Teaching Computer Vision using Primary Source Material

Kevin Novins and Brendan McCane

Department of Computer Science
University of Otago
Dunedin, New Zealand

E-mail for correspondence: novins@cs.otago.ac.nz

Abstract

Computer vision is a very broad and rapidly-changing field. A major challenge for teachers of this subject is to cover a large body of basic theory while at the same time giving a good sense of state-of-the-art research. Traditionally our fourth year undergraduate course in computer vision was taught using lectures and textbooks. We report on the results of altering the course so that it focused instead on our students' own research into recent primary source literature. This experiment was a qualified success. The students found the course interesting, gained a high-level understanding of the state of the art in computer vision, and developed critical analysis skills. However, despite a high course workload, student learning of details of techniques suffered. They were able to complete the course without a good understanding of the theoretical and mathematical foundations of computer vision. In response, this year, we will be adopting a hybrid approach.

1 Introduction and Motivation

Teaching computer vision in an undergraduate course is difficult because it is such a broad and rapidly changing subject. Because of its breadth, it is impossible to cover all topics in a single year course, so the lecturers must choose a subset of topics. Often depth must be sacrificed for breadth, as evidenced by the choices made by authors of recent textbooks [4]. The rapid rate of change of the field indicates that computer vision is a relatively immature subject and that most problems remain unsolved, or only partially solved. What topics, then, should we cover in such a course and which techniques should we teach?

To answer these questions we need to decide what we would like undergraduate computer vision students to learn. What are the most useful skills and
techniques for computer vision students? In the past our course attempted to teach students particular packets of knowledge. We found that this gave our students a skewed view of the field of computer vision. They learned only a selected subset of techniques but ended up believing that they had learned the most important ones. They didn’t learn how to go out on their own to discover new techniques to solve particular problems.

If such students continue their work in computer vision, they show a strong bias towards using only the techniques that they learned as undergraduates, even if these aren’t the best tools for the job. For example if line detection is called for, they might jump immediately to the Hough transform. From the lecturer’s point of view, the Hough transform was probably a useful example of the power of working in a dual space. From the student’s view, the Hough transform becomes a “solution” to the line-detection problem. In computer vision, most problems are still open problems, but undergraduate courses often give the impression that most problems can be solved using a small set of well-known tools and techniques.

According to Ramsden ([5], page 20), the most important educational objectives for any subject are:

- To teach students to analyze ideas or issues critically.
- To develop students’ intellectual/thinking skills.
- To teach students to comprehend principles or generalisations.

These objectives are exactly those that we believe are the most important attributes for a computer vision student to have. More specifically, we believe a successful computer vision student (or researcher for that matter) should be able to:

- Read and understand recent computer vision literature and find out about elements or techniques in that literature that they don’t know about.
- Implement techniques reported on in recent computer vision literature.
- Critically analyze recent techniques and determine their applicability for solving a problem.
- Place current research in the context of related computer vision research.

So, our idea of an ideal computer vision student is not one who knows the classical techniques of shape-from-shading, stereo, edge detection, segmentation, object recognition, etc. Although classical techniques are clearly important, it is not possible to cover all of them, or to decide which ones will be most significant in future research. Therefore, we decided that it was the skills of thinking, critical analysis, and self-directed learning in the discipline of computer vision that are far more important than specific knowledge about any given technique.

Given these attributes, the next problem is how to teach them. According to Gibbs [2], “it seems fairly clear that lecturing is rather poor at achieving most
of the higher level educational objectives that we aspire to: analysis, evaluation, attitude change and so on ...”. This quote struck a chord with us as we had experienced frustration in trying to teach computer vision in the traditional manner in previous years. We found that students in the main did not learn or understand much of the material, were not active during the lectures and typically found the lecturing mode somewhat boring. Even though readings of recent literature were also assigned, we found that many students did not read them prior to the class. Some only briefly skimmed them in the few minutes before class. This made for poor discussion sessions that often degenerated to the point where the lecturers were doing almost all of the talking.

In this paper, we report on the results of altering our course so that it focused instead on our students’ own research into recent primary source literature. This experiment was a qualified success: The students found the course interesting, and gained a high-level understanding of the state of the art in computer vision. They developed critical analysis skills and now know where to look for new algorithms. An added benefit is that we, as lecturers, also found the course to be extremely interesting since much of the recent literature which was covered was new to us. On the negative side, students’ knowledge of details of existing techniques suffered. Students were able to complete the course without a good understanding of the theoretical and mathematical foundations of computer vision. This year, we will be adopting a hybrid approach.

2 Structure of the Course

Initial enrollment of the course was six fourth-year undergraduate students. The course was scheduled to meet for one two-hour session each week, for fifteen weeks. Our university’s workload standards allowed us to assign eight hours of work for the course each week, including the two hour class session. Each session was devoted to a particular topic in computer vision. The students were given a “classic” paper in the field to read, and were expected to find a recent journal article of their own and review it in advance of each session. The marks for the best twelve (out of fifteen) weekly reviews were averaged to count for 15% of the final grade.

We also had 7 weeks of the year with no two-hour sessions, but for which we could assign 8 hours of work. For this period, each student was required to find a recent journal article that interested them, implement the technique, and test it. A written report on the project was worth 25% of the grade.

The remainder of the final grade was determined by a three hour final examination.

Each component of the course is described in the following sections.
3 Paper Reviews

We wanted to make sure that students learned how to go out into the computer vision literature on their own to find solutions to problems. We realized that we could effectively use the power of the group if we sent each class member on an individual quest each week: they all could benefit by sharing and comparing the solutions that they found. Our model was that each individual would prepare only one solution or technique each week, but that they would be able to learn about several others from their classmates during discussion sessions.

We decided to require the students to prepare for each weekly session by reading an assigned “classic” paper on a topic as well as by reviewing a journal article published in the past three years on the topic. In addition to giving students essential background, the classic paper ensured that the students had a common benchmark against which they could evaluate the new papers that they found.

We expected reviews that were 1–2 pages long and that were in the same style as reviews performed by referees for a journal. We distributed a modified form of the IEEE Transactions on Parallel and Distributed Systems review form as a basis of the student reviews (Appendix B). This review form employs a large checklist, which we thought would help focus the student’s writing.

We rated the initial sets of reviews that we received quite poorly. They lacked the critical analysis that we were looking for. We quickly discovered that part of the problem was in the IEEE review form. The questions were focused towards whether or not the paper should be accepted for publication. Students felt that the “right answer” must be that the paper was worthy. After all, the papers had to have passed peer review in order to be published. We also noted that the IEEE review form focused too much on presentation issues that were not of much interest to us.

After two weeks, we switched to a simpler form, based on five questions:

1. What is the specific problem (or aspect of a problem) that this research attempts to solve? That is, what is claimed about this particular technique that is not true of previously published techniques?

2. How does the technique solve the problem? That is, give a brief summary of the technique in your own words (one paragraph maximum). What is new about the technique? That is, how does it solve the problem it is supposed to address?

3. Is the testing of the technique adequate and do the results back up the claims made by the authors?

4. What are the limitations of the technique? Does the technique have any limitations which the authors do not report?

5. Is the paper well written, clearly organized, understandable, and of appropriate length?
We indicated that any review that fully answered the five questions would receive full credit, and suggested that students unsure of how to structure a review should aim toward writing a paragraph on each question as a starting point.

The new review form worked well. Students appreciated the more concrete nature of the instructions, and we were pleased to have the students focusing on summarizing and analyzing the papers that they found.

Students handed in their reviews along with a photocopy of the paper that they reviewed. These were returned within one week with, on average, a paragraph of feedback from the marker.

At first, we marked the reviews on a binary scale: each review received either full marks, or zero marks. We didn’t want students who never got the hang of understanding and analyzing papers to feel that their work was acceptable, so we decided that the pass threshold should be very high – essentially at the “A” level. But in order to allow the students to get used to the style of the course, we gave full marks to all reasonable efforts during the first two weeks. At the same time we gave feedback indicating how the review would need to be improved in order to reach the pass mark in later weeks.

The binary marking scheme proved to be too harsh. One student said it was “demoralizing” to work for several hours on a review, and then receive no credit at all for the work because it didn’t have enough analysis to reach an “A” level. In response to this, we changed to a trinary marking scheme. A zero was given for little or no effort, half marks were given for an honest effort that was not quite satisfactory, and full marks were set to the “A” level, as before. This did not interfere with our own goals because half marks corresponded to the university’s official definition of failure. But the students were satisfied to know that they would receive at least some credit, as long as they put in a reasonable effort.

4 Weekly Discussion Sessions

Discussion sessions first addressed any difficulties the students had in understanding the “classic” paper. Then, each student in turn would briefly present the paper that they reviewed. Students initially had difficulty adjusting to this style of learning. They didn’t seem interested in their classmates’ presentations and were reluctant to ask their classmates questions, even if they didn’t understand. In the early sessions we tended to fill the gap by asking questions of our own. Of course, we were interested in the papers that the students found, many of which we had not encountered before.

Part of the problem was a general shyness on the part of the students. But another major factor was that students seemed to genuinely misunderstand the purpose of discussing the techniques as a group. They worried about what would be on the final exam, and didn’t see how they could possibly be expected to know techniques that their fellow students studied.

Our response was two-fold. First, we had to learn to stop asking questions of the student presenters ourselves. Instead, we would stimulate discussion by
asking random “audience” students to summarize what the presenting student just said. If they couldn’t do it, they were asked to pose a question to the presenter that would aid in their understanding.

We also handed out an aims and objectives document (Appendix A) that explained to the students more clearly what we were trying to achieve with the discussion sessions and what expectations we had for the final exam. This information, along with a subsequent group discussion of the aims and objectives document seemed to have a very positive effect on student attitudes to the course. In retrospect, it’s clear that we should have had a discussion of aims and objectives at the outset.

Students tended to optimize their behavior by selecting papers with little or no mathematical content. There was no formal mathematics prerequisite for the course and most students had first-year calculus experience that had become somewhat rusty. We did not find a solution to the problem of student emphasis on easier papers. In a sense it didn’t matter, because even the easiest published journal articles still required a significant effort before they were understood.

We did try to get students to expand their own mathematical understanding by placing an emphasis on understanding the “classic” papers in discussion sessions. We began handing out particular focus questions along with each paper, and used these as the basis for class discussion. The questions often tried to highlight mathematical understanding. For example, one of the focus questions for Horn and Schunk’s optical flow paper [3] was: “Give a brief intuitive explanation of how the equations for $u^{n+1}$ and $v^{n+1}$ on Page 12 perform the update shown in Figure 4. What is the role of $\alpha$?”.

5 Computer Vision Project

The purpose of the computer vision project was to make sure that each student had hands-on experience in implementing a computer vision algorithm. We also wanted to be sure that each student understood at least one technique in depth.

Students had to select a paper from the recent literature and implement it and test it. They had to hand in a project proposal, worth 5% of the final grade and a final report, worth 20%. In keeping with university guidelines, a computer vision project was meant to take 56 hours of work. As a rough guide we suggested the following distribution of effort:

- Library research and project proposal (6 hours)
- Implementation (24 hours)
- Testing (16 hours)
- Final report (10 hours)

We emphasized that the main point of the project was to understand and analyze a single paper in detail. We indicated that 56 hours of honest work and a good report was acceptable, even if a working algorithm was not achieved.
The project proposals encouraged students to get an early start on their projects and afforded us an opportunity to steer students towards algorithms that could be implemented in a short time period. We also looked for experimental designs that were likely to result in interesting findings. Project proposals were due after one week and failing proposals were given a single opportunity for re-submission.

Four of the six students made a serious attempt at their projects. They seemed to enjoy the work, and learned a lot about both algorithms and how to analyze them. Two students in particular were able to write reports that very clearly and usefully supplemented the information published with the original paper. However, these students probably spent significantly more than the nominal 56 hours.

The main difficulty with running a computer vision project based on recent literature is that most published algorithms simply cannot be implemented in just 24 hours of work.

6 Final Exam

University regulations required that we hold a three-hour exam for the course, worth 60% of the final grade. Although variations to this structure are possible, they must be petitioned several months in advance of the start of the term. Our course redesign came too late to meet the course variation deadline.

We felt that a traditional closed-book exam would be inappropriate, given that the course focused on research skills, rather than memorizing packets of knowledge. In order to meet the letter of the university regulations, if not the spirit, we decided to release the questions in advance, and allow the students to bring in any written or printed materials to the exam. A wise exam strategy in this case would be for the student to write complete answers in advance, and merely copy the results into an exam book during the exam period. In effect, our exam became equivalent to an ordinary essay assignment, with the length limited by what could be hand written in three hours.

There were two exam questions. The first focused on understanding of a particular topic in computer vision and allowed the students some flexibility in terms of choosing an area of interest. The second was more open-ended and tested student understanding of the field as a whole. Both questions required library research. The questions were:

1. Choose a topic area from the following list and report on the state of the art in that area. Include descriptions of several recent published approaches and their limitations:
   - Edge Detection
   - Segmentation
   - Active Contours
   - Shape Representation
• Structure from Motion
• Object Recognition
• Optical Flow
• Stereo Correspondence

2. Suppose that you are employed by Cyberdyne corporation, which wants to improve security by installing an automatically opening door. The door is to open itself for authorised individuals only, who are to be recognised using a passive computer vision system. The personnel do not want to be constrained to wear special clothing or other accessories.

Write a report discussing possible approaches to solving this problem and include an overall assessment of its feasibility. Also include relevant references to the literature to convince your boss that you are up to date with current research.

Of the six students in the course, only three sat the final exam. Two students that missed the exam had effectively dropped out of the course early in the term. This was probably not a reflection on our course in particular; these same students dropped out of all of their courses that year. The third student that missed the exam did so due to illness.

Of the three exams that we marked, two were in the B/B+ range and one was in the C range. We found some excellent high level analysis, but we were disappointed that none of the students gave descriptions of algorithms in sufficient detail to convince us that they truly understood them. None of the individual answers to questions were in the A range.

Interestingly, the exam marks correlated very highly with the best twelve marks of the fifteen reviews. For summative purposes, the trinary review marks, averaged across the term could have served as an acceptable final mark.

7 Students’ and Lecturers’ Impressions

We performed a small student survey at the end of the course and received three responses. Overall, students rated the interest level, and the learning experience of the course higher than that of their other fourth year computer science courses. They found the weekly reviews to be the most interesting part of the course, followed closely by the computer vision project.

On the negative side, all students rated the workload of the weekly review as being too high. Our own experiments with writing reviews indicated that the students had a point; we found it difficult to find a good article and write a good review within six hours, the nominal weekly workload for the paper. The students also were disappointed that all their hard work in reviews counted for only 15% of the final grade. One student complained that the marking scheme for the reviews was too harsh.

On a scale from 1 (excellent) to 5 (poor), two students rated the course overall as a 2, and the third rated it a 3.
From the lecturers’ point of view, the course was more interesting to run than a lecture-based course, though the workload was similar. We didn’t have to design lectures, but marking the reviews and giving good feedback required significant weekly effort. But the marking workload wasn’t as dreary as it might seem: it required us to read and understand the papers that the students selected; an activity that we wanted to do anyway to keep up with research in the field. We were fortunate to have only six students to teach in this way; the demands on the teacher of a course like this one would rise rapidly with the number of students in the course.

8 Conclusion and Future Plans

Near the end of the year, we discovered the web page of a graduate seminar in “Advanced Perception” that is taught at Carnegie-Mellon University in a strikingly similar manner to the one reported in this paper [1]. They also used weekly written reviews of students’ own selections in combination with classic papers. Their review focus questions are very similar to ours, and they also used a trinary marking scheme. We have found no reports of the success of their course, but the fact that we independently came up with such similar features seems to indicate that the problems we encountered were not unique to our situation, and that the solutions we employed may prove to be quite general.

We felt that we succeeded in running a course in which the students were active, hard working participants. Students became very familiar with the computer vision literature, and developed important skills in critical analysis. We are confident that all students that passed the course have a good sense of the state of the art in computer vision, and would know where to turn if trying to solve a computer vision problem. They know how to distill the essential facts from complex papers and won’t easily be fooled by papers that make misleading claims.

We were disappointed, however, in students’ ability to understand the details of papers that they read. The extent of this problem did not become clear to us until we marked the final exam. It is impossible to measure in-depth understanding from 1-2 page reviews. Except for the computer vision project students were able to “get away” with only surface-level understanding.

At the outset, we felt that it was not essential to teach specific techniques, as long as we were confident that our graduates would be equipped to teach themselves specific techniques “on demand” later in their careers. We are not convinced that we met that goal. It seems that learning specific packets of knowledge may be useful in giving the students experience in delving into the details of complex algorithms. Even if the algorithms themselves become outdated, the basic reasoning skills and the confidence gained remain useful forever.

For this reason, we have redesigned the course for this year using a hybrid format. We eliminated the final exam and divided the course into six problem-based learning units on a particular topic. Each unit will be assessed
independently and will contain library research, writing, and programming components. Literature reviews of the type described above will count as part of the library research and writing requirements. The new structure will allow us to ensure that students learn the low-level details of particular techniques, while not losing sight of recent publications. The main sacrifice we will make with this format will be that we will cover fewer topics in computer vision. We will be forced to sacrifice some breadth for depth. The workshop will take place midway through the southern hemisphere academic year; we should be able to present preliminary results using our revised course at that time.

References


Appendix A: Course Aims and Objectives

Preamble

Computer vision is a large field that includes many problems and many attempted solutions to those problems. It would be impossible to cover everything in the space of a year, and even if we could, the knowledge you would gain would become out of date quickly. Instead of us deciding on a fixed curriculum and then proceeding to teach our view of that curriculum, we feel it is more beneficial for you to do much more self directed learning within the scope of the topics being discussed. We realise this is a different mode of teaching/learning to what you are used to, but it more closely reflects the type of learning you will have to do if you continue on in your studies to postgraduate level and even in the workplace (so called lifelong learning!). You are now facing the same problems that we, as researchers, face all the time. What are the good papers and techniques? Is this new technique significant? We are continually working to fill gaps in our own knowledge.

We expect you to be learning primarily through your individual forays into the computer vision literature, and through your interactions with each other over what you found there. As lecturers, our contribution to your learning is in the form of feedback both on the reviews and in the lecture timeslot rather than in the form of standing out the front of the class and lecturing to you.

Below are our intended learning outcomes, and how the assessment in the course relates to these outcomes. We also list a set of focus questions on this topic. We will both be available at the lecture on 26 May to discuss these issues with you, if that’s what you want.

Intended Learning Outcomes

1. Knowledge of the state-of-the-art in computer vision: What problems have been solved? What are the active research areas?

2. Knowledge of techniques used in computer vision systems: What problems are they applied to? What are their strengths? What are their weaknesses?

3. Possession of critical analysis skills necessary for research.

4. Possession of research skills for computer vision: What are the major sources of current information? What differentiates them?

5. Ability to maintain state-of-the-art knowledge in computer vision: What are the necessary steps between learning about a new technique and being able to understand it and implement it?
Assessment in Relation to Learning Outcomes

Weekly reviews of the literature

Through weekly assessment and feedback, students develop their skills in critical analysis (3). Since the reviews are on primary source material of the student’s own choosing, they develop library research skills (4). The review focus questions require knowledge of what problem is being solved and how it is solved (1) (2). Feedback from lecturers indicates whether they have done so to sufficient depth.

Computer vision project

Since the topic is based on a recent article of the student’s choosing, library research skills are developed (4). The student must understand and implement the technique, which may involve further study, such as following up background references (5). Students also gain insight into a particular research issue and a particular computer vision technique (1) (2). Through the design and execution of experiments, they learn to apply critical analysis to their own work (3).

Final Exam

Questions will focus on student’s general knowledge and appreciation of the state of the art in computer vision. Students will be required to generalise their knowledge of individual topics to include techniques that their classmates studied. They will also be required to integrate their knowledge from different topic areas in computer vision (1) (2) (3).

Students should display a more detailed understanding of the papers (background and weekly choices) that they read for the course (1) (2) (5).

The final exam will also test knowledge of the basic image processing techniques taught during the first five lectures (2).

Focus Questions

1. How well do our intended learning outcomes match with yours? Can you suggest others?

2. Has the effort that you have applied in COSC453 paid off in terms of the target learning outcomes?
Appendix B: Initial Review Form
The purpose of these literature reviews is give you practice reading research papers with a critical eye. In this context, we (IE Kevin and myself) have put together a review form which is largely based on the official review form of the journal "IEEE Transactions on Parallel and Distributed Systems". Hopefully, this will give you some idea of the review process as well as the literature itself. What we would like you to do is assess the quality of the paper as if you were a reviewer for an appropriate journal. In fact, you can consider yourself a reviewer for the COSC453 class. That is, you should indicate the quality of the paper and its likely interest to other members of the COSC453 class. Following is a description of what comprises a quality paper in terms of research.

For a paper to be acceptable for publication in any research journal or conference either as a paper or a brief contribution, it must comprise novel material not previously published. The novelty will usually lie in original results, methods, observations, concepts, or applications, but may also reside in syntheses of or new insights into previously reported research. The title, abstract, introduction, and summary should be sufficiently informative to make the contribution of the paper clear to the broadest possible audience, and to place them in context with related work.

In addition to these fundamental requirements, the quality of a paper depends on a number of important criteria relating to reader interest, technical content, and presentation. To assist you in addressing these criteria, the Review Form includes a short-answer overview (Section II). The principal intent of the overview is to raise the kind of questions that should be addressed in assessing the paper. In other words, the overview provides a list of the criteria referred to above and, in this sense, serves as a part of these instructions. In addition, the short answers to these questions provide a uniform synopsis of the review for the lecturers to determine the usefulness and relevance of the paper for possible inclusion as one of the required readings for subsequent years.

The essential part of the evaluation, however, is the information contained in the reviewer's detailed comments, and it is this section which will be marked by the lecturers. Particular attention should be given to details that guide possible revisions and reasons that other students should or should not be encouraged to read the paper. Even though the paper has already been published you should still include comments on how the paper could be improved as this will force you to look at the paper with a more critical eye.

A recommendation that other members of the class should read the paper should be reserved for papers of the very best quality and of general interest to the class. In general, it is expected that you will include basic concepts and ideas gained from reading the research papers into your answers on the end of year exam. You may choose to include ideas from the papers you have chosen, or from papers recommended to you by classmates. So the final recommendation for the paper is very important.

You should not regard the writing of these reviews as a particularly onerous task. Review lengths of one or two pages will usually suffice. Do not be tempted to plagiarise the paper you are reviewing because you don't really understand it. It will be quite obvious that you have plagiarised, so you are unlikely to get a good mark. Also, the critical evaluation of a paper is at least as important to us as actual understanding. If you don't understand the paper, then include in your review why you didn't understand it - it may be a fundamental flaw of the paper.

Please complete the following review form and hand this in with your detailed review as well as a copy of the paper you reviewed.
Section I. Summary and Recommendation

Summary of Evaluation – Overall quality is:
- Excellent
- Good
- Fair
- Poor

Recommendation:
- Definitely recommend others read
- Could be useful if interested in topic
- Probably not worth the effort
- How in Hell’s name did this get published?

Section II. Overview

Reader Interest
a) Is the paper of current interest to a reasonable segment of the Computer Vision class?
- Yes
- Perhaps
- No

b) Within its particular field of specialization, is the topic of the paper considered important?
- Yes, definitely
- Moderately so
- Not really

Content
a). Is the paper technically sound?
- Yes
- Appears to be, but didn't check completely
- Only partially
- No

b). How would you describe the technical depth of the paper?
- Expert level
- Appropriate for someone working in the field
- Suitable for the non-specialist
- Superficial

c). Does the paper make a tangible contribution to the state-of-the-art in its field?
- Yes, definitely
- To a limited extent
- No

d). Is the bibliography adequate?
- Yes
- Yes, after certain additions and/or deletions (see Section III)
- No

e). To what extent is the material in the paper likely to be used by other researchers and practitioners?
- Large
- Average
- Small

Presentation
a). Is the abstract an appropriate and adequate digest of the work presented?
- Yes
- No

b). Does the introduction clearly state the background and motivation in terms understandable to the non-specialist?
- Yes
- Probably
- No

c). How would you rate the overall organization of the paper?
- Satisfactory
- Could be improved
- Poor

d). Relative to the technical content, is the length of the paper appropriate?
- Yes
- No, should be lengthened
- No, should be shortened

e). Is the English satisfactory?
- Yes
- No

f). How readable is the paper for a computer scientist or engineer who is not a specialist in this particular field?
- Readable with ordinary effort
- Paper is self-contained, but considerable effort is needed
- If the definitions of certain concepts, terms and symbols were included (see Section IV), readability would be improved
- Less than half the paper is readable
- Unreadable

g). Disregarding the technical content, how would you regard the quality of presentation?
- Excellent
- Good
- Fair
- Poor