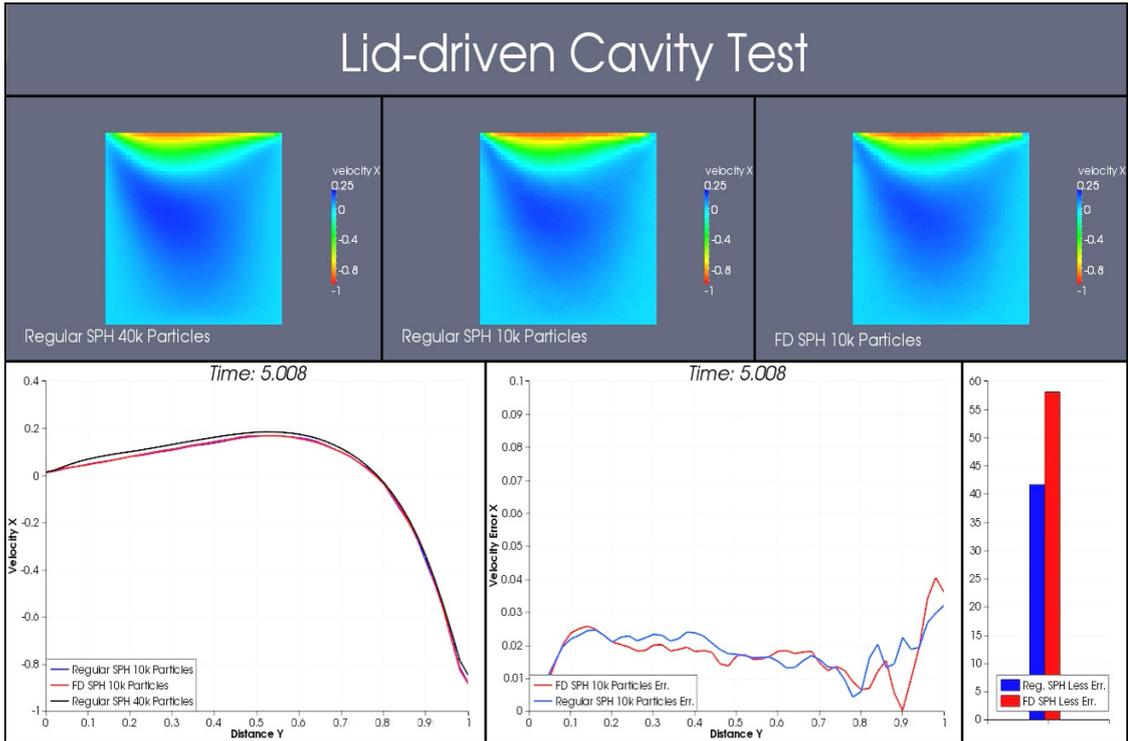


Figure 4.4: Lid-driven cavity test comparing the solution of a high resolution regular SPH simulation with low resolution regular SPH and Fractional SPH simulations.

the complexity or the running time of the algorithm. In Equation 4.8, the weights are always calculated based on the terms q and $n - i$. The value of q must always be equal to 0.5 to acquire the half derivatives. Given that we only use the three most recent terms of the history terms, $n - i$ terms always stay the same for all the three weights, except for the first three timesteps. Because Equation 4.8 makes use of the past particle velocities, we require some extra memory space to store the previous velocities. Therefore the weights can be precomputed and used in combination with pre-recorded velocities.

Fractional SPH also proves itself useful by allowing larger timesteps in the integration. It is observed that Fractional SPH simulations are more stable when using large timesteps especially for viscous fluids. In Figure 4.5, regular SPH and Fractional SPH simulations are compared for different sizes of timesteps. In Figure 4.5, Fractional SPH allows to use x2 larger timesteps, while regular SPH becomes unstable after a small increase. We also noticed that our method performed better in the early stages of the shear-driven cavity test, when the flows were not stabilized yet. This was observed in

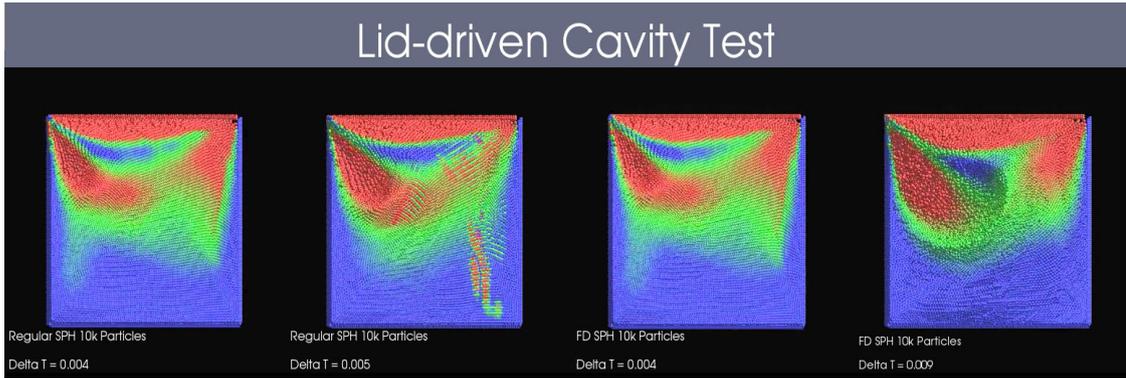


Figure 4.5: Lid-driven cavity test for comparing the stability of regular SPH and Fractional SPH simulations.

the 3D version of shear-driven cavity test and some results are presented in Figure 4.6.

4.2.3 Discussion

We have introduced a new methodology for fluid simulation, which is based on the use of Fractional Calculus with Smoothed Particle Hydrodynamics. We have also demonstrated in several experiments that our method can better simulate observed fluid behavior emerging from flow collisions. The fact that the memory-laden viscosity terms modeled by fractional derivatives are able to increase the accuracy of low resolution SPH simulations is promising as a technique to improve the quality and computational efficiency of SPH.

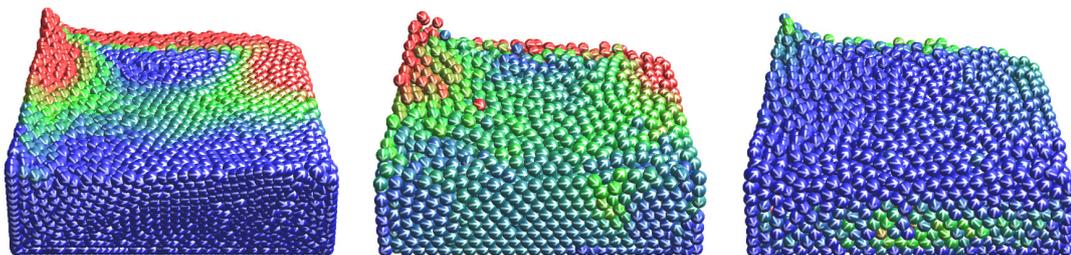


Figure 4.6: The figure shows the average velocities and the velocity directions of particles in the context of Shear Driven Cavity Test. The figures correspond to 21k standard SPH, 6k Fractional SPH and 6k standard SPH, respectively. Colors red, green and blue represent high, medium and low velocities, respectively. The color distribution and regional velocity directions of Fractional SPH simulation are similar to the ones of the high resolution reference simulation.

Flow fields we have generated also reveal that standard rendering techniques are inadequate to represent fluid behavior for unsteady flows in 3D. It's not possible to recognize colliding streams under the fluid surface due to occlusion, and it's very difficult to interpret direction and the magnitude of the flow even for the visible streams on the surface. For the analysis of flow fields produced by physical simulations, automatically created representative images would be useful to aid recognition.

5 VISUALIZATION & ANALYSIS OF MULTIVARIATE SPATIO-TEMPORAL DATA

The simulation results we have produced in Chapter 4 can be classified as multivariate spatio-temporal data. We used SPH to model fluid behavior in a container which is limiting our spatial domain, and each particle has coordinates in 2D or 3D vector space depending on the scenario. Whether the results are stored in a particle basis or interpolated to the equidistant grid cells, they all have a spatial component. The results are updated for each time step using a semi-implicit time integration solver. For each particle or a grid cell, a scalar is stored to define the pressure, and a vector is stored to define the velocity so that the simulation results are multivariate.

Although the flow fields are only a subset of multivariate spatio-temporal data, majority of the data studied in this category are provided from the simulations that are producing flow fields in practice. A detailed overview of methods for integration based geometric flow visualization are presented by McLoughlin et al. in their paper[30]. Po-bitzer et al. [40] published a state of art report on topology based flow visualization for unsteady flow. Reader should refer to Chapter 2 for previous work and a detailed literature review on flow visualization including integration and topology based visualization as well as information theory assisted methods.

In this chapter, we briefly overview direct and integration based visualization methods for flow fields with applications to sample data obtained from our SPH simulations as well as Hurricane Isabel data produced by Weather Research and Forecast (WRF) model, courtesy of NCAR and the U.S. National Science Foundation (NSF) [3]. Next, there will be a brief introduction to the use of Information Theory in visualization, and we go over our viewpoint selection method considering salient features on 3D polygonal

meshes as a case study. In an analogous manner, we mention the importance of locating salient features in multivariate spatio-temporal data in order to determine transfer functions for rendering, choosing viewpoints or apply hybrid methods involving direct and integration based flow visualization methods. In addition to the direct approaches for detecting salient features, we employ Entropy function used in Information Theory with a histogram based method similar to Xu et al. [65]. Then we improve their histogram binning method with a modification to take vector magnitudes into account in addition to the directions for diversity. At last but not least, we propose a new histogram generation method using singular value decomposition and principal components.

5.1 Visualizing Flow Fields

There are several methods in literature proposed for the visualization of flow fields as already reviewed in Chapter 2. Depending on the characteristics of the flow fields, simplistic direct visualization techniques would sometimes suffice, or even the most sophisticated methods might be inadequate to make a good interpretation of fluid behavior in some cases. Fluids are grouped into two categories depending on their behavior in time, and unsteady flows are much difficult to interpret by a visual representation in comparison to steady flows. Dimensionality is another important issue, and visualization of 3D flow fields are more challenging than 2D flow fields due to the occlusion. Size and complexity of the data also matters, as well as the number of vector and scalar components. In this section, generic methods for flow visualization are applied on SPH simulation results and Hurricane Isabel data sets to discuss all those aspects before proposing our approach for visualization.

5.1.1 Direct Methods for Visualization

Direct methods aim to present data as it is, without any modification to existing data or generation of new attributes. Those methods are usually applied in a simplistic manner, and easy to implement. They are also computationally inexpensive requiring less

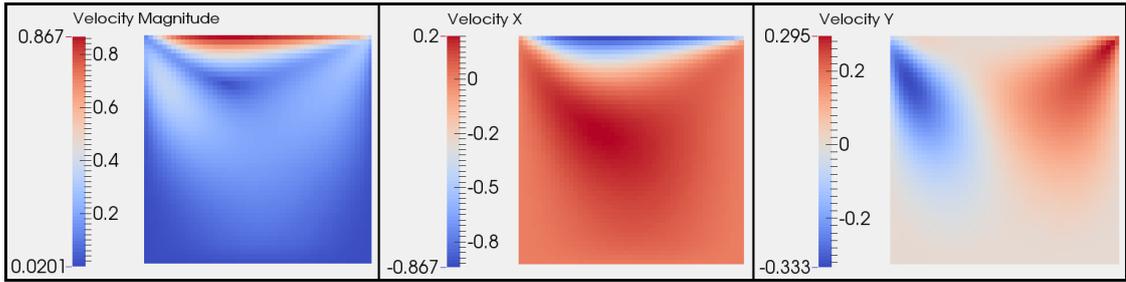


Figure 5.1: Direct visualization of 2D Lid-driven cavity test using color codes for velocity magnitude, and directional components separately.

resources. Although they might be sufficient for simple coherent flow fields in 2D, it's difficult have a good interpretation of large data sets.

Color Coding

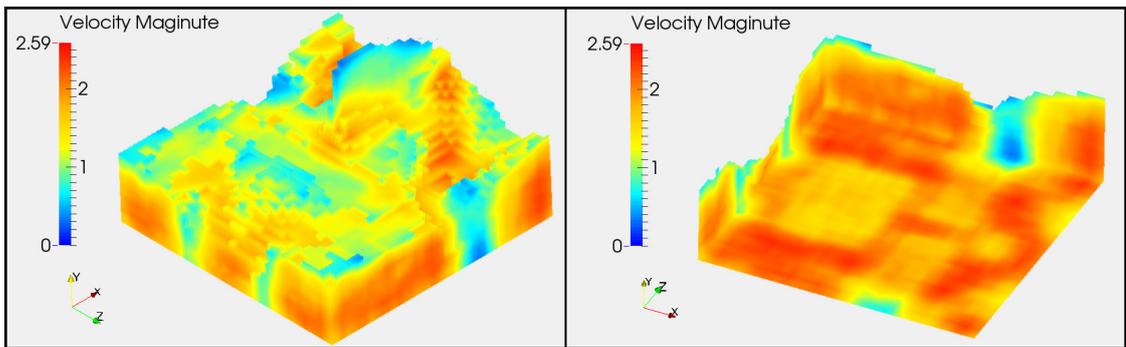


Figure 5.2: Direct visualization of 3D fluid flow generated by SPH simulations, colors are representing velocity magnitudes.

In color coding, scalar values are mapped to a color scale, and every pixels color is linearly interpolated based on the corresponding scalar value on the grid. For the vector fields, usually vector magnitudes are visualized and the direction information is omitted, or each component rendered separately. It's also possible to map each directional component of the vectors to a color value in RGB color space after scaling and shifting, so the visualization might give an insight for the directions and relative magnitudes from the apparent colors. Several examples of color coding of flow fields are presented in Figures 5.1, 5.2, and 5.3.

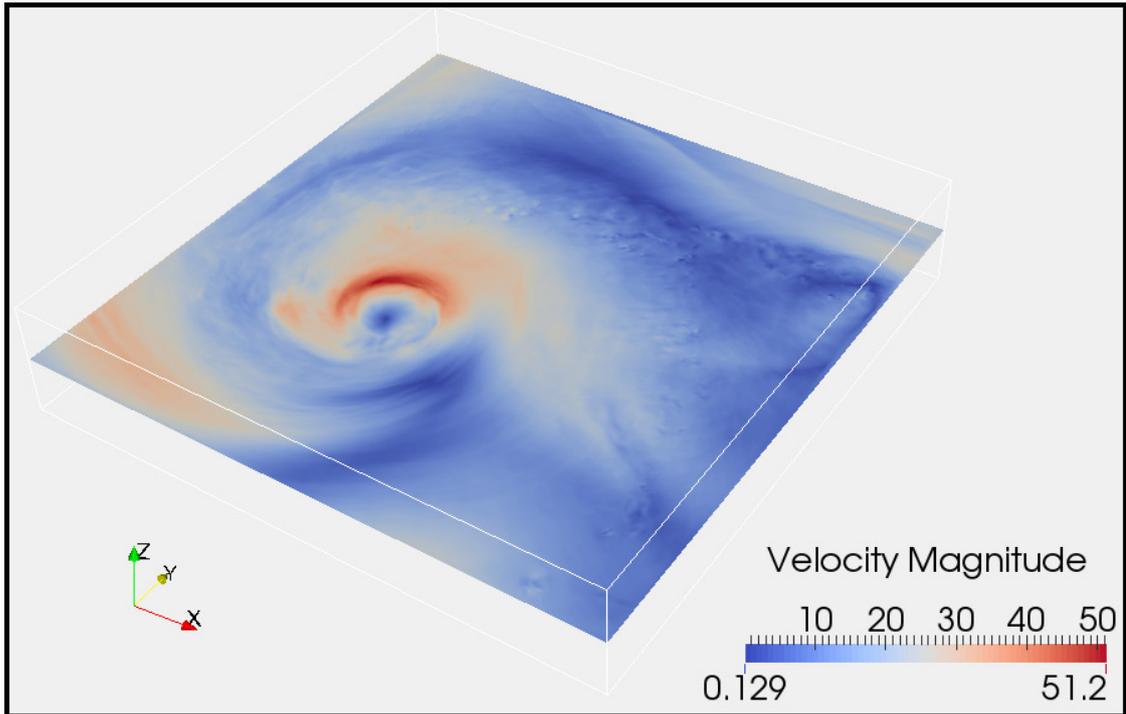


Figure 5.3: A 2D slice from the simulation data of Hurricane Isabel, rendered using color coding representing velocity magnitudes.

Arrow Glyph

Arrow glyph methods are able to represent directions as well as magnitudes. They're widely used in the community for simple flow fields having small data sets due to high comprehensibility and ease of implementation. However, arrow glyph technique is vulnerable to cluttering and occlusion for large datasets and unsteady flow fields. As it can be seen in Figure 5.4 and 5.5, use of arrow glyph produces more obscure results for flow fields in 3D, then the ones in 2D.

In this method, vector magnitudes can also be color coded on the glyph, and normalized unit vectors can be used instead of scaling to represent the relative size of the real vector. Glyph geometry and density can also be adjusted in order to produce less cluttered and more readable visualizations.

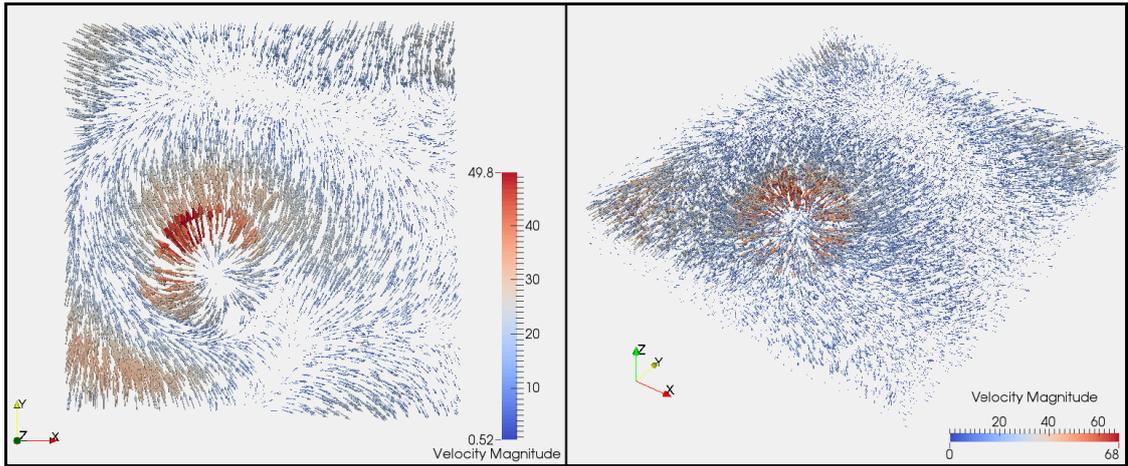


Figure 5.4: Arrow glyph visualization of a 2D slice from the simulation data of Hurricane Isabel on the left, and the use of arrow glyph technique in 3D for the whole dataset.

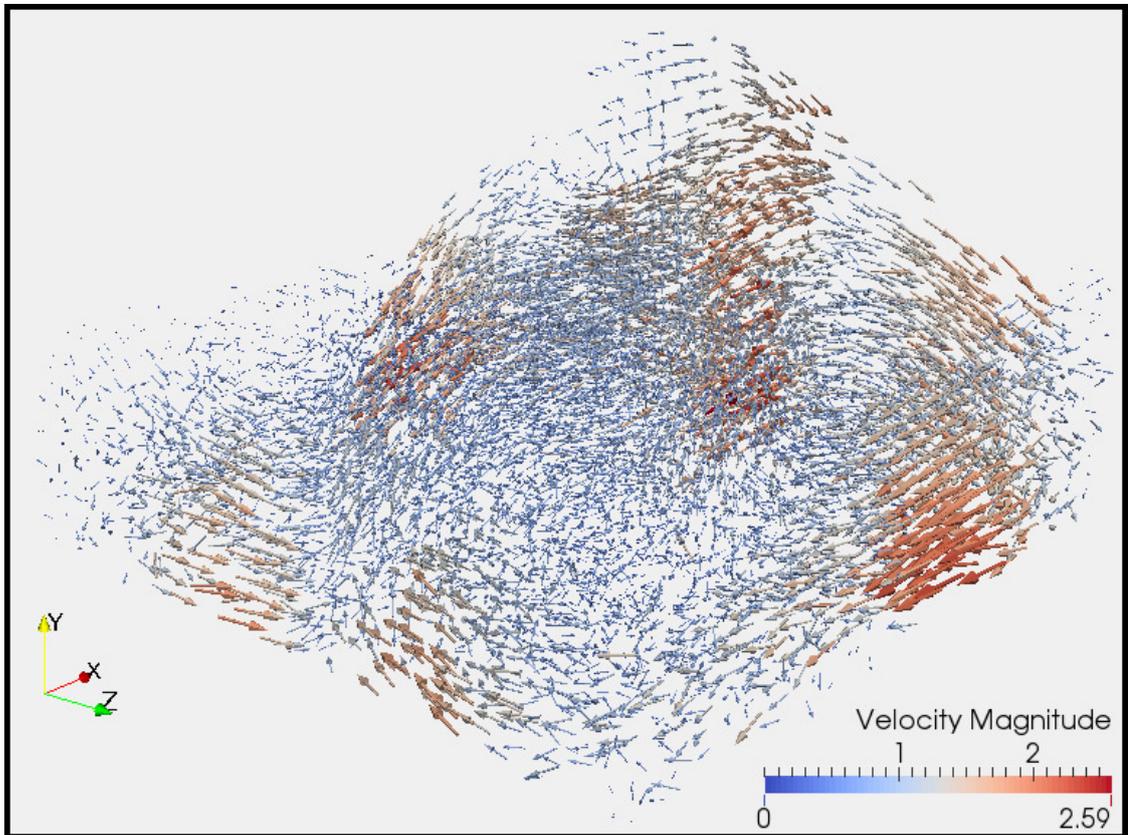


Figure 5.5: Visualization of 3D fluid flow generated by SPH simulations using arrow glyph method.

Simple Future Selection

There are several direct rendering methods after filtering data with simple methods such as applying thresholds or grouping data on selected intervals to represent features based

on values. Such techniques include generating iso-lines and iso-surfaces based on vector magnitudes or intervals of angles. These techniques are not suitable for unsteady flows, and iso-surfaces are vulnerable to occlusions in 3D.

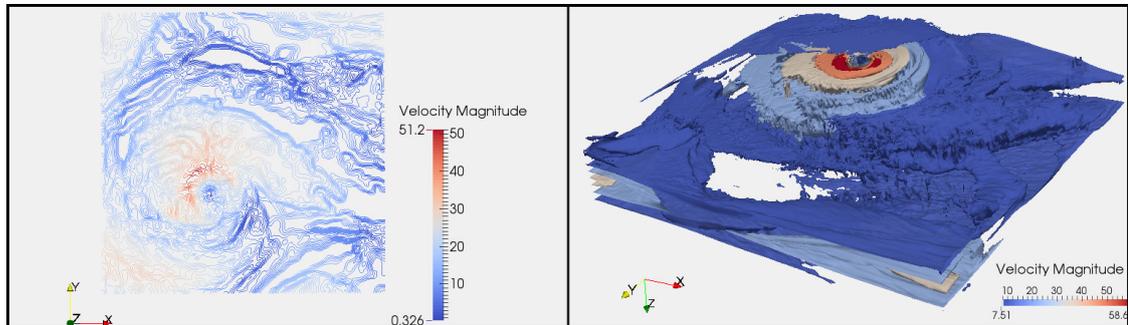


Figure 5.6: A 2D slice from the simulation data of Hurricane Isabel used to generate iso-lines representing vector magnitudes is shown on the left, and iso-surfaces representing the same dataset are given on the right.

Transfer functions used in volume rendering also serves as a filter to eliminate or emphasize some of the features. A scalar value can be mapped the transparency value, and depending on the ranges of values some parts of the data contribute more on the final rendering, while the rest might be completely hidden.

5.1.2 Integration Based Geometry Extraction Methods

There are many methods in literature for integration based geometry extraction, and the majority of those methods are originated from streamlines. Streamlines and other derivations are commonly used by Computational Fluid Dyanmics community, and many commercial and open source tools are already available.

Streamlines

Streamlines are the curves those are tangents of the vectors in the flow field. Pathlines are the particle trajectories a mesh-less particle takes in the flow field in time, and streak-lines are the curves connecting the particles which are seeded from the same location. All those methods require choosing spatial locations to start integration, and defining

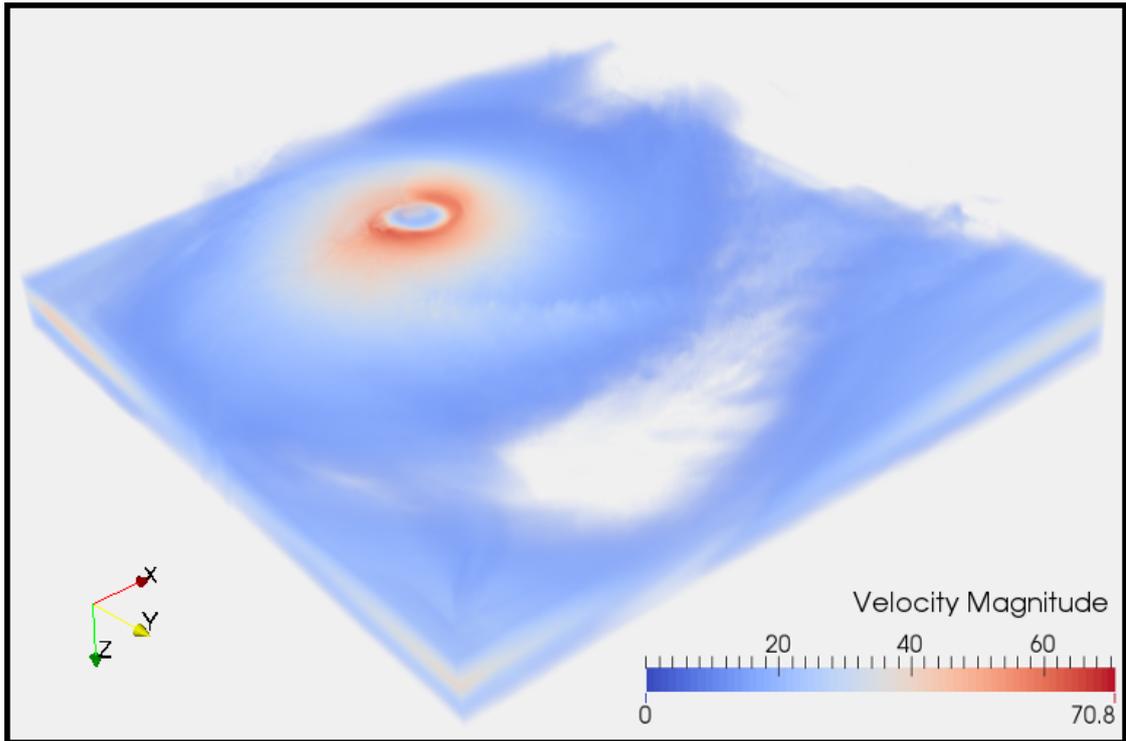


Figure 5.7: Hurricane Isabel data volume rendered in 3D, a linear transfer function scaled to vector magnitudes is used for transparency and color.

intervals in spatial or time domain. In Figures 5.8 and 5.9, SPH and Hurricane Isabel data sets are used to illustrate streamlines.

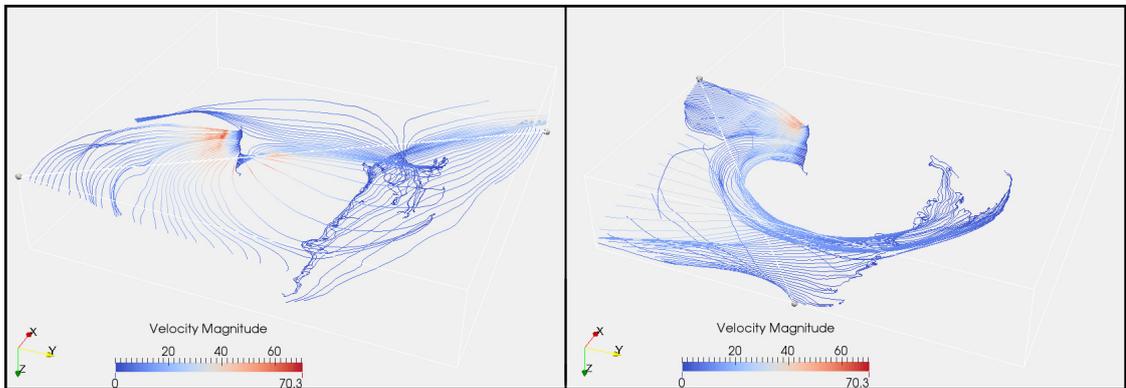


Figure 5.8: Streamline generation for Hurricane Isabel data set along two different lines used for seeding.

As it can be seen in Figures 5.8 and 5.9, geometry of the generated streamlines are heavily depending on the locations to start integration. It's shown that different locations to seed the streamlines on the same data set cause having streamlines with different

geometry and characteristics. Integration for the streamlines can be performed both on forward and backward directions, and it's also important to define the maximum length of the interval to continue integration.

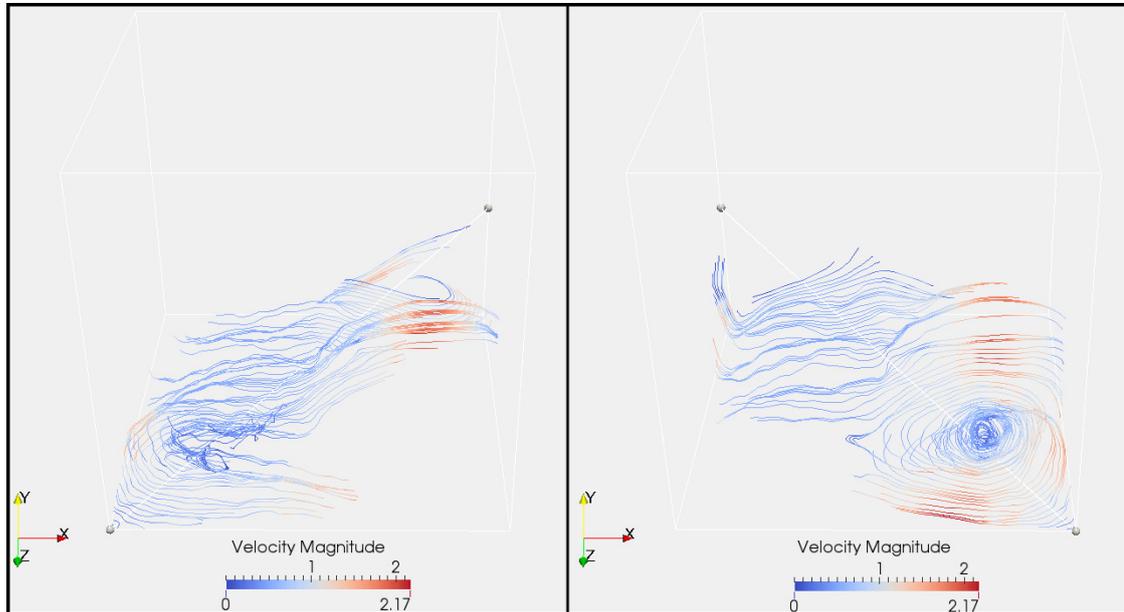


Figure 5.9: Streamline generation for SPH data set along two different lines used for seeding.

Streamlines are also vulnerable to occlusion and cluttering based the seeding locations and integration intervals. For an unexplored flow field, it's big challenge to find the right location for seeding streamlines without causing occlusion or cluttering.

5.2 Use of Information Theory in Visualization

Information Theory is able to aid visualization in Computer Graphics by defining measures to quantify the information content of the data itself, and the amount of information passed through the visualization pipeline after processing with or without losses. For the flow fields, we use an Information Theory based approach to determine the information content of data on a spatial domain and assist visualization. Existing approaches that are using Information Theory for the visualization of the flow fields are already reviewed in Chapter 2.

Our motivation for involving Information Theory in the evaluation of 3D vector data is emerged from the success of utilizing Information Theory in our approach used to select viewpoints for 3D polygonal meshes [50]. In this approach, we measure the information content before and after the visualization to find optimal viewpoints while still keeping the salient locations visible.

5.2.1 Case Study and Inspiration: Viewpoint Selection for 3D Polygonal Meshes

Information Theory helps us to quantify the two projected information of a 3D model, the faces of the model and its salient features. In our work [50] Viewpoint Kullback-Leibler distance introduced by Sbert et al. [48] is employed and *Viewpoint Saliency Kullback-Leibler (vSKL)* is presented as a novel view descriptor.

VKL distance is an informational theoretical measure which is interpreted as the difference between the normalized distribution of projected areas and the ideal projection. The ideal projection is given by the normalized distribution of the actual area of the triangles. Viewpoint Kullback–Leibler distance measure depends on the concept of *Viewpoint Entropy* introduced by Vazquez et al. [59] which is used to determine the amount of information from a viewpoint using the projected faces of the model.

We introduce a novel view descriptor, *vSKL* distance, based on the idea by Lee et al. [26]. It is an information of regional importance which is considered as the salient feature of the model or the graphics meshes. We map the good or best definition as a camera position where the perception of two defined information is maximized. Our contributions are introduction of a novel view descriptor called *vSKL* distance and integrating N-Best View Selection with *vSKL* distance descriptor in view selection for generating “good” representational images.

Mesh Saliency

Mesh Saliency is the concept of regional importance, which can be specified as distinction in pixel colors, or luminance or geometric attributes. In our approach we borrow the techniques [23], and [26] to calculate curvature based mesh saliency. Curvature is one of the important features of a vertex which can point-out its distinctiveness among the other vertices. We use the Gaussian filtered mean curvatures of vertices proposed by [26] using Taubin's procedure to calculate mean curvatures [57]. Meyer et al. [31] also provide a technique to calculate surface mean curvatures. Let $N(v, \sigma)$ be the set of points within a distance σ for vertex v therefore $N(v, \sigma) = \{x \mid \|x - v\| < \sigma, x \text{ is a meshpoint}\}$, and let $S(v)$ denote the surface mean curvature, hence Gaussian-weighted average of the surface mean curvature $G(S(v), \sigma)$ can be defined as;

$$G(S(v), \sigma) = \frac{\sum_{x \in N(v, 2\sigma)} S(x) \exp\left(-\frac{\|x-v\|^2}{2\sigma^2}\right)}{\sum_{x \in N(v, 2\sigma)} \exp\left(-\frac{\|x-v\|^2}{2\sigma^2}\right)} \quad (5.1)$$

In equation(5.1), a cut-off distance for the Gaussian filter is assumed to be 2σ . The saliency for vertex v is the absolute difference between coarse and fine scales, where the coarse scale standard deviation is twice of the fine scale. Then the saliency for vertex v for multiple scales is,

$$M_i(v) = |G(S(v), \sigma_i) - G(S(v), 2\sigma_i)| \quad (5.2)$$

where σ_i is the standard deviation of the Gaussian filter at scale i . We used five scales that are mentioned in [26] with this work. After the calculation of curvature saliency for five different scales we linearly added those feature maps after the normalization method proposed by Itti et al. [23] hence denote $M(v)$. The calculated feature map for a hand model is shown in Figure 5.10, where the hot colors show the high salient points.

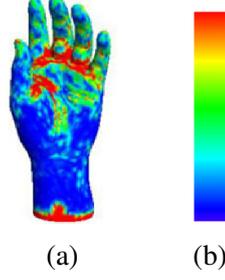


Figure 5.10: Mesh saliency for a hand model shown in (a). HSV color model shown in (b) is used to mark the saliency of the vertices. Hot colors (red) $Hue=0$ shows the highest saliency, and $Hue=240$ for the lowest. *Saturation* and *Value* are kept fixed in distribution.[50]

Viewpoint Saliency Kullback-Leibler

Viewpoint Saliency Kullback–Leibler distance is defined between *true* probability distribution $p = \{p(x)\}$ and target probability distribution $q = \{q(x)\}$. Using the saliency distribution, *vSKL* distance is denoted by

$$KL_{vs} = \sum_{i=1}^{N_f} \frac{a_i s_i}{s_t} \log_b \frac{\frac{a_i s_i}{s_t}}{\frac{A_i S_i}{S_t}} \quad (5.3)$$

where s_i is the projected saliency amount of polygon i , $s_t = \sum_{i=1}^{N_f} a_i s_i$. $A_i S_i$ is the actual saliency amount of polygon i and $S_t = \sum_{i=1}^{N_f} A_i S_i$ is the total saliency of the 3D object. In order to select high quality views using the saliency distribution KL_{vs} should be minimized. Because the minimization of *vSKL* distance is based on the visibility and captures the maximum number of polygons with the maximal uniformity of the projected areas as well as uniformity of the projected saliency.

The greedy best view selection technique [49] is extended to use newly introduced *vSKL* distance as view descriptor. In Figure 5.11, we compare the output of viewpoints for Stanford Bunny model are compared. The computed five viewpoints on top row are from the approach presented in [48] and [60], and the bottom row presents the images from our *vSKL*. This model has total 69743 faces. Figure 5.11.(a)-(e) cover the 68152

faces of the model, Figure 5.11.(f)-(j) can cover 68251 faces, and our method covers 68149 faces of the model. We can observe that the face coverage perturbation is minimal in $vSKL$ distance method. In our algorithm saliency information is also conveyed along with the face coverage due to the definition of probability mass function in $vSKL$ distance.

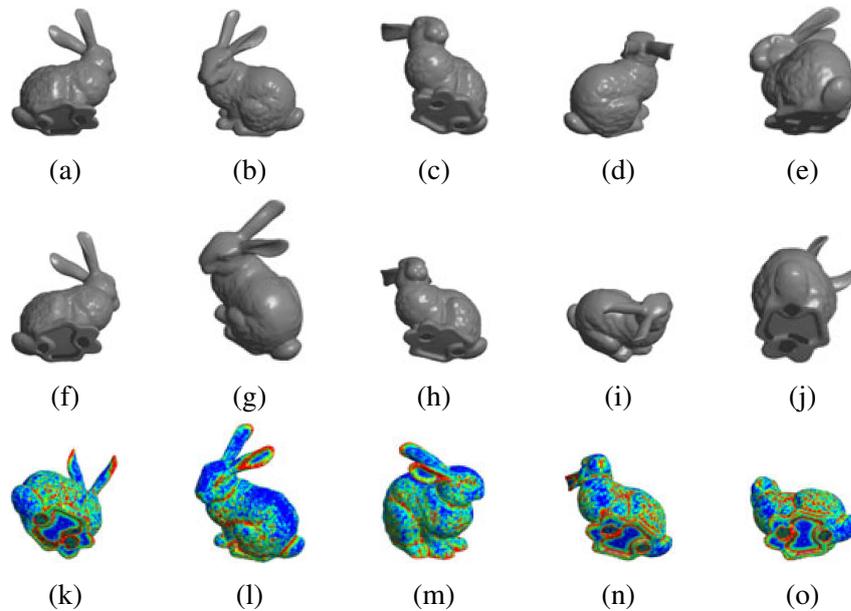


Figure 5.11: Stanford Bunny is displayed with five viewpoints using the approach from [48] and [60] compared to greedy integrated $vSKL$ method.

5.2.2 Entropy Guided Visualization of Multivariate Spatio-Temporal Data

Information Theory is a very powerful tool to quantify information content of the data, and in addition to assisting viewpoint selection by providing distance measures, we noticed that it's also used for flow visualization as well as determining salient features in 3D data fields. Although the use of entropy to highlight important features of volumetric vector fields is existing in the area of scientific visualization[62, 65], our approach differs in evaluating vector directions and magnitudes while creating histograms to calculate entropy. Revisiting Shannon's Entropy [59] in Equation 5.4;

$$H(x) = - \sum_{i=1}^n p(x_i) \log_b p(x_i) \quad (5.4)$$

Probability distribution function $p(x_i)$ is needed to be defined as the first step in order to calculate an entropy field. Characteristics of the entropy field strongly depend on underlying definition of probability distribution function. Having a variable x with a set of possible values from x_1 to x_n , let $C(x_i)$ give count of values x_i encountered in the space. Probability of having value x_i is defined as;

$$p(x_i) = \frac{C(x_i)}{\sum_{j=1}^n C(x_j)} \quad (5.5)$$

It's trivial to calculate $p(x)$ when the variable is a scalar integer, and there are enough samples in the subspace. For each bucket x_i , we simply count how many samples are falling inside the same bucket having equal integer values. When a variable is a floating point number, a threshold needed to be defined in order to place the samples into the same bucket within the threshold. Treating each distinct floating number separately would cause the sample space to be too sparse. So the general approach is placing the buckets uniformly, such as simply casting the floating point numbers to integer values. As a convention, bin is the synonym of bucket and those words are used interchangeably while counting number of samples into bins to determine probability distribution is referred as histogram generation.

For the vector fields, common approaches are converting 3 dimensional coordinates to angle/magnitude pairs, they ignore the magnitudes or calculating entropy fields separately to join them later with a linear interpolation based on user defined weights. Our method differs from the others in terms of histogram generation, and we propose two different approaches to utilize angles and magnitudes together before and after entropy calculation.

Histogram Generation for 3D Vector Fields

The general approach is to calculate entropy for each voxel within a local neighborhood determined by window size. For the boundary voxels, mirroring used to avoid problems due to not having enough samples. Let v_a to v_n represents a set of vectors representing buckets for possible vectors, number of vectors in each bucket v_a for the neighborhood W can be counted in Equation 5.6.

$$C(v_a) = \sum_{i,j,k} [W_{ijk} \cong v] \quad (5.6)$$

Then $C(v)$ or so called histogram function can be used in calculating probability distribution of the vectors in 3D space in Equation 5.7.

$$p(v_a) = \frac{C(v_a)}{\sum_{b=1}^n C(v_b)} \quad (5.7)$$

Equation 5.4 for entropy calculation remains intact, only probability distribution function $p(v)$ is replaced in Equation 5.8.

$$H(x) = - \sum_{i=1}^n p(x_i) \log_b p(x_i) \quad (5.8)$$

For the calculation of entropy, it's important to choose right set and amount of representative vectors v_a to v_n , and how the congruence relation \cong is defined to assign vectors in the sample space to the buckets of the histogram represented by v_a to v_n . Xu et al. [65] partition the range of vectors represented by polar angle Θ , $0 < \Theta \leq 2\pi$ for 2D vectors. For 3D vectors, they decompose unit sphere into 360 patches of equal area to use cones that connect sphere center and patches to quantize the vectors. This approach is ignoring the magnitudes and has several other drawbacks. Wang et al. [62] mention the concept

of multidimensional histograms, and offers calculating entropy fields separately to join them later.

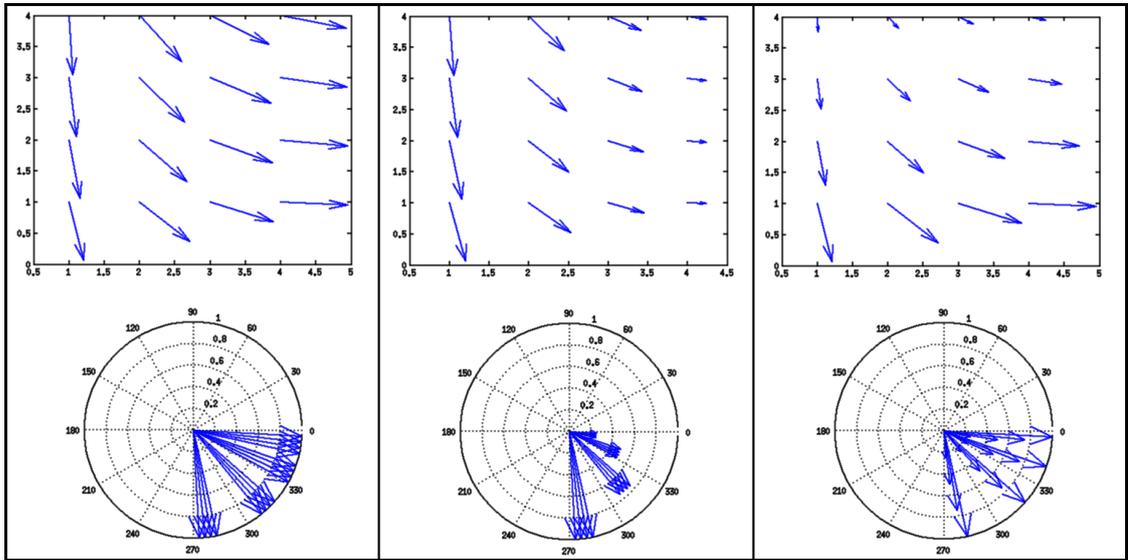


Figure 5.12: Sample vector field with varying directions and same magnitude is given on the left, and vector fields with the same directional component, but varying magnitudes are given on the middle and the right.

As it can be seen in Figure 5.12, ignoring vector magnitudes would cause all three distributions to have the same entropy. In case of calculating entropy of vector magnitudes separately, the distributions on the middle and on the right will have the same distribution since both have the variation in terms of the magnitudes. Our first approach is to rearrange the buckets so that we can group the vectors into the same bucket according to their directions and magnitudes. Instead of dividing the unit sphere into patches, we put equidistant points on the unit sphere. Number of points are empirically set, and each vector assigned to the closest point on the sphere like the other approaches. Instead of using one set of points on the unit sphere, we scale the sphere according to the global distribution of magnitudes and use the points on two spheres to represent the buckets in addition to a point on the center as in shown in Figure 5.13.

Comparison of our first method with the existing approach is given in Figure 5.14. As it can be seen from the figure, our approach can eliminate some of the high entropy fields appearing on the regular method due to high directional variation at low magnitude

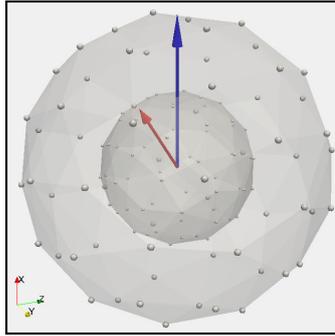


Figure 5.13: Blue vector is assigned to a point on the outer sphere, while red one is assigned to a point on the inner sphere.

fields. This might be desired behavior in some cases, while some low magnitude fields with high directional variation might be also canceled out. In order not to misinterpret the data, radiuses of the inner and outer spheres should be chosen carefully according to magnitude variation and focus of the analysis. In our first method, we empirically set this parameter, as well the bucket and window sizes, similar to the existing methods in the literature.

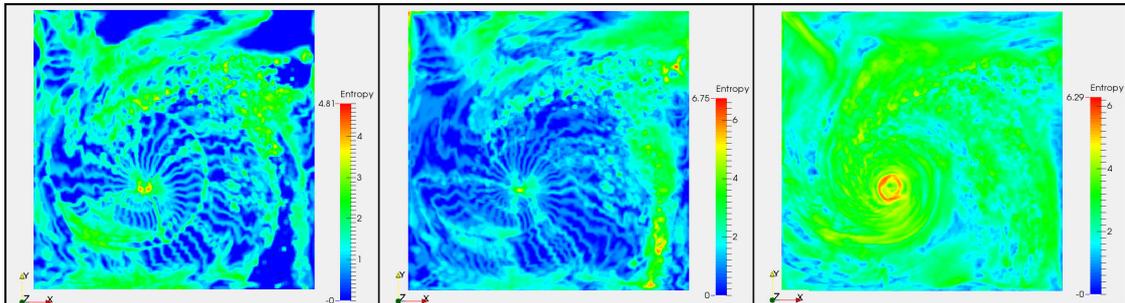


Figure 5.14: Entropy is calculated on Hurricane Isabel data using buckets of varying magnitude is on the left. Angular entropy field calculated with the regular approach is on the middle, and entropy field calculated from magnitudes is on the right.

Setting the window size depends on characteristics of data and desired level of detail. Unless there's too few samples falling inside the window, it's convenient to set windows size small in order to get as much spatial detail as possible, although high entropy fields might be less recognizable. During our calculations for 3D vector fields, we were able to cubic windows down to three samples per dimension. It's possible to make a visual comparison in Figure 5.15, and it can be seen that different window sizes would be

beneficial for different applications.

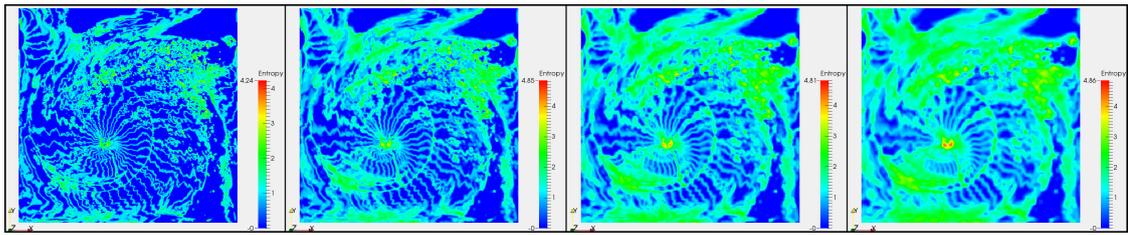


Figure 5.15: Entropy is calculated on Hurricane Isabel data using different window sizes. Each dimension of the cubic windows are 3,5,7,9 samples wide.

Another important aspect is setting the bucket size, therefore the number of histogram bins. Effects of varying bucket sizes can be seen in Figure 5.16. Note that the number of thin lines originated from the center of the storm are changing according to the bin size. Those separating lines are appearing on low entropy fields when most of the samples are grouped into the same bucket, few samples falling onto border of one bucket may fall into neighboring bucket for the new sampling window with a slight change of direction.

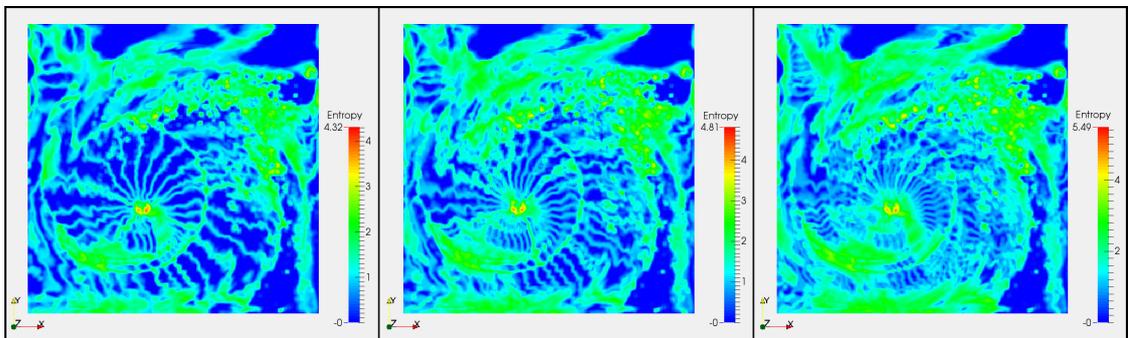


Figure 5.16: Entropy is calculated on Hurricane Isabel data using different histograms with varying bin sizes. We have used spheres with 180, 360 and 720 patches from left to right.

We propose another approach to generate histograms for 3D vectors using Principal Component Analysis (PCA). PCA is mainly used for transforming a set of observations into a set of linearly uncorrelated variables called principal components in a way that the first principal component has the highest variance, and next principal components have to be orthogonal to the preceding components while having the highest possible variance

in a decreasing order. Motivated from PCA, we use Singular Value Decomposition (SVD) to find a new coordinate system for each window, in which the projection of 3D vectors onto the new XY plane produces highest directional variation among all possible projections onto any 2D plane (Figure 5.17).

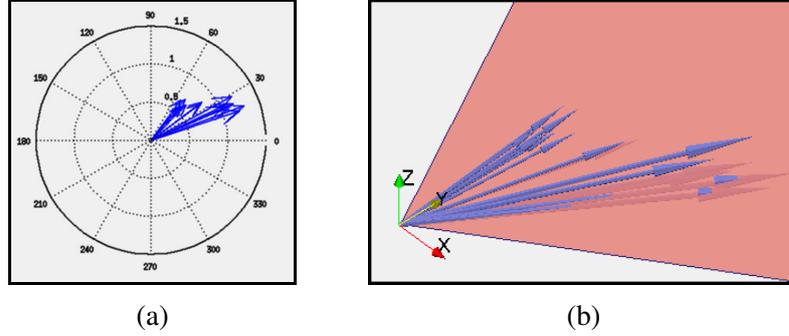


Figure 5.17: Sample set of vectors, and the projection plane found by principal component analysis.

In principal component analysis, eigenvectors corresponding to the largest eigenvalues are found which gives the directions of most variation in data. In order to find principal components, we first mean center the set of vectors V in each window, and take the transpose to find matrix X as given in Equation 5.9.

$$x_{ji} = v_{ij} - \frac{1}{n} \sum_i^n v_{sij} \quad (5.9)$$

Using Singular Value Decomposition, we find the values of matrix U which is the set of eigenvectors in the order of decreasing eigenvalues in Equation 5.10.

$$X = UTV^T \quad (5.10)$$

Then using first eigenvectors from U we find \bar{U} , as the transformation matrix for projecting every member of V to the new coordinate frame, where the most angular variation occurs on XY plane. Note that we're still on a three dimensional coordinate

system after applying Equation 5.11.

$$V_{projected} = \bar{U}V \quad (5.11)$$

At the end, we use polar coordinate transformations to calculate Θ for angles, r for radiuses and keep the transformed z values. After applying entropy calculation for each scalar value using properly sized histograms, results can be seen in Figure 5.18. In Figure 5.18 we have used windows 7 samples wide in each direction. Histogram used for angular component had 36 bins, and other components had histograms up to 50 bin according to the range of the data. As it can be seen from the figure, results are pretty consistent with the previously used entropy calculation methods. In addition, entropy maps produced with our method produces more distinctive regional boundaries, not interfered with thick lines due to histogram discretization as shown in Figure 5.19. Note that, artifacts due to discretization are still present, but they're not forming boundaries as straight lines. Artifacts are curvilinear due to rotating coordinate frames of the histogram bins, after performing PCA and applying coordinate transformation.

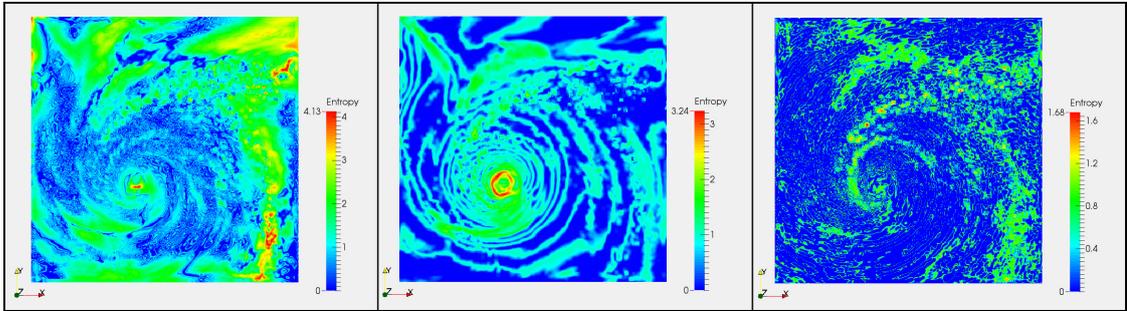


Figure 5.18: Entropy calculated with our method on Hurricane Isabel dataset; entropy calculated with angle of direction on the projected plane is given on the left, entropy calculated with the magnitudes, and z coordinate after projection are given on the middle and left.

Second entropy map is calculated from r for radiuses of vectors in corresponding windows, and reveals important information about vector magnitudes which can also be used in analysis as a whole, or weighted averaging with the directional entropy field as an existing approach in entropy calculation. Third entropy map is produced with z

values in transformed coordinate systems for variation maximization, and corresponds to the components with less directional variation. Although low entropy values observed in the map verifies this assumption, this entropy map can also be used in visualization and analysis.

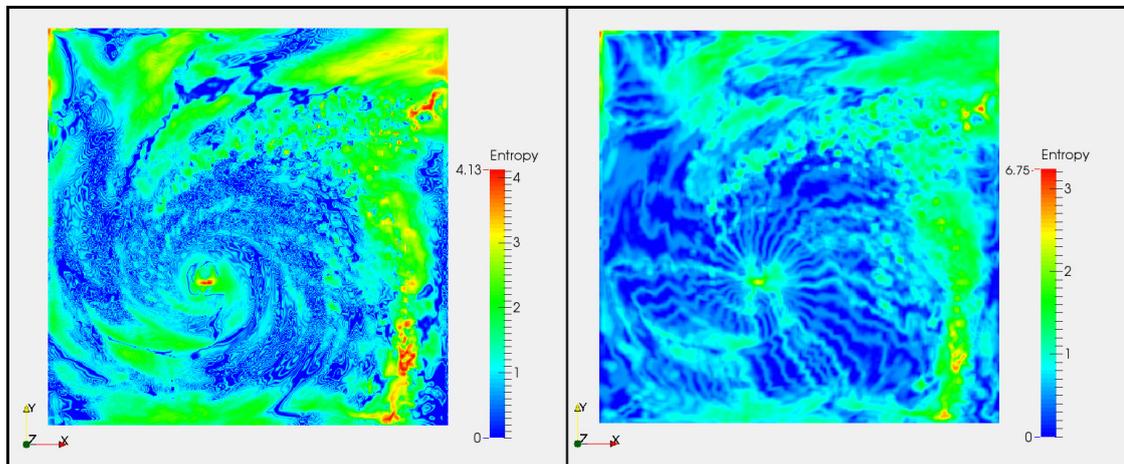


Figure 5.19: Directional entropy calculated after utilizing PCA is given on the left, and directional entropy calculated using unit sphere for discretization is given on the right.

We have also experimented with various bucket and window sizes. Like the previous methods, having too many buckets cause high entropy and less distinctive fields due to under sampling. On the opposite, having too few buckets cause vectors to fall into same buckets and produce low entropy fields with more artifacts due to discretization. Window size may also cause under sampling, and having small windows as much as possible without causing under sampling produce sharper boundaries.

Results and Discussion

We have utilized Information Theory to calculate entropy for the purpose of aiding visualization and analysis of vector flow fields. Entropy is used to measure the uncertainty of a random variable, and high entropy regions indicate the areas of high directional variation in vector flow field. On the opposite, low entropy field is an indicator of order, and entropy map of the directional component reveals the areas where the vector directions are coherent.

In order to demonstrate the utilization of entropy fields generated with the proposed approach, different direct rendering methods are merged to visualize SPH simulation and Hurricane Isabel data sets in 3D. In Figures 5.20 and 5.21, we have used volume rendering for the entropy field, and due to the transfer function linearly correlated with the entropy value, low entropy regions are transparent. Since we have used directional entropy map, high entropy regions are chaotic in terms of vector directions, while the directions are coherent in low entropy areas. Due to this reason, we utilized arrow glyph method on low entropy regions to reveal flow behavior without cluttering.

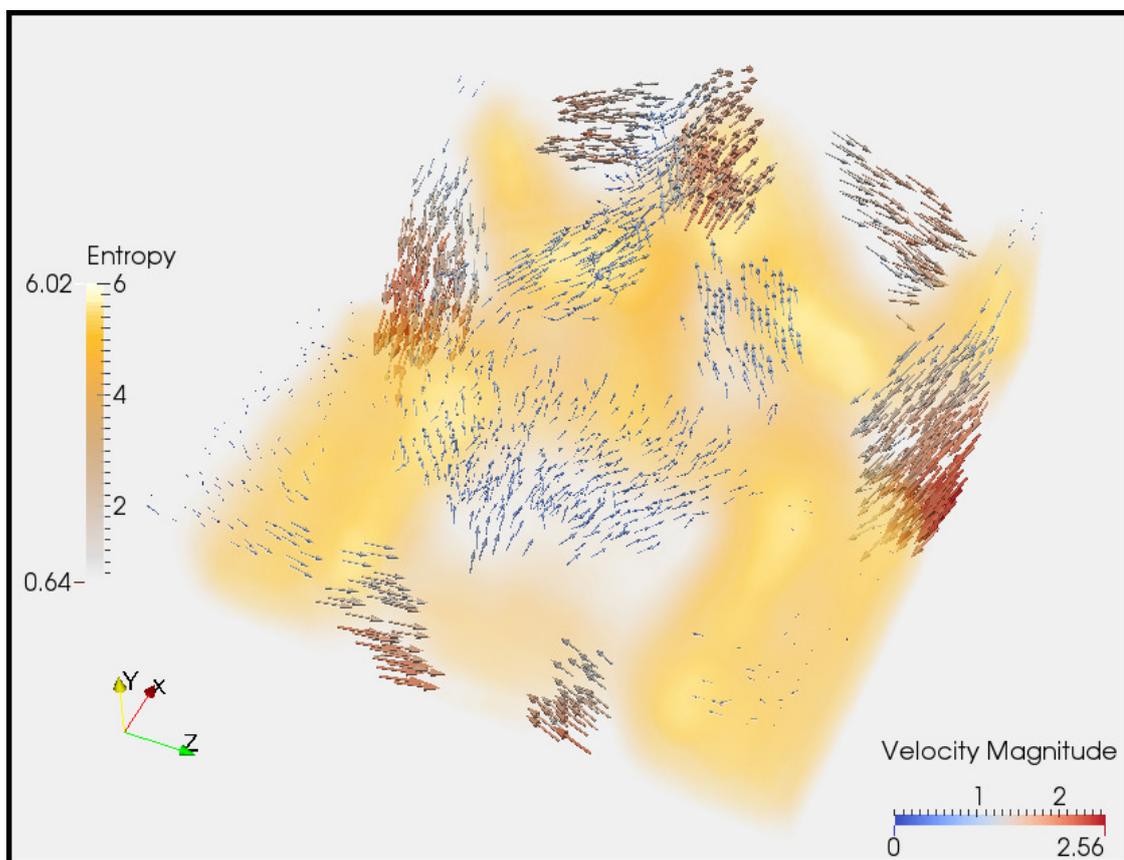


Figure 5.20: SPH simulation data is rendered using color coding for the entropy values and arrow glyph for the velocity vector where the entropy value is below the threshold to reveal fluid behavior.

Since SPH simulation data in figure 5.20 has very high variation in three dimensions, proposed approach produces much more readable and informative visualizations than the direct rendering techniques described in previous sections. Reader is advised to

revisit Figures 5.1 and 5.5 for a direct visual comparison.

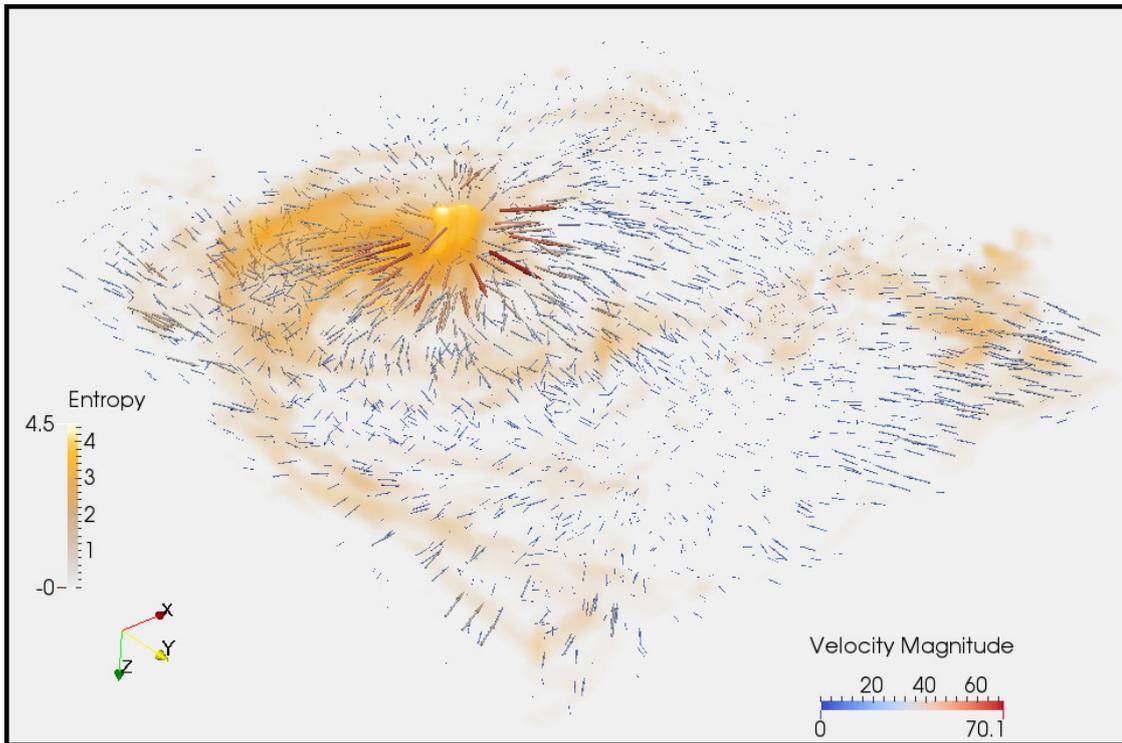


Figure 5.21: Simulation data of Hurricane Isabel is rendered using color coding for the entropy values and arrow glyph for the velocity vector where the entropy value is below the threshold to reveal fluid behavior.

Hurricane Isabel data rendered in Figure 5.21 with the proposed approach also reveals important information, such as high entropy areas like the center of vortex, as well as the directional information around it. Note that the resolution of the data set is much higher than the SPH simulation data set, and visualization technique used in our approach produces a representative view of the whole field in a single picture.

Further Investigations

Information Theory guided visualization of multivariate spatio-temporal data is a promising field to explore for further investigations using our approach. Our approach for generating entropy fields using principal component analysis can be used to fuse other attributes. Many scalars and vectors can be fused while finding principal components with a joint or cascaded manner, and/or entropy maps can be joined later. It would also

be interesting to investigate the use of Information Theory and entropy fields to navigate through different time frames. Creating a representative image with the joint information extracted from the whole data set might be useful as well. Although re-sampling of large data sets could be cumbersome, creating local histograms from higher resolution vector fields created by interpolation might avoid problems due to sparsity for small windows. Entropy based level-of-detail rendering for the visualization would be possible for large data sets.

6 CONCLUSION

In this thesis, we presented a framework for visualization of multivariate spatio-temporal data sets generated by physically based simulation, and we mostly focused on vector flow fields. Our main contribution is;

- Utilizing PCA to generate histograms of 3D vector fields by polar coordinate transformation.

To the best of our knowledge, this is the first work to use PCA for generating histograms of the 3D vector fields with polar coordinate transformation in order to calculate entropy fields, and perform entropy guided visualization. Our method is less prone to discretization errors than the previous methods used for histogram generation by grouping 3D vectors into the buckets with fixed geometry in 3D space, since dimensional reduction allows us to use less number of buckets oriented in space according to the variation of data, and avoids sparsity. After projection, entropy field from the directional component is generated as well as the entropy fields those are generated from vector magnitudes, and z-coordinates in cylinder coordinate system as the remaining components. Those entropy fields can be investigated separately or joined together for investigation.

We have also evaluated entropy fields generated by using our method, and we performed entropy guided visualization of flow fields for demonstration. Note that for entropy guided visualization, it's possible use any combination of the topology based flow visualization methods [40, 30] with our approach, and choose different methods according to characteristics of vector fields in local regions partitioned by entropy. Regions of interest can also be defined according to the entropy ranges, and entropy field itself can reveal the characteristics and behavior of underlying flow field.

In addition, entropy fields can be used to monitor information content of other data sets rather than flow fields, and multivariate data sets having 3D vector components can take advantage of using our method for histogram generation.

Our additional contributions involving Information Theory for the visualization can be listed as;

- A histogram generation method for 3D vector fields taking magnitudes and directions into account,
- Motivated from Information Theory, introduction of vSKL distance for generating representational images of 3D polygonal meshes.

For the previous methods performing histogram generation by using fixed buckets in 3D space, we proposed to involve directional components of 3D vectors. We revised the existing strategy for assigning vectors into buckets those are defined on a unit sphere as equi-areal patches. Instead, we used equidistant points on multiple spheres to represent buckets, and assigned vectors according to proximity. Like the previous methods, number of buckets and radiuses of spheres are set empirically according to the statistical variation of data.

While exploring Information Theory, we started from generating representational images of 3D polygonal meshes using Information Theoretical distance metrics. A well known metric for measuring information loss after 2D projection of a 3D polygonal mesh is called viewpoint Kullback-Leibler (vKL) distance, and previous approaches involve vKL to generate representative images by maximizing viewpoint information. We added curvature based mesh saliency as a local attribute to consider while finding best viewpoints, and proposed using viewpoint Saliency Kullback-Leibler (vSKL) distance as another contribution.

We developed our own framework to generate flow fields by physically based simulation in addition to using data sets available for research purposes. During the process of our work, we also contributed to;

- The development of an SPH framework utilizing fractional derivatives to improve performance and stability.

We proposed to introduce fractional derivatives to the viscosity term of governing equations, and produced results similar to the ones obtained with high-resolution SPH simulations. Our experiments with the results also revealed that standard visualization techniques are not adequate for the visualization of chaotic flow fields. On the other hand, we managed to apply our entropy guided visualization approach on the same data sets successfully, and we present representative images to reveal flow behavior.

6.1 Future Work

Visualization and analysis of spatio-temporal data is a challenging problem that has many other aspects for further investigation, and many improvements motivated from Information Theory might be possible in addition to our contributions.

In terms of utilizing PCA for histogram generation, our approach can be extended to allow fusing other vector and scalar attributes of the same data set. In addition, multi-modal data sets are in interest of many researchers, and our method can also be involved in fusing those data sets with vector or scalar attributes.

Although it can be seen by visual comparison that our histogram generation method produces fewer artifacts than the conventional methods, a quantitative analysis of our work can be performed as a future work. Our method might be improved by allowing variable bucket sizes instead of fixed intervals, in that case a quantitative analysis will be a necessity to investigate small differences.

Entropy fields are used for determining most informative time frames for navigation as well as compression in time. Our histogram generation method can be further evaluated by generating entropy fields to assist time navigation while visualization.

It will also be beneficial to evaluate entropy guided visualization of flow fields together with the conventional methods that are already being used by Computational

Fluid Dynamics (CFD) community to determine areas of interest.

We believe Information Theoretical approaches in scientific visualization will be a hot topic for a while because of having many existing and new application areas, producing interesting problems to investigate for many researchers.

Biography

Selcuk Sumengen received his B.S. and M.S. (with TUBITAK fellowship) in Computer Science and Engineering from Sabanci University in 2004, and in 2006. Currently, he is also a member of the image processing team and working as a R&D Software Engineer for the largest software company in Turkish Defense Industry. His main research interests are mostly in the area of Computer Graphics, particularly Physically Based Simulation, Parallel & Distributed Computing, and Information Visualization.

Refereed Journals

Ekrem Serin, Selcuk Sumengen, and Selim Balcisoy. Representational image generation for 3d objects. *The Visual Computer*, 29(6-8):675–684, 2013.

Oktar Ozgen, Selcuk Sumengen, Marcelo Kallmann, Carlos FM Coimbra, and Selim Balcisoy. Simulating colliding flows in smoothed particle hydrodynamics with fractional derivatives. *Computer Animation and Virtual Worlds*, 2013.

Serdar Cakici, Selcuk Sumengen, Ugur Sezerman, and Selim Balcisoy. Dockpro: a vr-based tool for protein-protein docking problem. *International Journal of Virtual Reality*, pages 19–23, 2009.

Book Chapters

Selcuk Sumengen, MustafaTolga Eren, Serhat Yesilyurt, and Selim Balcisoy. A multi-resolution mesh representation for deformable objects in collaborative virtual environments. In Jose Braz, Alpesh Ranchordas, HelderJ. Araujo, and JoaoMadeiras Pereira, editors, *Computer Vision and Computer Graphics. Theory and Applications*, volume 21 of *Communications in Computer and Information Science*, pages 75–87. Springer Berlin Heidelberg, 2009.

Refereed Conference Papers (International)

Aysegul Cayci, Selcuk Sumengen, Cagatay Turkey, Selim Balcisoy, and Yucel Saygin. Temporal dynamics of user interests in web search queries. *Advanced Information Networking and Applications Workshops, International Conference on*, 0:762–767, 2009.

Selcuk Sumengen, Mustafa Tolga Eren, Serhat Yesilyurt, and Selim Balcisoy. Real-time deformable objects for collaborative virtual environments. In Jose Braz, Pere-Pau Vazquez, and Joao Madeiras Pereira, editors, *GRAPP (AS/IE)*, pages 121–128. INSTICC - Institute for Systems and Technologies of Information, Control and Communication, 2007.

B. Alper, S. Sumengen, S. Balcisoy, B. Alper, S. Sumengen, and S. Balcisoy. Dynamic visualization of geographic networks using surface deformations with constraints. *Computer Graphics International, International Conference on*, 2007.

Refereed Conference Papers (Local)

Selcuk Sumengen, Caglar Senaras, and Ahmet Erdem. Yuksek cozunurluklu uydu goruntuleri icin pankromatik keskinlestirme yontemi. In *SAVTEK, Ulusal Savunma Teknolojileri Kongresi*, page 80, 2012.

Caglar Senaras, Selcuk Sumengen, and Ahmet Erdem. Uzaktan algilama uygulamalari icin uzamsal veri yonetim sistemi. In *SAVTEK, Ulusal Savunma Teknolojileri Kongresi*, page 11, 2012.

Billur Engin, Emre Koc, Selcuk Sumengen, and Selim Balcisoy. Yuksek cozunurluklu arazi verilerinin grafik islemcisi kullanilarak gorsellestirilmesi. In *USMOS, Second National Defence Applications Modelling and Simulation Conference*, 2007.

Selcuk Sumengen and Selim Balcisoy. Hava ve kara arac gruplarinin detayli arazi verisi uzerinde gercek zamanli simulasyonu. In *HITEK, Havacilikta Ileri Teknolojiler Sempozyumu*, 2004.

Bibliography

- [1] Bart Adams, Mark Pauly, Richard Keiser, and Leonidas J. Guibas. Adaptively sampled particle fluids. In *SIGGRAPH '07: ACM SIGGRAPH 2007 papers*, page 48, New York, NY, USA, 2007. ACM.
- [2] Markus Becker and Matthias Teschner. Weakly compressible sph for free surface flows. In *SCA '07: Proceedings of the 2007 ACM SIGGRAPH/Eurographics symposium on Computer animation*, pages 209–217, Aire-la-Ville, Switzerland, Switzerland, 2007. Eurographics Association.
- [3] Cindy Bruyere Tim Scheitlin Bill Kuo, Wei Wang and Don Middleton. Hurricane isabel data produced by the weather research and forecast (wrf) model, courtesy of ncar and the u.s. national science foundation (nsf)., 2004.
- [4] U.D. Bordoloi and Han-Wei Shen. View selection for volume rendering. In *Visualization, 2005. VIS 05. IEEE*, pages 487–494, 2005.
- [5] R. Bramon, M. Ruiz, A. Bardera, I. Boada, M. Feixas, and M. Sbert. Information theory-based automatic multimodal transfer function design. *Biomedical and Health Informatics, IEEE Journal of*, 17(4):870–880, 2013.
- [6] Raphael Burger and Helwig Hauser. Visualization of multi-variate scientific data. In *EuroGraphics 2007 State of the Art Reports (STARs)*, pages 117–134, 2007.
- [7] A. Chaudhuri, Teng-Yok Lee, Bo Zhou, Cong Wang, Tiantian Xu, Han-Wei Shen, T. Peterka, and Yi-Jen Chiang. Scalable computation of distributions from large scale data sets. In *Large Data Analysis and Visualization (LDAV), 2012 IEEE Symposium on*, pages 113–120, 2012.
- [8] Cheng-Kai Chen, Shi Yan, Hongfeng Yu, Nelson Max, and Kwan-Liu Ma. An illustrative visualization framework for 3d vector fields. *Computer Graphics Forum*, 30(7):1941–1951, 2011.

- [9] M. Chen and H. Jaenicke. An information-theoretic framework for visualization. *Visualization and Computer Graphics, IEEE Transactions on*, 16(6):1206–1215, 2010.
- [10] C.F.M. Coimbra. Mechanics with variable order operators. *Annalen der Physik*, (12):692–703, 2003.
- [11] C.F.M. Coimbra, D. L’Esperance, A. Lambert, J.D. Trolinger, and R.H. Rangel. An experimental study on the history effects in high-frequency Stokes flows. *Journal of Fluid Mechanics*, (504):353–363, 2004.
- [12] C.F.M. Coimbra and R.H. Rangel. General solution of the particle equation of motion in unsteady Stokes flows. *Journal of Fluid Mechanics*, (370):53–72, 1998.
- [13] Mathieu Desbrun and Marie-Paule Gascuel. Smoothed particles: a new paradigm for animating highly deformable bodies. pages 61–76, 1996.
- [14] M. B. Liu G. R. Liu. *Smoothed Particle Hydrodynamics: A Meshfree Particle Method*. World Scientific Publishing Company, 2003.
- [15] R. A. Gingold and J. J. Monaghan. Smoothed particle hydrodynamics - theory and application to non-spherical stars. *Royal Astronomical Society*, 181:375–389, November 1977.
- [16] Wu Guoqing, Cao Yi, Yin Junping, Wang Huawei, and Song Lei. Information assisted visualization of large scale time varying scientific data. In *Virtual Reality and Visualization (ICVRV), 2011 International Conference on*, pages 38 –45, nov. 2011.
- [17] Takahiro Harada, Seiichi Koshizuka, and Yoichiro Kawaguchi. Smoothed particle hydrodynamics on gpus. *Structure*, pages 1–8, 2007.
- [18] Mattijs Janssens Andy Heather Sergio Ferraris Graham Macpherson Helene Blanchonnet Jenya Collings Henry Weller, Chris Greenshields. Openfoam, April 2010.

- [19] R. Hilfer. *Applications of Fractional Calculus in Physics*. World Scientific, River Edge, NJ, 2000.
- [20] Y. Hu. Integral transformations and anticipative calculus for fractional brownian motions. In *Memoirs of the American Mathematical Society*, 2005.
- [21] Markus Ihmsen, Nadir Akinci, Markus Becker, and Matthias Teschner. A parallel sph implementation on multi-core cpus. *Computer Graphics Forum*, 0(0):1–12, 2011.
- [22] Markus Ihmsen, Nadir Akinci, Marc Gissler, and Matthias Teschner. Boundary handling and adaptive time-stepping for pcisph. In *VRIPHYS*, pages 79–88, 2010.
- [23] L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 20(11):1254–1259, 1998.
- [24] J. Kehrer and H. Hauser. Visualization and visual analysis of multifaceted scientific data: A survey. *Visualization and Computer Graphics, IEEE Transactions on*, 19(3):495–513, 2013.
- [25] A.A. Kilbas, H.M. Srivastava, and J.J. Trujillo. *Theory and Applications of Fractional Differential Equations*. Amsterdam, The Netherlands, 2006.
- [26] Chang Ha Lee, Amitabh Varshney, and David W. Jacobs. Mesh saliency. *ACM Trans. Graph.*, 24(3):659–666, July 2005.
- [27] D. L’Esperance, C.F.M. Coimbra, J.D. Trolinger, and R.H. Rangel. Experimental verification of fractional history effects on the viscous dynamics of small spherical particles. *Experiments in Fluids*, (38):112–116, 2005.
- [28] Jun Ma, Chaoli Wang, and Ching-Kuang Shene. Coherent view-dependent streamline selection for importance-driven flow visualization. pages 865407–865407–15, 2013.

- [29] S. Marchesin, Cheng-Kai Chen, C. Ho, and Kwan-Liu Ma. View-dependent streamlines for 3d vector fields. *Visualization and Computer Graphics, IEEE Transactions on*, 16(6):1578–1586, 2010.
- [30] Tony McLoughlin, Robert S. Laramée, Ronald Peikert, Frits H. Post, and Min Chen. Over two decades of integration-based, geometric flow visualization. *Computer Graphics Forum*, 29(6):1807–1829, 2010.
- [31] Mark Meyer, Mathieu Desbrun, Peter Schroder, and Alan H. Barr. Discrete differential-geometry operators for triangulated 2-manifolds, 2002.
- [32] K.S. Miller and B. Ross. *An Introduction to the Fractional Calculus and Fractional Differential Equations*. John Wiley and Sons, New York, NY, 1993.
- [33] J J Monaghan. Smoothed particle hydrodynamics. *Reports on Progress in Physics*, 68(8):1703, 2005.
- [34] Matthias Müller, David Charypar, and Markus Gross. Particle-based fluid simulation for interactive applications. In *Proceedings of the 2003 ACM SIGGRAPH/Eurographics symposium on Computer animation*, SCA '03, pages 154–159, Aire-la-Ville, Switzerland, Switzerland, 2003. Eurographics Association.
- [35] G. Oger, M. Doring, B. Alessandrini, and P. Ferrant. An improved sph method: Towards higher order convergence. *J. Comput. Phys.*, 225(2):1472–1492, 2007.
- [36] K.B. Oldham and J. Spanier. *The Fractional Calculus*. Academic Press, New York, NY, 1974.
- [37] Oktar Ozgen, Marcelo Kallmann, L.E.S. Ramirez, and C.F.M. Coimbra. Underwater cloth simulation with fractional derivatives. *ACM Transactions on Graphics (TOG)*, 29(3):1–9, 2010.
- [38] Oktar Ozgen, Selcuk Sumengen, Marcelo Kallmann, Carlos FM Coimbra, and Selim Balcişoy. Simulating colliding flows in smoothed particle hydrodynamics

- with fractional derivatives. *Computer Animation and Virtual Worlds*, pages n/a–n/a, 2013.
- [39] H.T.C. Pedro, M.H. Kobayashi, J.M.C. Pereira, and C.F.M. Coimbra. Variable order modeling of diffusive-convective effects on the oscillatory flow past a sphere. *Journal of Vibration and Control*, 14(9-10):1659–1672, 2008.
- [40] Armin Pobitzer, Ronald Peikert, Raphael Fuchs, Benjamin Schindler, Alexander Kuhn, Holger Theisel, Krešimir Matković, and Helwig Hauser. The state of the art in topology-based visualization of unsteady flow. *Computer Graphics Forum*, 30(6):1789–1811, 2011.
- [41] I. Podlubny. *Fractional Differential Equations*. Academic Press, San Diego, CA, 1999.
- [42] L.E.S. Ramirez and C.F.M. Coimbra. A variable order constitutive relation for viscoelasticity. *Annalen der Physik*, 16(7-8):543–552, 2007.
- [43] L.E.S. Ramirez and C.F.M. Coimbra. On the variable order dynamics of the non-linear wake caused by a sedimenting particle. *Physica D*, 240,13:1111–1118, 2011.
- [44] Karthik Raveendran, Chris Wojtan, and Greg Turk. Hybrid smoothed particle hydrodynamics. In *Proceedings of the 2011 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, SCA '11, pages 33–42, New York, NY, USA, 2011. ACM.
- [45] Mats Rudemo. Empirical choice of histograms and kernel density estimators. *Scandinavian Journal of Statistics*, pages 65–78, 1982.
- [46] M. Ruiz, I. Boada, M. Feixas, and M. Sbert. Viewpoint information channel for illustrative volume rendering. *Computers & Graphics*, 34(4):351 – 360, 2010.

- [47] Mateu Sbert, Miquel Feixas, Ivan Viola, Jaume Rigau, and Miguel Chover. Information theory in computer graphics and visualization. In *SIGGRAPH Asia 2011 Courses*, SA '11, pages 10:1–10:58, New York, NY, USA, 2011. ACM.
- [48] Mateu Sbert, Dimitri Plemenos, Miquel Feixas, and Francisco González. Viewpoint quality: Measures and applications. In Lásza Neumann, Mateu Sbert Casasayas, Bruce Gooch, and Werner Purgathofer, editors, *Computational Aesthetics*, pages 185–192. Eurographics Association, 2005.
- [49] Ekrem Serin, Candemir Doger, and Selim Balcisoy. 3d object exploration using viewpoint and mesh saliency entropies. In Erol Gelenbe, Ricardo Lent, and Georgia Sakellari, editors, *Computer and Information Sciences II*, pages 299–305. Springer London, 2012.
- [50] Ekrem Serin, Selcuk Sumengen, and Selim Balcisoy. Representational image generation for 3d objects. *The Visual Computer*, 29(6-8):675–684, 2013.
- [51] C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27:379–423, 623–656, July, October 1948.
- [52] B. Solenthaler and R. Pajarola. Predictive-corrective incompressible sph. In *SIGGRAPH '09: ACM SIGGRAPH 2009 papers*, pages 1–6, New York, NY, USA, 2009. ACM.
- [53] Barbara Solenthaler and Markus Gross. Two-scale particle simulation. *ACM Trans. Graph.*, 30(4):81:1–81:8, July 2011.
- [54] C.M. Soon, C.F.M. Coimbra, and M.H. Kobayashi. The variable viscoelasticity oscillator. *Annalen der Physik*, 14(6):378–389, 2005.
- [55] Jun Tao, Jun Ma, Chaoli Wang, and Ching-Kuang Shene. A unified approach to streamline selection and viewpoint selection for 3d flow visualization. *Visualization and Computer Graphics, IEEE Transactions on*, 19(3):393–406, 2013.

- [56] Yubo Tao, Hai Lin, Hujun Bao, Feng Dong, and G. Clapworthy. Structure-aware viewpoint selection for volume visualization. In *Visualization Symposium, 2009. PacificVis '09. IEEE Pacific*, pages 193–200, 2009.
- [57] Gabriel Taubin. Estimating the tensor of curvature of a surface from a polyhedral approximation. In *Computer Vision, 1995. Proceedings., Fifth International Conference on*, pages 902–907, 1995.
- [58] Xin Tong, Teng-Yok Lee, and Han-Wei Shen. Salient time steps selection from large scale time-varying data sets with dynamic time warping. In *Large Data Analysis and Visualization (LDAV), 2012 IEEE Symposium on*, pages 49–56, 2012.
- [59] Pere-Pau Vazquez, Miquel Feixas, Mateu Sbert, and Wolfgang Heidrich. Viewpoint selection using viewpoint entropy. In *Proceedings of the Vision Modeling and Visualization Conference 2001, VMV '01*, pages 273–280. Aka GmbH, 2001.
- [60] I. Viola, M. Feixas, M. Sbert, and M.E. Groller. Importance-driven focus of attention. *Visualization and Computer Graphics, IEEE Transactions on*, 12(5):933–940, 2006.
- [61] Chaoli Wang and Kwan-Liu Ma. Information and knowledge assisted analysis and visualization of large-scale data. In *Ultrascale Visualization, 2008. UltraVis 2008. Workshop on*, pages 1–8, 2008.
- [62] Chaoli Wang and Han-Wei Shen. Information theory in scientific visualization. *Entropy*, 13(1):254–273, 2011.
- [63] Chaoli Wang, Hongfeng Yu, R.W. Grout, Kwan-Liu Ma, and J.H. Chen. Analyzing information transfer in time-varying multivariate data. In *Pacific Visualization Symposium (PacificVis), 2011 IEEE*, pages 99–106, 2011.

- [64] Chaoli Wang, Hongfeng Yu, and Kwan-Liu Ma. Application-driven compression for visualizing large-scale time-varying data. *Computer Graphics and Applications, IEEE*, 30(1):59–69, 2010.
- [65] Lijie Xu, Teng-Yok Lee, and Han-Wei Shen. An information-theoretic framework for flow visualization. *Visualization and Computer Graphics, IEEE Transactions on*, 16(6):1216–1224, nov.-dec. 2010.