About me. I am a sixth graduate student in the School of Computational Science and Engineering working under the supervision of Professor Alexander G. Gray with research focus on algorithm development for data analysis and machine learning. My thesis is on developing new paradigms for approximate nearest-neighbor search. My home page is http://www.cc.gatech.edu/~pram.

Foundation: Reading and Assignments

The reading material prescribed for every topic generally consisted of a chapter (or chapters) from the textbook *Computer Vision: A modern approach*. While each chapter covered interesting topics and went into great details regarding each topic, it was generally hard for me to grasp the topics in a single reading before class. I managed to pick up pieces of concepts from the reading material. In my opinion, the assignments were usually easier to solve because they were more focussed in scope and hence were easier to tackle. Reading and understanding the whole chapter was harder for me. Surprisingly, there were cases where I did not require to read the material at all to solve the assignments. However, the classroom coverage of the topic helped a lot in clarifying (some of) the concepts presented in the book.

Most of the assignments that were given to us over the period of the course were fairly intuitive. However, I would not classify any of the assignments as my favorite. But I can state that in terms of topics, linear filters and related topics (Chapters 4 - 6) were my favorite. I found their representational power of these filters and their abilities to extract information from an image really amazing. Signal processing is new to me, so the parts covering them were very exciting to me. In terms of the assignment that taught me the most, I can say that the assignment on geometric camera models stands out. Prior to this class, I had no knowledge regarding camera models. This assignment helped me understand the geometry involved in the image formation in camera.

I believe that I performed fairly well in my assignments. This is also reflected in the grades I obtained for the assignments. While I did not get perfect scores on all assignments, but I think I managed to get good scores in general. I had a better handle on the assignment questions involving simple mathematics or linear algebra. Moreover, the later assignments on clustering and classification were less challenging to me because of my background in machine learning. Hence I did well on these later assignments and the assignment questions involving algebra and geometry. One assignment I did have a lot of trouble with was the question 7.7 in the chapter on Stereopsis. There was very little information in the chapter regarding this question and I had very little intuition. Another assignment I had trouble with was the pop quiz in class. This was an example of where the terms used in class were a little unfamiliar to me (probably because I have mostly worked in a different field). I cannot exactly remember the exact terms but I do remember them being a little confusing to me only because I had some preconceived notions on the meanings of those terms. I also never found out what question(s) I lost my point on.

Skills: Mini-projects

The mini-projects helped me learn a lot about some fields of Computer Vision research. The reading involved in coming up with the idea and the proposal helped me learn the most. Moreover, the projects made me aware of various tools and baseline data sets used in Computer Vision. My first project was on intrinsic image estimation [2]. This problem has been studied for quite some time and still lacks a very efficient solution with high accuracy. I am still very interested in this problem and would like to pursue it in my future. Through my work and research on this project, I found baseline datasets for this problem. My second project aimed towards performing multiview stereo with SIFT-matches [3]. This project allowed me to familiarize with tools available for obtaining SIFT features for images and the application of RANSAC to obtain SIFT matches. My third project [1] introduced me to the problem of optical flow estimation between images (or in a video). Again, it introduced me to various seminal as well as current work for this problem, and also made me aware of the baseline datasets for optical flow. All the mini-projects taught me a lot about the problems I addressed in the projects and allowed me to be fairly current with the existing solutions available for that problem.

Each of the mini-projects had their share of successes and failures. While some ideas worked well, some other
did not to fruition as expected. The most challenging project was my first mini-project on intrinsic image estimation. The hypothesis we were exploring was that the observed image is the intrinsic reflectance image corrupted with shading – if the intrinsic reflectance image is a signal, then the observed image is the original signal passed through a “shading” linear filter. We tried to validate this hypothesis without success and this could be attributed to our limited understanding of the concepts of linear filters and intrinsic images. It was only towards the conclusion of the mini-project that we realized that the shading is also part of the intrinsic and cannot be consider as a noise filter. The second and third mini-projects were more successful since we were just proposing a new method and we just implemented the method to evaluate its capability. Both these projects provided some favorable results, while having some caveats which seem obvious in retrospection. They were more limited in scope and concept compared to the idea for the first mini-project. The third mini-project appeared to be most successful in terms of actual results. It was a simple idea that has surprisingly favorable performance. I do feel like I spent equivalent amount of time on all the mini-projects and all my papers were positively received by the mock review panels.

Final Comments

My research focus has been on nearest-neighbor search and machine learning, more geared towards algorithms for large scale data. To me, Computer Vision is one of the best test beds for algorithms both in search and machine learning. Vision deals with large scale data and poses complex learning problems. My motivation to take this course was to learn about this field in a structured manner and understand the different aspects of this field. Not only is the material in Computer Vision immensely stimulating and useful to me, I found the classroom teaching in this course very unique and, in my humble opinion, just amazing. Not only did this course give me an opportunity to learn about the different aspects of Computer Vision, but also learn it from instructors who are/were actively involved in that field (or have deep knowledge of the field).

While giving me a overview on the field of Computer Vision, the mini-projects in this course has also gotten me interested in some of the problems in this field. Specifically, I am deeply interested in the problems of intrinsic image estimation and optical flow estimation. I believe that both these problems can possibly benefit from various statistical estimation methods prevalent in machine learning. I plan to pursue a couple of ideas I have in the future. Overall, I believe I got a lot more from this course than I anticipated and I am thankful for that.

CIOS Course Feedback Participation. I have completed the CIOS and TAOS surveys online.

References