

MINGZHOU (JOE) SONG

**Application for a tenure-track assistant professor position in computational biology**

194-40B 64th Ave. Apt. 1B  
Fresh Meadows, New York 11365, U.S.A.  
Phone: + 1 (718) 997-3584  
E-mail: [Song.Joe@Gmail.com](mailto:Song.Joe@Gmail.com)  
URL: <http://www.cs.qc.edu/~msong>

November 8, 2004

Professor Rob de Ruyter van Steveninck  
Chair, Faculty Search Committee  
Biocomplexity  
Indiana University  
Department of Physics  
Swain Hall West 117  
Bloomington, IN 47405-7105  
<http://biocomplexity.indiana.edu/recruit>

Dear Professor van Steveninck:

I am inviting you to interview me for the tenure-track assistant professor position in computational biology in Biocomplexity at Indiana University, which you advertised recently on professional magazines. I am a tenure-track assistant professor of computer science at Queens College. I am also appointed on the doctoral faculty at the Graduate Center of the City University of New York. I believe that the advertised position will allow me to grow substantially in scholarly research capacity. My qualifications will strengthen your research programs in statistical and computational learning methods for biochemical interaction network and statistical analysis of genome in bioinformatics/computer biology. My work will increase the multidisciplinary dimension of your research agenda as well. I consider myself an excellent match to the position for the following reasons.

First, my research is included in your targeted recruiting areas. My current focus is probabilistic modeling and statistical learning for evolving complex data sets, particularly biological and image data. My ongoing projects include spike sorting for neuronal signals, statistical analysis of genomic repetitive elements, mathematical modeling of gene regu-

latory network and microscopic image analysis. Through these projects, I am developing a theme of the three dependencies framework which models the structural, temporal and statistical dependencies of the mechanism underlying the data sets.

Second, I have a strong record of interaction with biomedical researchers. My statistical modeling of genomic repetitive elements is a joint work with a professor of biology. I also collaborate with a company manufacturing electronic recording devices for neuronal signals in Dallas. I am preparing for a possible collaboration opportunity with USDA on modeling yeast gene regulation. If hired, I anticipate major multidisciplinary involvement with faculty in either your medical school or other life science departments affiliated with your institution.

Last, I will have had 3.5 years undergraduate and graduate teaching experience by Fall 2005 at my present institution, where I introduced new curricula in bioinformatics and computational biology at both undergraduate and graduate levels. I had also been a teaching assistant for 2.5 years when I was a graduate student. Over the years, both the effectiveness and evaluation of my teaching have improved steadily. I have been helping students to discover the joy of learning within themselves. I also find myself enjoy teaching tremendously.

I am applying selectively and have not yet discussed my plans with others in my department, thus I would appreciate it if I were notified before my current department was contacted regarding the position.

Thank you for considering my application. I am looking forward to an on-campus interview at your earliest convenience. I will present, in much greater detail, myself and my future research and teaching plan to Biocomplexity at Indiana University.

Very truly yours,

A handwritten signature in black ink, appearing to read 'Mingzhou Song' or similar, written in a cursive style.

Mingzhou (Joe) Song

encl: Curriculum Vitae, Statement of Research, and Statement of Teaching Philosophy.

# MINGZHOU (JOE) SONG

---

194-40B 64<sup>th</sup> Ave. Apt. 1B, Fresh Meadows NY 11365  
Song.Joe@Gmail.com, <http://www.cs.qc.edu/~msong/>

## RESEARCH INTERESTS

Statistical pattern recognition, computational learning, computer vision, medical image analysis, modeling gene regulatory network, statistical and computational genomics, neuronal signal classification and clustering.

## EDUCATION

**University of Washington** Seattle, WA  
Ph.D. in Electrical Engineering March 2002  
Thesis: *Integrated Surface Model Optimization from Images and Prior Shape Knowledge.*

M.S. in Electrical Engineering June 1999  
Thesis: *Ultrasound Imaging Simulation and Echocardiographic Image Synthesis.*

**Florida International University** Miami, FL  
Graduate Study in Computer Science September 1996 – August 1997

**Beijing University of Posts and Telecommunications** Beijing, China  
Graduate Study in Electrical Engineering August 1992 – July 1996  
B.S. in Electrical Engineering July 1992

## EXPERIENCE

**Ph.D. Program in Computer Science** *Doctoral Faculty*  
**Graduate Center**  
**City University of New York**  
New York, NY. August 2002 – Present  
Teaching & Research in statistical pattern recognition, bioinformatics & computer vision.

**Department of Computer Science** *Assistant Professor*  
**Queens College**  
**City University of New York**  
Flushing, NY. February 2002 – Present  
Teaching undergraduate and graduate level computer science courses.

**Department of Electrical Engineering** *Graduate Teaching Assistant*  
**University of Washington**  
Seattle, WA. September 2000 – December 2001  
Taught electrical engineering courses.

**Informatics Research, Celera Genomics** *Consultant*  
Rockville, MD. June 2000 – September 2000

Computational gene finding by comparative genomics.

**Department of Electrical Engineering  
University of Washington**

Seattle, WA.

3-D left ventricle model reconstruction from multiple 2-D ultrasound heart images.

*Graduate Research Assistant*

September 1997 – May 2000

**School of Computer Science  
Florida International University**

Miami, FL.

Research in a compiler design and taught introductory computer science courses.

*Graduate Assistant*

September 1996 – August 1997

**Beijing University of Posts and  
Telecommunications**

Beijing, China.

Research in video compression.

*Graduate Research Assistant*

August 1992 – July 1996

**AWARDS**

**Department of Electrical Engineering, University of Washington**

*Graduate Poster Contest, Second Place, with Certificate*

Seattle, WA

May 1999

**Beijing University of Posts and Telecommunications**

- *Certificate of Honored Graduate*
- *Motorola Inc. Scholarship with Certificate*
- *Chinese Ministry of Posts and Telecommunications  
(now Ministry of Information Industry)  
Scholarship with Certificate*
- *Scholarship*

Beijing, China

July 1992

October 1991

May 1991

October 1990

**COURSES TAUGHT**

**Queens College and Graduate Center, CUNY**

*Instructor.*

Teaching load: 4 courses per year now. To teach 6-7 courses per year from fall 2005.

Duties: lecturing, holding office hours, grading homework, projects and exams, maintaining course web pages and email lists.

- Artificial Intelligence
- Computational Biology Journal Club
- Data Communications
- Design and Analysis of Algorithms
- Honors Problems
- Introduction to Bioinformatics
- Research Practicum (in Bioinformatics)

New York City, NY

- Software Development Practicum (in Medical Image Analysis)
- Topics in Computational Biology

**University of Washington**

Seattle, WA

*Teaching Assistant.* Duties: conducting weekly review sessions; teaching substitute lectures; typesetting lecture notes; holding weekly office hours; grading homework and exams; maintaining course web pages, class newsgroups and class e-mail lists.

- Continuous Time Linear Systems
- Digital Signal Processing
- Digital Signals and Filtering
- Introduction to Computer Networks

**Seattle University**

Seattle, WA

*Substitute Lecturer.* Duties: taught three lectures.

- Computer Vision and Computer Graphics

**Florida International University**

Miami, FL

*Lab Assistant.* Duties: gave lab instructions; graded projects; maintained a course web.

- Introduction to Microcomputer

**GRANTS**

**Howard Hughes Medial Institute**

Summer 2004 – 2008

2004 Undergraduate Science Education Program.

**PSC-CUNY Research Award 35**

July 1, 2004 – June 30, 2005

Statistical Learning Algorithms for Neural Signal Analysis.

**PSC-CUNY Research Award 34**

July 1, 2003 – June 30, 2004

Research in Optimal Quantization for Pattern Recognition.

**CUNY Institute for Software**

June 1, 2003 – December 31, 2003

**Design & Development**

Software for Extraction, Visualization and Pattern Discovery of Genomic Features Associated with Transposable Elements.

**Division of Natural Sciences &**

March 2003

**Mathematics, Queens College**

Equipment Grant.

**CUNY Institute for Software**

August 2002 – January 2003

**Design & Development**

Spike Sorting for Neural Signals.

**Queens College**

Spring 2002

Start-up Fund for Equipments.

## **RESEARCH MENTORING**

### **Doctoral Students:**

Christoforou Christoforos Summer 2004 – Present  
Topic: nonparametric statistical learning algorithms.

Christina Schweikert Summer 2004  
Topic: statistical modeling of LINE1 in human genome.

Negmat Mullodzhanov Spring 2004 – Present  
Topic: modeling gene regulatory network from gene expression.

Hongbin Wang Fall 2003 – Present  
Topic: data stream clustering algorithms and application on neuronal spike sorting.

Guanghai Weng Fall 2003 – Summer 2004  
Topic: extraction & visualization of features of repetitive genomic element LINE1.

### **Undergraduate Students:**

Sergey Averchenkov Summer 2003  
Topic: genomic sequence feature extraction.

Ha T. Lam Spring and Summer 2003  
Topic: user interface and algorithms for spike sorting.

## **UNIVERSITY SERVICES**

Member, Howard Hughes Grant Committee, Fall 2003 – Present  
Queens College

Member, Master's Program in Bioinformatics Planning Committee, Fall 2003  
Queens College

Undergraduate Advisor, Spring 2003 – Present  
Department of Computer Science, Queens College

Academic Senator, Spring 2003 – Present  
Queens College

Member, Self Study Curriculum Committee, Fall 2002  
Department of Computer Science, Queens College

## PROFESSIONAL MEMBERSHIPS

Member, International Society for Computational Biology (ISCB)	2003 – Present
Member, Institute for Mathematical Statistics (IMS)	2000 – Present
Member, Institute for Electrical and Electronic Engineers (IEEE) Engineering in Medicine and Biology Society	2000 – 2003
Student Member, Association for Computing Machinery (ACM)	2000 – 2001

## PUBLICATIONS

### Journal Articles:

M. Song and S. Boissinot. Statistical patterns of recombination rate associated with LINE1 elements in human genome. (in preparation for *Genome Research*)

M. Song and R. M. Haralick. Finding adaptive variable width bin histograms by optimal quantization and smoothing. (in preparation for *Pattern Recognition*)

M. Song, R. M. Haralick, F. H. Sheehan, and R. K. Johnson. 2002. Integrated Surface Model Optimization for Freehand 3-D Echocardiography, *IEEE Transactions on Medical Imaging, special issue on 3-D cardiac image analysis*, vol. 21, no.9, pages 1077-1090.

### Book Chapters:

M. Song and R. M. Haralick. Nonparametric Pixel Appearance Probability Model Using Grid Quantization For Local Image Information Representation, *Medical Imaging Systems: Theory and Applications*, World Scientific Publishing CO. (in press)

### Conference Proceedings:

M. Song and H. Wang. Very Efficient Incremental Clustering Algorithm for Online Data Stream. *Artificial Intelligence and Statistics 2005*. (submitted)

M. Song and N. Mullodzhanov. 2004. Modeling and Finding Gene Regulatory Network by Combining Generalized Logical Network with Probability Description for State Transition and Time Delay. Abstract. *The 7th International Meeting of the Microarray Gene Expression Data Society*. Toronto, Canada.

M. Song, S. Boissinot, R. M. Haralick, and I. T. Phillips. 2003. Estimate Recombination Rate Distribution by Optimal Quantization, *Proceedings of IEEE Computational Systems Bioinformatics Conference*, pages 403-406. Stanford, CA.

M. Song and H. T. Lam. 2003. A Non-parametric Bayesian Framework for Spike Sorting

Using Optimal Quantization, *Joint Statistical Meetings*. San Francisco, CA.

M. Song, R. M. Haralick, F. H. Sheehan, and R. K. Johnson. 2002. An Integrated Approach to Surface Modeling in Freehand 3-D Echocardiography, *Proceedings of 2nd Joint Meeting of the IEEE Engineering in Medicine and Biology Society and Biomedical Engineering Society*, pages 1082-1083. Houston, TX.

M. Song and R. M. Haralick. 2002. Optimal Multidimensional Quantization for Pattern Recognition, *Proceedings of SPIE Vol. 4875: 2nd International Conference on Image and Graphics*, Part 1, pages 15-30. Hefei, China.

M. Song and R. M. Haralick. 2002. Optimal Grid Quantization, *Proceedings of International Conference on Pattern Recognition*, vol. III, pages 444-447. Quebec City, Canada.

M. Song and R. M. Haralick. 2002. Optimally Quantized and Smoothed Histograms, *Proceedings of Joint Conference on Information Sciences, Computer Vision, Pattern Recognition and Image Processing*, pages 894-897. Durham, NC.

M. Song, R. M. Haralick, and F. H. Sheehan. 2000. Ultrasound Imaging Simulation and Echocardiographic Image Synthesis. *Proceedings of International Conference on Image Processing*, vol.III, pages 420-423. Vancouver, Canada.

M. Song, A. Guo, and R. M. Haralick. 2000. Single View Computer Vision in Polyhedral World: Geometric Inference and Performance Characterization. *Proceedings of International Conference on Pattern Recognition*, vol.I, pages 766-769. Barcelona, Spain.

S. Aksoy, M. Ye, M. Schauf, M. Song, Y. Wang, R.M. Haralick, J.R. Parker, J. Pivovarov, D. Royko, C. Sun, and G. Farneback. 2000. Algorithm Performance Contest. *Proceedings of International Conference on Pattern Recognition*, vol.IV, pages 870-876. Barcelona, Spain.

M. Song, A. Cai, and J. Sun. 1996. Motion Estimation in DCT Domain. *Proceedings of International Conference on Communication Technology*, vol.2, pages 670-674. Beijing, China.

## REFERENCES

*Robert M. Haralick*, Distinguished Professor of Computer Science  
Graduate Center, City University of New York  
365 Fifth Avenue, New York, NY 10016  
Phone: (212) 817-8192  
E-mail: haralick@ptah.gc.cuny.edu  
URL: <http://www.prl.cs.gc.cuny.edu/web/LabWebsite/Haralick/main.htm>

*Stéphane Boissinot*, Assistant Professor of Biology



Department of Biology  
Queens College, City University of New York  
65-30 Kissena Boulevard, Flushing, NY 11367  
Phone: (718) 997-3437  
E-mail: [stephane.boissinot@qc.edu](mailto:stephane.boissinot@qc.edu)  
URL: [http://qcpages.qc.edu/Biology/fac\\_stf/boissinot.html](http://qcpages.qc.edu/Biology/fac_stf/boissinot.html)

*Ihsin T. Phillips*, Professor of Computer Science  
Department of Computer Science  
Queens College, City University of New York  
65-30 Kissena Boulevard, Flushing, NY 11367  
Phone: (718) 997-3524  
E-mail: [yun@cs.qc.edu](mailto:yun@cs.qc.edu)  
URL: <http://www.cs.qc.edu/cgi-bin/faculty.cgi?fid=phillips>

*Linda G. Shapiro*, Professor of Computer Science and Electrical Engineering  
University of Washington  
Department of Computer Science and Engineering  
Box 352350, Seattle, WA 98195-2350  
Phone: (206) 543-2196, Fax: (206) 543-2969  
E-mail: [shapiro@cs.washington.edu](mailto:shapiro@cs.washington.edu)  
URL: <http://www.cs.washington.edu/homes/shapiro/>

*Florence H. Sheehan*, Research Professor of Medicine  
University of Washington  
School of Medicine, Division of Cardiology  
Box 356422, Seattle, WA 98195  
Phone: (206) 543-4535, Fax: (206) 685-9394  
E-mail: [sheehan@u.washington.edu](mailto:sheehan@u.washington.edu)  
URL: <http://depts.washington.edu/cvrtc/index.html>