

Athanasios Katsamanis

3740 McClintock Avenue, Room 427
Los Angeles CA 90089-2564, USA

nkatsam@cs.ntua.gr
<http://cvsp.cs.ntua.gr/nassos>

Research Interests

Speech Production and Synthesis Speech aeroacoustics and articulatory synthesis. Modeling the production of fricatives. Computational speech modeling for acoustic-to-articulatory inversion. Statistical articulatory modeling based on real data. Articulatory feature extraction from real data. Improved acoustic modeling for speech. Statistical speech synthesis. Biomedical applications and second language tutoring.

Speech Recognition Exploiting articulatory knowledge for speech recognition using dynamic Bayesian networks. Statistical speaker articulation adaptation. Dynamic speech modeling for speech enhancement and recognition. Applications for large vocabulary continuous speech recognition.

Multimodal Processing Statistical audiovisual speech synthesis and recognition. Statistical multimodal interaction and integration. Audiovisual speech inversion exploiting multimodal articulatory data. Multimodal human computer interfaces. Multimodal fusion for sign language recognition. Dynamic Bayesian networks for multimodal fusion.

Education

PhD in Electrical and Computer Engineering 2003-2009
National Technical University of Athens, School of E.C.E Athens, Greece
Subject: Non-linear computational speech modeling exploiting aerodynamics in the vocal tract
Advisor: Prof. Petros Maragos (URL: <http://cvsp.cs.ntua.gr/maragos>)

Diploma/M.Eng. in Electrical and Computer Engineering 1998 - 2003
National Technical University of Athens, School of E.C.E Athens, Greece
Final Grade: 9.49/10 (highest honors)
Diploma Thesis: Statistical Multiband Demodulation and Speech Analysis Using Kalman Filters and Energy Related Methods (Advisor: Prof. Petros Maragos)

Research Experience

Postdoctoral Research Associate 2009-today
University of Southern California, School of Electrical Engineering Los Angeles, California, USA
Member of the Signal Analysis and Interpretation Laboratory

Graduate Research Assistant 2004-2009
National Technical University of Athens, School of E.C.E. Athens, Greece

As a member of the Computer Vision, Speech Communication & Signal Processing (CVSP) Group, I participated in European and national research projects:

ΠΙΕΝΕΔ: Research on computational speech production modeling based on aeroacoustics.

ASPI (Audiovisual to Articulatory Speech Inversion): Research on statistical audiovisual to articulatory speech inversion exploiting multimodal articulatory data.

MUSCLE (Multimedia Understanding via Semantics, Computation & Learning): Research on multimodal fusion for audiovisual speech recognition and event detection in audio streams.

HIWIRE (Human Input that Works in Real Environments): Research on speaker normalization and adaptation for user-robust continuous speech recognition.

Protagoras: Research on speech demodulation and formant tracking.

Visiting Researcher

summer 2007

École Nationale Supérieure des Télécommunications

Paris, France

Project: Improved acoustic modeling for speech inversion.

Mentor: Prof. Shinji Maeda.

Research Assistant

July 2000-May 2002

Institute for Language and Speech Processing

Athens, Greece

Designed and implemented an educational environment in MATLAB for the “Signals and Systems” course taught at the E.C.E. School at NTUA (Mentor: Prof. G. Carayiannis.)

Worked on a web interface for text-to-speech synthesis (Mentor: Dr. Ath. Protopapas)

Participated in the design and preparatory group for the European research project **METIS** on machine translation based on monolingual corpora. (Mentor: Dr. G. Tampouratzis)

Research Assistant

summer 2002

Hong Kong Polytechnic University, School of Computer Science

Hong Kong, China

In the frame of the student exchange program IAESTE I worked under the supervision of Assoc. Prof. Robert Luk on the design and implementation of an automatic Cantonese speech recognizer using the Hidden Markov Model Toolkit (HTK). Worked on duration modeling and on general system optimization.

Awards-Scholarships

PhD Studies Scholarship

2005-2009

Onassis Public Benefit Foundation

PhD Studies Scholarship

2006-2008

Greek National Scholarship Foundation

Research Assistantship Scholarship

2000-2002

Institute for Language and Speech Processing

Papakyriakopoulou and Kritikou Awards for excellent performance in Mathematics 2000

National Technical University of Athens

Award for excellent performance at the E.C.E. School

2000

Technical Chamber of Greece

Undergraduate Studies Scholarship

1998-1999

Greek National Scholarship Foundation

Teaching Experience

Graduate Teaching Assistant

2004-2008

National Technical University of Athens, School of E.C.E

Athens, Greece

Served as a lab and teaching assistant (lab preparation, presentation and grading) for the following courses, instructed by Prof. Petros Maragos:

Pattern Recognition with Emphasis on Speech Recognition (9th Semester, 2004-2006),

Signals and Systems (4th Semester, 2007),

Digital Signal Processing (6th Semester, 2008).

Academic Activities

Reviewer

IEEE Transactions on Audio, Speech and Language Processing

IEEE Signal Processing Letters

Journal of the Acoustical Society of America

Invited Talks/ Seminars

- [1] **Audiovisual Speech: Recognition and Inversion**
at GIPSA-lab in Grenoble, France, Apr. 2009.
- [2] **Audio-Visual Speech Analysis and Recognition**
in MUSCLE joint with VITALAS Conference, Cannes, France, Feb. 2008.
- [3] **Towards Automatic Speech Recognition in Adverse Environments**
in the Nonlinear Speech Processing Workshop (WNSP05), Heraklion, Greece, 2005.
- [4] **Statistical Speech Analysis and Nonlinear Modeling**
in COST277 Seminar in Limerick, 2004.
- [5] **Speech Technologies for E-Commerce and Transactional Services**
in EFTEHNOS series of seminars on supportive information and telecommunication technologies for people with special needs, University of Athens, 2004.

Professional Experience in Industry

Voice-In S.A.

Developer

<http://www.voice-in.gr>

October 2002-July 2003

Designed and implemented an Interactive Voice Response (IVR) application optionally integrating automatic speech recognition and speech synthesis technologies.

Professional Memberships

Institute of Electrical and Electronics Engineers

Signal Processing Society

2003-present

Technical Chamber of Greece

2004-present

Software Development, System & Network Administration Experience

Programming Languages C/C++, Java, Perl, Python, MATLAB

Operating Systems Linux, Mac, Windows

Contributed in the design, development and administration of the CVSP group website and the CVSP network and computer system infrastructure (<http://cvsp.cs.ntua.gr/>).

Languages

- **English:** Proficient level (Cambridge Proficiency, CEFR level C2)
- **German:** Intermediate level (Mittelstufe, CEFR level B2)
- **French:** Basic level (CEFR level A2)
- **Greek:** Native

References

References available upon request

List of Publications

Journal Papers

- [1] A. Katsamanis, G. Papandreou, and P. Maragos,
Face Active Appearance Modeling and Speech Acoustic Information to Recover Articulation,
IEEE Trans. on Audio, Speech and Language Processing, vol. 17(3): 411-422, Mar. 2009.
- [2] G. Papandreou, A. Katsamanis, V. Pitsikalis, and P. Maragos,
Adaptive Multimodal Fusion by Uncertainty Compensation with Application to Audio-Visual Speech Recognition,
IEEE Trans. on Audio, Speech and Language Processing, vol. 17(3): 423-435, Mar. 2009.
- [3] A. Katsamanis, A. Roussos, and P. Maragos,
Audiovisual Speech Inversion Exploiting Multimodal Articulatory Data,
under preparation to be submitted to the Journal of the Acoustical Society of America.

Papers in International Refereed Conferences

- [1] S. Theodorakis, A. Katsamanis, P. Maragos,
Product-HMMs for automatic sign language recognition,
Proc. IEEE Int'l Conf. on Acoustics, Speech, and Signal Processing (ICASSP-2009), Taipei, Taiwan, Apr. 2009.
- [2] A. Katsamanis, T. Roussos, P. Maragos, M. Aron and M.-O. Berger,
Inversion from Audiovisual Speech to Articulatory Information by Exploiting Multimodal Data,
Proc. Int'l Seminar on Speech Production (ISSP 2008), Strasbourg, France, Dec. 2008.
- [3] A. Katsamanis, G. Ananthakrishnan, G. Papandreou, P. Maragos, O. Engwall
Audiovisual Speech Inversion by Switching Dynamical Modeling Governed by a Hidden Markov Process,
Proc. European Signal Processing Conf. (EUSIPCO 2008), Lausanne, Switzerland, Aug. 2008.
- [4] A. Katsamanis, G. Papandreou, and P. Maragos,
Audiovisual-to-Articulatory Speech Inversion Using Active Appearance Models for the Face and Hidden Markov Models for the Dynamics,
Proc. IEEE Int'l Conf. on Acoustics, Speech, and Signal Processing (ICASSP-2008), Las Vegas, NV, U.S.A., Mar.-Apr. 2008.

- [5] S. Lefkimmiatis, P. Maragos, A. Katsamanis,
Multisensor Multiband Cross-Energy Tracking for Feature Extraction and Recognition,
Proc. IEEE Int'l Conference on Acoustics, Speech, and Signal Processing (ICASSP-2008), Las Vegas, NV, U.S.A., Mar.-Apr. 2008.
- [6] G. Papandreou, A. Katsamanis, V. Pitsikalis, and P. Maragos,
Multimodal Fusion and Learning with Uncertain Features Applied to Audiovisual Speech Recognition,
Proc. IEEE Workshop on Multimedia Signal Processing (MMSP-2007), pp. 264-267, Chania, Greece, Oct. 1-3, 2007.
- [7] A. Katsamanis, G. Papandreou, and P. Maragos,
Audiovisual-to-Articulatory Inversion Using Hidden Markov Models,
Proc. IEEE Workshop on Multimedia Signal Processing (MMSP-2007), pp. 457-460, Chania, Greece, Oct. 1-3, 2007.
- [8] A. Katsamanis, P. Tsiakoulis, P. Maragos, A. Potamianos,
Investigations in Articulatory Synthesis,
Proc. Int'l Congress of Phonetic Sciences, ICPHS 2007, Saarbruecken, Germany, Aug. 2007.
- [9] V. Pitsikalis, A. Katsamanis, G. Papandreou, and P. Maragos,
Adaptive Multimodal Fusion by Uncertainty Compensation,
Proc. Int'l Conf. on Spoken Language Processing, Sep. 17-21, Pittsburgh PA, USA, Interspeech-ICSLP 2006.
- [10] A. Katsamanis, G. Papandreou, V. Pitsikalis, and P. Maragos,
Multimodal Fusion by Adaptive Compensation for Feature Uncertainty with Application to Audiovisual Speech Recognition,
Proc. 14th European Signal Processing Conf. (EUSIPCO-2006), Florence, Italy, Sept. 4-8 2006.
- [11] A. Katsamanis and P. Maragos,
Advances in Statistical Estimation and Tracking of AM-FM Speech Components,
Proc. Interspeech 2005 - Eurospeech – 9th European Conf. on Speech Communication and Technology, Lisbon, Portugal, Sept. 2005.

Book Chapters

- [1] G. Papandreou, A. Katsamanis, V. Pitsikalis and P. Maragos,
Adaptive Multimodal Fusion by Uncertainty Compensation with Application to Audio-Visual Speech Recognition,
in *Multimodal Processing and Interaction: Audio, Video, Text*, edited by P. Maragos, A. Potamianos, and P. Gros, Springer-Verlag, New York, 2008.
- [2] P. Maragos, P. Gros, A. Katsamanis and G. Papandreou,
Cross-Modal Integration for Performance Improving in Multimedia: A Review,
in *Multimodal Processing and Interaction: Audio, Video, Text*, edited by P. Maragos, A. Potamianos, and P. Gros, Springer-Verlag, New York, 2008.