A survey of computer vision education and text resources

BRUCE A. MAXWELL

Department of Engineering
Swarthmore College, Swarthmore, PA 19081, USA
maxwell@swarthmore.edu
http://www.palantir.swarthmore.edu/~maxwell

This paper provides a survey of the variety of computer vision [CV] and image processing [IP] courses being taught at institutions around the world. This survey shows that, in addition to classic survey courses in CV/IP, there are many focused and multi-disciplinary courses being taught that reportedly improve both student and faculty interest in the topic. The survey also demonstrates that students can successfully undertake a variety of complex lab assignments. In addition to the survey, this paper includes a comparative review of current textbooks and supplemental texts appropriate for CV/IP courses.

Keywords: computer vision, image processing, education, textbook survey

1. Introduction

Ten years ago, an undergraduate computer vision course was not a common feature of many engineering or computer science curricula. A digital image processing course was more common, but generally as an application of signal processing techniques within an EE/CompE program. Today, however, a larger number of institutions offer computer vision courses at the undergraduate level. In some cases the CV courses are offered as a complement or continuation of an IP course, in other cases as a standalone elective.
to be

Along with this increase in image computation-related courses is an attendant increase in new CV/IP educators. When an educator begins to plan a course they must select a text and then plan topics, a syllabus, and lab/programming assignments. Unless they spend a significant amount of time talking with other educators or searching for course web pages, the inspiration for texts, lab assignments, and topics comes largely from their own CV/IP education and interests.

There is a lot of good, successful CV/IP education going on, however, that we can and should draw upon to develop high quality courses. To that end, this paper presents a snapshot of CV/IP education around the world. This includes both a survey of the variety of courses being offered and a review of texts and supplemental materials.

As discovered through the survey--discussed in section 2--many undergraduate CV/IP courses are general survey courses that cover something approaching a standard list of topics. However, education has also begun to follow CV/IP research into interdisciplinary areas such as biological perception, perceptual grouping, and human-computer interfaces. Educators are also experimenting with adding a focus to their courses, modifying the order and list of topics they cover to highlight a particular area or application of CV/IP. These new directions are apparently improving both student and faculty interest.

The remainder of the paper is organized as follows. Section 2 attempts to provide a summary of the trends and new directions in CV education. Hopefully, the specific and summary information will provide inspiration to educators as they try to plan their courses. Section 3 then reviews some of the currently available textbooks and supplemental texts that are appropriate for a CV/IP course. This comparative review is meant to be complementary to the author's previous comparative reviews[11, 12].

2. Survey of computer vision education

The author used two methods to collect information about computer vision courses currently being taught at both the undergraduate and graduate level. The first was to send an email survey to several mailing lists--ACM SIGCSE, 1997 IEEE Workshop on Undergraduate Education and Image Computation--and selected individuals; the second was to locate computer vision course web pages and review syllabi, lab assignments, and reading assignments.

As of this writing, the author has received 18 surveys and reviewed 25 course web pages--including 13 of the 18 who returned a survey--for a total of 30 different courses at 26 different institutions. This section summarize the findings and hopefully provides a starting point for educators developing new courses or revising existing ones.

The survey--which readers are requested to fill out at http://www.palantir.swarthmore.edu/~maxwell/visionCourses.htm--contains twelve questions. They are included here for reference.

(1) Please list the title of the course and give a 1 sentence description.
(2) Please list all textbooks for the course. Also, please indicate whether you used a significant number of papers/handouts in addition to/instead of a textbook.
(3) For each textbook, please rate it on a scale of 1-4 (1: hated it, 2: OK but would like something else, 3: pretty good, 4: would recommend it and use it again).
(4) Would you consider the course a survey course in either image processing or computer vision?
(5) Did you have a focus in your course?
(6) Did you feel that having a focus improved the interest level of the students?
(7) Did you feel that having a focus improved your interest in teaching the course?
(8) How many labs/assignments did you have for the course?
(9) Please give a 1 sentence description of the most successful lab/assignment.
(10) Please give a 1 sentence description of the least successful lab/assignment.
(11) Has any of the students' work during the course resulted in a publication?
(12) If you have a web page for the course, please give the URL below.

For the purposes of discussion, the survey and web search results are integrated for questions 1, 2, 4, 5, 8, and 12 where the answers exist on or can be synthesized from the course web pages. Links to all of the course web pages used in this survey exist at the URL above.

2.1. Questions 1,4

The courses show a wide variety of focuses and topics. They group roughly into five categories: classic IP, classic CV, application oriented, focused CV courses, and high-level vision. The classic IP and CV courses--which constitute 12 of the 28 courses-- present a survey of topics designed to provide breadth of knowledge in computer vision/image processing. Variation in these courses appears primarily in the lab work and assignments that vary in complexity and scope. Most of these are intended for upper-level undergraduates.

The application oriented courses--either self-described or inferred from the web pages--focus on functioning vision systems and labs that apply vision to existing problems. The implied trade-off is that more theoretical topics in CV are covered in less depth if at all. Three of the courses identified themselves in this category. The specific applications varied from mechanical and production engineering tasks to optical character recognition to detection of calcification in mammograms.

The focused CV courses, of which there were six, present the CV material as it applies to a specific focus task. Examples of the different foci include: object recognition [OR], motion analysis and content-based image retrieval [CBIR], robot perception, and human detection and analysis. The author's CV course falls in this category. In the author's course the material covered in lectures does not, in general, differ greatly in scope from a classic CV course. However, the order of presentation is somewhat different, and there is an attempt to focus on applying the CV techniques to the selected focus. In the author's course, for example, color vision is covered more in-depth and each of the three main labs had students implement a different object recognition system. The two OR focused courses in this section included more AI/machine learning topics than the classic CV courses, the robot perception course focused on extracting 3D information and models, and the motion/CBIR course included significantly more optical flow analysis and grouping and segmentation when compared to classic CV courses. Both the applications-oriented and focused courses tended to be upper-level undergraduate and/or first-year graduate courses.

Finally, the high-level vision courses focused on advanced topics in CV. There were nine in this category, most of them graduate courses. Five focused on current research topics--monitoring, 3D imaging, HCI systems, image-based modeling and rendering, and vision for computer graphics--three focused on perceptual grouping for segmentation and OR, and the seventh course focused on a combination of biological and computer perception. Most of these are good examples of hybrid courses that are merging computer vision with computer graphics, cognitive science, biology, and psychology. This seems to be a new direction in CV education (reflecting an existing trend in research), and one that is hopefully the harbinger of more interdisciplinary CV curricula.

2.2. Questions 2-3

The 28 courses used eight different textbooks, not including paper based courses, with *Image Processing Analysis and Machine Vision* by Sonka *et. al*.[14] being the most common (7 of 27). Everyone using the book gave it the highest rating (4), indicating that they would recommend it to other educators. *Computer Vision* by Stockman & Shapiro[15] was the only other text to receive the highest rating, but it has only been used in two courses so far.

The other books to receive ratings were--listed in order of their average rating--*Digital Image Processing* by Gonzalez & Woods[5], *Digital Image Processing* by Castleman[3], *A*

Guided Tour of Computer Vision by V. Nalwa, *Machine Vision* by Jain *et. al*[10], and *Introductory Techniques for 3-D Computer Vision* by Trucco & Verri[16]. Note that the image processing books were usually listed as supplementary materials. Horn's *Robot Vision*[9], in use by only one survey respondent, received a rating of 4, but was supplemented with a significant number of handouts.

2.3. Questions 5-7

While not a particularly scientific set of questions, two things were common among the answers. First, having a focus improved interest in the course for both students and faculty. Second, it increased the amount of work required for the faculty member in order to develop lectures, labs, and student reference materials. Hopefully, in the future this paper and other CV/IP educational resources can improve the availability of educational materials and help faculty to focus their courses without as large an overhead.

2.4. Questions 8-10

The number of labs in a CV/IP course ranged from 1 to 11, with the average being 3-4. The most successful labs reported by survey respondents were: Hough transform detection of circles in noisy images, detection of calcification in mammograms, face recognition using eigenfaces, face detection, visual combination lock (gesture recognition), quantifying blood flow from images, character enhancement for optical character recognition [OCR], image segmentation (two courses), building a 3D model from a single image, and image mosaics. From an educator's point of view, it is nice to see such a wide variety of successful projects. All of these labs appear to have two things in common. First, they deal with real-world images and problems, and second, they have immediate and intuitive results. Note that success in the task is not necessarily a requirement--but obviously helps retain student interest. At least one respondent indicated that the method used in the lab did not work extremely well, but the students enjoyed the lab anyway and learned a significant amount.

Labs that reportedly did not work as well included: object recognition [OR] by alignment, tracking soil grains in consecutive images, segmentation with histogram thresholding, connected component extraction, and face recognition using eigenface and Gaussian pyramids to allow multi-scale recognition. Two out of four of these--the face recognition and OR by alignment--reportedly did not work as well because of the complexity of the task, debugging time, and the number of components required to get the system to work. Poor results are also not helpful in motivating student interest, as noted by other respondents to this question.

In the author's experience with OR by alignment, it is important to pick a data set that is commensurate with the amount of time the students have to deal with the task. In other words, the educator must minimize the number of modules required for successful implementation of the main method. In the case of OR by alignment, this means that foreground/background separation and feature finding should be extremely simple to implement so that students can focus on the alignment and classification issues. Even with a mildly complex data set--multi-colored 3D objects on a black background--the segmentation and feature location issues dominated the problem and reduced the amount of time students spent on the OR issues such as model development and matching techniques.

By contrast, the data set the author generated for a lab on face recognition using eigenfaces was extremely well-behaved (all the faces were approximately the same size and centered in the image), did not require any preprocessing, and allowed students to focus on the details of eigenspace recognition. It was by far the students' favorite lab.

2.5. Question 11

As perhaps would be expected, few of the undergraduate courses produced published work, while most of the graduate courses did. However, the important point here is that at least two of the undergraduate courses did produce publishable quality work, albeit with extra time put in by the students after the course had ended. Clearly this will depend upon factors such as undergraduate student quality and preparation, but it is not the case that CV education and research are independent of one another.

2.6. Question 12

Almost all of the courses included in this survey have web pages providing the syllabus, lab assignments, readings, and possibly lectures notes for the course. The following web page links all of these courses together in the hopes that it will be a useful educational resource.

URL: http://www.palantir.swarthmore.edu/~maxwell/visionCourses.htm

The author, for example, has a complete set of 35 lectures in a single Adobe PDF document that are freely available for use by other CV/IP educators.

2.7. Summary of survey results

Hopefully this review of computer vision education will provide both inspiration and resources for computer vision/image processing educators. There are a wide variety of courses being offered and, while the most common is still a CV/IP survey course, interdisciplinary hybrids, application oriented courses, and focused CV courses exist that can be templates for curriculum development. In particular, the author urges educators to look at lab and homework assignments at other institutions to see the richness and complexity of the assigned tasks. What the survey and the author's research clearly demonstrate is that our students are capable of undertaking interesting and complex projects so long as appropriate tools and data sets are available for them to use.

3. Educational resources

One of the most important educational resources for both educators and students is a good textbook. It provides a framework for tying the material together, a reference and guide for complex material, and a source of both paper and hands-on exercises and study problems. Two of the ways you can tell the maturity of a field are A) how many textbooks are there, and B) how many of them are past their first edition. CV/IP has apparently reached middle age since there are now at least four current comprehensive CV texts--comprehensive IP texts have been around a bit longer--and at least one of them is a significantly revised second edition (Sonka *et. al.*).

In addition, there are a number of supplemental texts that may be appropriate for undergraduate or graduate vision and image computation courses. These supplemental texts are focused on a particular aspect of CV/IP and provide more depth in a particular area than is available in any other single source. Educators and students may find it useful either to have one of these books available as a strongly suggested text or to have one or more of them on reserve reading.

Finally, there are a class of applications-oriented texts that may be appropriate for CV/IP courses aimed at students who are not necessarily engineers or computer scientists. These texts are aimed at practitioners--potentially in the art and computer graphics realm--who need to use CV/IP tools, but do not need as in-depth an understanding of the algorithms and mathematics involved.

Based on these divisions, this section consists of four parts. Part one looks at the two most recent comprehensive computer vision texts: Sonka *et. al.*, and Stockman & Shapiro (based on their internet distribution). Part two briefly covers two standard IP texts: Gonzalez & Woods and Castleman. Part three looks at two of the application-oriented books: Baxes and Umbaugh. Finally, part four reviews two potential supplemental texts: Ullman and Fairchild.

For a comparative review of the more classic computer vision texts--*Robot Vision* by B. K. P. Horn, *Computer and Robot Vision* by R. Haralick, and L. Shapiro, *Machine Vision* by R. Jain, R. Kasturi, and B. Schunck, *Machine Vision* by E. R. Davies, and *A Guided Tour of Computer Vision* by V. S. Nalwa--please refer to Maxwell '98[11].

3.1. Comprehensive Computer Vision Textbooks

3.1.1. Image Processing, Analysis, and Machine Vision, 2nd ed., M. Sonka, V. Hlavac, and R. Boyle, 1998

This textbook is a comprehensive computer vision and image processing text. Its depth in some topics approaches that of the well-known *Computer and Robot Vision* by Haralick & Shapiro[8], but with an algorithmic and conceptual focus rather than a rigorous mathematical one. Like many other texts, *Image Processing, Analysis, & Machine Vision* begins with filtering, edge-finding, segmentation, and 2-D shape representation. The text then looks at various approaches to object recognition, including artificial neural networks, graph matching, and fuzzy logic. One of the strengths of this book is that it provides enough background on these topics for students to follow along. However, having an artificial intelligence, neural networks, or fuzzy logic course as prerequisite would improve students' ability to focus on the computer vision applications rather than the tools.

After object recognition, *Image Processing, Analysis, & Machine Vision* moves on to 3D vision including chapters on image understanding, 3D vision (calibration, stereo, and physics-based vision), and motion analysis. The book does not follow the same order as other texts, however, and later chapters include mathematical morphology, linear discrete transforms, and image compression, chapters that are often placed earlier in an image processing text. These chapters are almost as comprehensive on these subjects as the texts in the IP section, which makes this book appropriate for a CV, IP or hybrid course. This arrangement of topics is convenient for an undergraduate instructor, because these later chapters contain much of the material that is challenging for non-mathematically inclined students. Their placement later in the book, and the resulting implication that previous chapters do not depend on them for understanding, means an instructor can more easily pick and choose which of these topics to cover. Thus, they would be covered in a course focused on IP topics and possibly left out in a more CV-focused course.

The real strength of *Image Processing, Analysis, & Machine Vision* is its comprehensive, in-depth coverage of both CV and IP. The material is well-written and, while the mathematical formulation of methods is still the focus of the narrative, the text includes algorithms for many of the methods it covers. The text also seems to contain more example images and image comparisons than other texts, making it easier to obtain an intuitive understanding of the material. Its approach relies on a mixture of EE and CS concepts, and there is even a chapter on data structures for image analysis. Thus, with intelligent topic selection and adequate instruction this text is appropriate for students with a variety of backgrounds.

The book is written at a higher level than the other texts in this review, however, and instructors should carefully consider their prerequisites for a computer vision or image processing course before selecting this text. Based on its level and breadth, *Image Processing, Analysis, & Machine Vision* will be more accessible to students who have already had an artificial intelligence course, or related EE course that introduces them to artificial neural networks, search, and fuzzy logic. Students without this background could be overwhelmed with a range of new concepts that, while they are useful in computer vision, are not specific to the field.

One departure from other textbooks worth noting is that some methods that receive central coverage in other texts are de-emphasized, not mentioned, or placed in different contexts in *Image Processing, Analysis, & Machine Vision*. For example, the Hough transform is covered in the chapter on segmentation rather than within the chapter covering edge and line-finding. While line-finding is presented as an application of the Hough transform, this presentation comes after the line-finding chapter and in the middle of a discussion of segmentation. These issues are not a major drawback to the text, but an instructor should be aware of them and should structure the lectures and reading assignments appropriately.

Overall, *Image Processing, Analysis, & Machine Vision* is comprehensive CV/IP text and a good compromise with respect to level and target audience. Through strategic selection of chapters, it would be appropriate for either a senior level CS or EE undergraduate computer vision course, or an advanced graduate level course.

3.1.2. Computer Vision, G. Stockman & L. Shapiro, 2001

The book presents a nice complement to *Image Processing, Analysis and Machine Vision* [IPAMV]. As the difference in names implies, *Computer Vision* is not appropriate as an image processing textbook. It contains sufficient information on image processing to implement computer vision algorithms, but the focus of the book is on image analysis and high-level vision. The result is that the combination of IPAMV and *Computer Vision* cover the spectrum from intensive image processing and manipulation to high level analysis, object recognition and content-based image retrieval.

Computer Vision contains sixteen chapters that fall into roughly four categories: overview, 2-D CV topics, 3D CV topics, and special CV topics. Since it was written with the intent of reaching a broader audience than IPAMV, this book is appropriate as a primary text or reference for a wider variety of courses. For example, it would be appropriate for courses ranging from an introduction to imaging for non-scientists to a sophomore-junior elective to a first-year graduate seminar.

The overview chapters (chapters 1-4) include a summary of problems in CV, imaging and image representations, simple binary image analysis and a survey of pattern recognition concepts. The 2-D processing topics (chapters 3, 5-7, and 11) include thresholding and binary image analysis, filtering and enhancement, edge detection, Fourier Transforms, color, texture, segmentation, and 2-D matching and pose calculation. The 3-D computer vision topics (chapters 9-10, and 12-14) include motion detection and analysis, range image analysis, stereo, calibration, intrinsic image analysis and line labeling, shape from X, and camera models. The special topics (chapters 6-8, 15-16) include color and shading, texture, content-based retrieval, virtual reality, and a set of case studies of CV systems. Different combinations of these are appropriate for different types of courses.

In comparison with other texts, the coverage of color and shading in *Computer Vision* is the best available without consulting a color reference such as Fairchild's *Color Appearance Models* (described below). However, it still does not contain adequate coverage of physical models of reflection or color appearance. The texture chapter is comparable to Sonka *et. al.*, and the CBIR and VR chapters are unique. It is these latter two areas that give *Computer Vision* a nice high-level flavor and provides a reference for these growing areas of CV.

Like IPAMV, *Computer Vision* contains a large number of example images, diagrams, and algorithms. The writing is clear and the mathematics--when it is necessary to present it--is complete and accessible. Since the book is designed with multiple audiences in mind, the heavy mathematical sections are flagged and the book can be used effectively with or without them.

Of particular interest to CV practitioners and students dealing with issues of calibration, chapter 13 contains a nice description of Roger Tsai's camera calibration algorithm, complete with an example. Note that Trucco and Verri (see below) also cover Tsai's calibration algorithm.

Overall, the choice between *Computer Vision* and IPAMV should be based on personal preference, the focus of your course, and the background of your students. IPAMV will be more accessible to engineers and contains more in-depth coverage of image processing techniques. *Computer Vision* is more accessible to computer scientists and covers a number of higher-level aspects of CV that are either not covered or briefly covered in IPAMV. In a number of areas--texture, stereo, motion, calibration, and segmentation--the two books are quite similar and the differences are mainly in style and emphasis.

3.1.3. Introductory Techniques for 3-D Computer Vision, E. Trucco and A. Verri, 1998

Whether or not you feel the “introductory” adjective applies to this book, it is largely about 3-D computer vision. According to the authors, their intention in writing it was to make a text that addressed the fundamental problems of computer vision in a way that would age well. One interesting parallel here is that E. R. Davies’ *Machine Vision* was written with exactly the same goal in mind[4, 11]. The curious thing is that this led Davies to write about 2-D computer vision with a focus on the Hough transform and only one or two chapters on 3-D issues. Conversely, this same intention led Trucco and Verri to write about 3-D computer vision, with a chapter or two on 2-D issues.

All of the authors are probably correct in their assessment of the state of the field. 3-D computer vision is arguably less integrated into industrial situations, but there is a significant amount of mathematics underlying the field that appears at this juncture to be fundamental.

The topics in Trucco and Verri fall into three categories: 2-D processing, 3-D structure, and object recognition. The section on 2-D processing encompasses four chapters: introduction to imaging and optics, noise and filtering, edges and features, and line and curve detection. The only edge detector covered in detail is the Canny edge detector; the Sobel and Roberts edge filters are covered in a single page.

The section on 3-D structure is the strength of the book and includes calibration, stereopsis, motion, and shape-from-X. The calibration section develops the extrinsic and intrinsic parameters and walks through two different calibration methods (Tsai[17] and Faugeras[7]). Stereopsis covers both area-based and feature-based stereo, but in a terse manner. The motion chapter mentions most of the major approaches to optical flow and structure-from-motion, but focuses on those that use linear algebra techniques, in particular singular-value decomposition [SVD].

Note that there is a clear bias on the part of the authors towards methods that use overconstrained sets of linear equations. While this does make many of the problems tractable for teaching purposes, you the educator will have to decide whether to introduce other types of techniques or approaches to the problem.

Finally, the shape-from-X chapter looks at shape from shading and shape from texture. Unfortunately, the shape-from-texture only looks at the case of regular geometric decal textures, and does not consider other kinds.

The section on object recognition includes two chapters: an image-based recognition chapter, and a range-based recognition chapter. The image-based recognition chapter does include a nice development of the eigenimage method of recognition (note the use of overconstrained linear equations and SVD), but as with a number of other topics the coverage and development is terse. The range-based recognition chapter is unique among current texts, and sets Trucco and Verri apart from the others if you need to cover range-based material.

The book ends with a nice appendix that covers a number of numerical analysis techniques, including experimental design, numerical differentiation, the sampling theorem, projective geometry, differential geometry, SVD, robust estimators, Kalman filtering, and 3-D rotations. Do not expect students to be able to read and understand these sections as they are lacking in detailed description. However, as a refresher or reference it works reasonably well.

Overall, the text is math-heavy, but someone with a linear algebra background should be able to deal with it, so long as they understand SVD. The authors assume you have access to a numerical programming suite.

The weakness of this text is that segmentation, color, and texture are not dealt with in any significant way. The text does not mention content-based retrieval or deal with pattern recognition concepts (which the authors declare in the beginning is not what the book is about). Finally, there are no case studies, as in the other two comprehensive texts compared above.

The strength of this text is the accessible mathematics and the coverage of fundamental 3-D topics. The book also has of the more accessible descriptions of the use of a Kalman filter in image-based tracking.

If the focus of your course is on obtaining 3-D structure from images, this may be the textbook to consider. You should definitely read a chapter or two before making a decision, however, because of the style of the writing. If you are teaching a survey course or one focused on object recognition, for example, this text does not have sufficient coverage of a number of fundamental topics and would have to be heavily supplemented by readings from other texts and articles.

3.2. Image Processing

3.2.1. Digital Image Processing, R. Gonzalez and R. Woods, 1992

Gonzalez & Woods' book is a standard textbook for an image processing (IP) course for engineers. It has also been used by educators as a supplemental text for CV courses (see section[2]). Like all of the texts in this section, it covers digital image fundamentals, image transforms, image enhancement, image restoration, image compression, and image segmentation. Gonzalez & Woods (G&W) also provide a chapter on shape & region descriptors and a chapter on low-level recognition. The coverage of image transforms focuses on the Fourier transform, but also includes short sections on the Walsh, Hadamard, discrete cosine, Haar, slant, and Hotelling transforms. Because of the age of the book, however, wavelet transforms are not mentioned. (A new edition may be forthcoming in 2000-2001).

With respect to color image processing, G&W provide a 25 page section on the topic, including color plates that aid the presentation of the material. Castleman's book also provides a chapter on color and multi-spectral images.

3.2.2. Digital Image Processing, 2nd Ed. K. Castleman, 1996

Castleman's book attempts to reach a broader audience than G&W, but is still most appropriate for an image processing course aimed at engineers or computer scientists with a strong linear systems background. Unlike G&W, who divide their book based on the major IP topics, Castleman divides the book into three parts based on complexity and overall theme. Part I presents basic IP concepts that do not rely heavily on mathematics. Part II covers IP techniques based on mathematical tools such as the Fourier and wavelet transforms, and the part III touches on computer vision topics and applications of IP.

The strength of Castleman's book is its extended coverage of both linear and discrete transforms, as well as the addition of wavelets. Part III also sets Castleman apart from G&W and the two other IP texts in this section. It contains chapters on segmentation, edge detection, line-finding, binary processing, texture and shape analysis, pattern matching, color image processing, and 3D imaging, including a short section on stereo. Because of this extended coverage, it could be an appropriate text for a two-semester EE course on image processing and computer vision. However, it is important to note that, when compared to other computer vision or comprehensive CVIP texts, it does not have the same breadth of CV topics or depth within each topic.

3.3. Applications oriented textbooks & supplements

3.3.1. Digital Image Processing: Principles and Applications, G. Baxes, 1994

Unlike G&W and Castleman's texts, Baxes' text, like the following one by Umbaugh, is not aimed at engineers and students with a strong mathematical background. Instead, Baxes' and Umbaugh's texts address the needs of application programmers and people who need to understand and apply digital image processing techniques. From an educators point of view, these texts would be appropriate for a general education course on digital image processing and visual information systems. They would also be appropriate for an IP course for computer scientists who may, or may not have an adequate background to delve into linear and discrete transforms. Given the existence of digital cameras, cheap scanners, and the explosion of digital images available on the world-wide web, such a course is going to play an important role in the future of CVIP education.

Baxes divides his book into four parts. Part I presents an introduction to the field. Part II covers image enhancement and restoration, segmentation, feature extraction, simple object classification, image compression, and image synthesis. Part III looks at image processing systems, video formats, and image data handling, and part IV is a long list of image processing examples, demonstrating most of the concepts covered in the book. As a supplement to the text, the book contains a disk with implementations of a number of the algorithms.

The strength of this text, compared to Umbaugh, is the broader coverage of system topics. Both texts cover image compression, restoration, analysis, and segmentation. Baxes, however, also includes the visual information system topics and a chapter on image synthesis. The latter is particularly relevant to applications programmers and computer science students, and is unique to this book compared to any of the others in this review.

Where an educator may have to supplement the text are topics such as segmentation, edge and line finding, and filtering, which are not covered in detail.

3.3.2. Computer Vision and Image Processing: A Practical Approach Using CVIPtools, S. Umbaugh, 1997

Umbaugh's book is the most limited in scope of the IP textbooks. The title is, unfortunately, somewhat misleading as this text is definitely focused on IP. The organization of the first section of this book is similar to G&W, with chapters on IP fundamentals, image analysis, restoration, enhancement, and compression. The second section is a reference for the CVIP code library and CVIP applications provided on a CD with the book.

The target audience for this text falls somewhere between Baxes and G&W, as it presents most concepts conceptually and algorithmically, while still providing more mathematical detail than Baxes. Like Baxes, this text would be appropriate for a computer science IP course, where not all students would be comfortable with a mathematical presentation of linear transforms.

The strength of Umbaugh's text is the CVIP library and applications provided with the book. In the author's experience it is the most extensive library of routines currently available with a textbook. It also has the most extensive documentation, which both students and educators should appreciate.

3.3.3. Algorithms for Image Processing and Computer Vision, J. R. Parker, 1997

Parker's book takes the reader close to the pixels. It focuses on implementations of a variety of CV/IP tools, provides lots of source code (in C) both in the book and on a CD, and encourages the reader to explore the world of image manipulation on their own. There are ten chapters: edge-detection, digital morphology, gray-level segmentation, texture, skeletonization, image restoration, wavelets, optical character recognition, symbol recognition, and evolutionary computing. In most cases, each chapter focuses on the theory and function of one or two techniques rather than providing a survey of the topic. This limited scope allows the description of each technique to be reasonably detailed. Whether or not you agree with the chosen methods, they are well explained.

While much of the book would be considered image processing, there are a number of topics--texture analysis, segmentation, edge detection, and symbol recognition--that would also be taught in a traditional CV course. The book is intended for someone with a calculus background--but walks a fine line between intuition and mathematical rigor. As expressly pointed out by the author (page 258[13]), when the two are in contention, intuition wins. The result is a readable introduction to a number of fairly advanced topics--wavelet transforms, filtering and image restoration, texture analysis, and edge detection--that would be appropriate for students without an engineering-style background.

The focus on low-level tools and a strong connection to pixels makes Parker's book a potentially useful supplement to a more theoretical textbook. It would also be an appropriate supplemental text for a robotics, or AI course that emphasized visual computing. The included source code would let an educator focus more on what the tools can do than on the tool development (which is usually the focus on an IP/CV course). If nothing else, Parker's book is a good one to have on reserve in the library. If you happen to be using one of the algorithms in the book, it may help students to get through the details of it more quickly and get something working.

3.4. Supplemental texts

3.4.1. High-level Vision: Object Recognition and Visual Cognition, S. Ullman, 1996

If the focus of your CV course is on object recognition--not an uncommon occurrence given the survey results--High-level Vision may be an appropriate supplemental text. The focus of this book is on object recognition by alignment but it covers a wider range of topics that nicely complement the CV texts.

The first two chapters of the book are a well-written introduction to object recognition. Chapter 2, in particular, summarizes the major approaches to recognition and provides a gentle introduction to the remainder of the book.

The primary OR by alignment section (chapters 3-5) covers alignment of rigid 3-D objects using affine transformation estimation, smooth boundary alignment, and view combination. These are the technical chapters of the book, but are written in an accessible manner for upper-class undergraduates. Detailed proofs and algorithms are reserved for the appendices, which contain sufficient detail to implement alignment by affine transformation estimation.

The remainder of the book looks at more general issues in OR and relates them to biological cognitive processing. These topics include segmentation, saliency and region selection, classification, correspondence, and visual cognition. These chapters contain less technical detail and are more an exploration of the major issues in each area. For a seminar or course with a small number of students, they provide a good basis for discussion of fundamental issues in OR.

Overall, if you want to focus your course on object recognition, *High-level Vision* will provide more in-depth coverage of OR by alignment and a stronger connection to biological processing than any currently available text. As a supplemental text or reserve reading it nicely complements other sources in a well-written and approachable manner.

3.4.2. Visual Information Retrieval, A. Del Bimbo, 1999.

If your course is focused on content-based image retrieval, this would be an appropriate supplemental text. It would also be possible to use this book as a primary text for a graduate course on the subject, so long as the it was supplemented with primary source material.

The book contains six chapters. The first is a decent introduction to the field, touching on interactivity, representation, similarity models, indexing methods, data structures, and performance evaluation. Chapters 2 through 5 cover each of the major methods of visual information retrieval: color, texture, shape, and spatial relationships. The final chapter looks at video retrieval with a focus on transition detection, video segmentation, annotation, and browsing.

Overall, Visual Information Retrieval is closer to an annotated bibliography--albeit a well-written one--than a textbook or traditional monograph on a specific field. Hence the comment above about supplementing it with primary source material. The chapters contain numerous sub-headings, and each topic gets only a few pages.

It does provide a useful snap-shot of the field at the time of publication, however, and is pretty good as a reference. Since the author is focused on the applications of vision techniques to CBIR, however, some of the chapter bibliographies are missing papers that would be relevant, but have not been applied directly to CBIR.

3.4.3. Color Appearance Models, M. Fairchild, 1998

One of the resources most lacking in current CV resources is comprehensive coverage of color vision. Fairchild's book provides a comprehensive discussion of what color means in terms of human perception, and how we can develop computational models that predict color perception. These models are relevant to color computer vision because they provide clues as to what characteristics of color are important from a biological standpoint--which is our working model.

The book divides into four sections: definitions and terminology, color phenomena, color appearance models, and applications or colorimetry. The introductory section (chapters 1-5) give an overview of the human visual system, psycho-physics, colorimetry, color appearance terminology, and color-order systems. These topics are directly relevant to any CV/IP course with a focus on color vision and will help students to think more clearly and precisely about color.

The second section (chapters 6-8) looks at well-known color phenomena in psychophysical experiments. While perhaps not directly relevant to a machine vision system, it gives the reader a better understanding of some of the issues in color perception.

The third section (chapters 9-15) looks at a variety of computational color appearance models that predict human responses to various stimuli. Chapter 10 is particularly relevant to color computer vision and content-based image retrieval as it presents the CIELAB and CIELUV color spaces that attempt to match color distances with perceptual distances. All of the color appearance models are presented in sufficient mathematical detail that it is possible to implement them.

The final section of the book (chapters 16-18) is a summary of colorimetric applications and a discussion of device-independent color imaging. The latter topic is essential for any color machine vision person to know, since color depends so heavily on device characteristics.

Overall, like *High-level Vision*, *Color Appearance Models* is a useful supplemental text or reserve reading. It is well-written and approachable, especially the introductory and phenomena sections of the book, and there is no better resource for color science currently available.

4. Summary

Overall, the text resources available for computer vision educators are improving. In addition to two high quality main textbooks, there are now a number of useful supplemental texts that should enable educators to fine-tune their reading list. For a quick summary of the textbooks, see Table 1.

In addition, the web resources available to educators are growing. Lab assignments, syllabi, reading lists, and lectures notes posted on the web create a wealth of inspiration and ideas when trying to decide what labs or readings to use during the upcoming semester. Even if you don’t actually use anything you find, looking over the variety that exists will help you to come up with tasks that are both fun and challenging for your students.

Finally, in an effort to continue to improve the distribution of knowledge about CV education, the author encourages all CV/IP educators to submit a brief survey on their course(s) at the following web address.

http://www.palantir.swarthmore.edu/~maxwell/visionCourses.htm

Table 1. Summary of the Computer Vision and Image Processing TextBooks

Author(s)	Sonka, Hlavac, and Boyle	Stockman & Shapiro	Trucco & Verri	Gonzalez & Woods	Castleman
Coverage	Computer Vision: Broad	Computer Vision: Broad	Computer Vision: Focus on 3-D Techniques	Image Processing: Broad	Image Processing: Broad (some CV)
Audience	Jr/Sr undergrad or graduate	Undergrad or 1st-year grad	Jr/Sr undergrad	Undergrad	Undergrad
Suggested background	EE, ok for CS	CS, ok for EE	CS or EE with good linear algebra	EE	EE
Strengths	Case studies, depth in most areas, and AI algorithms	Good examples, covers CBIR, good stereo section	Depth in 3-D analysis of images	Well-written and broad coverage	Includes an introduction to computer vision

5. References

1. G. Baxes, *Digital Image Processing*, John Wiley & Sons, New York, NY, 1994.
2. A. Del Bimbo, *Visual Information Retrieval*, Morgan Kaufmann Publishers, Inc, San Francisco, CA, 1999.
3. K. Castleman, *Digital Image Processing*, Prentice-Hall, Upper Saddle River, NJ, 1996.
4. Davies, E. R., *Machine Vision: Theory, Algorithms, Practicalities, 2nd Ed*., Academic Press, San Diego, CA, 1997.
5. R. Gonzalez and R. Woods, *Digital Image Processing*, Addison-Wesley, Reading, MA, 1992.
6. M. Fairchild, *Color Appearance Models*, Addison Wesley Longman Inc., Reading MA, 1998.
7. O. D. Faugeras, *Three-Dimensional Computer Vision: A Geometric Viewpoint*, MIT Press, Cambridge (MA), 1993.
8. R. Haralick and L. Shapiro, *Computer and Robot Vision*, Addison-Wesley, Reading, MA, 1992.
9. B. K. P. Horn, *Robot Vision*, MIT Press, Cambridge, MA, 1985.
10. R. Jain, R. Kasturi, and B. G. Schunck, *Machine Vision*, McGraw-Hill, New York, NY, 1995.
11. B. A. Maxwell, “Teaching computer vision to computer scientists: issues and a comparative textbook review”, in *Int l Journal of Pattern Recognition and Artificial Intelligence*, Vol. 12, No. 8, pp. 1035-1051, August 1998.
12. B. A. Maxwell “Comparative review of image processing and computer vision textbooks”, in *Proceedings of SPIE Medical Imaging ’98,* February, 1998.
13. J. R. Parker, *Algorithms for Image Processing and Computer Vision*, John Wiley & Sons, Inc, New York, 1997.
14. M. Sonka, V. Hlavac, and R. Boyle, *Image Processing, Analysis, and Machine Vision*, PWS Publishing, Pacific Grove, 1999.
15. G. Stockman and L. Shapiro, *Computer Vision*, Prentice-Hall, Upper Saddle River, NJ, in press.

16. E. Trucco, A. Verri, *Introductory Techniques for 3-D Computer*, Prentice-Hall, Upper Saddle River, NJ, 1998.
17. R. Y. Tsai, “A Versatile Camera Calibration Technique for High Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses”, in *IEEE Trans. on Robotics and Automation*, Vol RA-3 (4), pp. 323-344, 1987.
18. S. Ullman, *High-level Vision: Object Recognition and Visual Cognition*, MIT Press, Cambridge, MA, 1996.
19. S. Umbaugh, *Computer Vision and Image Processing: A Practical Approach Using CVIPtools*, Prentice-Hall, Upper Saddle River, NJ, 1998.