

Vishwakarma Singh

CONTACT INFORMATION	<i>E-mail:</i> vsingh014@gmail.com <i>Voice:</i> (805) 315-8633 <i>WWW:</i> www.cs.ucsb.edu/~vsingh	
INTERESTS	Developing data structures, high performance algorithms, and scalable applications for analyzing, querying, and mining very large multi-modal datasets of high dimensions.	
EDUCATION	PhD., Department of Computer Science University of California, Santa Barbara	CGPA: 3.99/4.0 9/2006 - 12/2011
	M.S., Department of Computer Science University of California, Santa Barbara	CGPA: 3.99/4.0 9/2006 - 8/2011
	B.Tech., Department of Computer Engineering Institute of Technology, Benaras Hindu University	DGPA: 9.52/10.0 7/1999 - 5/2003
WORK EXPERIENCE	Principal Applications Engineer Oracle, CA, USA	12/2011-Present
	Senior Software Engineer CenturyLink, Bangalore, India	3/2005 - 6/2006
	Applications Engineer Oracle, Hyderabad, India	1/2004 - 2/2005
	Associate Consultant Oracle Financial Services Software Limited, Mumbai, India	6/2003 - 1/2004
RESEARCH EXPERIENCE	Querying and Mining Large Scale High Dimensional Data Department of Computer Science, UCSB Research Assistant, Advisor: Prof. Ambuj K. Singh	6/2006 - Present
	Duplicate Video Detection Dolby Laboratories (Research), San Francisco Summer Intern, Advisors: Dr. Wenyu Jiang and Dr. Claus Bauer	6/2009 - 9/2009
	Natural Language Query Processing (Start Mobile) Nokia Research Center, Cambridge, MA 02142 Summer Intern, Advisors: Dr. Raimondas Lencevicius and Dr. Alexander Ran	6/2007 - 9/2007
	Role Based Access Control in SELinux Kernel School of Computing, NUS, Singapore Summer Intern, Advisor: Prof. Roland Yap	5/2002 - 7/2002
	Segmentation and Skeletonization of Images Using Neural Network Models Indian Statistical Institute, Kolkata, India Summer Intern, Advisor: Prof. Amitava Dutta	5/2001 - 7/2001

TECHNICAL SKILLS Specializations: Multi-dimensional Indexing and Search, Data Mining, Machine Learning, Information Retrieval, Predictive Analytics

Languages: Java Core, Java Swing, PL/SQL, C, C++, Python, Matlab

Technology: J2EE, XML, Oracle database

Platforms: Linux and Windows

PUBLICATIONS **Peer Reviewed Conferences and Journals**

“SIMP: Accurate and Efficient Near Neighbor Search in Very High Dimensional Spaces”, Vishwakarma Singh and Ambuj K. Singh, to appear in EDBT, 2012, Berlin, Germany (Recommended for the best paper award and the best student paper award).

”Unified Hypergraph For Image Ranking In a Multimodel Context”, Jiejun Xu, Vishwakarma Singh, Ziyu Guan, and B. S. Manjunath, to appear in ICASSP 2012.

“Finding Skyline Nodes in Large Networks”, Arijit Khan, Vishwakarma Singh, and Jian Wu, to appear in the 3rd International Workshop on Graph Data Management: Techniques and Applications 2012 (GDM 2012), co-located with ICDE 2012.

“Querying Spatial Patterns”, Vishwakarma Singh, Arnab Bhattacharya, and Ambuj K. Singh, in EDBT, 2010.

“Efficient and Robust Detection of Duplicate Videos in a Large Database”, Anindya Sarkar, Vishwakarma Singh, Pratim Ghosh, B. S. Manjunath, and Ambuj K. Singh, in IEEE CSVT, 2010.

“Geo-Clustering of Images With Missing GeoTags”, Vishwakarma Singh, Sharath Venkatesha, and Ambuj K. Singh, in IEEE Granular Computing, 2010.

“An Algorithm and Hardware Design for Very Fast Similarity Search in High Dimensional Space”, Vishwakarma Singh and Wenyu Jiang, in IEEE Granular Computing, 2010.

“Efficient Computation of Statistical Significance of Query Results in Databases”, Vishwakarma Singh, Arnab Bhattacharya, and Ambuj K. Singh, in SSDBM, 2008.

Technical Reports and Works in Submission

“ProMiSH: Nearest Keyword Set Search in Very High Dimensional Spaces”, Vishwakarma Singh and Ambuj K. Singh.

“Profile Based Sub-Image Search in Image Databases”, Vishwakarma Singh and Ambuj K. Singh, UCSB Tech. Report, 2010-20.

“QUIP: Querying Significant Patterns from Image Databases”, Vishwakarma Singh, Arnab Bhattacharya, Ambuj K. Singh, Chris Banna, Geoffrey P. Lewis, and Steven K. Fisher, UCSB CS-Technical Report, 2007-10.

ACHIEVEMENTS AND AWARDS

- 2nd Place for web based education portal “Click2School”, ICS Technology HITEC Entrepreneurship Competition-2008, University of California at Irvine.
- Outstanding employee award, CenturyLink, 2006.
- All India Rank 945 in IIT-JEE among more than 150 thousand students.
- All India Rank 630 in Rorkee-EE among more than 150 thousand students.
- Top 1% of more than 400 undergraduate students at IT-BHU.
- Merit Award in B.Tech. Part3, 2001-2002.
- UGC Merit Scholarship for continuous 4 years in undergrad.
- BSL Merit(SAIL) Scholarship.
- 2nd prize for PC-controlled Security Door, Technex-2001, IT-BHU.
- 3rd prize for PC-controlled Chess Board model, Technex-2001, IT-BHU.

PROFESSIONAL SERVICES

- Reviewer for VLDB 2011, ICDM 2011, KDD 2011, ICDE 2010, VLDB 2010, SIGMOD 2010, ICDE 2009, VLDB 2009, ICDE 2008, VLDB 2008, KDD 2008, ICDE 2007
- Talk at IIT-Kanpur in November, 2010.
- PC member of UCSB Graduate Student Workshop.
- Summer student mentor at UCSB.
- Member IEEE.
- Graduate Student Representative for the undergraduate curriculum committee.

RESEARCH OVERVIEW

Nearest keyword set search in very high dimensional spaces: Keyword based queries on text rich spatial data provide new tools for exploring the data and building novel applications and services. In this project we developed an efficient solution for querying the closest keyword set in very high dimensional spaces.

Nearest neighbor search in very high dimensional spaces: Nearest neighbor search for very high dimensional spaces is an open classical problem. In this project we developed a scalable and efficient method called SIMP based on a novel space partitioning technique. We achieve impressive performance results on very large datasets having very large dimension. We also developed an elegant and simple algorithm for high dimensional similarity search which can be implemented in hardware for real time query.

Querying patterns in images and spatial datasets: This project proposed a novel and flexible score based similarity measure for matching a given region of an image to regions of images in a large image repository. Similarity measure takes spatial structure into consideration, can incorporate domain knowledge, and is independent of the choice of an image feature. Proposed access methods are general enough to be used for structured querying in other spatial datasets.

Duplicate video detection: We developed a new similarity measures to detect video duplicates with high precision in large video datasets. We also proposed scalable techniques that search duplicates in milliseconds and can be used in real time systems.

Computing significance of query results: Scientists using querying techniques for obtaining matches in a medical or biological image repositories or chemical compound repositories are interested to know the statistical significance of a match score. We developed a scalable technique for measuring the p-value of the results of a multi-object

query in a large dataset.

Geo-Clustering of images with missing geo tags: There is obvious need of tools to exploit information from vast repositories of images like Flickr for building new applications. This work proposed a graph based technique to fuse temporal, tag similarity, and image similarity scores with geo information available for a small training set of images to cluster all the images in a large dataset based on spatial proximity. This is useful in predicting geo location of images which miss geo tags.

Profile based sub-image search: Quality and scalability for sub-image search is a challenging problem. We design a new neighborhood profile based feature vector for each interest point in an image. Similarity between a pair of feature vectors capture similarity between the descriptors as well as the spatial layout of the descriptors. It achieves high precision for sub-image search results.

INDUSTRY JOB
DESCRIPTION

Natural Language Based Job Querying: I worked on developing a natural language query based job search engine as an intern at Nokia research lab.

Telecommunication software: As a senior software engineer at CenturyLink, I developed modules for supporting CenturyLink customers network connections and devises. My responsibility involved designing, developing, and finally deploying the modules in a production environment.

Oracle e-business suite: I was involved in developing modules for human resource management system of Oracle e-business suite. My responsibility involved module design and development.

REFERENCES

Available on request