

# Geng Zhao

Department of Computer Science  
Swansea University  
Swansea, SA2 8PP, UK

Phone: (+44) 07872551593  
Email: [cszg@swan.ac.uk](mailto:cszg@swan.ac.uk)  
Homepage: <http://cs.swan.ac.uk/~cszg/>

## Education

Ph.D. **Perceptive Information Visualization on Large, Multi-Dimensional Data Set**, Computer Science, Swansea University, May 2009 - present.

M.S.c. **Computer Science, (Merit, the 1st place in examination)**, Swansea University, 2007 - 2009.

B.S.c **Computer Science**, Swansea University, 2004 - 2007.

## Research

My research interest is focused on visualization designs and techniques for perceiving and understanding large, multi-dimensional data sets. In my first project, I have a collaboration with researchers from HCI Design Center of City University, London. We are trying to evaluate how well the free existing off-the-shelves information visualization software can benefit the document triage study. In my second project, we propose a frequency-based angular histograms to avoid the clutter and overplotting problem when rendering large data sets in parallel coordinates. The data we experiment with is collected by biologists from the multiple sensors which track the activities of marine wildlife. After that, we further propose another frequency-based model using Markov Chain. Based on this model, a probability value is computed for each N-dimensional data tuple. This probability value can be utilized to further guide the user for selection and filtering for interesting patterns. During the internship in Thomson Reuters, I work on a novel structure-aware treemap for large, time-varying hierarchical business data. The proposed visualization system is evaluated on real world sales force management data sets. In my fourth project, I work together along with the Arts and Humanities Department. The project is about exploring the variations of translations for Shakespeare's work. Every translation is a different interpretation of the play. These large quantities of translations reflect changing culture and express individual thought by the authors. Researchers from Arts and Humanities at Swansea University are collecting a large number of translations of William Shakespeare's. In this project, I have developed an interactive visualization system which is composed of two parts: the structure-aware Treemap for document selection and meta data analysis, and Focus + Context parallel coordinates for in-depth document comparison and exploration.

## Personal

Born on August 13, 1986.

Chinese Citizen.

## Academic Experience

### *Swansea University*

Teaching Assistant, CS-337/CS-Mo7 Data Visualization, Sep – Dec 2009, Sep – Dec 2010

Teaching Assistant, CS-254 Software Engineering, Jan – May 2010.

Lab Demonstrator, CS-141 Programming Principles and Practice and CS-M01 Distributed Programming in Java , September – December 2009.

Lab Demonstrator, CS-242 Algorithms , Jan – May 2010,

Coursework Marking, CS-335 Foundations of AI , Sep – Dec 2009.

## Publication

### *Journal Papers*

Zhao Geng, Zhenmin Peng, Robert S.Laramee, Rick Walker, and Jonathan C. Roberts, **Angular Histograms: Frequency-Based Visualizations for Large, High-Dimensional Data**, IEEE Transactions on Visualization and Computer Graphics (IEEE TVCG), Vol. 17, No. 6, November/December 2011, (VisWeek 2011 Proceedings)

### *Conference Papers*

James Walker, Zhao Geng, Mark Jones and Robert S. Laramee, **Visualization of Large, Time-Dependent, Abstract Data with Integrated Spherical and Parallel Coordinate**, Proceedings of EuroVis 2012 Short Papers, 5 - 8 June 2012, Vienna, Austria

Tom Cheesman, Stephan Thiel, Zhao Geng, Robert S.Laramee, David M.Barry, Alison Ehrmann, **Translation Arrays: Exploring Cultural Heritage Texts Across Languages**, Digital Humanities 2012 , 16-22 July, 2012, Hamburg, Germany, forthcoming

Zhao Geng, Robert S.Laramee, Tom Cheesma, Alison Ehrmann and David M. Berry, **Visualizing Translation Variation: Shakespeare's Othello**, 7th International Symposium on Visual Computing(ISVC) , Sep. 26-28, 2011, Las Vegas, Nevada, USA, forthcoming (Oral Presentation ~ 30%)

Zhao Geng, Robert S. Laramee, Fernando Loizides, and George Buchanan, **Visual Analysis of Document Triage Data**, International Conference on Information Visualization Theory and Applications (IVAPP), pages 151-163, 5-7 March, 2011, Algarve, Portugal, forthcoming (Full paper acceptance rate ~ 15%)

### *Extended Abstract*

Zhao Geng, Robert S.Laramee, Tom Cheesman, Andrew Rothwell, David M. Berry, and Alison Ehrmann, **Visualizing Translation Variation of Othello: A Survey of Text Visualization and Analysis Tools**, Refereed Abstract at the 3rd International Symposium for Humanities and Technology, InterFace 2011, University College London, UK, 27-29 July 2011

Max L. Wilson, Zhao Geng, Robert S.Laramee, Tom Cheesman, Andrew Rothwell, David M. Berry, and Alison Ehrmann, **Studying International Variations in Culture and Literature: Visualizing Translation Variations in Shakespeare's Othello**, Refereed Abstract at the ACM Web Science 2011 Conference (ACM WebSci '11), Koblenz, Germany, 14-17 June 2011

Zhenmin Peng, Zhao Geng, and Robert S.Laramee, **Design and Implementation of A System for Interactive Higher Dimensional Vector Field Visualization**, Eurographics Posters 2011, 11-15 April 2011, Llandudno, Wales

Zhao Geng, Gaurav Gathania, Robert S.Laramee, and Zhenmin Peng, **Visual Analysis of Hierarchical Management Data**, Eurographics Posters 2011, 11-15 April 2011, Llandudno, Wales

### Technical Reports

Richard Walker, Serban Pop, Zhao Geng, Robert S. Laramee, and Jonathan C. Roberts, **Force-Directed Parallel Coordinates**

Zhenmin Peng, Zhao Geng, and Robert S. Laramee, **Design and Implementation of A System for Interactive Higher Dimensional Vector Field Visualization**, Technical Report, Department of Computer Science, University of Wales, Swansea, UK, July 2011

### Work-In-Progress

Zhao Geng, James Walker, and Robert S. Laramee, **Markov Chain Driven Multi-Dimensional Visual Pattern Analysis With Parallel Coordinates**, Technical Report, Department of Computer Science, University of Wales, Swansea, UK, December 2011 ( Submitted to the IEEE InfoVis 2012)

Zhenmin Peng, Zhao Geng, Robert S. Laramee, and Nick Croft, **Visualization of Flow Past a Marine Turbine: The Search for Sustainable Energy**, ( Submitted to the Computer & Graphics )

### Talks

Angular Histograms: Frequency-Based Visualizations for Large, High-Dimensional Data, VisWeek 2011, Providence, RI, USA

Visualizing Translation Variation: Shakespeare's Othello, ISVC 2011, Las Vegas, USA

Visualizing Translation Variation of Othello: A Survey of Text Visualization and Analysis Tools, Interface 2011, University College London, UK

Visual Analysis of Document Triage Data, International Conference on Information Visualization Theory and Applications (IVAPP), 5-7 March, 2011, Algarve, Portugal

Visualization of Document Triage Data, *Visual Computing Research Away Day, Computer Science Department, Swansea University, 17th June 2009.*

EuroVis 2009 Conference Report - Scatter Plots and Parallel Coordinates, *The Visible Lunch, Computer Science Department, Swansea University, 30th July 2009.*

IEEE Visualization 2009 Conference Report - Multidimensional Visualization, *The Visible Lunch, Computer Science Department, Swansea University, 29th Oct 2009.*

Visual Analysis of Sales Force Management Data, *Project Demonstration, Thomson Reuter, Pinnacle #15, Bangalore, India, 21th Sept 2010.*

Visual Data Mining: Concepts and Techniques, *Invited talk on module CS-337/CS-Mo7 Data Visualization, Computer Science Department, Swansea University, 14th Nov 2010*

Clustering and Classification Techniques in Data Mining , *Invited talk on module CS-337/CS-Mo7 Data Visualization, Computer Science Department, Swansea University, 14th Oct 2011*

Visualizing Textual Corpora, *Digital Shakespeare Monday 16 May 2011, Workshop and Talks, Swansea University*

## Work Experience

### **Interactive Visualization System for Sales Force Management Data**

**Management Information System Team, Thomson Reuters, July 2010 – Sept 2010**

**Description** The sales force management data at Thomson Reuters often contains multiple hierarchies and dependencies. Conventional business graphics are not sufficient for analyzing, presenting and exploring such data sets. The Treemap is effective for depicting hierarchical data, although it lacks structural clarity. Thus, distinguishing different levels within the treemap is difficult. In this project, I'm responsible for developing an interactive system for business analysts which enables structure-aware visualization of hierarchical data using multiple coordinated views. The design of this system includes hierarchy extraction, structure tracings and the direct manipulation of the treemap. A demonstration video is available at <http://cs.swan.ac.uk/~cszg/VASM/VASM.mpeg>.

### **Research Assistant : Visualizing Variations of Othello**

**Swansea University, Feb 2011 - July 2011**

**Description:** Recognized as great works of world literature, Shakespeare's poems and plays have been translated into dozens of languages for over 300 years. Also, there are many re-translations into the same language, for example, there are more than 60 translations of *Othello* into German. Every translation is a different interpretation of the play. These large quantities of translations reflect changing culture and express individual thought by the authors. They demonstrate wide connections between different world regions today, and reveal a retrospective view of their cultural, intercultural, and linguistic histories. Researchers from Arts and Humanities at Swansea University are collecting a large number of translations of William Shakespeare's *Othello*. In this project, we have developed an interactive visualization system to present, analyze and explore the variations among these different translations. Our system is composed of two parts: the structure-aware Treemap for document selection and meta data analysis, and Focus + Context parallel coordinates for in-depth document comparison and exploration. In particular, we want to learn more about which content varies highly with each translation, and which content remains stable. We also want to form hypotheses as to the implications behind these variations. Our visualization is evaluated by the domain experts from Arts and Humanities.

### **Research Assistant : Advanced Visualization of Large, High-Dimensional Data**

**Swansea University, August 2011 - Feb 2012**

**Description:** Parallel coordinates are a widely used visualization technique for presenting, analyzing and exploring multi-dimensional data. However, like many other visualizations, it can suffer from an overplotting problem caused by rendering large data sets. Until now, quite a few methods are proposed to extract and illustrate the major data trends in the cluttered parallel coordinates. Among them, the frequency-based approaches using binning and histograms are widely adopted. The traditional binning method, which records line-segment frequency, only considers data in a two-dimensional subspace and the multi-dimensional features are not taken into account for cluster and outlier analysis. This is mainly because the multidimensional binning can suffer from the curse of dimensionality. In this project, we utilize the Markov Chain model to compute the n-dimensional joint probability for each data tuple based on the two-dimensional binning method. This probability value can be utilized to guide the user for selection and filtering. Then the original data can be decomposed into the multidimensional principal data trend and outliers. We provide various interaction techniques for the user to control the parameters during the filtering process. Filtered data with a high probability

measure often explicitly illustrates cluster profiles. The data with low probability value is able to preserve outliers. We demonstrate the results on a wide variety of data sets, including real-world, high-dimensional biological and stock market data.

## Academic Duties

Reviewer for IEEE Information Visualization

Reviewer for Computer Graphics Forum

Reviewer for IEEE Transactions on Visualization and Computer Graphics

Student Volunteering in VisWeek 2011

Student Volunteering in Eurographics 2011

Student Volunteering in Rivic Visual Computing Graduate School, 2010

## Computer Skills

Programming Languages: Java, C++, C#, C, Matlab, Prolog, Haskell, Clips.

Applications: Graphic packages: *Qt*, *OpenGL*, *Java Swing*, *Prefus*, *D3* and Common office packages: *Database*, *Word Processors - Office*, *LATEX*.

Operating Systems: Linux, Mac OS X, Windows.

## Hobbies

Basketball, Literature, History, Philosophy

## References

Dr. Robert S. Laramee  
Senior Lecturer of Computer Science  
Swansea University, UK  
+44 (0) 1792 602 609  
r.s.laramee@swansea.ac.uk  
Homepage: <http://cs.swan.ac.uk/~csbob/>

Dr. Jonathan C. Roberts  
Senior Lecturer of Computer Science  
Bangor University, UK  
(+44) 01248 382725  
j.c.roberts@bangor.ac.uk  
Homepage: [http://www.bangor.ac.uk/cs/Staff/jonathan\\_roberts.php?](http://www.bangor.ac.uk/cs/Staff/jonathan_roberts.php?)