**Hedco Neuroscience Building**
3641 Watt Way
Los Angeles, CA 90089-2520

Research in iLab

The main fundamental research focus of the lab is in **using computational modeling to gain insight into biological brain function**. Thus, we study biologically-plausible brain models, and we compare the predictions of model simulations to empirical measurements from living systems. The brain subsystem towards which most of our efforts are focused is the visual system. Our modeling efforts range from fairly detailed models of small neuronal circuits, such as a single hypercolumn of orientation-selective neurons in primary visual cortex, to large-scale models embodying several million highly-simplified neurons to explore mechanisms of visual attention, gaze control, object recognition, and goal-oriented scene understanding. Further, we strive to employ **modeling principles which are mathematically optimal** in some task- and goal-dependent sense. Thus, we are interested in investigating the tasks and conditions for which the biological brain approaches the theoretical limits of information processing.
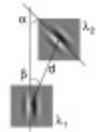
Our fundamental research activity includes experimental work with human subjects. One experimental technique used in the lab is **visual psychophysics**, with which we probe the mechanisms underlying basic visual perception by asking observers to quickly report on some attributes of simple visual patterns flashed on a computer screen. This is complemented by **eye-tracking** research, where we highly accurately monitor the gaze of human participants to provide an implicit behavioral response, in complement to possible explicit responses such as pressing a response button. A second experimental focus is to employ *in vivo* **functional neuroimaging** techniques to correlate brain activity to psychophysical performance, for example using functional magnetic resonance imaging (fMRI) to measure local changes in brain blood oxygenation correlated with mental activity. This neuroimaging focus is interested not only in the basic science of normal brain function, but also in the medical investigation of how such function may be altered in disease conditions. Finally, a third upcoming experimental focus is to employ **electrophysiological recording** to probe the activity of single neurons or small groups of neurons in the living monkey brain as well as in slice preparations from rodent brains.
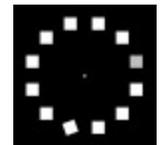
Directly complementing our modeling and experimental focus on the basic science of brain function, our lab also explores a number of engineering applications of this basic research work, mainly in the fields of **machine vision**, **image processing**, **robotics**, and **artificial intelligence**. The underlying vision is our belief that, to be proven truly useful and insightful, computational neuroscience models should not only be tested against neural or behavioral data in the context of specialized laboratory experiments, but should also be exercised in the context of more general applications which confront the models to the real world. For example, we investigate whether our biologically-inspired visual models can be extended to solve problems such as automatic target detection in cluttered natural scenes, video compression, autonomous robotic navigation on land or under water, or animation of virtual agents. We also investigate how learning and knowledge representation techniques derived from research in artificial intelligence could be used to make our models more performant at solving
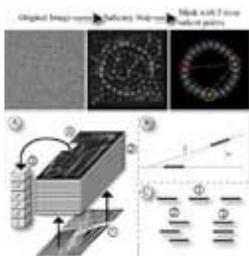
given machine vision tasks.

			Our generous sponsors:			
USC / Eng.	DARPA	NIH / NEI	NSF	NGA	USC / Zumberge Fund	HFSP

Research Highlights and Projects

Contour Integration and Complex Features

How do humans connect simple unconnected elementary visual features of objects in the visual field into a *Whole Object*? At some point our brain takes what seems to be several unconnected object parts and links them into an object such as a circle. Complex interaction among neurons in V1 and V2 seem to help us connect these items. Here we create a realistic biologically-plausible model of human contour integration. The model predicts many behaviors we would expect to see in human subjects and also sheds some light on what the contour integration mechanism might be like.

Web page: See our **Contour Integration Home Page**.

Selected Publications:

- **T. N. Mundhenk, L. Itti, Biological Cybernetics, 2005.**
- **R. J. Peters, T. N. Mundhenk, L. Itti, C. Koch, Proc. Society for Neuroscience Annual Meeting (SFN'03), 2003.**
- **T. N. Mundhenk, L. Itti, CINNIC, Neurocomputing, 2003.**
- **T. N. Mundhenk, L. Itti, In: Lecture Notes in Computer Science, 2002.**
- More on the **iLab publication server** and on **Contour Integration Research**

Modeling Goal-Oriented Scene Understanding

How do we understand and interpret complex visual environments in a manner that depends on our higher intentions and goals?
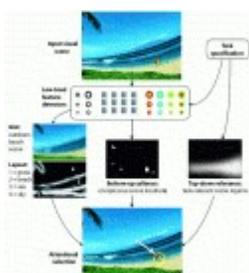
Our modeling efforts emphasize four aspects that are important in biological vision: determining task-relevance of an entity, biasing attention for the low-level visual features of desired targets, recognizing these targets using the same low-level features, and incrementally building a visual map of task-relevance at every scene location. Given a task definition in the form of keywords, our model first determines and stores the task-relevant entities in symbolic working memory, using prior knowledge stored in symbolic long-term memory. The model then biases its saliency-based visual attention system for the learned low-level visual features of the most relevant entity. Next, it attends to the most salient location in the scene, and attempts to recognize the attended object through hierarchical matching against stored object representations in a visual long-term memory. The task-relevance of the recognized entity is computed and used to update the symbolic working

memory. In addition, a visual working memory in the form of a topographic task-relevance map is updated with the location and relevance of the recognized entity.

Web page: See our **Goal-Oriented Scene Understanding Home Page**.

Selected Publications:
- **V. Navalpakkam & L. Itti, Lecture Notes in Computer Science, 2002**
- **V. Navalpakkam & L. Itti, Proc. WAPCV, 2003**
- **C. Siagian & L. Itti, Proc. IEEE International Workshop on Face Processing in Video, 2004**
- **V. Navalpakkam & L. Itti, Vision Research, 2005**
- **C. Siagian & L. Itti, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006**
- More on the **iLab publication server** on **Scene Understanding**

Eye Movement Research
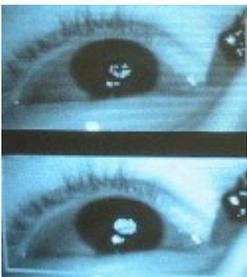
Where do you look when confronted with complex natural environments, and why?

Using an **eye-tracking machine** we can record gaze locations of human observers watching or interacting with a wide range of dynamic stimuli. This reference data then allows us to test whether computational models of attention and scene understanding can correctly predict where humans look while doing such tasks as watching TV, walking outdoors, or playing video games.

Web page: See our **Eye-Tracking Home Page**.

Selected Publications:
- **R. Carmi & L. Itti, Proc. VSS, 2004**
- **L. Itti, Visual Cognition, 2005 (1)**
- **L. Itti, Visual Cognition, 2005 (2)**
- **R. Carmi & L. Itti, Proc. ETRA, 2006**
- **R. Carmi & L. Itti, Journal of Vision, 2006**
- More on the **iLab publication server** on **Eye-Tracking Research**

Quantifying the Wow! - A formal Bayesian theory of surprise

Why do some events immediately catch our attention while others just become ignored as background clutter? To address this question, we have developed and tested a new formal Bayesian theory of surprise, in collaboration with Prof. Pierre Baldi at the University of California at Irvine. Crucially, our theory emphasizes how, for something to become surprising, it much change your beliefs about the world. This new definition provides for the first time a formal mathematical framework with which surprise may be quantitatively measured. This theory complements Shannon's theory of information by emphasizing the effects an event may have onto the subjective beliefs of an observer, while Shannon's theory emphasizes measuring the intrinsic objective complexity and predictability of the event. In experiments with human subjects, we found that surprise as defined by our theory is the strongest known attractor of human attention and gaze: observers looked towards surprising events in television

and video game clips significantly more reliably than they looked towards simply colorful, contrasted, moving, or informative events. This work has widespread applications ranging from video surveillance to web search and advertising design.

**Web page:**   See our **Bayesian Theory of Surprise Home Page**.

Selected Publications:
- **L. Itti & P. Baldi, Proc. IEEE-CVPR, 2005**
- **L. Itti & P. Baldi, Proc. NIPS, 2006**
- More on the **iLab publication server** on **Bayesian Theory of Surprise**

Components of Bottom-Up Gaze Allocation

How are eye movements influenced by known computational mechanisms in early vision?

We use our model of **bottom-up attention** along with **eyetracking psychphysics** to evaluate the roles of several types of non-linear interactions known to exist in visual cortex, and of eccentricity-dependent processing. For each of these, we add a component to the salience model, including richer interactions among orientation-tuned units, both at spatial short range (for clutter reduction) and long range (for contour facilitation), and a detailed model of eccentricity-dependent changes in visual processing. Psychophysics subjects free-view naturalistic and artificial images while their eye movements are recorded, and the resulting fixation locations are compared with the models' predicted salience maps. We find that the proposed interactions indeed play a significant role in the spatiotemporal deployment of attention in natural scenes; about half of the observed inter-subject variance can be explained by these different models. This suggests that attentional guidance does not depend solely on local visual features, but must also include the effects of interactions among features. As models of these interactions become more accurate in predicting behaviorally-relevant salient locations, they become useful to a range of **applications in computer vision** and **human-machine interface design**.
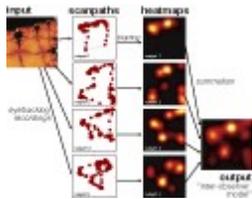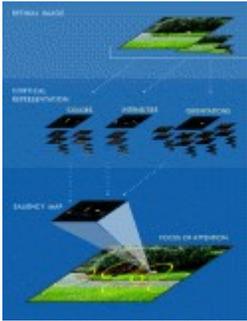
Web page: **Source code for computational models**
GroovX C++ toolkit for psychophysics experiments

Selected Publications:
- **R. J. Peters, L. Itti & C. Koch, SFN Annual Meeting, 2002**
- **R. J. Peters, T.N. Mundhenk, L. Itti & C. Koch, SFN Annual Meeting, 2003**
- **R. J. Peters, A. Iyer, L. Itti & C. Koch, Vision Research, 2005**
- **R. J. Peters & L. Itti, Proc. ETRA, 2006**

Modeling Bottom-Up, Saliency-Based Visual Attention

What are the cues that attract your visual attention to some locations in your visual environment rather than to others? Far from being a passive system comparable to a video camera, your visual system performs complex computation on incoming information such as to enhance those locations which will appear as perceptually salient.

How does it work? We have it all figured out at the link below!

Web pages: See our **Bottom-Up Visual Attention Home Page**.
Also see the **iLab Neuromorphic Vision C++ Toolkit** for all the source code.

Selected Publications:
- **L. Itti, C. Koch & E. Niebur, IEEE PAMI, 1998**
- **L. Itti & C. Koch, Vision Research, 2000**
- **L. Itti, Ph.D. Thesis, California Institute of Technology, 2000**
- **L. Itti & C. Koch, Nature Reviews Neuroscience, 2001**
- More on the **iLab publication server** on **Model of Bottom-Up Saliency-Based Visual Attention**

Neurobiology of Attention

A key property of neural processing in higher mammals is the capability to focus resources, by selectively directing attention towards the most important sensory inputs of the moment. Attention research has shown rapid growth over the past two decades, as new techniques have become available to study higher brain function in humans, non-human primates, and other mammals. **Neurobiology of Attention** is the first encyclopedic volume to summarize the latest developments in attention research. An authoritative collection of 111 concise articles organized into thematic sections provides both broad coverage and access to focused, up-to-date research findings. The volume presents a state-of-the-art multidisciplinary perspective on psychological, physiological and computational approaches to understanding the neurobiology of attention. Ideal for students, as a reference handbook, or for rapid browsing, the book has a wide appeal to anybody interested in attention research.

Web page: See our **Neurobiology of Attention Page** for the complete table of contents.

Modeling Basic Pattern Perception
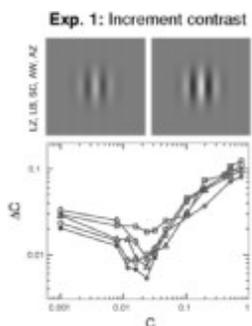
This research focuses on quantitatively modeling early visual processing, and on linking the activity in a small population of visual neurons to human psychophysical data. To this end, we developed a theoretical framework which allows quantitative predictions of human basic pattern discrimination performance. A unifying model of basic pattern vision is also being developed, which simultaneously predicts many psychophysical observations.
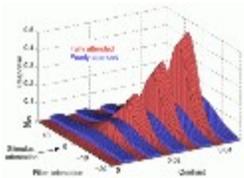
Web page: See our **Basic Pattern Perception Page** for examples.

- **L. Itti, J. Braun, D. K. Lee & C. Koch, NIPS 1998**
- **L. Itti, C. Koch & J. Braun, JOSA-A, 2000**
- More on the **iLab publication server** on **Computational Modeling** and **Human Psychophysics**

Selected Publications:

Modeling Attentional Modulation of Early Visual Processing

There is converging experimental evidence that attention modulates, top-down, early sensory processing. That is, neural activity in early visual (or other sensory) areas is enhanced at the currently attended location, compared to the rest of the visual field. Our work aims at understanding this modulatory effect of attention in computational and quantitative terms. Based on a detailed model of early visual processing and its prediction of attentional modulation observed in five different pattern discrimination tasks, we have recently proposed that attention activates a winner-take-all competition among early sensory neurons.

Web page: See our **Top-Down Attentional Modulation Page** for examples.

Selected Publications:

- **L. Itti, J. Braun, D. K. Lee & C. Koch, NIPS 1999**
- **D. K. Lee, L. Itti, C. Koch & J. Braun, Nature Neuroscience, 1999**
- **L. Itti, Ph.D. Thesis, California Institute of Technology, 2000**
- **L. Itti, J. Braun & C. Koch, NIPS 2002**
- More on the **iLab publication server** on **Model of Top-Down Attentional Modulation**

Medical Image Processing
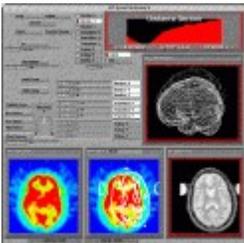
We are developing a number of image processing algorithms applied to the analysis of functional brain scans. This includes software for anatomical/functional co-registration (aligning two brain scans obtained with two different machines), correction for partial volume effects, automatic scan prescription, perfusion imaging, automatic segmentation of white-matter lesions, and many more.

Follow the link below for an overview!

Web page: See our **C.N.S. Overview Page** for examples.

Selected Publications:

- **L. Itti, L. Chang, J. F. Mangin, J. Darcourt & T. Ernst, Human Brain Mapping, 1997**
- **L. Itti, L. Chang, T. Ernst & F. Mishkin, Human Brain Mapping, 1997**
- **T. Ernst, O. Speck, L. Itti & L. Chang, Magnetic Resonance in Medicine, 1999**
- **L. Itti, L. Chang & T. Ernst, Magnetic Resonance in Medicine, 2001**
- **L. Itti, L. Chang & T. Ernst, Journal of Neuroimaging, 2001**
- More on the **iLab publication server** on **Medical Image Processing**

Functional Neuroimaging and Medical Research
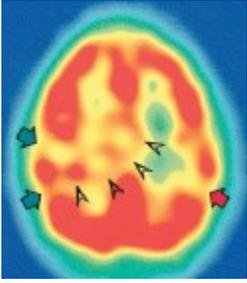
We are investigating a number of medical and disease conditions using functional brain scans. These includes investigating the effects of drug abuse on the brain (cocaine, ecstasy), HIV and other dementia, myotonic dystrophy, Alzheimer's disease, progressive multifocal leukoencephalopathy, postinfectious encephalitis, Klinefelter's syndrome, and more.

Follow the link below for an overview!

Web page: See our **C.N.S. Overview Page** for examples.

Selected Publications:

- **D. F. Benson, A. Djenderedjian, B. L. Miller, N. A. Pachana, L. Chang, L. Itti & I. Mena, Neurology, 1996**
- **D. I. Kaufer, B. L. Miller, L. Itti, L. A. Fairbanks, J. Li, J. Fishman, J. Kushi, J. L. Cummings, Neurology, 1997**
- **T. Ernst, E. Itti, L. Itti & L. Chang, Journal of Magnetic Resonance Imaging, 2000**
- **E. Itti, K. Huff, M. E. Cornford, L. Itti, K. Poruri & F. S. Mishkin, Clinical Nuclear Medicine, 2002**
- **E. Itti, I. T. Gaw Gonzalo, K. B. Boone, D. H. Geschwind, N. Berman, A. Pawlikowska-Haddal, L. Itti, F. S. Mishkin & R. S. Swerdloff, Annals of Neurology, 2003**
- **E. Itti, I. T. Gaw Gonzalo, A. Pawlikowska-Haddal, K. B. Boone, A. Mlikotic, L. Itti, F. S. Mishkin & R. S. Swerdloff, Journal of Clinical Endocrinology and Metabolism, 2006**
- More on the **iLab publication server** on **Functional Neuroimaging** and **Medical Research**

Neuromorphic Engineering and Robotics

We are starting a new research effort which aims at developing the next generation of highly-capable robots, based on neuromorphic vision algorithms. Because our vision algorithms require high computational power, our first effort is in developing a new high-speed, outdoors robotics platform with a 4-CPU **Beowulf** on board. We call this new breed of robots the **Beobots**.

See the link below for more information on this new, project, let by a team of undergraduate students.

Another robotics project also led by an undergraduate team is dedicated to designing and building an autonomous visually-guided submarine, with intended participation to the autonomous underwater-vehicle competition held by the Association for Unmanned Systems International (AUVSI). The competition tests aspects of autonomous control, planning, vision, and target localization. Our submarine combines disciplines of mechanical engineering, electrical engineering, and computer science to meet the challenges of the competition.

See the link below for more information.

Web pages:	See our **Beobots Page** for further information on the land beobots. See our **USCR Page** for further information on the submarine robots. Also see the **iLab Neuromorphic Vision C++ Toolkit** for all the source code.
Selected Publications:	• **D. Chung, R. Hirata, T. N. Mundhenk, J. Ng, R. J. Peters, E. Pichon, A. Tsui, T. Ventrice, D. Walther, P. Williams, L. Itti, Lecture Notes in Computer Science, 2002T. N. Mundhenk, C. Ackerman, D. Chung, N. Dhavale, B. Hudson, R. Hirata, E. Pichon, Z. Shi, A. Tsui & L. Itti, Proc. SPIE, 2003** • **C. Ackerman & L. Itti, IEEE Transactions on Robotics, 2005** • More on the **iLab publication server** on **Beobots**

Computer Graphics, Animation, and Video Games

Enabling the next generation of immersive virtual environments and video games necessitates that virtual agents develop human-like visual attention and gaze behaviors. A critical step is to devise computationally-tractable visual processing heuristics, to allow agents to rapidly find locations that would attract human gaze in complex dynamic environments. Here we evaluate image processing heuristics derived from biological vision against eye movement recordings from five humans playing video games or watching complex stimuli, so as to derive better models that can endow virtual agents with a sense of vision.

Web page:	See our **Saliency-Based Virtual Agents** for examples.
Selected Publications:	• **L. Itti, N. Dhavale & F. Pighin, Proc. SPIE, 2003** • **L. Itti, N. Dhavale & F. Pighin, Proc. IEEE Int'l Conference on Multimedia and Expo, 2006** • More on the **iLab publication server** on **Computer Vision**

Saliency-based Video Compression

We have developed a new method for automatically finding regions of interest in static images or video clips, using a neurobiological model of visual attention.

This model computes a topographic saliency map which indicates how conspicuous every location in the input image is, based on the responses from simulated neurons in primary visual cortex, sensitive to color, intensity, oriented edges, flicker and oriented motion. The method is applied as a front-end filter for image and video compression, where a spatially-variable blur is applied to the input images prior to compression. Image locations determined interesting (or salient) by the attention model receive no or little blur, while image locations increasingly farther from those hot spots are increasingly blurred. Locations which are more highly blurred will be compressed more efficiently and hence will yield a smaller overall compressed file size. To the extent that the algorithm indeed automatically marked as hot spots regions that human observers would find interesting, the blur around those regions should be tolerable when viewing the resulting clips.

Web page:	See our **Saliency-based Video Compression Home Page** for additional information.

Selected Publications:

- **N. Dhavale & L. Itti, Proc. IEEE International Symposium on Signal Processing and its Applications, 2003**
- **L. Itti, Proc. SPIE HVEI, 2004**
- **L. Itti, IEEE Transactions on Image Processing**
- More on the **iLab publication server** on **Computer Vision**