Vision-Based Interfaces

Ubiquitous Computing and the EasyLiving Intelligent Environment Project
Steve Shafer, Microsoft Research

Ubiquitous computing is not only about wireless access to the network, it is also about bringing computing into everyday life through its embodiment in "intelligent environments." In this talk we look at EasyLiving, a project at Microsoft Research aimed at developing a new prototype architecture and technologies for ubiquitous computing. It includes the extensibility of "home automation" and the novel UI style of "intelligent rooms," but goes beyond these by incorporating explicit models of geometry, UI devices and the location of things and people in the world. In this way, the payoff of EasyLiving can be much greater than previous automated environment architectures. Our current demo shows the ability to track people moving in a room, control a media player and integrate diverse devices into a seamless system. Even at this early stage, some profound questions about the UI and socialization of ubiquitous computing are beginning to surface.

BraMBLe: A Bayesian Multiple-Blob Tracker
John MacCormick, Compaq

Blob trackers have become increasingly powerful in recent years, largely due to the adoption of statistical appearance models that allow effective background subtraction and robust tracking of deforming foreground objects. It has been standard, however, to treat background and foreground modelling as separate processes—background subtraction is followed by blob detection and tracking—which prevents a principled computation of image likelihoods. This talk presents two theoretical advances that address this limitation and lead to a robust multiple-person tracking system suitable for single-camera real-time surveillance applications.

The first innovation is a multi-blob likelihood function that assigns directly comparable likelihoods to hypotheses containing different numbers of objects. This likelihood function has a rigorous mathematical basis: it is adapted from the theory of Bayesian correlation, but uses the assumption of a static camera to create a more specific background model while retaining a unified approach to background and foreground modeling. Second, I will describe a Bayesian filter for tracking multiple objects when the number of objects is unknown and varies over time. A particle filter can be used to perform joint inference on both the number of objects present and their configurations, and the resulting system runs comfortably in real time on a modest workstation when the number of objects is small. This is joint work with Michael Isard.

Flexible 3D Tracking for Faces and Other Blobs
Matthew Brand, MERL

Face tracking has applications in performance animation, low bit-rate video coding and unobtrusive interfaces. We approach the problem by relating nonrigid 3D motion directly to intensity changes in image sequences. A key idea is the use of matrix transforms to propagate image uncertainty all the way through calculations so that it can interact with the problem's geometric invariants. This yields a tracker that works robustly even in low-res, low-quality video, typically with sub-pixel accuracy.

Frontiers in Attentive Environments: An Overview of IBM's BlueEyes Project
Myron Flickner, IBM Almaden Research Center

IBM's BlueEyes research project is chartered to explore and define attentive environments—environments that are user and context aware. As computers become invisible, pervasive, location/context aware, a shift from explicit actions to implicit actions is taking place in the user interface. As evidence of this shift, MAGIC pointing (gaze enabled pointing), our SUITOR (Simple User Interest Tracker) system, the Emotion mouse, and Pong, our attentive robot, will be demonstrated.

The Open Source Computer Vision Library— New Developments
### Gary R. Bradski, Intel Corporation

The Open Source Computer Vision Library (OpenCV) is a collection of C and light C++ routines running under Linux and Windows intended to support research and to help create a market in consumer computer vision. As such, the source is open and is free for research or commercial use under a BSD- style license (user's source need not be open). An active user group of over 1500 users may be joined by sending email to *OpenCV-Subscribe@yahoogroups.com*. We have recently posted OpenCV to SourceForge to allow greater community contributions.

I will overview the library; demonstrate automatic camera calibration, embedded-HMM-based face recognition, and real-time performance analysis. I will end with new results in extremely efficient and lighting stable image segmentation and demonstrate its use in background differencing.

Speaker Detection Using Boosted Dynamic Bayesian Networks
Jim Rehg, Compaq

Advanced user interfaces based on speech and vision pose a challenging inference problem: The actions and intentions of multiple people must be estimated from sequences of noisy and ambiguous sensor data. In this talk I describe some recent applications of dynamic Bayesian network models to user interface for a Smart Kiosk. The Smart Kiosk provides information and entertainment to multiple people in public spaces and is based on vision, speech and touch sensing. I present a DBN architecture for speaker detection, inferring when a user is speaking to the kiosk. This architecture fuses off-the-shelf visual and audio sensors (face, skin, texture, mouth motion and silence detectors) with contextual cues from the interface itself. Experimental results confirm the importance of temporal duration and context in accurate classification. A novel application of boosting is shown to improve classifier performance. This is joint work with Vladimir Pavlovic and Ashutosh Garg.

❖❖❖
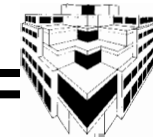
SCHEDULE

Time	Event
8:30	**B R E A K F A S T** and Registration 4th floor, CIT Building
9:15	Introduction and overview Tom Dean, CS Chairman John Savage, IPP Director Michael Black, Host
9:40	**Ubiquitous Computing and the Easy-Living Intelligent Environment Project** Steven Shafer, Microsoft Research
10:30	**B R E A K**
10:45	**BraMBLe: A Bayesian Multiple-Blob Tracker** John MacCormick, Compaq
11:35	**Flexible 3D Tracking for Faces and Other Blobs** Matthew Brand, MERL
12:30	**B U F F E T L U N C H**
2:00	**Frontiers in Attentive Environments: An Overview of IBM Almaden Research Center's BlueEyes Project** Myron Flickner, IBM
2:50	**The Open Source Computer Vision Library—New Developments** Gary Bradski, Intel Corporation
3:40	**B R E A K**
4:00	**Speaker Detection Using Boosted Dynamic Bayesian Networks** Jim Rehg, Compaq
4:50	**P A N E L D I S C U S S I O N**
5:30	**R E C E P T I O N** (5th floor)

EMAIL REGISTRATION

To: sjh@cs.brown.edu
By: April 23
Please include the following:
Name
Title
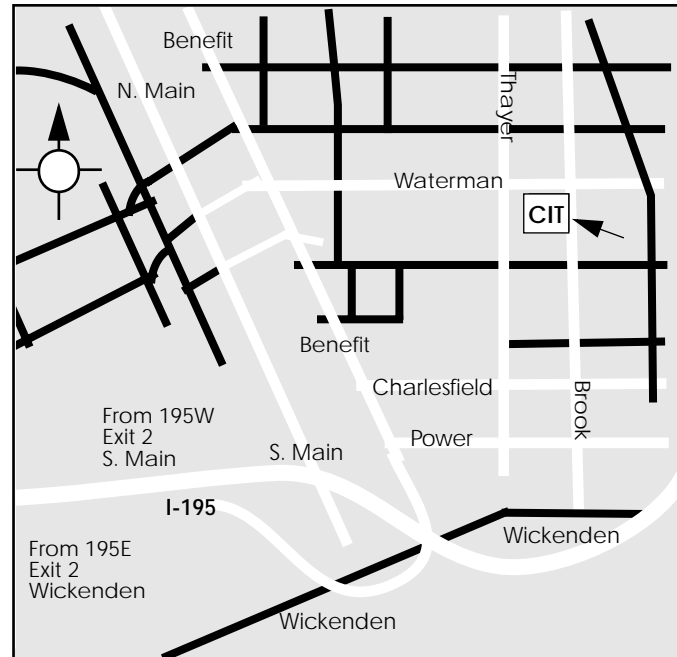*Company*
Department
Postal address
Phone/Fax

This symposium is for members of our Industrial Partner companies. *Member companies are: Compaq, EMC, Foxboro, Gtech, IBM, InterTrust, Latitude, MERL, Microsoft and Sun.* There is no charge.

❖❖❖

LODGING

Rooms have been reserved at the Inn at Brown (the Brown guest facility), corner of Thayer and Charlesfield for May 2nd ($100/night); parking is included. Please make reservations *by April 8,* by calling (401•863•7500).

Please refer to the CS Department's Industrial Partners Program when registering. Participants are responsible for their own lodging expenses. We look forward to seeing you soon.


Benefit
N. Main
Thayer
Waterman
CIT
Benefit
Charlesfield
Brook
From 195W
Exit 2
S. Main
Power
S. Main
I-195
From 195E
Exit 2
Wickenden
Wickenden
Wickenden

DIRECTIONS TO THE CIT BUILDING

- From I-95 N or S, take Exit 20 to I-195E.
- From I-195E take Exit 2, St.
- Go LEFT on Wickenden, LEFT again at the 2nd light onto Brook St.
- The red-brick CIT Building (Center for Information Technology) is on the left at the intersection of Brook and Waterman (1st light).
- Registration is on the 4th floor.

PARKING

Because most of the visitor parking has been assigned to University employees, I'm afraid we're unable to provide parking. Street parking is usually available for early birds, but watch out for newly-designated 2- and 3-hour zones, which used to be all-day spots. You might try the residential area NW of the CIT.

The 27th IPP Symposium
Department of Computer Science
BROWN UNIVERSITY
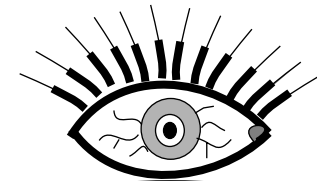
VISION-BASED INTERFACES

Thursday
May 3, 2001
Host: Professor Michael Black

INDUSTRIAL
PARTNERS
PROGRAM