Research Statement of Alex Leykin

http://cgi.cs.indiana.edu/~oleykin

We are surrounded by large amounts of recorded visual data, but these data are difficult to process and use without computers. This rather simple concern became a starting point in my career as a computer vision researcher. As soon as I began my doctoral studies, I realized that developing intelligent algorithms that can process and interpret visual data can have enormous implications for the understanding of how humans see as well as for the ways we search, analyze, and interpret visual information. In my research I focus primarily on applied computer vision problems and strive to find theoretical and algorithmic solutions for modeling and understanding vision-related actions and human behavior.

My first graduate research project was a development of an automated classifier of images based on the high-level semantic content. I built a system aimed at distinguishing painted art from photography. By studying various techniques used by painters and employing pattern recognition methods to extract color, edge and texture information, I was able to develop a classifier based on the conglomerate of neural networks that provided 93% classification accuracy. The results of this project were published in a journal and presented at the IEEE Conference on Computer Vision and Pattern Recognition in 2003.

Two other graduate research projects involved methods of image processing and pattern recognition to solve practical problems. I developed a machine learning framework based on support vector machine classifiers that maximized the readability of augmented reality screens by determining conditions under which texts were unreadable and optimally adjusting color and position of the text label to best fit a given background texture. I also worked on a distributed collaborative project "The Indiana Robotic Vehicle Initiative," aimed at meeting the DARPA Grand Challenge of racing fully autonomous vehicles in the off-road conditions. I was part of the team that developed a real-time algorithm for detecting the road outline, road vanishing point and the horizon line to improve road following. I used texture analysis methods, such as Gabor wavelets, to extract compact descriptors from the video and model curve road outline to probabilistically assess navigation scenarios in the decision making.

As I continued to look for challenging applications of computer vision research, I approached one of the leading researchers in the field of marking intelligence, Dr. Raymond R. Burke, and learned that computer vision can provide indispensable tools for understanding human cognition in general and for improving people’s shopping experiences in particular. The domain of marketing intelligence posed a challenge for computer vision methodology and required creativity and innovation that would benefit and advance both fields.

My dissertation titled "Visual Human Tracking and Group Activity Analysis: A Video Mining System for Retail Marketing" focused on two tasks - tracking human movements and analyzing group behavior in crowded store environments. These tasks are challenging primarily because the level of entropy generated by multiple body occlusions is high and because the dimensionality of the model increases exponentially with the increasing number of people.

As the first part of my dissertation I built a probabilistic monocular and panoramic video human tracking system designed to operate robustly on a large number of tracked objects. To provide
a solid starting point for the tracker, I developed a method for adaptive background subtraction, suited for high-noise compressed videos. The tracking process begins by detecting human head candidates, and then continues by estimating body positions of each individual based both on the color histogram distance as well as foreground mask consistency. To impose temporal consistency I used sequential Markov Chain Monte Carlo methods to probabilistically estimate the next state of the tracking system. Based on the tracking data, for the second task of analyzing and modeling group activities I modeled human activities as a swarm behavior with multiple independent agents acting according to number of simple rules, such as avoiding collisions, maintaining fixed or coordinating velocity vector with the neighbors. My method proved to be successful in detecting whether people congregate, ask for assistance or proceed to checkout in a retail store solely from video material. As an example, in three hours of retail store shopping sequence 78% of the shopper groups were correctly detected with only 1.8% in false positives.

For the next step in my research career I looked for opportunities to work on interdisciplinary, application-driven problems and develop methods that would work in real-world unconstrained environments. As a postdoctoral fellow in the Customer Interface Lab at the Kelley School of Business, Indiana University I am involved in a series of projects aimed at understanding consumer behavior by estimating visual attention. Attention is modeled as a combination of low-level salience measures of the objects in a shopping scene and high-level customer intentionality descriptors. To supplement this research I developed a cognitive vision system to analyze human visual attention and attribute it to specific items of merchandise or navigational signs in the retail store. The system includes of a full 3D interactive store. This allows the user to participate in a simulated shopping trip. Along the trip, the position of the user in the store and the exact eye gaze coordinates are recorded. This provides an unprecedented capability to map gaze to specific items in the scene. The visual properties of each item that affect its visual salience are then combined along with other non-visual characteristics in a dynamic system to estimate its potential for attracting visual attention. Moreover, I proposed a Neural Network based model which can be trained using human eye tracking data as a benchmark to enhance the qualities of the system. The outcome of this system is the detailed analysis of attention consumers pay for each item, brand, product category or a navigational sign in the store.

A postdoctoral position in the Customer Interface Lab provided me with an invaluable experience of working in an interdisciplinary environment and communicating with corporations capable of supporting cutting-edge computer vision research. My work elicited significant interest among the marketing research departments form the manufacturers and retailers. The companies who have contributed to the Customer Interface Lab in the past are Procter & Gamble, NCR, American Eagle Outfitters. Specifically, part of my work on visual attention has been funded by Sears Holdings Corporations. In the future I plan to continue to collaborate with my colleagues from the Kelley School of Business and work on computer vision methods that will help researchers to automatically process marketing intelligence data and analyze consumer behavior at the point of purchase. Some intriguing problems I want to address are the influence of scale on attention, motion cues and parallax, fusion of tracking and attention estimation to create temporal attention maps. I would also like to explore the applications of my research in security and automotive navigation and safety, areas in which I have actively participated in the past.

More information about my research including publications and video samples of my work can be found at http://cgi.cs.indiana.edu/~oleykin