Irfan Aziz Essa

Associate Professor

College of Computing
Georgia Institute of Technology
Atlanta, GA 30332-0280, USA

Table of Contents

EDUCATIONAL BACKGROUND

Degree	Year	University	Field
Ph.D.	1995	Massachusetts Institute of Technology, Cambridge, MA., USA.	*Media Arts & Sciences*
S.M.	1990	Massachusetts Institute of Technology, Cambridge, MA., USA.	*Engineering*
B.S.	1988	Illinois Institute of Technology, Chicago, IL., USA.	*Engineering*

EMPLOYMENT HISTORY

Title	Organization	Years
Associate Professor	College of Computing Georgia Institute of Technology, Atlanta, GA., USA	*2002-present*
Adjunct Professor	School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA., USA	*1998-present*
Assistant Professor	College of Computing Georgia Institute of Technology, Atlanta, GA., USA	*1996-2002*
Research Scholar/Scientist	Media Laboratory, Perceptual Computing Section, Massachusetts Institute of Technology, Cambridge, MA., USA	*1994-1996*
Research Assistant	Media Laboratory, Perceptual Computing Section, Massachusetts Institute of Technology, Cambridge, MA., USA.	*1988-1994*
Research Intern	Vision, Graphics and Visualization Group, Thinking Machines Corporation, Cambridge, MA., USA	*Summer 1990*

CURRENT FIELDS OF INTEREST

Computer Vision, Computer Graphics, Human-computer Interaction, Machine Learning, Numerical Modeling, Computational Perception and Artificial Intelligence.

Goals:

My research focuses on the analysis, interpretation, and synthesis of video. In particular, I am interested in developing methods for dynamic (and sometimes causal) analysis of the spatio-temporally variable video signals. Such methods allow for interpretation of what is happening in a scene and for plausible synthesis of similar scenes. I am motivated by the desire to build machines that can recognize complex behaviors and mimic (synthesize) similarly complex actions by observing (seeing) them. My work to-date has been aimed at making state-of-the-art contributions to computer vision and computer graphics, relying heavily on numerical modeling and learning, image processing, and control

theory. My research will have potential impact on work in artificial intelligence and human-computer interaction.

I. TEACHING

A. Courses Taught

Quarter/Year	Course Number & Title	Number of Students	Comments
	Media Laboratory, Massachusetts Institute of Technology.		
Spring, 1995	MAS 961 Machine Understanding of Video	15	New
	College of Computing, Georgia Institute of Technology.		
Winter, 1997	CS 7321 Low-level Computer Vision	26	Revised
Spring, 1997	CS 7322 High-level Computer Vision	15	Revised
Fall 1997	CS 7100 Introduction to Grad Studies	40	Revised
Fall 1997	CS 8113 Computational Perception	15	New
Winter 1998	CS 7321 Low-level Computer Vision	15	
Spring 1998	CS 7322 High-level Computer Vision	20	
Fall 1998	CS 7100 Introduction to Grad Studies	41	
Spring 1999	CS 4803/8113 Digital Special Effects	22	New
Spring 2000	CS 4480 Digital Video Special Effects	32	New
Spring 2001	CS 4480 Digital Video Special Effects	26	
Spring 2002	CS 4480 Digital Video Special Effects	25	
Spring 2002	CS 4496/7496 Computer Animation	33	Revised
Spring 2003	CS 4480 Digital Video Special Effects	37	
Spring 2003	CS 7496 Computer Animation	18	
Seminars			
Winter 97-Fall 98	Software Agents	7-10	Seminar
Winter-Spring, 98	Technology and Society	10	Seminar
Fall 97-Spring 00	Future Computing Environments	15-25	Seminar

B. Curriculum Development

Media Laboratory, Massachusetts Institute of Technology.

MAS 961 Machine Understanding of Video: A new graduate course dealing with applications of computer vision and audition techniques for automatic annotation of video. A project-based course in which groups of 4 students each designed and implemented systems for processing and analysis of video and audio signals.

College of Computing, Georgia Institute of Technology, Atlanta, GA.

CS 7321 Low-level Computer Vision (1997-98): A graduate introductory course for computer vision. Revised and updated old content. Hands-on work on different techniques of computer vision and image understanding in the form of projects and computer labs was added. In addition to theoretical concepts, emphasis was also given to real world problem solving.

CS 7322 High-level Computer Vision (1997-98): Continuation of low-level computer vision concepts from CS7321. Revised and updated old content. Emphasis laid on recognition and modeling methods for computer vision. Hands-on work on different techniques of computer vision and image

understanding in the form of projects and computer labs was added. In addition to theoretical concepts, emphasis was also given to real world problem solving.

CS 8113 Computational Perception (1997): Designed and taught a completely new class aimed at exploring various issues of machine perception of signals (video and audio) for interaction and analysis. Material covered included study of both the engineering and psychological issues of perception. Students worked hands-on towards the design and development of projects, in teams, which dealt with machine perception.

A curriculum for Digital Imaging in Science and Engineering Education (In Progress): Working with faculty from the Department of Electrical and Computer Engineering's Center for Signal and Image Processing and the College of Computing's Graphics, Visualization and Usability Center to develop and institute an Undergraduate/Graduate Certificate Program for Digital Imaging in Science and Engineering Education.

Seminar on Technology and Society (1998): (with J. Berman and Dr. T. Harpold) Started a new seminar class aimed at addressing the issues relating to the effects of modern (computing) technology to past, present, and future societies. A weekly seminar with participants from different units at Georgia Tech, in the format of round table discussions and invited speakers to bring to bear ideas of where technology has taken us, and where it will take us.

Conversion to Semesters (1998-99): In preparation for the semester conversion, redesigned the computer vision and computational perception courses for semesters. The two quarter series of computer vision courses was merged into a semester long computer vision course. The computer vision course will be offered as a 4000 level undergraduate course as well as a 7000 level graduate course. The computational perception course will be offered as a 7000 level graduate course.

CS 4480 Digital Video Special Effects (1999-2002): Developed a new class aimed at teaching undergraduates the basics of computing with video and audio. Students learn about digital representations of video and study manipulation, processing, interpretation and synthesis of video under the guise of generating special effects. Student in groups produce videos to showcase their technical skills of dealing with digital video, which are shown in an open public forum. Updated the class for the semester version. Technology Licensing funds (about $ 150,000) were acquired to upgrade the classrooms and to purchase audio/video equipment to support this class. Hewlett Packard and Intel donated computer for teaching of this class. Received *Georgia Tech. Outstanding use of Innovative Technologies in Teaching Award, 2000* for designing and teaching of this course.

CS 4496/7496 Computer Animation (2002): Redesigned and updated the Computer Animation course, and offered it for the first time after 1999. Adapted all of the newer material that has appeared in this field in the last few years. Emphasized traditional animation principles while covering fundamental approaches that are key for computer methods. Equal coverage of keyframing, motion capture, and simulation methods of animation. Students (both grad and undergrad) did assigned problem sets and did a final project.

C. Individual Student Guidance

Post-Doctoral Fellows

Lionel Reveret
Spring 2000 - Fall 2001.

Publications: *E.4.4, and E.2.24.*
Research on building Conversational Agents.
***At INRIA France, IMAGIS Group in Grenoble as* Charge de recherches**.

Ph.D. Students Supervised (in process as well as graduated)

Darnell J. Moore (ECE, with Professor M. Hayes)
Winter 1997 - Spring 2000.
Publications: *E.3.7, E.2.11, E.2.23 and E.2.26.*
Vision-Based Recognition of Actions using Context.
Member of Research Staff Texas Instrument Corp. Dallas, TX. USA. 2000 - Present.

Arno Schödl (CoC)
Fall 1997 - Summer 2002.
Publications: *E.2.9, E.2.10, E.2.15, E.2.17, E.2.21, E.2.25, F.1.1, and E.2.32.*
PhD Dissertation Title: "Multi-Dimensional Exemplar-Based Texture Synthesis."
Awarded Microsoft Research Fellowship 2000-2002.
Awarded Georgia Tech College of Computing's Outstanding Graduate Student Research Award 2002.
Awarded Georgia Tech College of Computing's Outstanding Thesis Dissertation Award 2003.
Founder and Technical Director, Think-Cell Software, Berlin, Germany, 2002 - Present

Gabriel J. Brostow (CoC)
Fall 1997 - Present.
Publications: *B.0.7, E.2.12, E.2.20 and F.2.1.*
Current Research on Finding Structure a Multi-view Data Sets
Awarded NSF HCI Trainee Fellowship 1998-2000.
Awarded Intel PhD Student Fellowship 2000-2001.
Awarded Georgia Tech College of Computing's Outstanding Graduate Student Teaching Award 2002.

Antonio Haro (CoC)
Fall 1997 - Fall 2003.
Publications: *E.2.9, E.3.8, E.2.14, E.2.18, E.2.28, E.2.29, F.2.3 and F.2.4*
PhD Dissertation Title: "Example-based Image and Video Synthesis".
Awarded AT& T Graduate Fellowship 1998-2002.

Drew Steedly (CoC)
Spring 1998 - Present.
Publications: *E.2.19 and E.2.33.*
Current Research on Optimal Structure from Motion from Video.
Transferred from Georgia Tech's School of ECE in Fall 1998.
Awarded Intel PhD Student Fellowship 2001-2002.

Vivek Kwatra (CoC, with Professor Aaron Bobick)
Fall 2001 - Present.
Publications: *E.2.32*
Current Research on Video-based Rendering and Animation

David Minen (CoC, with Professor Thad Starner)
Fall 2001 - Present.
Publications: *E.2.30.*
Current Research on Human Activity Recognition.
Awarded NSF Graduate Student Fellowship 2002 - 2005.

Stephanie Wojtkowski (CoC)
Fall 2002 - Present.
Current Research on Video-based Rendering.
Awarded NSF Graduate Student Fellowship 2003 - 2006.

Yan Huang (CoC)
Fall 2002 - Present.
Publications: *E.2.31 and F.2.2*
Current Research on Probabilistic Tracking.

Pei Yan (CoC)
Fall 2002 - Present.
Current Research on Facial Tracking and Modeling with Audio.

Mitch Parry (CoC)
Fall 2002 - Present.
Current Research on Audio Textures.

Mohammed Raffay Hamid (CoC)
Fall 2002 - Present.
Publications: *E.2.31 and F.2.2*
Research on Activity Recognition

Ph.D. Special Problems students.

Andrew Gardner (ECE)
Summer 1997 - Fall 1997.
Publications: *E.3.4.*
Title: "Audio in Interactive Environments."

Alexander Stoychev (CoC)
Fall 1998 - Summer 1999.

Scott Stillman (ECE)
Winter 1997 - Spring 2002.
Publications: *E.3.6 and E.2.22.*
Audio-visual tracking of multiple people, using multiple sensors.
Member of the Bell South SNT Center, Atlanta, GA.

Byungmoon Kim (CoC)
Spring 2002 - Present.

M.S. Thesis Students supervised.

Ken Miller (MIT)
Fall 1994 - Spring 1996.
Graduated: Spring 1996
Thesis title: "Eigenmethods for Representing 3D Face Data."

Sumit Basu (MIT, Co-advised with Dr. Alex Pentland)
Fall 1994 - Spring 1996.
Publications: *E.2.7, and E.2.8.*
Post Doc at Microsoft Research 2002.

Arno Schödl (CoC)
Fall 1997 - Present.
Publications: *E.2.9, E.2.10.*
Completed MS June 1999. Continuing towards a Ph.D.

Katherine Sukel (Psychology, with Prof. Richard Catrambone)
Winter 1998 - Present.
Publications: *B.0.7.*
Completed MS June 1999. Admitted to Ph.D. Program in CoC. Currently on Leave.

Sunil Mishra (CoC)
Fall 1997 - Fall 1999.
Research on analyzing audio and video streams to extract relevant communicative features.
Working at SRI in California.

Rohit Varma (CoC)
Summer 2000 - Present.
Publications: *F.1.1.*
Research on Alpha matting and analyzing videos of crowds.
Currently a PhD student at Stanford.

J Austin Hijar (CoC)
Summer 2001 - Fall 2001.
Research on Sound Spaces with Vision Support.

Alan Chen (CoC)
Fall 2000 - Fall 2002.
Research on Generating Non-photorealistic Renderings of Motion.

Ramprasadh Ramanarayanan (CoC)
Fall 2001 - Present.
Research on Fusion of Audio and Video for Interactive Agents.

Jeannie Lee (CoC)
Spring 2002 - Present.
Research on Fusion of Audio and Video for Distributed Computing.

MS. Special Problems students.

Junhoy Kim (Math)
Winter 1997 - Summer 1997.
Title: "Hidden Markov Models for Vision-based Recognition of Action."

Alexander Gdalevich (ECE)
Spring 2000 - Fall 2000.
Research on Embedded Vision Sensors.

Undergraduate Research Students.

Thad Starner (MIT)
Spring 1989 - Spring 1992
Title: "Efficient C implementation of the ThingWorld Modeling System."
Finished Ph.D. from MIT and joined Georgia Tech / College of Computing as faculty (January 1999).

Roberto Peon (CoC)
Spring 1997 - Spring 2000.
Awarded the Georgia Tech Undergraduate Research Internship Award, 1999.
Awarded the Undergraduate Research in Computing Research Award, 2000.
Currently on staff at Sportvision Inc.

Charles Hubbard (CoC)
Winter 1998- Fall 1999.

Paul Bennett (CoC)
Winter 1999-Fall 1999.

Robbie Walia (CoC)
Fall 2000 - Spring 2001.

Tim Keenan (CoC)
Fall 2000 - Spring 2001.
Awarded Undergraduate Research in Computing Award 2001.
on Staff at PDI Dreamworks Studios.

James Hays (CoC)
Spring 2001 - Present.
Publications: *F.2.5*
Awarded Intel Undergraduate Research Fellowship Award 2001.
First Prize, Georgia Tech College of Computing's Undergraduate Research in Computing (UROC) Symposium 2002.
Attending PhD Program at CMU CS Dept, 2003.
Georgia Tech Outstanding Undergraduate Award 2003.
NSF Graduate Student Fellowship Honorable Mention 2003.

Scott Carter (CoC)
Spring 2000 - Spring 2002.
Awarded Undergraduate Research in Computing Award 2001.
On Staff at Sammy Studios, Los Angeles, CA.

Ravikrishna Ruddarajju (ECE)
Spring 2002 - Present.
Publications: *E.2.29*
Awarded Intel Undergraduate Research Fellowship Award 2002.
Awarded Georgia Tech Presidential Undergraduate Research Award 2003.
Second Prize, Georgia Tech College of Computing Undergraduate Research in Computing (UROC) Symposium 2003

Siddhartta Maddi (CoC)
Fall 2002 - Present.
Working on Plan Recognition in the Aware Home.
Awarded Georgia Tech's Presidential Undergraduate Research Award 2003.

Ali Kamil (CoC)
Fall 2002 - Present.
Working on Phased-Array Microphones in the Aware Home.

John Hable (CoC)
Spring 2003 - Present.
Working on Facial Expression Cloning from Data.

About 15 other Undergraduate students from 1990 - Present.

D. Teaching Honors and Awards

1. Georgia Tech. Outstanding use of Innovative Technologies in Teaching Award, 2000.
2. Georgia Tech. College of Computing's William A. "Gus" Baird Faculty Teaching Award, 2002.

II. RESEARCH AND CREATIVE SCHOLARSHIP

A. Thesis

S.M. Thesis

Title: *"Contact Detection, Collision Response and Friction for Physically-based Virtual World Modeling and Vision Systems."*
Date Completed: *June 1990,*
Advisors: *Dr. Alex Pentland and Dr. John Williams,*
University: *Massachusetts Institute of Technology*.

Ph.D. Thesis

Title: *"Analysis, Interpretation, and Synthesis of Facial Expressions."*
Date Completed: *February 1995*,
Advisor: *Dr. Alex Pentland*,
University: *Massachusetts Institute of Technology*.

B. Published Journal Papers (refereed)

B.0.1 Pentland A., I. Essa, M. Friedmann, B. Horowitz, and T. Starner, "The ThingWorld Modeling System: Virtual Sculpting by Modal Forces", In *ACM Computer Graphics*, 24(2), pp. 143-144, ACM Press, March 1990.

B.0.2 Essa, I., S. Sclaroff, and A. Pentland. "A Unified Approach for Physical and Geometric Modeling for Graphics and Animation." In *Computer Graphics Forum, The International Journal of the Eurographics Association*, Vol. 11 (3), C129-C138, C470–C471, Cambridge, England, September 1992.

B.0.3 Darrell, T., I. Essa, and A. Pentland. "Task-specific Gesture Modeling using Interpolated Views." In*, IEEE Transactions on Pattern Analysis and Machine Intelligence*, Volume 18 (12), pp. 1236–1242, IEEE Computer Society Press, December 1996.

B.0.4 Essa. I., and A. Pentland. "Coding, Analysis, Interpretation and Recognition of Facial Expressions." In*, IEEE Transactions on Pattern Analysis and Machine Intelligence*, Volume 19 (7), pp. 757–763, IEEE Computer Society Press, July, 1997.

B.0.5 Essa, I. "Computers Seeing People", in *AI Magazine,* Volume 20 (1), pp. 69–82, AAAI Press, Summer 1999.

B.0.6 Essa, I. "Ubiquitous Sensing for Smart and Aware Environments", In *IEEE Personal Communications, Special Issue on Networking the Physical World*, Volume 7(5), pp. 47–49, IEEE Press. October 2000.

B.0.7 Sukel, K. E., R. Catrambone, I. Essa, and G. Brostow, "Presenting movement in a computer-based dance tutor." In *International Journal of Human-Computer Interaction (IJHCI).* To Appear, Final Manuscript Submitted August 2001.

[ACM SIGGRAPH Proceedings papers ([E.2.15], [E.2.20]) are considered final archival publications, with resubmission to a Journal not permitted. Several journal papers are under review. See F.1.]

C. Published Books and Parts of Books (Refereed)

C.1. Chapters in Books

C.1.1 Essa, I., S. Sclaroff, and A. Pentland. "Physically-based Modeling for Graphics and Vision." Chapter in *Directions in Geometric Computing*, pp. 161–196, R. Martin (Editor), Information-Geometers, UK., 1993.

C.1.2 Pentland, A., S. Sclaroff, B. Horowitz, and I. Essa. "Modal Descriptions for Modeling, Recognition and Tracking." *In Three-Dimensional Object Recognition Systems*, Vol. I, A K. Jain and P. J. Flynn (Editors), pp. 423–445, Elsevier Science Publishers, 1993.

C.1.3 Pentland, A., I. Essa, T. Darrell, A. Azarbayejani and S. Sclaroff. "Visually Guided Animation." In *Interactive Computer Animation*, N. Thalmann and D. Thalmann (Editors), pp. 143–164, Prentice-Hall, 1996

C.1.4 Essa, I. and Alex Pentland, "Facial Expression Recognition Using Image Motion." In *Motion Based Recognition*, M. Shah and R. Jain (Editors), Chapter 12 , Kluwer Academic Publishers, Computational Imaging and Vision Series, 1997.

D. Edited Proceedings and Collections

D.0.1 Proceedings of Second International Conference on Automatic Face and Gesture Recognition (Program Chair), Killington, VT., USA. IEEE Computer Society Press, October, 1996.

D.0.2 Proceedings of IEEE Workshop on Nonrigid and Articulated Motion, (Program Co-chair), San Juan, Puerto Rico, IEEE Computer Society Press, June 1997.

E. Conference Presentations

E.1. Invited Keynote Addresses

E.1.1 "Non-invasive Motion Analysis", at *International Symposium on Human Movement Analysis*, Stockholm, Sweden, July 1994.

E.1.2 "Aware Home: Sensing, Interpretation, and Recognition of Everyday Activities", at *IEEE Signal Processing Soceity, DSP Workshop*, Callaway Gardens, Pine Mountain, GA, USA, October 2002..

E.1.3 "Building and Aware Home: Technologies for the way we may live", at *First International Workshop on Man-Machine Symbiosis*, Kyoto, Japan, November 2002.

E.2. Conference Presentations with Proceedings (refereed and archival)

[Refereed publications in respected conferences, with extensive reviewing, and appearing in archival proceedings.]

E.2.1 Pentland, A., I. Essa, T. Darrell, and S. Sclaroff. "Visually Guided Animation.", In *Proceedings of Computer Animation 1994 Conference,* pp. 129–138, IEEE Computer Society Press, Geneva, Switzerland, May 1994.

E.2.2 Essa, I. and A. Pentland. "A Vision System for Observing and Extracting Facial Action Parameters", In *Proceedings of IEEE Computer Vision Pattern Recognition Conference 1994,* pp. 76–83, IEEE Computer Society Press, Seattle, WA., June 1994.

E.2.3 Essa, I., T. Darrell, and A. Pentland. "Tracking Facial Motion", In *Proceedings of IEEE Nonrigid and Articulated Motion Workshop 1994*, pp. 36–42, Austin, TX., IEEE Computer Society Press, November 1994.

E.2.4 Pentland, A., I. Essa, T. Darrell, A. Azarbayejani and S. Sclaroff. "Visually Guided Interaction and Animation" *Proceedings of the Twenty-eighth Asilomar Conference on Signals, Systems & Computers 1994,* pp. 1287–1291, Asilomar, CA., November 1994.

E.2.5 Darrell, T., I. Essa, and A. Pentland. "Correlation and Interpolation Networks for Real-time Expression Analysis/Synthesis", In G. Tesauro, D. S. Touretzky, and T. K. Leen (Editors), pp. 909–916, *Advances in Neural Information Processing Systems (Proceedings of NIPS) 7,* Denver, CO., MIT Press 1995.

E.2.6 Essa, I. and A. Pentland. "Facial Expression Recognition using a Dynamic Model and Motion Energy", In, *Proceedings of the IEEE International Conference on Computer Vision 1995,* pp. 360–367, IEEE Computer Society Press, Cambridge, MA., May 1995.

E.2.7 Basu, S., I. Essa, and A. Pentland."Motion Regularization for Model-based Head Tracking." In *Proceedings of International Conference on Pattern Recognition*, pp. C8A.3, Vienna, Austria, August 1996.

E.2.8 Essa, I., S. Basu., T. Darrell, and A. Pentland, "Modeling, Tracking and Interactive Animation of Facial Expressions and Head Movements using Input from Video", In *Proceedings of Computer Animation 1996 Conference,* pp. 68–79, IEEE Computer Society Press, Geneva, Switzerland, June 1996.

E.2.9 Schödl, A., A. Haro, and I. Essa, "Head Tracking using a Textured Polygonal Model", In *Proceedings of Perceptual User Interfaces Workshop*, (held in Conjunction with ACM UIST 1998), pp. 43–48, November 1998.

E.2.10 Schödl, A., K. Schwan, and I. Essa, "Adaptive Parallelization of Model-based Head Tracking", In Proceedings of *1999 International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA'99*) (7 pages), Monte Carlo Resort, Las Vegas, Nevada, USA, June 1999.

E.2.11 Moore, D., I. Essa, and M. Hayes, "Exploiting Human Actions and Object Context for Recognition Tasks", In *Proceedings of IEEE International Conference on Computer Vision 1999 (ICCV'99)*, pp. 80–86, IEEE Computer Society Press, Corfu, Greece, September 1999.

E.2.12 Brostow, G., and I. Essa, "Motion-based Video Decompositing." In *Proceedings of IEEE International Conference on Computer Vision 1999 (ICCV'99)*, pp. 8–13, IEEE Computer Society Press, Corfu, Greece, September 1999.

E.2.13 Kidd. C., G. Abowd, C. Atkeson, I. Essa, B. MacIntyre, E. Mynatt, T. Starner, "The Aware Home: A Living Laboratory for Ubiquitous Computing Research", In *Proceedings of Second International Workshop on Cooperative Buildings 1999*, Editors, Streitz, J. Siegel, V. Hartkopf, S. Konomi, Pittsburgh. LNCS 1670. Springer: Heidelberg, 1999.

E.2.14 Haro, A., M. Flickner and I. Essa, "Detecting and Tracking Eyes by Using their Physiological Properties, Dynamics and Appearance", In *Proceedings of IEEE Computer Vision and Pattern Recognition 2000 Conference*, pp. I163–I168. IEEE Computer Society Press, Hilton Head, SC, June 2000.

E.2.15 Schödl, A., R. Szeliski, D. Salesin, and I. Essa. "Video Textures", In *Proceedings of ACM SIGGRAPH 2000 Conference*, pp. 489-498, ACM Press, New Orleans, LA, August 2000.

E.2.16 Mynatt, E., I. Essa, and W. Rogers, "Increasing the Opportunities for Aging in Place", In *Proceedings of First ACM Universal Usability 2000* Conference (Inaugural Conference), pp. 65–71, Arlington, VA, November 2000.

E.2.17 Schödl, A., and I. Essa. "Learning Methods for Video-based Animation", In *Advances Neural Information Processing Systems 13 (Proceedings of NIPS 2000 Conference)*, T. K. Leen, T. .G. Dietrterich, and V. Tresp (Editors), MIT Press, pp. 1002–1008, May 2001.

E.2.18 Haro, A., B. Guenter, I. Essa. "Real-time, Photorealistic, Physically Based Rendering of Human Skin Microstructure" In *Proceedings of Eurographics Rendering Workshop*, pp. 53–62, London, England June 2001.

E.2.19 Steedly, D., and I. Essa. "Propagation of Innovative Information in Non-Linear Least-Squares Structure from Motion" In *Proceedings of IEEE International Conference on Computer Vision 2001 Conference*, Volume II, pp. 223–229, IEEE Computer Society Press, Vancouver, BC, Canada, July 2001.

E.2.20 Brostow, G., and I. Essa. "Image-based Motion Blur for Stop Motion Animation" In *Proceedings of ACM SIGGRAPH 2001 Conference*, Los Angeles, CA., USA. August 2001.

E.2.21 Schödl, A. and I. Essa, "Depth Layers from Occlusions" To appear in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, IEEE Computer Society Press, Kauaui, Hawaii, December 2001.

E.2.22 Stillman, S. and I. Essa, "Towards Reliable Multimodal Sensing in Aware Environments" In *Perceptual User Interfaces (PUI 2001) Workshop* (held in Conjunction with ACM UIST 2001 Conference, ACM Publishers, November 2001.

E.2.23 Moore, D. and I. Essa, "Recognizing Multi-tasking Activities using Context-Free Stochastic Grammar" In *Proceedings of Workshop on Models versus Exemplars in Computer Vision*, (held in Conjunction with IEEE Conference on Computer Vision and Pattern Recognition Conference), December 2001.

E.2.24 Reveret, L. and I. Essa, "Visual Coding and Tracking of Speech Related Facial Motions" In *Proceedings of Cues in Communications Workshop*, (held in Conjunction with IEEE Conference on Computer Vision and Pattern Recognition), December 2001.

E.2.25 Schödl, A. and I. Essa, "Controlled Animation of Video Sprites" In *Proceedings First ACM Symposium on Computer Animation 2002,* held in conjunction with ACM SIGGRAPH 2002 Conference, San Antonio, TX, USA, July 2002.

E.2.26 Moore, D. and I. Essa, "Recognizing Multitasked Activities from Video using Recognizing Multitasked Activities from Video using Stochastic Context-Free Grammar" To Appear in *Proceedings of American Association of Artificial Intelligence (AAAI) Conference 2002,* Alberta, Canada, July 2002.

E.2.27 Abowd, G. A. Bobick, I. Essa, E. Mynatt, and W. Rogers "The Aware Home: Developing Technologies for Successful Aging" To Appear in *Proceedings of AAAI Workshop and Automation as a Care Giver*, held in conjunction with American Association of Artificial Intelligence (AAAI) Conference 2002, Alberta, Canada, July 2002.

E.2.28 Haro, A. and I. Essa, "Learning Video Processing by Example" In *Proceedings of International Conference on Pattern Recognition 2002,* pp I:487–491, Quebec City, Canada, August 2002.

E.2.29 Ruddarraju, R., A. Haro, and I. Essa, "Fast Multi Camera Head Pose Tracking," In *Proceedings of Visual Interface Conference 2003,* Halifax, Canada, June 2003.

E.2.30 Minnen, D., I. Essa, and T. Starner, "Expectation Grammars: Leveraging High-Level Expectations for Activity Recognition" In *IEEE Proceedings of Computer Vision and Pattern Recognition Conference 2003,* Madison, Wisconsin, June 2003.

E.2.31 Hamid, M. R., Y. Huang, I. Essa, "ARGmode: Activity Recognition using Graphical Models" In *Proceedings of IEEE Workshop on Event Mining, Event Detection, and Recognition in Video*, held in Conjunction with IEEE Computer Vision and Pattern Recognition Conference 2003, Madison, Wisconsin, June 2003.

E.2.32 Kwatra, V., A. Schödl, I. Essa, G. Turk, and A. Bobick, "GraphCut Textures: Image and Video Synthesis using Graph-cuts," In *Proceedings of ACM SIGGRAPH 2003,* San Diego, CA, USA 2003.

E.2.33 Steedly, D., I. Essa, "Spectral Partitioning for Structure from Motion," In *Proceedings of International Conference on Computer Vision 2003*, Nice, France, October 2003.

E.2.34 Xu, J., R. Lipton, I. Essa, M. Sung and Y. Zhu, "Mandatory Human Participation: A New Authentication Scheme for Building Secure Systems," In *Proceedings of 12th International Conference on Computer Communications and Networks 2003*, Dallas, Texas, October 2003.

E.2.35 Ruddarraju, R., A. Haro, K. Nagel, Q. Tran, I. Essa, G. Abowd, E. Mynatt, "Perceptual User Interfaces Using Vision-based Eye Tracking." In *Proceedings of Fifth International Conference on Multimodal Interfaces (ICMI-PUI)*, Vancouver, B.C. November 2003.

E.2.36 A. Haro, and I. Essa, "Exemplar Based Surface Texture"", In *Proceedings of Vision, Modeling, and Visualization 2003 Conference*, Munich, Germany November 2003.

E.3. Conference Presentations with Proceedings (refereed)

[Publication (as a full paper, a short paper or a technical note) in peer refereed conference/workshop proceedings.]

E.3.1 Sclaroff, S., I. Essa, and A. Pentland. "Vision-based Modeling: An Application of a Unified Approach for Physical and Geometric Modeling for Graphics and Animation." in *Proceedings of the Eurographics Workshop on Animations and Simulations 1992,* (16 pages) Cambridge, England, September 1992.

E.3.2 Essa, I. and A. Pentland. "Facial Expression Recognition using Visually Extracted Facial Action Parameters", In, *Proceedings of the International Workshop on Automatic Face and Gesture Recognition,* pp. 35–40, Zurich, Switzerland, June 1995..

E.3.3 Brand, M. and I. Essa. "Causal Analysis for Visual Gesture Understanding", in *Proceedings of AAAI Fall Symposium on Computational Models for Integrating Language and Vision*, pp. 23–28, Cambridge, MA., November 1995.

E.3.4 Gardner, A. and I. Essa. "Prosody Analysis for Speaker Affect Determination." In *Proceedings of Perceptual User Interfaces Workshop*, (held in conjunction with ACM UIST 1997 Conference), pp 45–46, Banff, Canada, October 1997.

E.3.5 Abowd, G., C. Atkeson, and I. Essa. "Computational Perception in Future Computing Environments", In *Proceedings of Perceptual User Interfaces Workshop* (held in conjunction with ACM UIST 1997 Conference), pp. 24–25, Banff, Canada, October 1997.

E.3.6 Stillman, S., R. Tanawongsuwan, and I. Essa, "A System for Tracking and Recognizing Multiple People with Multiple Cameras", In *Proceedings of Second International Conference on Audio- Vision-based Person Authentication*, April 1999.

E.3.7 Moore, D., I. Essa, and M. Hayes, "Object Spaces: Context Management for Human Activity Recognition", In *Proceedings of Second International Conference on Audio- Vision-based Person Authentication*, April 1999.

E.3.8 Haro, A., I. Essa, and M. Flickner, "A Non-invasive Computer Vision System For Reliable Eye Tracking", In Proceedings of *ACM CHI 2000 Conference*, (Late Breaking Short Paper) pp. 167–168. The Hague, Netherlands, April 2000.

E.3.9 Abowd, G., C. Atkeson, A. Bobick. I. Essa, B. MacIntyre, E. Mynatt, T. Starner, "Living Laboratories: The Future Computing Environments Group at Georgia Institute of Technology", In *Proceedings of ACM CHI 2000 Conference*, (Organizational Overview) pp. 215–216. The Hague, Netherlands, April 2000.

E.3.10 Essa, I, G, Abowd, A. Bobick, B. Mynatt, W. Rogers, "Building and Aware Home: Technologies for the way we may live", in *Proceedings of First International Workshop on Man-Machine Symbiosis*, Kyoto, Japan, November 2002.

E.4. Conference Presentations without Proceedings (abstract refereed)

E.4.1 Friedmann, M., I. Essa, B. Horowitz, S. Sclaroff, T. Starner, and A. Pentland. "Distributed ThingWorld" Live Demonstration of the system at *G-Tech, ACM SIGGRAPH Conference, '92,* Chicago, IL., August 1992.

E.4.2 Pentland, A., I. Essa, M. Friedmann, B. Horowitz, S. Sclaroff, and T. Starner. "ThingWorld" Live Demonstration of the system at *ACM Computer Graphics, Second Interactive 3D Graphics Symposium,* Snowbird, UT., March 1990.

E.4.3 Essa, I. "Ubiquitous Sensing for Smart and Aware Environments", A Position paper at *DARPA / NIST / NSF Workshop on Smart Environments*, Atlanta, GA., July 1999.

E.4.4 Essa, I. and L. Reveret "Machine Perception of Human Activity: Recognizing and Modeling Facial Expressions and Affect.", Invited Presentation at NIPS workshop on Affective Computing. Beckenridge, CO., USA, December, 2000.

E.5. Conference Presentations: Tutorial and Courses

E.5.1 "Computer for Special Effects in Research and Education" at *IEEE Computer Vision and Pattern Recognition Conference, Course on Computer Vision for Special Effects.* (with 4 other instructors).

F. Other

F.1. Submitted Journal Papers

F.1.1 Varma, R., A. Schödl, I. Essa. "Fast Video Matting using Color Distance Sets", In preparation for *ACM Transactions of Computer Graphics*, March 2003.

F.2. Submitted Conference Papers

F.2.1 Brostow, G., I. Essa, and D. Steedly, "Automatic Armatures for Articulated Meshes", Submitted to *ACM SIGGRAPH 2003 Conference*, January, 2003.

F.2.2 Huang, Y., M. R. Hamid, I. Essa, "Towards a General Framework for Recognition of Activities and Anamolies," Submitted to *International Conference on Computer Vision 2003*, Beijing, China, 2003.

F.2.3 Haro, A., I. Essa, "Example-based Synthesis using Belief Propogation," Submitted to *International Conference on Computer Vision 2003*, Beijing, China, 2003.

F.2.4 Haro, A., and I. Essa, "Examplar-based Bidirectional Texture Functions," Submitted to *Eurographics Symposium on Rendering 2003*, April, 2003.

F.2.5 Hays, J., and I. Essa, "Non-photorealistic Moving Images", Submitted to *ACM Symposium on Computer Animation 2003*, April, 2003.

F.3. Technical Reports

F.3.1 Tanawongsuwan, R. A. Stoytchev, and I. Essa "Robust Tracking of People by a Mobile Robotic Agent" *Georgia Institute of Technology, Technical Report # GIT-GVU-99-19, 1999.*

F.3.2 Xu, J., D. Lipton, I. Essa, M-H. Sung "Mandatory Human Participation: A New Scheme for Building Secure Systems." *Georgia Institute of Technology, Technical Report # GIT-CC-01-09, 2001.*

F.4. Software

F.4.1 "The ThingWorld System", a solid modeling, sculpting software developed at the MIT Media Laboratory (1988-1992). Used in over 25 Research Institutions. Developed by a team of six people.

F.4.2 "DFace: A Dynamic Facial Modeling and Analysis Package", a facial analysis package developed at MIT Media Laboratory (1993-1995). At present ported and being further developed for Teleconferencing applications by BT (British Telecom) UK.

F.5. Video

F.5.1 Essa, I., T. Darrell, and A. Pentland. "Modeling and Tracking Facial Expression.", in *Standards of Facial Expressions, SIGGRAPH Video Review* 1995.

G. Research Proposals and Grants (Principal Investigator)

1. Approved and Funded

G.1.1 **Interaction, Collaboration, and Application Building within a 3D Virtual Environment.**
Sponsor: DURIP Program, Office of Naval Research.
Investigator(s): L. Hodges (PI), B. Ribarsky, I. Essa, J. Rossignac and three other faculty members from the Graphics, Visualization and Usability Center.
Amount: *$ 125,928 (Equipment)*
Submitted: September 1996. Funded: January 1997.

G.1.2 **Building an Intelligent, Adaptive User-Friendly Agent**
Sponsor: Yamaha Corporation, Japan. Research and Development Division
Investigator(s): A. Ram and I. Essa.
Amount: *$ 310,000, for 2 years.*
Submitted: December 1996. Funded: January 1997 - March 1999.

G.1.3 **Digital Imaging Systems for Science and Engineering.**
Sponsor: Hewlett Packard Corporation.
Investigator(s): R. Schafer, N. Ezquerra, I. Essa, and T. Barnwell, and six other faculty members from the Graphics, Visualization and Usability Center (College of Computing) and the Center for Signal and Image Processing (School of ECE).
Amount: *$ 479,000, (Equipment).*
Submitted: May 1997. Funded: June 1997.

G.1.4 **The Summer Internship Program for Socio-economically Disadvantaged Undergraduates.**
Sponsor: Office of Naval Research.
Investigator(s): I. Essa.

Amount: *$ 50,000, Annually for 2 years (1998 and 1999).*
Funded: June 1998 - May 2000.

G.1.5 **Automated Understanding of Captured Experiences.**
Sponsor: National Science Foundation, Program on Experimental Systems 1997.
Investigator(s): I. Essa (PI), C. Atkeson, G. Abowd, and K. Ramachandran.
Amount: *$ 861,216 for 3 years, ($ 60,000 matching from GT).*
Submitted: December 1997.Funded: September 1998.

G.1.6 **Data-driven Modeling for Real-Time Interaction and Animation.**
Sponsor: National Science Foundation. CISE Research Infrastructure Program 1998.
Investigator(s): J. Hodgins (PI), I. Essa, and C. Atkeson.
Amount: *$ 120,000 ($ 60,000 Matching from GT) (Equipment).*
Submitted: July 1998. Funded: February 1999.

G.1.7 **Augmenting the Capture and Understanding of Everyday Experiences**
Sponsor: National Science Foundation, CISE Research Infrastructure Program 1998.
Investigator(s): G. Abowd (PI), C. Atkeson, I. Essa, T. Starner, . . .
Amount: *$ 120,000 ($ 60,000 Matching from GT), Equipment.*
Submitted: July 1998. Funded: February 1999.

G.1.8 **Vision-based Transmedia**
Sponsor: Mitsubishi Electric Research Labs.
Investigator(s): I. Essa.
Amount: *$ 20,000 for 1999-2000.*
Funded through the GVU Affiliates Program (1/1/199).

G.1.9 **Ubiquitous Sensing in an Aware Home.**
Sponsor: Microsoft Research.
Investigator(s): I. Essa.
Amount: *$ 50,000 for 1 year.*
Funded: May 1999.

G.1.10 **Vision Technology Research**
Sponsor: Microsoft Research.
Investigator(s): I. Essa.
Amount: *$ 25,000 for 1 year.*
Funded: January 2000.

G.1.11 **A Spatio-temporal Representation for Analysis and Modeling of Facial Expressions from Video.**
Sponsor: National Science Foundation, CAREER Program.
Investigator(s): I. Essa.
Amount: *$ 300,000 for 4 years.*
Funded: September 2000.

G.1.12 **Technology First: Teaching Theory in the Context of Practice**
Sponsor: Hewlett Packard Corporation
Investigator(s): T. Barnwell (ECE) and several ECE and CoC Faculty.
Amount: *~$ 800,000 Equipment Grant ($ 100,000 for Imaging and Video Effects Classroom)*
Funded: February 2000.

G.1.13 **Video Textures.**
Sponsor: Microsoft Research.
Investigator(s): I. Essa.

Amount: *$ 25,000 for 1 year.*
Funded: March 2000.

G.1.14 **Mechanisms for Securing Emerging Applications.**
Sponsor: National Science Foundation, ITR Program 2000.
Investigator(s): M. Ahamad, I. Essa, and H. Venkateswaran.
Amount: *$ 484,000 for 3 years.*
Funded: November 2000.

G.1.15 **Human Identification by Movement.**
Sponsor: DARPA HumanID Program (IAO)
Investigator(s): A. Bobick, I. Essa, and J. Hodgins.
Amount: *$1,000,000 for 4 years (renewed 2002).*
Funded: August 2000.

G.1.16 **Aware Home Research Initiative.**
Sponsor: A Consortium of Companies that includes, Intel, MERL, Motorola, Accenture, Hewlett Packard.
Investigator(s): G, Abowd, A. Bobick, I. Essa. B. Macintyre, B. Mynatt, T. Starner.
Amount: *∼$445,000 per year.*
Consortium Established: Summer 2000.

G.1.17 **The Aware Home: Sustaining the Quality of Life**
Sponsor: National Science Foundation, ITR Program 2001
Investigator(s): G, Abowd, A. Bobick, I. Essa, B. Mynatt, and W. Rogers (Psych)
Amount: *$1,600,000 for 5 years*
Funded, July 2001 - June 2006.

G.1.18 **A Distributed Programming Infrastructure for Integrated Smart Sensors.**
Sponsor: National Science Foundation ITR 2001 Program
Investigator(s): U. Ramachandran, S. Deweerth, I. Essa, K. MacKenzie, T. Starner
Amount: *∼ $1,450,000 for 5 years.*
Funded, July 2001 - June 2006.

G.1.19 **Mandatory Human Participation: A New Paradigm for Building Secure Applications**
Sponsor: National Science Foundation, ITR Program 2001
Investigator(s): J. Xu, I. Essa, and D. Lipton
Amount: *$288,000*
Funded, August 2001 - July 2004 .

G.1.20 **Analysis of Complex Audio-Visual Events Using Spatially Distributed Sensors.**
Sponsor: National Science Foundation, ITR Program 2002
Investigator(s): J. Rehg and I. Essa
Amount: *$1,000,000*
Funded, August 2002 - July 2006.

G.1.21 **Human Activity Inference.**
Sponsor: Central Intelligence Agency, via a Subcontract from MIT
Investigator(s): A. Bobick and I. Essa.
Amount: *$500,000 for 1 year, renewable.*
Funded: August 2002.

G.1.22 **Enduring Personal Cognitive Assistant.**
Sponsor: SRI via a contract from DARPA IPTO, via a Subcontract from MIT

Investigator(s): A. Bobick. I. Essa, and J. Rehg
Amount: *$600,000/year for 3 years, renewable.*
Funded: April 2003.

2. Pending

3. Not Funded

G.3.1 **An Interactive System for Pediatric Cardiac Surgery**
Sponsor: Whitaker Foundation Biomedical Engineering Research Grant 1997
Investigator(s): I. Essa.
Amount: *$ 208,330 for 3 years*
Submitted: October 1997.

G.3.2 **Development of a Large Scale Ubiquitous Computing Interface.**
Sponsor: NSF Major Research Instrumentation Program 1998.
Investigator(s): C. Atkeson, G. Abowd, I. Essa, and K. Ramachandran.
Amount: *$ 1,984,964 for 3 years.*
Submitted: January 1998.

G.3.3 **Recognition of Group and Crowd Activity from Visual Information.**
Sponsor: DARPA MURI 2001 Program.
Investigator(s): A. Bobick, I. Essa, E. Grimsson (MIT), T. Poggio (MIT), T. Darrell (MIT), L. Guibas (Stanford), D. Koller (Stanford).
Amount: *~$3,000,000 for 3 years*
Submitted: November 2000. Under Review

H. Research Proposals and Grants (Contributor)

1. Approved and Funded

H.1.1 **Advanced Media-Oriented Systems Research: Ubiquitous Capture, Interpretation, and Access.**
Sponsor: National Science Foundation CISE Instrumentation Program
Investigator(s): U. Ramachandran and several CoC faculty.
Amount: *$ 1,636,194 for 5 years.*
Submitted: November 1998. Funded: July 1999.

2. Pending

3. Not Funded

H.3.1 **Enhancing the Quality of Life with Context-Aware Computing, Personalized Information Processing, and Secure Broadband Communications: Research at the Georgia Tech Residential Laboratory.**
Sponsor: National Science Foundation ITR 1999 Program
Investigator(s): N. Jayant, and faculty from ECE and CoC.
Amount: *$5,000,000 for 5 years.*
Submitted: March 2000. Not Funded.

H.3.2 **Engineering Research Center for Proactive Healthcare**
Sponsor: National Science Foundation
Investigator(s): Fauchett (U of Rochestor), Abowd, Larson (MIT), Helal (U of Florida) PIs., Essa listed as leader of one of the four thrusts

Amount: *$ 5,000,000, for 5 years*
Submitted: December 2002, Not Funded

I. Research Honors and Awards

- Imlay Fellowship, 1996-98.
- Edenfield Faculty Fellowship 1997.
- National Science Foundation, CAREER Award, 2000.
- Georgia Tech. College of Computing, Outstanding Junior Faculty Research Award, 2000.

III. SERVICE

A. Professional Activities

A.1. Memberships and Activities in Professional Societies

- Associate Member, Institute of Electrical and Electronics Engineers (IEEE).
  - IEEE Computer Society.
  - IEEE Computer Society, Technical Committee on Pattern Analysis and Machine Intelligence.
- Associate Member, Association for Computer Machinery.
  - SIGGRAPH.
  - SIGCHI.
- Member, Sigma-Xi, Tau Beta Pi.

A.2. Conference Committee Activities

1. **Invited Panelist**, *NSF Workshop on Facial Modeling and Animations*, Philadelphia, PA, USA, November 1993.
2. **Member, Program Committee**, *International Conference on Multi-modal Interaction*, Beijing, China, October 1996.
3. **Chair, Program Committee**, *Second International Conference on Automatic Face and Gesture Recognition 1996*, Woodstock, VT., USA, October 1996.
4. **Invited Panelist**, *NSF/DARPA Workshop on Perception of Action*, Brewster, MA., USA, May 1997
5. **Co-chair, Program Committee**, "Workshop on Nonrigid and Articulated Motion", In Conjunction with *IEEE Computer Vision and Pattern Recognition Conference 1997*, Puerto-Rico, June 1997.
6. **Member, Program Committee**, "Workshop on Recent Advances in Computer Vision", In Conjunction with *IEEE International Conference on Computer Vision 1998*, Bombay, India, January 1998.
7. **Member, Program Committee**, International Conference on Automatic Face and Gesture Recognition 1998, Nara, Japan, April 1998.
8. **Member, Program Committee**, "Workshop on Interpretation of Visual Motion", In Conjunction with *IEEE Computer Vision and Pattern Recognition Conference 1998*, Santa Barbara, CA., June 1998.
9. **Member, Program Committee**, *IEEE Computer Vision and Pattern Recognition Conference 1998*, Santa Barbara, CA., June 1998.
10. **Member, Program Committee**, *Computer Animation Conference 1998*, Philadelphia, PA, June 1998.
11. **Senior Reviewer & Member, Program Committee**, *ACM SIGGRAPH 1998 Conference 1998*, Orlando, FL., July 1998.

12. **Member, Program Committee**, *Computer Animation Conference 1999*, Geneva, Switzerland, June 1999.

13. **Member, Program Committee**, *International Conference on Automatic Face and Gesture Recognition 2000*, Grenoble, France, April 2000.

14. **Member, Program Committee**, *Computer Animation Conference 2000*, Philadelphia, PA, June 2000.

15. **Member, Program Committee**, *IEEE Computer Vision and Pattern Recognition Conference 2000*, Hilton Head, SC., June 2000.

16. **Member, Program Committee**, "Workshop on Modeling People" at the *IEEE Computer Vision and Pattern Recognition Conference 2000*, Hilton Head, SC., June 2000.

17. **Member, Program Committee**, *IEEE Workshop on Human Modeling* in Austin, TX, December 2000.

18. **Member, Program Committee**, *IEEE International Conference of Computer Vision* in Vancouver, BC, Canada, July 2001.

19. **Member, Program Committee**, "IEEE Workshop on Detection and Recognition of Events in Video" at the *IEEE International Conference of Computer Vision* in Vancouver, BC, Canada, July 2001.

20. **Member, Program Committee**, *Computer Animation Conference 2001*, Korea, June 2001.

21. **Member, Program Committee**, *ACM Workshop on Perceptual User Interfaces Workshop (held in Conjunction with UIST 2001)*, Orlando, FLorida, October 2001.

22. **Member, Program Committee**, *IEEE Workshop on Cues in Communications, (held in Conjunction with CVPR 2001)*, Kauai, Hawaii, December 2001.

23. **Member, Program Committee**, *American Association of Artificial Intelligence Conference 2002*, Alberta, Canada, July 2002.

24. **Member, Program Committee**, *First ACM Symposium on Computer Animation 2002*, San Antonio, TX, USA, July 2002.

25. **Member, Program Committee**, *International Conference on Pattern Recognition 2002*, Quebec City, Quebec, Canada, June 2002.

26. **Member, Program Committee**, *IEEE Workshop on Applications of Computer Vision 2002*, Orlando, Florida, December 2002.

27. **Member, Program Committee**, *IEEE Conference on Computer Vision and Pattern Recognition 2003*, Madison, Wisconsin, June 2003.

28. **Senior Reviewer and Member Program Committee**, *ACM SIGGRAPH 2003*, San Diego, CA, July 2003.

29. **Member, Program Committee**, *IEEE International Conference on Computer Vision*, Beijing, China, October 2003.

30. **Member, Program Committee**, *Neural Information Processing Systems (NIPS) 2003*, Vancouver, BC, CANADA, December 2003.

B. Awards

1. 1999 Dean's Award for Service to the College of Computing, GA Tech.

C. On-Campus Georgia Tech Committees

1. Member, Georgia Tech's College of Computing Awards Committee 1997-1998.
2. Member, Georgia Tech's GVU Center Ad-hoc Committee for HCI Recruiting 1997-1998.
3. Director, Georgia Tech's College of Computing Undergraduate Summer Internship Program. 1998-2000.
4. Faculty Advisor, Georgia Tech's Pakistan Student Association. 1999-present.
5. Member, Georgia Tech's College of Computing Faculty Recruiting Committee (for last few months only), 2000-2001.
6. Member (Elected), Georgia Tech's College of Computing Dean's Advisory Committee, 2001-2002.
7. Chair (Elected), Georgia Tech's College of Computing Dean's Advisory Committee, 2002-2003 (and also a member of College of Executive Committee and College Management Teams).

D. Member of Ph.D. Examining Committees

Ph.D. Examining Committee – Georgia Tech.

1. Fu-Huei Liu, Department of ECE, College of Engineering, Georgia Tech., December 1996.
   Thesis Title: "Rate Quality-based Video Coding."
   Principal Advisor: Dr. Russell M. Merserau.
2. Sadik Bayrakeri, Department of ECE, College of Engineering, Georgia Tech., March 1997.
   Thesis Title: "Scalable Video Coding using Spatio-Temporal Interpolation."
   Principal Advisor: Dr. Russell M. Merserau.
3. G. Drew Kessler, College of Computing, Georgia Tech., July 1997.
   Thesis Title: "Simple Virtual Environment."
   Principal Advisor: Dr. Larry Hodges.
4. Alan Docef, Department of ECE, College of Engineering, Georgia Tech., September 1997.
   Thesis Title: "Tele-medicine Applications of Sub-band Image Coding at Very Low Bit Rates"
   Principal Advisor: Dr. Mark Smith.
5. Ram Rao, College of Computing, Georgia Tech., July 1998.
   Thesis Title: "Multi-modal coding of lip movements."
   Principal Advisor: Dr. Russell M. Merserau.
6. Xue Li, College of Computing, Georgia Tech., July 1998.
   Thesis Title: "Scalable and Adaptive Video Multi-cast."
   Principal Advisor: Dr. Mostafa Ammar.
7. Clinton Knight, School of Electrical and Computer Engineering, Georgia Tech., March 1999.
   Thesis Title: "WWW-based Automatic Testing of Analog Circuits."
   Principal Advisor: Dr. Stephen DeWeerth.

8. Jack Tumblin, College of Computing, Georgia Tech., May 1999.
Thesis Title: "Three Detail-Preserving Contrast Reduction Methods for Displayed Images."
Principal Advisor: Dr. Gregory Turk.

9. Ara Nefian, School of Electrical and Computer Engineering, Georgia Tech., August, 1999.
Thesis Title: "A Hidden Markov Model-Based Approach for Face Detection and Recognition."
Principal Advisor: Dr. Monson Hayes.

10. James O'Brien, College of Computing, Georgia Tech., November, 1999.
Thesis Title: "Modeling Brittle Fracture."
Principal Advisor: Dr. Jessica Hodgins.

11. Peter Lindstrom, College of Computing, Georgia Tech., February, 2000.
Thesis Title: "Image-driven Polygon Simplification and Optimization."
Principal Advisor: Dr. Gregory Turk.

12. Charles Wilson, School of Electrical and Computer Engineering, Georgia Tech., June 2001.
Thesis Title: "Neuromorphic Implementations of Selective Attention."
Principal Advisor: Dr. Stephen DeWeerth.

13. David Stahl, College of Computing, Georgia Tech., August 2001.
Thesis Title: "Bag-of-Particles: An Interactive, Physically-based Deformable Model."
Principal Advisor: Dr. Norberto Ezquerra.

14. Ron Metoyer, College of Computing, Georgia Tech., August 2001.
Thesis Title: "Combining Character Intelligence and 2D Direction to Control 3D Animations."
Principal Advisor: Dr. Jessica Hodgins.

15. Victor Zordan, College of Computing, Georgia Tech., February 2002.
Thesis Title: "Motion Capture Simulations that Hit and React."
Principal Advisor: Dr. Jessica Hodgins.

16. Amos Johnson, School of Electrical and Computer Engineering, Georgia Tech., May 2002.
Thesis Title: "Human Identification: A Method for Gait Recognition Using Static, Activity Specific Body Parameters."
Principal Advisor: Dr. Aaron Bobick.

17. Huong Quynh Dinh, College of Computing, Georgia Tech., May 2002.
Thesis Title: "Implicit Shapes: Reconstruction and Transformation"
Principal Advisor: Dr. Greg Turk.

Ph.D. Area Exam Committee – Georgia Tech.

1. David Stahl, College of Computing, Georgia Tech., Fall 1996.
Principal Advisor: Dr. Norberto Ezquerra..

2. Victor Zordon, College of Computing, Georgia Tech., Fall 1997.
Principal Advisor: Dr. Jessica Hodgins..

3. Ron Metoyer, College of Computing, Georgia Tech., Fall 1997.
Principal Advisor: Dr. Jessica Hodgins..

4. Sunil Mishra, College of Computing, Georgia Tech., Fall 1997.
Principal Advisor: Dr. Chris Atkeson and Dr. Ashwin Ram..

5. Anind Dey, College of Computing, Georgia Tech., Spring 1998.
Principal Advisor: Dr. Gregory Abowd..

6. Huong Quynh Dinh, College of Computing, Georgia Tech., Spring 1999.
   Principal Advisor: Dr. Greg Turk..
7. Elliot Moore, School of Electrical Engineering, Georgia Tech., Spring 2000.
   Principal Advisor: Dr. Mark Clements..
8. Eugene Zhang, College of Computing, Georgia Tech., Fall 2001.
   Principal Advisor: Dr. Greg Turk..
9. Jun Xiao, College of Computing, Georgia Tech., Spring 2002.
   Principal Advisor: Dr. John Stasko..
10. Mark Carlson, College of Computing, Georgia Tech., Spring 2002.
    Principal Advisor: Dr. Greg Turk..
11. Mel Ericson, College of Computing, Georgia Tech., Spring 2003.
    Principal Advisor: Dr. Jarek Rossignac..

E. Consulting, Advisory, and Other External Appointments

- Limited Partner, Midway Airport Partnership, Chicago, IL, USA. 1996 - Present.
- Limited Partner, Yusuf Midway Partnership, Chicago, IL, USA. 1996 - Present.
- Member Local Advisory Board, DVT Sensors Inc. Norcross, GA, USA. 1997 - 1998.
- Consultant, Microsoft Research, Redmond, WA, USA, 2000 - 2001.
- Lead Technical Advisor, Accusense Inc. Atlanta, GA, USA, 1999 - Present.
- Member, Piedmont Hospitality Group, Mebane, NC, USA, 2000 - Present.
- Advisor/Consultant, Logic Junction Inc. Atlanta, GA, USA, 2002 - Present.
- Advisor/Consultant, DVT Sensors, Norcross, GA, USA, 2002 - Present.

IV. NATIONAL AND INTERNATIONAL PROFESSIONAL RECOGNITION

A. Invited Conference Session Chair

1. IEEE Workshop of Nonrigid and Articulated Motion 1994, Austin, TX.
2. SIGGRAPH 1998, Orlando FL. Session on Facial Animation.

B. Editorial and Reviewer Work for Technical Journals and Publishers

1. Reviewer, IEEE International Conference on Computer Vision, 1993, 1995, 2000.
2. Reviewer, IEEE Computer Vision and Pattern Recognition Conference, 1992, 1993, 1994, 1996, 1997, 1999 and 2000
3. Reviewer, ACM SIGGRAPH 1996, 1997, 1999, 2000.
4. Reviewer, ACM UIST 1999, 2000.

5. Reviewer, International Journal for Computer Vision (IJCV), Image and Visual Computing Journal (IVCJ), IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), IEEE Transactions on Robotics and Automation (R& A), Telepresence, IEEE Transaction on Multimedia, and several other leading journals for Computer Vision and Computer Graphics.

6. Guest Editor, Image and Visual Computing Journal, Special Issue on Face and Gesture Recognition 2001.

C. Patents

- Brostow. G., I. Essa "Motion Blur Generator for Stop Motion Animation." Provisional Patent Application Filed through GTRC (ID# 2449), May 2001.

V. OTHER CONTRIBUTIONS

A. Seminar Presentations (Invited Papers and Talks at Meetings and Symposia)

1. Over 10 more additional seminar/colloquia presentations.

2. "Nonrigid and Rigid Dynamic Modeling for Animation & Graphics", at *Ecole Normale Superieure, Department de Mathematiques et Informatique,* Paris, France, (Host: Dr. Marie-Paul Gascuel), September 1992.

3. "Physically-based Modeling for Vision, Graphics and Control.", at *Academy of Media Arts,* Cologne, Germany, (Host: Dr. Bernd Girod), September 1992.

4. "Visualization using Vision-based Modeling", at *GMD Super computing Facility, Visualization Group,* Bonn, Germany, (Host: Dr. Wolfgang Kreuger), September 1992.

5. "Visual Interpretation of Facial Expressions", at *NSF Workshop on Facial Animations* Philadelphia, PA., November 1993.

6. "Looking at People: Extracting Human Movements." at *International Workshop on Computer Vision and Parallel Processing 1995*, (Invited Speaker) Islamabad, Pakistan, January 1995.

7. "Looking at People: Facial Expressions", Department of Electrical Engineering Seminar, at *The Beckman Institute, University of Illinois, Urbana-Champaign,* Urbana, IL., (Host: Dr. Thomas Huang), February 1995.

8. "Looking at People: Expressions and Gestures", at Microsoft Research, Seattle WA., (Host: Dr. Matthew Turk), May 1995.

9. "Control and Estimation in Systems and Environments that Include People", Intelligent Engineering Systems Laboratory Colloquium, at *Massachusetts Institute of Technology, Department of Civil and Environmental Engineering,* Cambridge, MA., (Host: Dr. Robert Logcher), November 1995.

10. "Control and Estimation in Systems and Environments that Include People", Computer Science Colloquium, at *New York University, Department of Computer Science,* Cambridge, MA., (Host: Dr. Jack Schwartz), February 1996.

11. "Perceiving and Understanding Human Actions for Interactive Video Environments", at *Electrotechnical Laboratory*, Ibaraki, Japan, March 1996.

12. "Extracting Perceiving and Modeling of Human Activity ", in *ATR Research Laboratories Workshop on Intelligent Agents.*, (Invited Speaker) Kyoto, Japan, March 1996.

13. "Computational Perception of Human Activities in Scenes and Environments", at *Graphics, Visualization and Usability Center*, *Georgia Tech* Atlanta, GA, November 1996.

14. "Computational Perception of Human Activities in Scenes and Environments", at *Center for Signal and Image Processing*, *Georgia Tech*. Atlanta, GA, February 1996.

15. "Intelligent and Aware Environments", at CHI 1997, Ubiquitous Computing Workshop, Atlanta, GA., March 1997.

16. "Importance of Dynamic Models in Perception of Action", at NSF/DARPA Workshop on Perception of Action, Brewster, MA, May 1997.

17. "Ubiquitous Video and Audio", at *Broadband Telecommunications Center, Georgia Tech.* Atlanta, GA, March, 1999.

18. "Aware Home", at *Mitsubishi Electric Research Laboratory*, Cambridge, MA, (Host: Dr. Dick Waters), August 1999.

19. "Intelligent and Aware Environments", at *Phillips Research,* Briarcliff Manor, NY, (Host: Jim Schmidt), January, 2000.

20. "Perceiving and Modeling People, Environments and Interactions", at *Microsoft Research*, Seattle, WA (Host: P. Anandan), August 2000.

21. "Video-based Rendering and Animation" at *Graphics, Visualization and Usability Center*, *Georgia Tech* Atlanta, GA, November 2001.

22. "Temporal Reasoning from Video to Temporal Synthesis of Video" at *CMU's Robotics Institute Colloquium*, Pittsburgh, PA, February 2002.

23. "Temporal Reasoning from Video to Temporal Synthesis of Video" at *Georgia Tech's Cognitive Science Colluquium*, Atlanta, GA, February 2002.

24. "Aware Home: Sensing and Perception of People" at *University of Rochestor's Center for Future Health Colloquium*, (Host: Phillipe Fauchetti) Rochestor, NY, September 2002.

25. "Aware Home: Ubiquitous Sensing and Perception of People" at *University of Michigan, Computer Science Department*, Ann Arbor, MI, September 2002.

26. "Temporal Reasoning from Video to Temporal Synthesis of Video" at *University of Tokyo, Computer Vision Seminar*, (Host: Katsushi Ikeuchi) Tokyo, Japan, November 2002.

VI. PERSONAL DATA

Born:	11 December 1965, Karachi, Pakistan.
Home Address:	2655 Canna Ridge Circle NE Atlanta, GA 30345-1409.
Telephone:	(770) 414-8816.
Wife:	Shanila Essa; married 30 December 1990.
Children:	Gibran Essa; born 7 December 1998, Faiz Essa; born 15 April 2000.
Citizenship:	USA (Naturalized September 1996).
Email:	`irfan@cc.gatech.edu`
WWW:	`http://www.cc.gatech.edu/~irfan`

Statement of Accomplishments 2002-03

Irfan Aziz Essa

irfan@cc.gatech.edu
http://www.cc.gatech.edu/~irfan/

Research

Fundamentally, I am interested in building machines that can analyze what is happening in a scene and generate a plausible synthesis of such a scene. This follows my long-term quest of building computers that can see people, recognize them, interpret and model their actions, respond to them, and have a natural interaction with them (as I discuss in my AI Magazine article "Computers Seeing People" in 2000 [B.0.5]). My research is aimed at making fundamental contributions in *computer vision*, and *computer graphics*, with the potential intellectual impact in *artificial intelligence* and *human-computer interaction*.

The primary focus of my research is on the analysis, interpretation, and synthesis of video. In the last few years, my students and I have introduced various methods for spatio-temporal analysis of the video streams and demonstrated their viability for both interpretation and synthesis of video. Along these lines, we have extended the image-based rendering paradigm to *video-based rendering*, where analysis of video is used to synthesize dynamically changing scenes. We have developed methods for *decomposing* a 2D video of a 3D scene into a set of its constituent depth-ordered layers. We have also presented novel methods for *recognizing actions* by observing the interaction between actions and objects and by using detailed semantics for complex action recognition. We have developed various approaches for robust *processing of faces*, which includes tracking, modeling, recognition, and synthesis of facial motion. In all our work, we have heavily relied numerical modeling methods, machine learning techniques, image and signal processing approaches, and optimal control theory as the primary technical areas, allowing us to achieve strong research results, with potential long-term impact.

All these research results form the core of my fundamental contribution in the area of video interpretation and synthesis. I describe, in brief, a few of these below that my group and I made significant progress in during the last year.

Video-based Rendering & Animation: We continue to make very important and ground-breaking progress in the area of video-based rendering and animation. Our earlier paper in this area [E.2.15] is considered one of the earliest and primary contributions, and our additional publication [E.2.17, E.2.20] pretty much have established us as leaders in this field. The Video Textures papers [E.2.15] has opened the doors to much work in data-driven synthesis, and at ACM SIGGRAPH 2002, there were about 5 papers that applied the method to non-video data for animation. Overall, about 15 or so papers in ACM SIGGRAPH 2002 (of a total of 40 papers) were citing this paper.

In the last year, we have introduced two more concepts in this area that should also serve as important scholarly contributions to computer graphics and vision. In [E.2.25] we extend our approach to allow for controlled animation of video sprites (segmented regions of video that represent a specific character). This unique approach permits animation of characters, by simply video taping them and reordering the segmented video of subject, using various optimization techniques. Many researchers have explored these concepts of other data sources, but this is the first instance of applying such control for animation to video data.

We have also proposed learning and exemplar-based methods to generate novel videos and scenes. This includes methods for learning video processing by example [E.2.28,F.2.3] and modeling material and lighting properties for rendering realistic objects by using examples [F.2.4].

Finally, our most important contribution in the area of video-based rendering is our most recent ACM SIGGRAPH paper [E.2.32] on using graph-cuts optimization and search method for synthesizing

images and videos from examples. This paper introduces a completely new way of synthesizing videos from data, which generalizes to multiple dimensions and allows for stitching and merging of videos.

Action and Activity Recognition: The above-mentioned contributions are more in the space of interpreting video to aid in extracting temporal and spatial information. My second main thread of research is aimed at using video analysis to infer the context of a situation. Specifically, my contributions are in the interpretation of video to recognize what a person is doing. This is a crucial technology for building aware environments and will serve as a core technical approach for use in recognition of short-term and long-term activities in the Aware Home.

Following our earlier work in this area, which is quite well-cited [E.2.11, E.3.7], we are now focusing on extended activities within the context of the AWARE HOME effort. I am also looking into at research that can support surveillance applications, which has opened several important funding opportunities from DARPA and CIA [G.1.15,G.1.21,G.1.22]. The approaches we are exploring allow for use of context and language models to aid on interpretation of activities [E.2.26, E.2.30]. We have also started exploring methods to detect anomalies in actions to detect problems in specific actions (like cooking or blood glucose monitoring tasks) or in security applications (like person X did not complete task Y) [E.2.31,F.2.2]

[Further information on my research can be provided by Aaron Bobick, Gregory Abowd, and Kishore Ramachandran from within CoC and by Ramesh Jain in ECE.]

Teaching

My educational efforts continue to be both exciting and enriching. I describe both the classroom teaching and research advising briefly.

Classroom Teaching: Obviously, teaching the two popular undergrad specialization courses during the last year, has added to this excitement. In addition to my Digital Video Special Effects course (CS 4480), which is being offered for the fourth time (1999-2003) this year, I am now also teaching the Computer Animation (CS 4496/7496) courses.

The Digital Video Special Effects (DVFX) course, continues in its popularity both with students and with educators at other schools. I have several places who have asked me for help in designing a similar course and have even suggested sabbatical options towards that end. I have also offered course on this topic with industrial experts from production houses [E.5.1] at a leading vision conference. None of this of course compares to the very positive feedback I get from the students who have been taking this class with me. I am still in touch with several of them and they all share with me their excitement for the class. Some of them are now working in the film and games production area.

Last year, to cover the missing specialization courses in the area of computer animation (a UG specialization and a one of the sub areas for Graphics PhD core), I decided to redesign and teach the computer animation course. This course had not been offered for 2 years and there was a need both for a graduate and an undergraduate version. I taught a co-listed undergrad/grad course in spring 2002 with completely updated content that started with the basics of animation and ended with studying and implementation of some of the state of the art papers. In spring 2003 I offered only the grad version and plan on offering the undergrad version in spring 2004.

[Kurt Eiselt and Jim Foley should be able to talk about my classroom teaching efforts.]

Mentorship and Advising: I continue to work actively with students of all levels. I have a very good team of PhD students working with me. Arno Schödl finished his PhD last year and is now working as a CTO for startup in Berlin that he founded. I expect 3 more PhD students (Brostow, Haro,

Steedly) to finish this year. All of them are interested in academic positions, but the market is quite tight, so I am working with them to find Post-Doctoral positions. I am also now working with some new PhD students to bring up the number of students I am working with to eleven. Luckily, our strong reputation in graphics and vision research brings to us some of the best students in the area and it is a pleasure to work with this students (2 of them have the NSF Graduate Research Fellowships). I am also working with several

Teaching Statement

Irfan Aziz Essa

irfan@cc.gatech.edu
http://www.cc.gatech.edu/~irfan/

Classroom teaching and interacting with students on daily basis has been perhaps the most exciting and rewarding aspect of my last five years at Georgia Tech. A particular reason for my excitement in teaching is being able to combine my research with my teaching. In all my teaching interactions (classroom and advising), I take a very hands-on approach laden with real life examples and practical problems. I strongly believe in maintaining a very high-level of excitement in what I do and in all instances, I try to convey this excitement to the students. I believe that teachers with the right motivation and desire to leverage the different opportunities afforded by thought and instruction, can make for a very exciting learning environment. Here, I briefly describe my teaching philosophy and how I have put it to practice.

Curriculum Development

Five years ago, when I joined Georgia Tech's College of Computing, there were no courses offered in computer vision. I took it upon myself to (1) design a series of courses in computer vision and computational perception, and (2) to integrate the vision/perception courses with the existing courses in computer graphics, intelligent systems and robotics, human computer interaction, and the signal image processing courses offered by ECE, and (3) to provide all CoC students a background in processing, analysis, and synthesis of digital video. I introduced a three-course sequence, which included two quarter-length classes on computer vision (CS 7321 and CS 7322), followed by an advanced course on computational perception. I also opened these classes to undergraduates. I taught this sequence in 1996-97, and 1997-98 academic years. I also modified these courses into two courses, CS 4495 Computer Vision and CS 7695 Computational Perception for semester conversion. In addition to making the material in these classes accessible to undergraduates, I also designed it so that graphics, HCI, and IS students from CoC and signal/image processing students from ECE can enroll in these classes. Now that we have few more computer vision faculty (Bobick and Starner), I have turned these courses over to them.

In addition to developing the above-mentioned courses, I also worked with faculty from the Center for Signal and Image Processing (CSIP) and the Graphics, Visualization, and Usability (GVU) Center, primarily Schafer, Ezquerra, Barnwell Merserau, and Hayes towards a development of a complete curriculum in Digital Imaging in Science and Engineering Education (DISSE). The basic premise of this program was to allow for fusion of engineering and computer science concepts towards the analysis and synthesis of images. This was building on the existing strengths in computer graphics and vision (GVU) and image processing (CSIP). For this program, we pursued various forms of funding (both government and industry) and received a sizeable equipment grant from Hewlett Packard Corporation. This grant was used to establish a DISSE lab on the third floor of the College of Computing Building for used by CoC and ECE students.

My role in this curriculum planning was to design special course that would combine the basic and advanced principles of digital video analysis, processing, and synthesis with hands-on exposure to technical development in these areas. The resulting class is called **Digital Video Special Effects (DVFX)** and I have taught it for the past three spring terms.

In this class, we cover the technical details of computing with digital video and audio. Students learn about digital representations of video and study manipulation, processing, interpretation and synthesis of video under the guise of generating special effects. Students in groups produce videos to showcase their technical skills of dealing with digital video, which are then screened in an open public forum. In addition to designing this course, I also worked towards getting funds to support

the extensive video and computing needs of this course. Technology Licensing funds (about $50,000) were acquired to upgrade the classrooms and to purchase audio/video equipment to support this class. Hewlett Packard and Intel donated computer for teaching of this class. Additional funds were provided by the College of Computing. Several institutions are now modeling a similar course as this one; MIT offered one in Spring 2001, and invited me as a guest lecturer.

More information on this class, including the video projects of the students for classes in 1999, 2000, and 2001 is available from `http://www.cc.gatech.edu/dvfx/`.

Classroom Teaching

Someone once told me that teaching a class is more than just getting up in front of students and lecturing. This piece of advice pretty much defines my philosophy of teaching. Even though I do work very hard at writing my lectures for the class, it is the interactions that I support and encourage both between students and students and myself that makes my classroom teaching effective. These interactions are not limited to the classroom and are continue in out-of-class meetings and discussions on the net (via Cowebs, emails, C2000 notes). Almost all the students who have taken classes with me have found me very accessible and easy to interact with both in and out of class.

My lecturing style is very informal, allowing for extensive discussions. Most students note that they have to come prepared to discuss the material. In my lectures, I emphasize intuition about the concepts followed by deep technical depth. I am sometimes known to start a class with a very difficult problem and getting the students involved with the discussions that eventually results in a solution. I was told once that the best way to learn something is to teach it. I have learnt a lot about the field I work in by teaching courses in it.

I have strong preference for project-based classes and I have found it very useful to get students involved in projects very early in the term. This way the technical material that is covered in the class has special significance to the students as they can relate to it and are in fact dealing with it hands-on. My DVFX class is a perfect example of such a class and as the student comments from the course survey are evident, they find the class and my teaching style very exciting.

The popularity of my classroom teaching is perhaps best measured by the scores from the course survey conducted by CETL at the end of every class. For the last five years, I have average scores (out of 5) of 4.5 on C1, 4.8 on C2, 4.5 on C3, and 4.8 on the question of my effectiveness as a teacher. I was also awarded the **Georgia Tech Outstanding Use of Innovative Educational Technology Award** in 2000 for my DVFX 2000 offering.

Student Advising

I thoroughly enjoy working on research projects with both undergraduate and graduate students. I consider working with students an essential part of being a professor. Primarily because it allows for (a) to tackle very challenging research problems that would hard to undertake alone, (b) work with very bright and energetic students, and (c) mentor the students and teach them how to go about researching in the area that you have been working in for last so many years. I follow these principles in working with both graduate and undergraduate students at Georgia Tech.

I have been very lucky to find a group of very sharp and talented graduate students interested in working in the areas of computer vision and computer graphics. I have worked with them to improve their technical knowledge in the area that they profess to become experts in and then mentor them in how to go about tackling relatively open questions, and how to endure frustrating times, inherent in most research quests. I also spend considerable time with them discussing methodologies, sometime even implementing them, and then trying to get the required results to support a proposed hypothesis. Then I help them identify a possible publication option and write the paper with them. All my students

know that I like to publish in the most prestigious venues in our research area and that I will not support submitting a paper that is not ready. In some cases, my group is very deadline driven, and we do succeed in submitting some very good papers to prestigious conferences, and luckily, we can boast a very good acceptance ratio.

In addition to working on the research with my students, I also like to aid them with their writing and presentation styles. We have regular meetings where students present their current work or a new idea. I encourage very deep and technical (no-holds-barred) discussions. In fact several outside visitors to my group meetings have commented on the high-level of energy at our meetings.

Georgia Tech has very bright and talented undergraduates; a resource that is not very well tapped into for research. Most of our bright students take classes and don't undertake any research. I, and many other CoC faculty, are working hard to fight this trend and I think we are meeting with great success (with support from the administration). At any time, I usually have 3-4 undergrads working with me on small research projects. It is amazing to see their eyes light up when they get exposed to research and learn that there is something for them to do outside of the standard classroom environment.

My students have been very successful in finding excellent career opportunities. Darnell Moore, my fist PhD student at Georgia Tech now works for Texas Instruments (TI) in their research and advanced technologies division. He was extensively recruited by other companies, but wanting to live in Dallas and his background in DSP attracted him to TI. Lionel Reveret, who has been my Post-Doc for the last two years is joining INRIA, a premiere academic research institution in France as *Charge de recherches* (equivalent to Assistant Professor) in October 2001. Arno Schödl, who should finish his dissertation within a year already has options for position at the Microsoft Cambridge Research Labs, and as post-doc at MIT. All my other PhD students have been in considerable demand every year for internships. Most of my PhD students have also been very successful in winning very competitive fellowships (see my CV for complete details). In addition, some of my undergraduate students have gone on for PhD at other top schools or are working for very exciting companies doing advanced development.

The College of Computing does annual reviews of all graduate students. At this time of the year, I ask my students to conduct an evaluation of my advising skills. They provide very thorough feedback to me at that time as a group, and I have found that very useful in improving my advising skills. This idea is borrowed from Jessica Hodgins that I encourage all advisors to consider.

Research Statement

Irfan Aziz Essa

irfan@cc.gatech.edu
http://www.cc.gatech.edu/~irfan/

Fundamentally, I am interested in building machines that can analyze what is happening in a scene and generate a plausible synthesis of such a scene. This follows my long-term quest of building computers that can see people, recognize them, interpret and model their actions, respond to them, and have a natural interaction with them (as I discuss in my AI Magazine article "Computers Seeing People" in 2000 [B.0.5]). My research is aimed at making fundamental contributions in *computer vision*, and *computer graphics*, with the potential intellectual impact in *artificial intelligence* and *human-computer interaction*.

The primary focus of my research is on the analysis, interpretation, and synthesis of video. In the last few years, my students and I have introduced various methods for spatio-temporal analysis of the video streams and demonstrated their viability for both interpretation and synthesis of video. Along these lines, we have extended the image-based rendering paradigm to *video-based rendering*, where analysis of video is used to synthesize dynamically changing scenes. We have also developed methods for *decomposing* a 2D video of a 3D scene into a set of its constituent depth-ordered layers. We have presented novel methods for *recognizing actions* by observing the interaction between actions and objects and by using detailed semantics for complex action recognition. We have developed various approaches for robust *processing of faces*, which includes tracking, modeling, recognition, and synthesis of facial motion. In all our work, we have heavily relied numerical modeling methods, machine learning techniques, image and signal processing approaches, and optimal control theory as the primary technical areas, allowing us to achieve strong research results, with potential long-term impact.

All these research results form the core of my fundamental contribution in the area of video interpretation and synthesis. I will describe, in brief, a few of these below.

Video-based Rendering

Recently there has been considerable excitement in the vision and graphics communities on using real imagery in the generation of realistic synthetic scenes. This approach, referred to as "Image-based Rendering" led to many significant contributions to both visions and graphics domains in the late 1990s. My students and I have focused on expanding the concept of image-based rendering to include the temporal dimension, or *video-based rendering*.

Our approach to video-based rendering, we determine a temporal model of repeating information in a video sequence. Just as texture patches can be stitched together to generate larger textures, novel videos can be generated by stitching together subsequences once the temporal structure is known. We propose a mathematical framework based on Markov chains to link the spatio-temporally similar subsequences of a short video sequence. Once the temporal representation is known, we can generate novel videos by stitching subsequences in a similar fashion as texture patches are stitched together to generate larger textures. For this form of representation of video, we propose a mathematical framework based on Markov chains to link spatio-temporally similar subsequences of video contained in the original short bounded video sequence. Optimization algorithms are used to prune the search space, and then methods for blending, fading, and morphing, are used to synthesize novel, unbounded videos. We have also developed methods of analyzing the video stream to yield segmented video subregions, called sprites, which can be stitched and composited to yield even more complex synthesis as shown in Video #1, and #2 (see Figure 1). This work was done in collaboration with my PhD student Arno Schödl, Rick Szeliski at Microsoft Research, and David Salesin at the University of Washington and was presented at SIGGRAPH 2000 [E.2.15].

Arno and I continuing to develop numerical machine learning methods for analysis of video and dy-

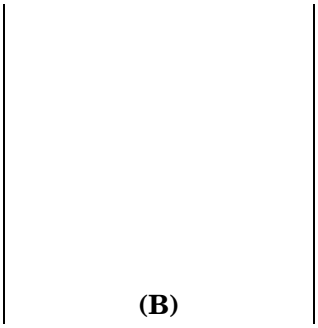

Figure 1: Examples of Video-based Rendering. (A) shows a flame sequence synthesized by merging subsequences. The red lines show the transitions. (B) Shows a fish tank synthesized by compositing 2 swaying water plants, 2 bubble streams, and 4 fishes that follow a user-specified path. (C) Shows user-control of the speed of a person on the treadmill (middle image) between a slow run (left image) and fast run (right image). See [E.2.15, E.2.17] and Videos #1, #2, and #3.

namic pruning of the space of transitions between video subsequences. We use reinforcement learning methods and optimized search algorithms for traversing through a transition space with constraints on entropy [E.2.17]. We are now working on novel control frameworks that will allow the motions of these sprites to be scripted. This will enable the generation of video-based animations by simply videotaping our subjects. We are focusing on animation of animals and insects, as it is considerably harder to generate controlled animations of such uncooperative subjects. Video #3 is a good demonstration of our work in progress along these lines. A measure of the impact of this work is that several important extensions to these methods were proposed this year in ICCV 2001 (2 papers) and Eurographics 2001 (1 paper).

Another contribution of mine in the area of video-based rendering is synthesizing motion and movement in videos to emphasize activity. Both psychologists and imaging experts consider the appearance of motion blur in video as significant for perception of motion. In computer graphics, animators make extensive use of motion blur to show motion. As the true motion of the object being moved is known, generating motion blurred computer animations is relatively easy. However, stop motion animators (responsible for among other things, feature films like *Chicken Run, Nightmare before Christmas, King Kong*) do not have such information available. Stop motion animation relies on the painstaking processing of photographing a static subject, animating it very slightly and then photographing it again. Sequencing of these crisp photographs to generate video results in a lack of motion-blur. To address this problem, my PhD student Gabriel Brostow and I have developed a fully automatic method to render motion-blur by analyzing the moving objects in video. In our method, motion blur is synthesized as a post-processing step, and it is not limited to stop motion footage, but can be applied to any video footage. Our algorithm involves extracting out a moving object, computing its motion and blurring the pixels where motion was using an exposure setting. We have used this method to generate motion blur for sequences from well-known stop motion animations as well as for real-life video sequences. Video #4 shows the results of our method. This work was presented at SIGGRAPH 2001 [E.2.20]. We have a patent pending on this method and are communicating with various production houses for possible licensing of our technology.

One of the eventual goals of my work in the area of video-based rendering is modeling of complex 3D scenes, analyzed by videos from multiple-viewpoints. This is an area where structure-from-motion methods from computer vision are merging with image-based modeling methods from computer graphics. However, making accurate scene models by dynamically evaluating thousands of images is still not tractable, as it requires complete global reoptimization. My PhD student Drew Steedly and I have developed a novel technique that enables efficiently solving structure-from-motion for extended

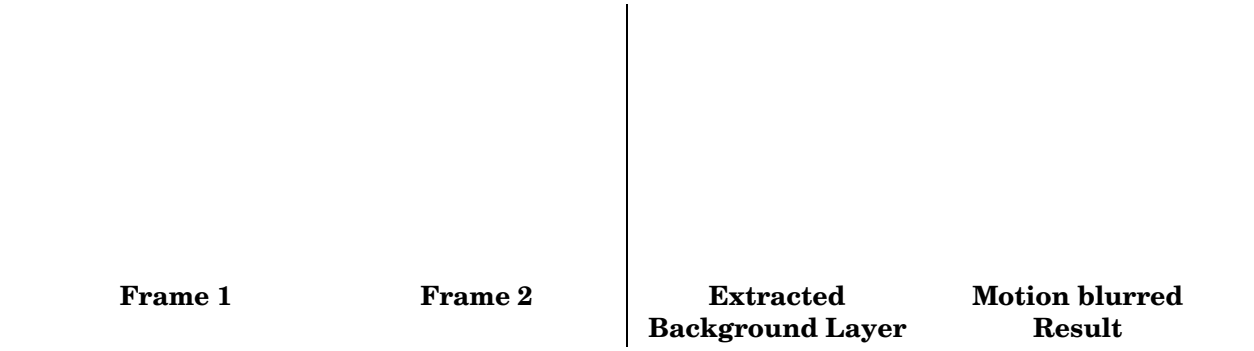

Figure 2: A simple pictorial explanation of the image-based motion blur approach. Frames 1 & 2 are two snapshots of a stop motion footage, generated by simply moving the propeller. Note the missing propeller in the extracted background layer and the motion blur in the resulting image. See [E.2.20], Figure 2, and Video #4.

sequences. Our algorithm allows for the incremental addition of new information from newly added viewpoints and optimizes the solution locally first, before expanding to a global solution. We presented this approach to the vision community in an ICCV 2001 paper [E.2.19]. Now we are concentrating on hierarchical representations of cameras and scenes to allow for even more robust and efficient solution techniques.

Decompositing and Compositing Video

The above-mentioned methods are more aimed at analyzing the temporal nature of video for synthesis and for capturing temporal structure. Next I discuss methods that allow for extracting spatial information by analyzing the temporal changes in video. Such spatial information is essential in complete scene analysis and synthesis. Video can be considered as a composite of a various possibly separable layers that are stacked on top of each other to form the complete scene image. The extraction of such depth ordered layers is significant as it provides important cues about the relative motions and the occlusions in a scene. Gabriel Brostow and I have developed a new way to decompose a video into its set of constituent layers. Our approach relies on the observation that a moving object can be used to measure how it occludes the static objects in the scene. To achieve this, first we use color segmentation for determining visible object boundaries in scene that can be occluded. Observing the changes in boundaries of segmented regions caused by the motion of a moving object then allows for decomposition of the scene from a monocular, uncalibrated camera. We demonstrated the extraction depth layers of very complex scenes using this technique in a paper at ICCV 1999 [E.2.12] (see Figure 3 and Video #5).

Arno Schödl and I have since extended this method to be independent of the initial color segmentation using an information theoretic approach, which makes the analysis even more robust. We will present this method at CVPR2001 in December [E.2.21]. In addition, My students and I have also introduced a new method for segmenting video for compositing application, where a user provide exemplars of foreground and background [F.1.1].

Action Recognition

The above-mentioned contributions are more in the space of interpreting video to aid in extracting temporal and spatial information. Now I present some of my work in using video analysis to infer the context of a situation. Specifically, my contributions are in the interpretation of video to recognize

(A) **(B)** **(C)**

Figure 3: Method extracting depth ordered layers from a video sequence. (A) Shows the motivation, note how the chair is segmented out by the person moving behind it. (B) Shows a person move in a construction scene and (C) shows the extracted depth ordered layers. See [E.2.12] and Video #5.

what a person is doing. This is a crucial technology for building aware environments and will serve as a core technical approach for use in recognition of short-term and long-term activities in the Aware Home.

Darnell Moore, a graduated PhD student, and I have proposed a new paradigm for action recognition. We combine the relationship of an appearance of the object that someone is interacting with and the appearance of interaction itself. This allowed for recognition of an object if we know the action associated with it and vice versa. We use hidden Markov model to characterize hand actions and we use simple vision algorithms to extract low-level object features. A Bayesian belief network then evaluates these labelled actions and object features to differentiate the type, or class, of the unknown object. This contribution is especially unique, as this is a major step toward functional recognition of both actions and objects. This approach was presented at ICCV 1999, and AVBPA 1999 [E.2.11, E.3.7].

We have recently extended this approach to characterize complex, multi-tasked activities [E.2.23]. Here we combine object context, image features, and heuristics to label domain-specific events. By representing each event with a unique symbol, a sequence of interactions can be described as an ordered symbolic string. A stochastic context-free grammar, which is developed using the underlying rules of an activity, provides the structure for recognizing semantically meaningful behavior over extended periods. To undertake extensive experimental evaluation of our approach, we chose the game of blackjack since we could input all the rules of the game as a grammar. These experiments show that high-level narratives of multi-player games as well as identification of player strategies and behavior can be extracted in real-time using vision. One result of this work is that after having the system observe an expert and a novice blackjack player, the system was able to quantify the "abilities" of other players.

Face Processing: Modeling, Analysis, and Synthesis

I have done a considerable amount of work on modeling and tracking faces during my career. This includes my dissertation work on facial expression recognition and modeling [E.2.2, E.2.3, E.2.6, E.2.5, B.0.3, B.0.4, E.2.8, C.1.4], as well as methods for tracking heads and eyes from videos [E.2.7, E.2.9, E.2.10, E.3.8, E.2.14], and approaches for detailed facial modeling [E.2.18]. For this reason, I am among the leading experts in the area of face processing and was recently asked to co-organize a DARPA workshop of Face Processing (Jan 2002) with the intent of discussing new problems in this area.

Facial modeling, analysis, and synthesis is an interesting and challenging domain that crosses the technical fields of computer vision, graphics, learning, and human-computer interaction. My recent focus is on face processing is on developing a relationship between speech and facial motion for ma-

Figure 4: A six state HMM used to recognize the action of flipping pages while reading a book. See [E.3.7, E.2.11, E.2.23] and Video #6.

chine perception of visual speech. I hope this will lead to ways of mediating communications between humans and between humans and machines.

Lionel Reveret, a Post-Doctoral researcher, and I have developed a new method to capture all the variations of facial motions inherent with speaking. Our method proposes a set of six Facial Speech Parameters (FSP): jaw opening, lip rounding, lip closure, lip raising, jaw advance, and pharynx raise to represent the primary visual gestures in speech articulation. These parameters are extracted by analyzing video of an expert phonetician and generating a canonical model. This model is then used to accurately track facial motions with speech. The alignment of the speech using phonemes allows us to compare the lip shapes for different subjects. We have found considerable uniformity between subjects and at present we are collecting a large database of subjects to further evaluate our approach. This work is especially exciting as it combines, vision, graphics, and learning methods to provide a robust solution to the face-processing problem [**??**, E.2.24] (Video #7).

Above is a brief description of some of the important research results on the analysis, interpretation, and synthesis of video. Further details are available in the various publications on from my web page at `http://www.cc.gatech.edu/~irfan`.

Future Research

In the upcoming years, I intend to continue on the research path that combines vision, graphics, human-computer interaction, and numerical modeling and learning. Here is a brief list of my upcoming research topics.

- In the graphics arena, I plan to pursue more work in the video-based rendering area. At present, I am working with several students on exploring methods for nonphotorealistic rendering of video. The key emphasis in this work will be on rendering motions. Then, combining this work and our earlier work on video textures, I intend to see if I can develop a representation of motion that can be used for motion-based recognition from video. I am also hoping to collaborate more with Greg Turk on some video-based modeling ideas that we have discussed.
- I intend to continue pursuing research in action and activity recognition. The newer question that I would like to address is tracking and modeling of crowds. For several years now I have been thinking that crowds can be modelled as viscous liquids, especially their motion through environments with obstacles. This idea also links back to graphics and animation, as animators use a similar principle to model crowds. Aaron Bobick and I have already written a proposal along these lines and I am also in discussions with Professor Leo Guibas at Stanford University about possible collaborations in this area.
- I am going to focus more of my effort on the Aware Home research. I find this domain very challenging and, more importantly the concept of designing technologies for elderly in the home setting very rewarding. Working with the Aging-in-place faculty (especially Mynatt, Rogers, Bobick, Abowd, and several researchers from the Atlanta Veteran's Administration) I intend to make perception and activity recognition a reality in the aware home. With Aaron Bobick, I will concentrate on getting multi-modal tracking of inhabitants working and will concentrate on building short-term and long-term behavior models of the elderly residents, which in turn can help with the recognition of their activities.
- I intend to continue my interactions with the systems group. The reason for this twofold; firstly, if any of the rich multimedia stream processing work that I work on is to become a reality, then we need researchers who deal with the complexity of computing and processing to work with us. Secondly, and perhaps more importantly, the potential of developing advanced processing paradigms by this interaction could let us develop complex algorithms that are not possible today. I am especially excited about this area because of the implications vision and sensing research may have on the embedded systems and on the security of emerging systems.

These are exciting times for research in computing, especially for me, since I thrive on multidisciplinary research. I intend to continue pursuing my diverse interests and to reach out for collaborations where necessary. I have been very lucky to find some amazing collaborators and mentors at Georgia Tech. It is my hope to let my intellectual hunger drive me to address very challenging research questions, and to use creativity and intellect to answer them, in concert with sharp students and proficient collaborators.

Funding for Research

I have been very successful in acquiring funds for my research efforts. All of my funding successes are in the areas of convergence of computer vision research with computer graphics, human-computer interaction, computer systems and security, and intelligent systems research. The budget for running my research group for 2000-2001 over $450,000, which includes my summer salary, a course buy-out, funding for Post-Docs, students, equipment, and travel. A brief overview of my research funding, especially as it relates to my interactions with other faculty at Georgia Tech, follows. A more comprehensive list is available from my CV (see Section II.G.).

Figure 5: An overview of our Animated Speakers Toolkit, used to develop a model for representing facial motions associated with speech. Applied to tracking of facial motions and for generating animations. See [E.2.24, **??**] and Video #7.

- Received a grant from the NSF's

vision, graphics, and intelligent systems research.

- Awarded the NSF CAREER Award for research on developing and evaluating machine perception of human expressions ($300k, 2000).
- Worked with Limb, Jayant, Abowd, and Atkeson to acquire Georgia Research Alliance (GRA) funds for the building and continued operation of the Broadband Institute Residential Laboratory on campus (1999)
- Actively contributed to the CoC Systems Group's Proposal for NSF CISE Instrumentation Grant (PI: Ramachandran, $1,636k, 1999).
- Received DARPA Human Identification at a Distance (HumanID) grant with Bobick and Hodgins (PI: Bobick, $500k, 2000).
- Supported Bobick and Abowd (with other FCE Faculty) in the forming of the Aware Home Initiative to acquire industrial funding for the Aware Home Project ($500k/year, 2000).
- Awarded the NSF ITR grant on securing emerging applications (PI: Ahamad, $484k, 2000).
- Awarded three NSF ITR grants; Elder Care in Aware Home (PI: Abowd, $1,600k, 2001), Distributed Smart Sensors (PI: Ramachandran, $1,350k, 2001), and a small grant on securing applications (PI: Xu, $288k, 2001).

Selected Publications for Review

Computer Graphics: ACM SIGGRAPH

Schödl, A., R. Szeliski, D. Salesin, and I. Essa. "Video Textures", In *Proceedings of ACM SIGGRAPH 2000 Conference*, pp. 489-498, ACM Press, New Orleans, LA, August 2000.

Brostow, G., and I. Essa. "Image-based Motion Blur for Stop Motion Animation" In *Proceedings of ACM SIGGRAPH 2001 Conference*, Los Angeles, CA., USA. August 2001.

Computer Vision: IEEE ICCV

Brostow, G., and I. Essa, "Motion-based Video Decompositing." In *Proceedings of IEEE International Conference on Computer Vision 1999 (ICCV'99)*, pp. 8–13, Corfu, Greece, September 1999.

Moore, D., I. Essa, and M. Hayes, "Exploiting Human Actions and Object Context for Recognition Tasks", In *Proceedings of IEEE International Conference on Computer Vision 1999 (ICCV'99)*, pp. 80–86, Corfu, Greece, September 1999.

Machine Learning: NIPS

Schödl, A., and I. Essa. "Learning Methods for Video-based Animation", In *Advances Neural Information Processing Systems 13 (Proceedings of NIPS 2000 Conference)*, T. K. Leen, T. .G. Dietrterich, and V. Tresp (Editors), MIT Press, pp. 1002–1008, May 2001.

Videos

All the videos referenced in this document are available from the webpage:

`http://www.cc.gatech.edu/~irfan/videos/.`

Direct links to the 7 videos mentioned in the research statement are listed below. Additional details on these projects are available from:

`http://www.cc.gatech.edu/~irfan/research/`.

Video #1: Video Textures (SIGGRAPH 2000) video.
`http://www.cc.gatech.edu/~irfan/videos/video1.mpg`

Video #2: Video Textures (SIGGRAPH 2000) video, shorter but with imported fish tank.
`http://www.cc.gatech.edu/~irfan/videos/video2.mpg`

Video #3: Character Animation from Video (work in progress).
`http://www.cc.gatech.edu/~irfan/videos/video3.mpg`

The results in Videos 1-3 are described in [E.2.15, E.2.17]. More information is available from the project website at `http://www.cc.gatech.edu/cpl/projects/videotexture/`

Video #4: Image-based motion blur.
http://www.cc.gatech.edu/~irfan/videos/video4.mpg

The results in Videos 4 are described in [E.2.20]. More information is available from the project website at `http://www.cc.gatech.edu/cpl/projects/blur/`

Video #5: Decompositing into Layers.
`http://www.cc.gatech.edu/~irfan/videos/video5.mpg`

The results in Videos 5 are described in [E.2.12, E.2.21]. More information is available from the project website at `http://www.cc.gatech.edu/cpl/projects/layering/.`

Video #6: Action Recognition (Blackjack!).
`http://www.cc.gatech.edu/~irfan/videos/video6.mpg`

The results in Videos 6 are described in [E.2.11, E.3.7, E.2.23]. More information is available from the project website at `http://www.cc.gatech.edu/cpl/projects/objectspaces/.`

Video #7: Animated Speakers.
`http://www.cc.gatech.edu/~irfan/videos/video7.mpg`

The results in Videos 7 are described in [E.2.24, **??***]. More information is available from the project website at* `http://www.cc.gatech.edu/cpl/projects/animated-speakers/.`

[video#.mpg works on both (Apple) Quicktime and (Microsoft) Windows Media Player]

Teaching Statement

Irfan Aziz Essa

irfan@cc.gatech.edu
http://www.cc.gatech.edu/~irfan/

Classroom teaching and interacting with students on daily basis has been perhaps the most exciting and rewarding aspect of my last five years at Georgia Tech. A particular reason for my excitement in teaching is being able to combine my research with my teaching. In all my teaching interactions (classroom and advising), I take a very hands-on approach laden with real life examples and practical problems. I strongly believe in maintaining a very high-level of excitement in what I do and in all instances, I try to convey this excitement to the students. I believe that teachers with the right motivation and desire to leverage the different opportunities afforded by thought and instruction, can make for a very exciting learning environment. Here, I briefly describe my teaching philosophy and how I have put it to practice.

Curriculum Development

Five years ago, when I joined Georgia Tech's College of Computing, there were no courses offered in computer vision. I took it upon myself to (1) design a series of courses in computer vision and computational perception, and (2) to integrate the vision/perception courses with the existing courses in computer graphics, intelligent systems and robotics, human computer interaction, and the signal image processing courses offered by ECE, and (3) to provide all CoC students a background in processing, analysis, and synthesis of digital video. I introduced a three-course sequence, which included two quarter-length classes on computer vision (CS 7321 and CS 7322), followed by an advanced course on computational perception. I also opened these classes to undergraduates. I taught this sequence in 1996-97, and 1997-98 academic years. I also modified these courses into two courses, CS 4495 Computer Vision and CS 7695 Computational Perception for semester conversion. In addition to making the material in these classes accessible to undergraduates, I also designed it so that graphics, HCI, and IS students from CoC and signal/image processing students from ECE can enroll in these classes. Now that we have few more computer vision faculty (Bobick and Starner), I have turned these courses over to them.

In addition to developing the above-mentioned courses, I also worked with faculty from the Center for Signal and Image Processing (CSIP) and the Graphics, Visualization, and Usability (GVU) Center, primarily Schafer, Ezquerra, Barnwell Merserau, and Hayes towards a development of a complete curriculum in Digital Imaging in Science and Engineering Education (DISSE). The basic premise of this program was to allow for fusion of engineering and computer science concepts towards the analysis and synthesis of images. This was building on the existing strengths in computer graphics and vision (GVU) and image processing (CSIP). For this program, we pursued various forms of funding (both government and industry) and received a sizeable equipment grant from Hewlett Packard Corporation. This grant was used to establish a DISSE lab on the third floor of the College of Computing Building for used by CoC and ECE students.

My role in this curriculum planning was to design special course that would combine the basic and advanced principles of digital video analysis, processing, and synthesis with hands-on exposure to technical development in these areas. The resulting class is called **Digital Video Special Effects (DVFX)** and I have taught it for the past three spring terms.

In this class, we cover the technical details of computing with digital video and audio. Students learn about digital representations of video and study manipulation, processing, interpretation and synthesis of video under the guise of generating special effects. Students in groups produce videos to showcase their technical skills of dealing with digital video, which are then screened in an open public forum. In addition to designing this course, I also worked towards getting funds to support

the extensive video and computing needs of this course. Technology Licensing funds (about $50,000) were acquired to upgrade the classrooms and to purchase audio/video equipment to support this class. Hewlett Packard and Intel donated computer for teaching of this class. Additional funds were provided by the College of Computing. Several institutions are now modeling a similar course as this one; MIT offered one in Spring 2001, and invited me as a guest lecturer.

More information on this class, including the video projects of the students for classes in 1999, 2000, and 2001 is available from `http://www.cc.gatech.edu/dvfx/`.

Classroom Teaching

Someone once told me that teaching a class is more than just getting up in front of students and lecturing. This piece of advice pretty much defines my philosophy of teaching. Even though I do work very hard at writing my lectures for the class, it is the interactions that I support and encourage both between students and students and myself that makes my classroom teaching effective. These interactions are not limited to the classroom and are continue in out-of-class meetings and discussions on the net (via Cowebs, emails, C2000 notes). Almost all the students who have taken classes with me have found me very accessible and easy to interact with both in and out of class.

My lecturing style is very informal, allowing for extensive discussions. Most students note that they have to come prepared to discuss the material. In my lectures, I emphasize intuition about the concepts followed by deep technical depth. I am sometimes known to start a class with a very difficult problem and getting the students involved with the discussions that eventually results in a solution. I was told once that the best way to learn something is to teach it. I have learnt a lot about the field I work in by teaching courses in it.

I have strong preference for project-based classes and I have found it very useful to get students involved in projects very early in the term. This way the technical material that is covered in the class has special significance to the students as they can relate to it and are in fact dealing with it hands-on. My DVFX class is a perfect example of such a class and as the student comments from the course survey are evident, they find the class and my teaching style very exciting.

The popularity of my classroom teaching is perhaps best measured by the scores from the course survey conducted by CETL at the end of every class. For the last five years, I have average scores (out of 5) of 4.5 on C1, 4.8 on C2, 4.5 on C3, and 4.8 on the question of my effectiveness as a teacher. I was also awarded the **Georgia Tech Outstanding Use of Innovative Educational Technology Award** in 2000 for my DVFX 2000 offering.

Student Advising

I thoroughly enjoy working on research projects with both undergraduate and graduate students. I consider working with students an essential part of being a professor. Primarily because it allows for (a) to tackle very challenging research problems that would hard to undertake alone, (b) work with very bright and energetic students, and (c) mentor the students and teach them how to go about researching in the area that you have been working in for last so many years. I follow these principles in working with both graduate and undergraduate students at Georgia Tech.

I have been very lucky to find a group of very sharp and talented graduate students interested in working in the areas of computer vision and computer graphics. I have worked with them to improve their technical knowledge in the area that they profess to become experts in and then mentor them in how to go about tackling relatively open questions, and how to endure frustrating times, inherent in most research quests. I also spend considerable time with them discussing methodologies, sometime even implementing them, and then trying to get the required results to support a proposed hypothesis. Then I help them identify a possible publication option and write the paper with them. All my students

know that I like to publish in the most prestigious venues in our research area and that I will not support submitting a paper that is not ready. In some cases, my group is very deadline driven, and we do succeed in submitting some very good papers to prestigious conferences, and luckily, we can boast a very good acceptance ratio.

In addition to working on the research with my students, I also like to aid them with their writing and presentation styles. We have regular meetings where students present their current work or a new idea. I encourage very deep and technical (no-holds-barred) discussions. In fact several outside visitors to my group meetings have commented on the high-level of energy at our meetings.

Georgia Tech has very bright and talented undergraduates; a resource that is not very well tapped into for research. Most of our bright students take classes and don't undertake any research. I, and many other CoC faculty, are working hard to fight this trend and I think we are meeting with great success (with support from the administration). At any time, I usually have 3-4 undergrads working with me on small research projects. It is amazing to see their eyes light up when they get exposed to research and learn that there is something for them to do outside of the standard classroom environment.

My students have been very successful in finding excellent career opportunities. Darnell Moore, my fist PhD student at Georgia Tech now works for Texas Instruments (TI) in their research and advanced technologies division. He was extensively recruited by other companies, but wanting to live in Dallas and his background in DSP attracted him to TI. Lionel Reveret, who has been my Post-Doc for the last two years is joining INRIA, a premiere academic research institution in France as *Charge de recherches* (equivalent to Assistant Professor) in October 2001. Arno Schödl, who should finish his dissertation within a year already has options for position at the Microsoft Cambridge Research Labs, and as post-doc at MIT. All my other PhD students have been in considerable demand every year for internships. Most of my PhD students have also been very successful in winning very competitive fellowships (see my CV for complete details). In addition, some of my undergraduate students have gone on for PhD at other top schools or are working for very exciting companies doing advanced development.

The College of Computing does annual reviews of all graduate students. At this time of the year, I ask my students to conduct an evaluation of my advising skills. They provide very thorough feedback to me at that time as a group, and I have found that very useful in improving my advising skills. This idea is borrowed from Jessica Hodgins that I encourage all advisors to consider.