Computer vision

Computer vision is an interdisciplinary scientific field that deals with how computers can be made to gain high-level understanding from digital images or videos. From the perspective of engineering, it seeks to automate tasks that the human visual system can do.[1][2][3]

Computer vision tasks include methods for acquiring, processing, analyzing and understanding digital images, and extraction of high-dimensional data from the real world in order to produce numerical or symbolic information, *e.g.*, in the forms of decisions.[4][5][6][7] Understanding in this context means the transformation of visual images (the input of the retina) into descriptions of the world that can interface with other thought processes and elicit appropriate action. This image understanding can be seen as the disentangling of symbolic information from image data using models constructed with the aid of geometry, physics, statistics, and learning theory.[8]

As a scientific discipline, computer vision is concerned with the theory behind artificial systems that extract information from images. The image data can take many forms, such as video sequences, views from multiple cameras, or multi-dimensional data from a medical scanner. As a technological discipline, computer vision seeks to apply its theories and models for the construction of computer vision systems.

Sub-domains of computer vision include scene reconstruction, event detection, video tracking, object recognition, 3D pose estimation, learning, indexing, motion estimation, and image restoration.[6]

Contents

Definition

Computer vision is an interdisciplinary field that deals with how computers can be made to gain high-level understanding from digital images or videos. From the perspective of engineering, it seeks to automate tasks that the human visual system can do.[1][2][3] "Computer vision is concerned with the automatic extraction, analysis and understanding of useful information from a single image or a sequence of images. It involves the development of a theoretical and algorithmic basis to achieve automatic visual understanding."[9] As a scientific discipline, computer vision is concerned with the theory behind artificial systems that extract information from images. The image data can take many forms, such as video sequences, views from multiple cameras, or multi-dimensional data from a medical scanner.[10] As a technological discipline, computer vision seeks to apply its theories and models for the construction of computer vision systems.

History

In the late 1960s, computer vision began at universities that were pioneering artificial intelligence. It was meant to mimic the human visual system, as a stepping stone to endowing robots with intelligent behavior.[11] In 1966, it was believed that this could be achieved through a summer project, by attaching a camera to a computer and having it "describe what it saw".[12][13]

What distinguished computer vision from the prevalent field of digital image processing at that time was a desire to extract three-dimensional structure from images with the goal of achieving full scene understanding. Studies in the 1970s formed the early foundations for many of the computer vision algorithms that exist today, including extraction of edges from images, labeling of lines, non-polyhedral and polyhedral modeling, representation of objects as interconnections of smaller structures, optical flow, and motion estimation.[11]

The next decade saw studies based on more rigorous mathematical analysis and quantitative aspects of computer vision. These include the concept of scale-space, the inference of shape from various cues such as shading, texture and focus, and contour models known as snakes. Researchers also realized that many of these mathematical concepts could be treated within the same optimization framework as regularization and Markov random fields.[14] By the 1990s, some of the previous research topics became more active than the others. Research in projective 3-D reconstructions led to better understanding of camera calibration. With the advent of optimization methods for camera calibration, it was realized that a lot of the ideas were already explored in bundle adjustment theory from the field of photogrammetry. This led to methods for sparse 3-D reconstructions of scenes from multiple images. Progress was made on the dense stereo correspondence problem and further multi-view stereo techniques. At the same time, variations of graph cut were used to solve image segmentation. This decade also marked the first time statistical learning techniques were used in practice to recognize faces in images (see Eigenface). Toward the end of the 1990s, a significant change came about with the increased interaction between the fields of computer graphics and computer vision. This included image-based rendering, image morphing, view interpolation, panoramic image stitching and early light-field rendering.[11]

Recent work has seen the resurgence of feature-based methods, used in conjunction with machine learning techniques and complex optimization frameworks.[15][16]

Related fields

Artificial intelligence

Areas of artificial intelligence deal with autonomous planning or deliberation for robotical systems to navigate through an environment. A detailed understanding of these environments is required to navigate through them. Information about the environment could be provided by a computer vision system, acting as a vision sensor and providing high-level information about the environment and the robot.

Artificial intelligence and computer vision share other topics such as pattern recognition and learning techniques. Consequently, computer vision is sometimes seen as a part of the artificial intelligence field or the computer science field in general.

Information engineering

Computer vision is often considered to be part of information engineering.[17][18]

Solid-state physics

Solid-state physics is another field that is closely related to computer vision. Most computer vision systems rely on image sensors, which detect electromagnetic radiation, which is typically in the form of either visible or infra-red light. The sensors are designed using quantum physics. The process by which light interacts with surfaces is explained using physics. Physics explains the behavior of optics which are a core part of most imaging systems. Sophisticated image sensors even require quantum mechanics to provide a complete understanding of the image formation process.[11] Also, various measurement problems in physics can be addressed using computer vision, for example motion in fluids.

Neurobiology

A third field which plays an important role is neurobiology, specifically the study of the biological vision system. Over the last century, there has been an extensive study of eyes, neurons, and the brain structures devoted to processing of visual stimuli in both humans and various animals. This has led to a coarse, yet complicated, description of how "real" vision systems operate in order to solve certain vision-related tasks. These results have led to a subfield within computer vision where artificial systems are designed to mimic the processing and behavior of biological systems, at different levels of complexity. Also, some of the learning-based methods developed within computer vision (*e.g.* neural net and deep learning based image and feature analysis and classification) have their background in biology.

Some strands of computer vision research are closely related to the study of biological vision – indeed, just as many strands of AI research are closely tied with research into human consciousness, and the use of stored knowledge to interpret, integrate and utilize visual information. The field of biological vision studies and models the physiological processes behind visual perception in humans and other animals. Computer vision, on the other hand, studies and describes the processes implemented in software and hardware behind artificial vision systems. Interdisciplinary exchange between biological and computer vision has proven fruitful for both fields.[19]

Signal processing

Yet another field related to computer vision is signal processing. Many methods for processing of one-variable signals, typically temporal signals, can be extended in a natural way to processing of two-variable signals or multi-variable signals in computer vision. However, because of the specific nature of images there are many methods developed within computer vision which have no counterpart in processing of one-variable signals. Together with the multi-dimensionality of the signal, this defines a subfield in signal processing as a part of computer vision.

Other fields

Beside the above-mentioned views on computer vision, many of the related research topics can also be studied from a purely mathematical point of view. For example, many methods in computer vision are based on statistics, optimization or geometry. Finally, a significant part of the field is devoted to the implementation aspect of computer vision; how existing methods can be realized in various combinations of software and hardware, or how these methods can be modified in order to gain processing speed without losing too much performance. Computer vision is also used in fashion ecommerce, inventory management, patent search, furniture, and the beauty industry.

Distinctions

The fields most closely related to computer vision are image processing, image analysis and machine vision. There is a significant overlap in the range of techniques and applications that these cover. This implies that the basic techniques that are used and developed in these fields are similar, something which can be interpreted as there is only one field with different names. On the other hand, it appears to be necessary for research groups, scientific journals, conferences and companies to present or market themselves as belonging specifically to one of these fields and, hence, various characterizations which distinguish each of the fields from the others have been presented.

Computer graphics produces image data from 3D models, computer vision often produces 3D models from image data.[20] There is also a trend towards a combination of the two disciplines, *e.g.*, as explored in augmented reality.

The following characterizations appear relevant but should not be taken as universally accepted:

- Image processing and image analysis tend to focus on 2D images, how to transform one image to another, *e.g.*, by pixel-wise operations such as contrast enhancement, local operations such as edge extraction or noise removal, or geometrical transformations such as rotating the image. This characterization implies that image processing/analysis neither require assumptions nor produce interpretations about the image content.
- Computer vision includes 3D analysis from 2D images. This analyzes the 3D scene projected onto one or several images, *e.g.*, how to reconstruct structure or other information about the 3D scene from one or several images. Computer vision often relies on more or less complex assumptions about the scene depicted in an image.
- Machine vision is the process of applying a range of technologies & methods to provide imaging-based automatic inspection, process control and robot guidance[21] in industrial applications.[19] Machine vision tends to focus on applications, mainly in manufacturing, *e.g.*, vision-based robots and systems for vision-based inspection, measurement, or picking (such as bin picking[22]). This implies that image sensor technologies and control theory often are integrated with the processing of image data to control a robot and that real-time processing is emphasised by means of efficient implementations in hardware and software. It also implies that the external conditions such as lighting can be and are often more controlled in machine vision than they are in general computer vision, which can enable the use of different algorithms.
- There is also a field called imaging which primarily focuses on the process of producing images, but sometimes also deals with processing and analysis of images. For example, medical imaging includes substantial work on the analysis of image data in medical applications.
- Finally, pattern recognition is a field which uses various methods to extract information from signals in general, mainly based on statistical approaches and artificial neural networks. A significant part of this field is devoted to applying these methods to image data.

Photogrammetry also overlaps with computer vision, e.g., stereophotogrammetry vs. computer stereo vision.
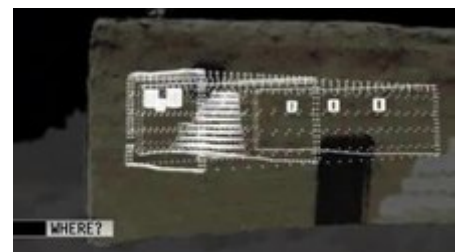
Applications

Applications range from tasks such as industrial machine vision systems which, say, inspect bottles speeding by on a production line, to research into artificial intelligence and computers or robots that can comprehend the world around them. The computer vision and machine vision fields have significant overlap. Computer vision covers the core technology of automated image analysis which is used in many fields. Machine vision usually refers to a process of combining automated image analysis with other methods and technologies to provide automated inspection and robot guidance in industrial applications. In many computer-vision applications, the computers are pre-programmed to solve a particular task, but methods based on learning are now becoming increasingly common. Examples of applications of computer vision include systems for:

- Automatic inspection, *e.g.*, in manufacturing applications;
- Assisting humans in identification tasks, e.g., a species identification system;[23]
- Controlling processes, *e.g.*, an industrial robot;
- Detecting events, *e.g.*, for visual surveillance or people counting;
- Interaction, *e.g.*, as the input to a device for computer-human interaction;
- Modeling objects or environments, *e.g.*, medical image analysis or topographical modeling;
- Navigation, *e.g.*, by an autonomous vehicle or mobile robot; and
- Organizing information, *e.g.*, for indexing databases of images and image sequences.
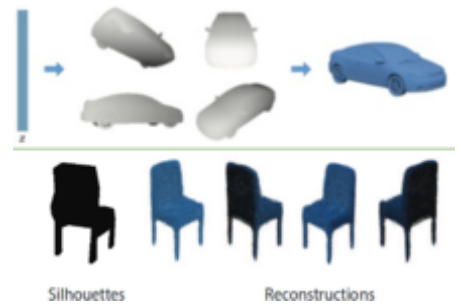
Learning 3D shapes has been a challenging task in computer vision. Recent advances in deep learning has enabled researchers to build models that are able to generate and reconstruct 3D shapes from single or multi-view depth maps or silhouettes seamlessly and efficiently [20]

One of the most prominent application fields is medical computer vision, or medical image processing, characterized by the extraction of information from image data to diagnose a patient. An example of this is detection of tumours, arteriosclerosis or other malign changes; measurements of organ dimensions, blood flow, etc. are another example. It also supports medical research by providing new information: *e.g.*, about the structure of the brain, or about the quality of medical treatments. Applications of computer vision in the medical area also includes enhancement of images interpreted by humans—ultrasonic images or X-ray images for example—to reduce the influence of noise.

A second application area in computer vision is in industry, sometimes called machine vision, where information is extracted for the purpose of supporting a manufacturing process. One example is quality control where details or final products are being automatically inspected in order to find defects. Another example is measurement of position and orientation of details to be picked up by a robot arm. Machine vision is also heavily used in agricultural process to remove undesirable food stuff from bulk material, a process called optical sorting.[24]

Play media
DARPA's Visual Media Reasoning concept video

Military applications are probably one of the largest areas for computer vision. The obvious examples are detection of enemy soldiers or vehicles and missile guidance. More advanced systems for missile guidance send the missile to an area rather than a specific target, and target selection is made when the missile reaches the area based on locally acquired image data. Modern military concepts, such as "battlefield awareness", imply that various sensors, including image sensors, provide a rich set of information about a combat scene which can be used to support strategic decisions. In this case, automatic processing of the data is used to reduce complexity and to fuse information from multiple sensors to increase reliability.

One of the newer application areas is autonomous vehicles, which include submersibles, land-based vehicles (small robots with wheels, cars or trucks), aerial vehicles, and unmanned aerial vehicles (UAV). The level of autonomy ranges from fully autonomous (unmanned) vehicles to vehicles where computer-vision-based systems support a driver or a pilot in various situations. Fully autonomous vehicles typically use computer vision for navigation, *e.g.* for knowing where it is, or for producing a map of its environment (SLAM) and for detecting obstacles. It can also be used for detecting certain task specific events, *e.g.*, a UAV looking for forest fires. Examples of supporting systems are obstacle warning systems in cars, and systems for autonomous landing of aircraft. Several car manufacturers have demonstrated systems for autonomous driving of cars, but this technology has still not reached a level where it can be put on the market. There are ample examples of military autonomous vehicles ranging from advanced missiles to UAVs for recon missions or missile guidance. Space exploration is already being made with autonomous vehicles using computer vision, *e.g.*, NASA's Mars Exploration Rover and ESA's ExoMars Rover.

Other application areas include:

- Support of visual effects creation for cinema and broadcast, *e.g.*, camera tracking (matchmoving).
- Surveillance.
- Tracking and counting organisms in the biological sciences[25]

Typical tasks

Each of the application areas described above employ a range of computer vision tasks; more or less well-defined measurement problems or processing problems, which can be solved using a variety of methods. Some examples of typical computer vision tasks are presented below.

Artist's concept of a Mars Exploration Rover, an example of an unmanned land-based vehicle. Notice the stereo cameras mounted on top of the rover.

Computer vision tasks include methods for acquiring, processing, analyzing and understanding digital images, and extraction of high-dimensional data from the real world in order to produce numerical or symbolic information, *e.g.*, in the forms of decisions.[4][5][6][7] Understanding in this context means the transformation of visual images (the input of the retina) into descriptions of the world that can interface with other thought processes and elicit appropriate action. This image understanding can be seen as the disentangling of symbolic information from image data using models constructed with the aid of geometry, physics, statistics, and learning theory.[8]

Recognition

The classical problem in computer vision, image processing, and machine vision is that of determining whether or not the image data contains some specific object, feature, or activity. Different varieties of the recognition problem are described in the literature:

- **Object recognition** (also called **object classification**) – one or several pre-specified or learned objects or object classes can be recognized, usually together with their 2D positions in the image or 3D poses in the scene. Blippar, Google Goggles and LikeThat provide stand-alone programs that illustrate this functionality.
- **Identification** – an individual instance of an object is recognized. Examples include identification of a specific person's face or fingerprint, identification of handwritten digits, or identification of a specific vehicle.
- **Detection** – the image data are scanned for a specific condition. Examples include detection of possible abnormal cells or tissues in medical images or detection of a vehicle in an automatic road toll system. Detection based on relatively simple and fast computations is sometimes used for finding smaller regions of interesting image data which can be further analyzed by more computationally demanding techniques to produce a correct interpretation.

Currently, the best algorithms for such tasks are based on convolutional neural networks. An illustration of their capabilities is given by the ImageNet Large Scale Visual Recognition Challenge; this is a benchmark in object classification and detection, with millions of images and hundreds of object classes. Performance of convolutional neural networks, on the ImageNet tests, is now close to that of humans.[26] The best algorithms still struggle with objects that are small or thin, such as a small ant on a stem of a flower or a person holding a quill in their hand. They also have trouble with images that have been distorted with filters (an increasingly common phenomenon with modern digital cameras). By contrast, those kinds of images rarely trouble humans. Humans, however, tend to have trouble with other issues. For example, they are not good at classifying objects into fine-grained classes, such as the particular breed of dog or species of bird, whereas convolutional neural networks handle this with ease.

Several specialized tasks based on recognition exist, such as:

- **Content-based image retrieval** – finding all images in a larger set of images which have a specific content. The content can be specified in different ways, for example in terms of similarity relative a target image (give me all images similar to image X), or in terms of high-level search criteria given as text input (give me all images which contains many houses, are taken during winter, and have no cars in them).

- **Pose estimation** – estimating the position or orientation of a specific object relative to the camera. An example application for this technique would be assisting a robot arm in retrieving objects from a conveyor belt in an assembly line situation or picking parts from a bin.
- **Optical character recognition** (OCR) – identifying characters in images of printed or handwritten text, usually with a view to encoding the text in a format more amenable to editing or indexing (*e.g.* ASCII).
- **2D code reading** – reading of 2D codes such as data matrix and QR codes.
- **Facial recognition**
- **Shape Recognition Technology** (SRT) in people counter systems differentiating human beings (head and shoulder patterns) from objects
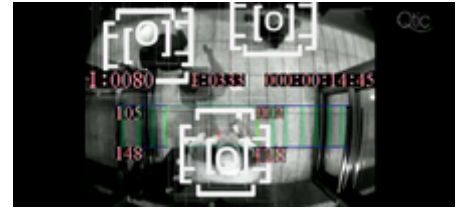
Computer vision for people counter purposes in public places, malls, shopping centres

Motion analysis

Several tasks relate to motion estimation where an image sequence is processed to produce an estimate of the velocity either at each points in the image or in the 3D scene, or even of the camera that produces the images . Examples of such tasks are:

- **Egomotion** – determining the 3D rigid motion (rotation and translation) of the camera from an image sequence produced by the camera.
- **Tracking** – following the movements of a (usually) smaller set of interest points or objects (*e.g.*, vehicles, humans or other organisms[25]) in the image sequence.
- **Optical flow** – to determine, for each point in the image, how that point is moving relative to the image plane, *i.e.*, its apparent motion. This motion is a result both of how the corresponding 3D point is moving in the scene and how the camera is moving relative to the scene.

Scene reconstruction

Given one or (typically) more images of a scene, or a video, scene reconstruction aims at computing a 3D model of the scene. In the simplest case the model can be a set of 3D points. More sophisticated methods produce a complete 3D surface model. The advent of 3D imaging not requiring motion or scanning, and related processing algorithms is enabling rapid advances in this field. Grid-based 3D sensing can be used to acquire 3D images from multiple angles. Algorithms are now available to stitch multiple 3D images together into point clouds and 3D models.[20]

Image restoration

The aim of image restoration is the removal of noise (sensor noise, motion blur, etc.) from images. The simplest possible approach for noise removal is various types of filters such as low-pass filters or median filters. More sophisticated methods assume a model of how the local image structures look, to distinguish them from noise. By first analysing the image data in terms of the local image structures, such as lines or edges, and then controlling the filtering based on local information from the analysis step, a better level of noise removal is usually obtained compared to the simpler approaches.

An example in this field is inpainting.

System methods

The organization of a computer vision system is highly application-dependent. Some systems are stand-alone applications that solve a specific measurement or detection problem, while others constitute a sub-system of a larger design which, for example, also contains sub-systems for control of mechanical actuators, planning, information databases, man-machine interfaces, etc. The

specific implementation of a computer vision system also depends on whether its functionality is pre-specified or if some part of it can be learned or modified during operation. Many functions are unique to the application. There are, however, typical functions that are found in many computer vision systems.

- **Image acquisition** – A digital image is produced by one or several image sensors, which, besides various types of light-sensitive cameras, include range sensors, tomography devices, radar, ultra-sonic cameras, etc. Depending on the type of sensor, the resulting image data is an ordinary 2D image, a 3D volume, or an image sequence. The pixel values typically correspond to light intensity in one or several spectral bands (gray images or colour images), but can also be related to various physical measures, such as depth, absorption or reflectance of sonic or electromagnetic waves, or nuclear magnetic resonance.[24]
- **Pre-processing** – Before a computer vision method can be applied to image data in order to extract some specific piece of information, it is usually necessary to process the data in order to assure that it satisfies certain assumptions implied by the method. Examples are:

  - Re-sampling to assure that the image coordinate system is correct.
  - Noise reduction to assure that sensor noise does not introduce false information.
  - Contrast enhancement to assure that relevant information can be detected.
  - Scale space representation to enhance image structures at locally appropriate scales.

- **Feature extraction** – Image features at various levels of complexity are extracted from the image data.[24] Typical examples of such features are:

  - Lines, edges and ridges.
  - Localized interest points such as corners, blobs or points.

    More complex features may be related to texture, shape or motion.

- **Detection/segmentation** – At some point in the processing a decision is made about which image points or regions of the image are relevant for further processing.[24] Examples are:

  - Selection of a specific set of interest points.
  - Segmentation of one or multiple image regions that contain a specific object of interest.
  - Segmentation of image into nested scene architecture comprising foreground, object groups, single objects or salient object parts (also referred to as spatial-taxon scene hierarchy),[27] while the visual salience is often implemented as spatial and temporal attention.
  - Segmentation or co-segmentation of one or multiple videos into a series of per-frame foreground masks, while maintaining its temporal semantic continuity.[28][29]

- **High-level processing** – At this step the input is typically a small set of data, for example a set of points or an image region which is assumed to contain a specific object.[24] The remaining processing deals with, for example:

  - Verification that the data satisfy model-based and application-specific assumptions.
  - Estimation of application-specific parameters, such as object pose or object size.
  - Image recognition – classifying a detected object into different categories.
  - Image registration – comparing and combining two different views of the same object.

- **Decision making** Making the final decision required for the application,[24] for example:

  - Pass/fail on automatic inspection applications.
  - Match/no-match in recognition applications.
  - Flag for further human review in medical, military, security and recognition applications.

Image-understanding systems

Image-understanding systems (IUS) include three levels of abstraction as follows: low level includes image primitives such as edges, texture elements, or regions; intermediate level includes boundaries, surfaces and volumes; and high level includes objects, scenes, or events. Many of these requirements are really topics for further research.

The representational requirements in the designing of IUS for these levels are: representation of prototypical concepts, concept organization, spatial knowledge, temporal knowledge, scaling, and description by comparison and differentiation.

While inference refers to the process of deriving new, not explicitly represented facts from currently known facts, control refers to the process that selects which of the many inference, search, and matching techniques should be applied at a particular stage of processing. Inference and control requirements for IUS are: search and hypothesis activation, matching and hypothesis testing, generation and use of expectations, change and focus of attention, certainty and strength of belief, inference and goal satisfaction.[30]

Hardware

There are many kinds of computer vision systems; however, all of them contain these basic elements: a power source, at least one image acquisition device (camera, ccd, etc.), a processor, and control and communication cables or some kind of wireless interconnection mechanism. In addition, a practical vision system contains software, as well as a display in order to monitor the system. Vision systems for inner spaces, as most industrial ones, contain an illumination system and may be placed in a controlled environment. Furthermore, a completed system includes many accessories such as camera supports, cables and connectors.

Most computer vision systems use visible-light cameras passively viewing a scene at frame rates of at most 60 frames per second (usually far slower).

A few computer vision systems use image-acquisition hardware with active illumination or something other than visible light or both, such as structured-light 3D scanners, thermographic cameras, hyperspectral imagers, radar imaging, lidar scanners, magnetic resonance images, side-scan sonar, synthetic aperture sonar, etc. Such hardware captures "images" that are then processed often using the same computer vision algorithms used to process visible-light images.

While traditional broadcast and consumer video systems operate at a rate of 30 frames per second, advances in digital signal processing and consumer graphics hardware has made high-speed image acquisition, processing, and display possible for real-time systems on the order of hundreds to thousands of frames per second. For applications in robotics, fast, real-time video systems are critically important and often can simplify the processing needed for certain algorithms. When combined with a high-speed projector, fast image acquisition allows 3D measurement and feature tracking to be realised.[31]

Egocentric vision systems are composed of a wearable camera that automatically take pictures from a first-person perspective.

As of 2016, vision processing units are emerging as a new class of processor, to complement CPUs and graphics processing units (GPUs) in this role.[32]

See also

- Computational imaging
- Computational photography
- Machine vision glossary
- Space mapping
- Teknomo–Fernandez algorithm
- Visual system
- Visual perception
- Vision science
- Egocentric vision
- Visual agnosia

Lists

- List of computer vision topics
- List of emerging technologies
- Outline of artificial intelligence

References

1. Dana H. Ballard; Christopher M. Brown (1982). *Computer Vision* (http://homepages.inf.ed.ac.uk/rbf/BOOKS/BAN DB/bandb.htm). Prentice Hall. ISBN 978-0-13-165316-0.
2. Huang, T. (1996-11-19). Vandoni, Carlo, E (ed.). *Computer Vision : Evolution And Promise* (http://cds.cern.ch/rec ord/400313/files/p21.pdf) (PDF). 19th CERN School of Computing (http://cds.cern.ch/record/300250). Geneva: CERN. pp. 21–25. doi:10.5170/CERN-1996-008.21 (https://doi.org/10.5170%2FCERN-1996-008.21). ISBN 978-9290830955.
3. Milan Sonka; Vaclav Hlavac; Roger Boyle (2008). *Image Processing, Analysis, and Machine Vision*. Thomson. ISBN 978-0-495-08252-1.
4. Reinhard Klette (2014). *Concise Computer Vision*. Springer. ISBN 978-1-4471-6320-6.
5. Linda G. Shapiro; George C. Stockman (2001). *Computer Vision*. Prentice Hall. ISBN 978-0-13-030796-5.
6. Tim Morris (2004). *Computer Vision and Image Processing*. Palgrave Macmillan. ISBN 978-0-333-99451-1.
7. Bernd Jähne; Horst Haußecker (2000). *Computer Vision and Applications, A Guide for Students and Practitioners*. Academic Press. ISBN 978-0-13-085198-7.
8. David A. Forsyth; Jean Ponce (2003). *Computer Vision, A Modern Approach*. Prentice Hall. ISBN 978-0-13-085198-7.
9. http://www.bmva.org/visionoverview The British Machine Vision Association and Society for Pattern Recognition Retrieved February 20, 2017
10. Murphy, Mike. "Star Trek's "tricorder" medical scanner just got closer to becoming a reality" (https://qz.com/95728 5/star-treks-tricorder-handheld-medical-scanner-just-got-closer-to-becoming-a-reality/).
11. Richard Szeliski (30 September 2010). *Computer Vision: Algorithms and Applications* (https://books.google.com/ books?id=bXzAlkODwa8C). Springer Science & Business Media. pp. 10–16. ISBN 978-1-84882-935-0.
12. Papert, Seymour (1966-07-01). "The Summer Vision Project". *MIT AI Memos (1959 - 2004)*. hdl:1721.1/6125 (htt ps://hdl.handle.net/1721.1%2F6125).
13. Margaret Ann Boden (2006). *Mind as Machine: A History of Cognitive Science* (https://books.google.com/books?i d=yRyETy43AdQC). Clarendon Press. p. 781. ISBN 978-0-19-954316-8.
14. Takeo Kanade (6 December 2012). *Three-Dimensional Machine Vision* (https://books.google.com/books?id=e0jt BwAAQBAJ&pg=PA191). Springer Science & Business Media. ISBN 978-1-4613-1981-8.
15. Nicu Sebe; Ira Cohen; Ashutosh Garg; Thomas S. Huang (3 June 2005). *Machine Learning in Computer Vision* (https://books.google.com/books?id=lemw2Rhr_PEC). Springer Science & Business Media. ISBN 978-1-4020-3274-5.
16. William Freeman; Pietro Perona; Bernhard Scholkopf (2008). "Guest Editorial: Machine Learning for Computer Vision". *International Journal of Computer Vision*. **77** (1): 1. doi:10.1007/s11263-008-0127-7 (https://doi.org/10.10 07%2Fs11263-008-0127-7). ISSN 1573-1405 (https://www.worldcat.org/issn/1573-1405).
17. "Information Engineering | Department of Engineering" (http://www.eng.cam.ac.uk/research/academic-divisions/in formation-engineering). *www.eng.cam.ac.uk*. Retrieved 2018-10-03.
18. "Information Engineering Main/Home Page" (http://www.robots.ox.ac.uk/). *www.robots.ox.ac.uk*. Retrieved 2018-10-03.
19. Steger, Carsten; Markus Ulrich; Christian Wiedemann (2018). *Machine Vision Algorithms and Applications* (http s://books.google.com/?id=tppFDwAAQBAJ&printsec=frontcover#v=onepage&q&f=false) (2nd ed.). Weinheim: Wiley-VCH. p. 1. ISBN 978-3-527-41365-2. Retrieved 2018-01-30.
20. "Soltani, A. A., Huang, H., Wu, J., Kulkarni, T. D., & Tenenbaum, J. B. Synthesizing 3D Shapes via Modeling Multi-View Depth Maps and Silhouettes With Deep Generative Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 1511-1519)" (https://github.com/Amir-Arsalan/Synthesize3Dvia DepthOrSil). 2019-01-25.

21. Turek, Fred (June 2011). "Machine Vision Fundamentals, How to Make Robots See". *NASA Tech Briefs Magazine*. **35** (6). pages 60–62

22. https://www.robots.com/blog/viewing/the-future-of-automated-random-bin-picking

23. Wäldchen, Jana; Mäder, Patrick (2017-01-07). "Plant Species Identification Using Computer Vision Techniques: A Systematic Literature Review" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6003396). *Archives of Computational Methods in Engineering*. **25** (2): 507–543. doi:10.1007/s11831-016-9206-z (https://doi.org/10.1007%2Fs11831-016-9206-z). ISSN 1134-3060 (https://www.worldcat.org/issn/1134-3060). PMC 6003396 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6003396). PMID 29962832 (https://www.ncbi.nlm.nih.gov/pubmed/29962832).

24. E. Roy Davies (2005). *Machine Vision: Theory, Algorithms, Practicalities*. Morgan Kaufmann. ISBN 978-0-12-206093-9.

25. Bruijning, Marjolein; Visser, Marco D.; Hallmann, Caspar A.; Jongejans, Eelke; Golding, Nick (2018). "trackdem: Automated particle tracking to obtain population counts and size distributions from videos in r". *Methods in Ecology and Evolution*. **9** (4): 965–973. doi:10.1111/2041-210X.12975 (https://doi.org/10.1111%2F2041-210X.12975). ISSN 2041-210X (https://www.worldcat.org/issn/2041-210X).

26. O. Russakovsky et al., "ImageNet Large Scale Visual Recognition Challenge (https://arxiv.org/abs/1409.0575)", 2014.

27. Barghout, Lauren. "Visual Taxometric Approach to Image Segmentation Using Fuzzy-Spatial Taxon Cut Yields Contextually Relevant Regions (http://www.lirmm.fr/~lafourcade/pub/IPMU2014/papers/0443/04430163.pdf)." Information Processing and Management of Uncertainty in Knowledge-Based Systems. Springer International Publishing, 2014.

28. Liu, Ziyi; Wang, Le; Hua, Gang; Zhang, Qilin; Niu, Zhenxing; Wu, Ying; Zheng, Nanning (2018). "Joint Video Object Discovery and Segmentation by Coupled Dynamic Markov Networks" (https://qilin-zhang.github.io/_pages/pdfs/Joint_Video_Object_Discovery_and_Segmentation_by_Coupled_Dynamic_Markov_Networks.pdf) (PDF). *IEEE Transactions on Image Processing*. **27** (12): 5840–5853. doi:10.1109/tip.2018.2859622 (https://doi.org/10.1109%2Ftip.2018.2859622). ISSN 1057-7149 (https://www.worldcat.org/issn/1057-7149). PMID 30059300 (https://www.ncbi.nlm.nih.gov/pubmed/30059300).

29. Wang, Le; Duan, Xuhuan; Zhang, Qilin; Niu, Zhenxing; Hua, Gang; Zheng, Nanning (2018-05-22). "Segment-Tube: Spatio-Temporal Action Localization in Untrimmed Videos with Per-Frame Segmentation" (https://qilin-zhang.github.io/_pages/pdfs/Segment-Tube_Spatio-Temporal_Action_Localization_in_Untrimmed_Videos_with_Per-Frame_Segmentation.pdf) (PDF). *Sensors*. **18** (5): 1657. doi:10.3390/s18051657 (https://doi.org/10.3390%2Fs18051657). ISSN 1424-8220 (https://www.worldcat.org/issn/1424-8220). PMC 5982167 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5982167). PMID 29789447 (https://www.ncbi.nlm.nih.gov/pubmed/29789447).

30. Shapiro, Stuart C. (1992). *Encyclopedia of Artificial Intelligence, Volume 1*. New York: John WIley & Sons, Inc. pp. 643–646. ISBN 978-0-471-50306-4.

31. Kagami, Shingo (2010). *High-speed vision systems and projectors for real-time perception of the world*. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops*. **2010**. pp. 100–107. doi:10.1109/CVPRW.2010.5543776 (https://doi.org/10.1109%2FCVPRW.2010.5543776). ISBN 978-1-4244-7029-7.

32. Seth Colaner (January 3, 2016). "A Third Type Of Processor For VR/AR: Movidius' Myriad 2 VPU" (http://www.tomshardware.com/news/movidiud-myriad2-vpu-vision-processing-vr,30850.html). *www.tomshardware.com*.

Further reading

- David Marr (1982). *Vision*. W. H. Freeman and Company. ISBN 978-0-7167-1284-8.
- Azriel Rosenfeld; Avinash Kak (1982). *Digital Picture Processing*. Academic Press. ISBN 978-0-12-597301-4.
- Barghout, Lauren; Lawrence W. Lee (2003). *Perceptual information processing system*. U.S. Patent Application 10/618,543. ISBN 978-0-262-08159-7.
- Berthold Klaus Paul Horn (1986). *Robot Vision*. MIT Press. ISBN 978-0-262-08159-7.
- Michael C.Fairhurst (1988). *Computer Vision for robotic systems*. Prentice Hall. ISBN 978-0-13-166919-2.
- Olivier Faugeras (1993). *Three-Dimensional Computer Vision, A Geometric Viewpoint*. MIT Press. ISBN 978-0-262-06158-2.

- Tony Lindeberg (1994). *Scale-Space Theory in Computer Vision* (http://www.nada.kth.se/~tony/book.html). Springer. ISBN 978-0-7923-9418-1.
- James L. Crowley and Henrik I. Christensen (Eds.) (1995). *Vision as Process*. Springer-Verlag. ISBN 978-3-540-58143-7.
- Gösta H. Granlund; Hans Knutsson (1995). *Signal Processing for Computer Vision*. Kluwer Academic Publisher. ISBN 978-0-7923-9530-0.
- Reinhard Klette; Karsten Schluens; Andreas Koschan (1998). *Computer Vision – Three-Dimensional Data from Images* (http://www.cs.auckland.ac.nz/~rklette/Books/SpringerCV98/Springer98.html). Springer, Singapore. ISBN 978-981-3083-71-4.
- Emanuele Trucco; Alessandro Verri (1998). *Introductory Techniques for 3-D Computer Vision*. Prentice Hall. ISBN 978-0-13-261108-4.
- Bernd Jähne (2002). *Digital Image Processing*. Springer. ISBN 978-3-540-67754-3.
- Richard Hartley and Andrew Zisserman (2003). *Multiple View Geometry in Computer Vision*. Cambridge University Press. ISBN 978-0-521-54051-3.
- Gérard Medioni; Sing Bing Kang (2004). *Emerging Topics in Computer Vision*. Prentice Hall. ISBN 978-0-13-101366-7.
- R. Fisher; K Dawson-Howe; A. Fitzgibbon; C. Robertson; E. Trucco (2005). *Dictionary of Computer Vision and Image Processing*. John Wiley. ISBN 978-0-470-01526-1.
- Nikos Paragios and Yunmei Chen and Olivier Faugeras (2005). *Handbook of Mathematical Models in Computer Vision* (http://www.mas.ecp.fr/vision/Personnel/nikos/paragios-chen-faugeras/). Springer. ISBN 978-0-387-26371-7.
- Wilhelm Burger; Mark J. Burge (2007). *Digital Image Processing: An Algorithmic Approach Using Java* (http://www.imagingbook.com/). Springer. ISBN 978-1-84628-379-6.
- Pedram Azad; Tilo Gockel; Rüdiger Dillmann (2008). *Computer Vision – Principles and Practice* (http://ivt.sourceforge.net/book.html). Elektor International Media BV. ISBN 978-0-905705-71-2.
- Richard Szeliski (2010). *Computer Vision: Algorithms and Applications* (http://szeliski.org/Book/). Springer-Verlag. ISBN 978-1848829343.
- J. R. Parker (2011). *Algorithms for Image Processing and Computer Vision (2nd ed.)*. Wiley. ISBN 978-0470643853.
- Richard J. Radke (2013). *Computer Vision for Visual Effects*. Cambridge University Press. ISBN 978-0-521-76687-6.

External links

- USC Iris computer vision conference list (http://iris.usc.edu/Information/Iris-Conferences.html)
- Computer vision papers on the web (http://www.cvpapers.com/index.html) A complete list of papers of the most relevant computer vision conferences.
- Computer Vision Online (http://www.computervisiononline.com/) News, source code, datasets and job offers related to computer vision.
- Keith Price's Annotated Computer Vision Bibliography (http://iris.usc.edu/Vision-Notes/bibliography/contents.html)
- CVonline (http://homepages.inf.ed.ac.uk/rbf/CVonline/) Bob Fisher's Compendium of Computer Vision.
- British Machine Vision Association (http://www.bmva.org/) Supporting computer vision research within the UK via the BMVC and MIUA conferences, *Annals of the BMVA* (open-source journal), BMVA Summer School and one-day meetings