International Journal of Research in Engineering, Science and Management
Volume-2, Issue-11, November-2019
www.ijresm.com | ISSN (Online): 2581-5792

360° Overview of Computer Vision

Parth Joshi[1], Nishant Jethwa[2], Rohit Pandey[3]
[1,2,3]*Student, Department of Computer Engineering, NMIMS Mukesh Patel School of Technology Management and Engineering, Mumbai, India*

***Abstract*: This paper will give an overview of what is computer vision, the base which lead to the creation of computer vision, what a computer vision system is and what basic tasks it performs and a few applications of computer vision. Getting a base knowledge on any subject allows a better understanding of it, especially when you try to implement something related to it or conduct some deep research regarding the topic.**

***Keywords*: computer vision**

1. Introduction

Vision is a significant contribution of the human senses. It offers us, seemingly effortlessly, with a detailed three-dimensional (3-D) description of a complex and rapidly changing world. Computer vision is the study of enabling computers to understand and interpret visual information in a way similar to a human eye. Computer vision in computer science and engineering fields is becoming a mainstream topic of research and study. With the rapid explosion of multimedia and the extensive use of video and image-based communications over the World Wide Web, every student in computer science and engineering should receive some basic education related to computation with images. Currently, there is strong industry demand for computer vision scientists and engineers well versed in this technology - people who understand computer vision technology and know how to apply it in real-world problems. Section III and IV cover give a basic idea on Computer vision and the base that lead to its creation. Section V, VI and VII cover what a Computer Vision system is, its required hardware, mentioning of algorithms used and what functions it performs. Section VIII covers the applications of computer vision.

2. Related Work

Deep Learning for Computer Vision: A brief review by Athanasios Voulodimos [2] talks about what Computer Vision is and a detailed information about Neural Networks, the science behind it and how it is used in a computer vision system. Computer Vision for 3D Perception by Niall O' Mahony [3] informs how 3D images are perceived and a few algorithms and equipment that are preferred to do so. This paper provides information regarding many different topics not covered together in any paper and also simplifies information in form of tables and diagrams.

3. Overview

Computer Vision allows learning and representation of data with the help of multiple layered computational models which helps to represent the way a human brain would perceive visual information. Computer Vision is strongly related to artificial intelligence, since the computer must interpret what it sees and then analyze or behave accordingly. Computer vision tasks include methods for acquiring, processing, analyzing and understanding digital images. These tasks are the most basic part of computer vision which are used to perform other bigger tasks like converting a 2D image into a 3D image, image classification, object localization, object detection, image segmentation etc. Computer vision has grown over the past ten years from a research area to a widely accepted technology that can deliver a dramatic increase in productivity and improve living standards. A computer vision system can be used to perform operations on videos or images and the system uses various algorithms

4. Deep Learning: The base of Computer Vision

Convolutional Neural Networks (CNNs) were the base used in the beginning of deep learning which were later on used for various applications one of which being Computer Vision. CNN focuses on the input comprising of images [6]. CNNs were inspired by the visual system's structure, and in particular by the models of it proposed as mentioned in [2].

A CNN comprises three main types of neural layers, namely - convolutional layers, pooling layers, and fully connected layers. Each type of layer plays a different role. Convolutional layers: A CNN uses different kernels in the convolutional layers to convert the entire image as well as the intermediate feature maps, creating various feature maps.

Pooling Layers: Pooling layers are in charge of reducing the spatial dimensions of the input volume for the next convolutional layer. The pooling layer does not affect the depth dimension of the volume, but reduction in size leads to loss of information. Such a reduction is advantageous to the network, however, because the size decrease leads to less computational overhead for the network's future layers and also protects against overfitting.

Fully Connected Layers: Following several convolutional and pooling layers, the high-level reasoning in the neural network is performed via fully connected layers. Such a reduction is advantageous to the network, however, because the size decrease leads to less computational overhead for the network's future layers and also protects against overfitting. The derived vector either could be fed forward into a certain number of categories for classification or could be considered as a feature vector for further processing.

5. Computer Vision System: Components

A computer vision system is a collection of specific hardware which are used to make the software perform specific tasks in the field of image and video processing. A computer vision system uses different computer vision algorithms to implement them on images or videos for further work.

Necessary hardware requirements for a Computer vision system include are mentioned in Table 1.

Table 1
Necessary hardware requirements for a Computer vision system

Component	Usage
Camera & optics	Usually contains one or more cameras and lensing to acquire an image
Lighting	Illuminate parts to acquire the best of the image. ex- LED
Frame Grabber	A video capture card which interfaces the camera to the host computer
PC Platform	To keep data in check and perform operations on an acquired image
Inspection Software	Computer vision software to create and execute programs, process incoming image information, and make PASS / FAIL choices.
Digital I/O & Network connection	Output information or commands to devices once inspection part is complete.

Besides the mentioned equipment's additions can be made based on requirements. A detailed explanation, regarding how to build a machine vision system, is provided in the white paper [1] by Gregory Robert. It is possible to implement a computer vision system as an online service as shown in [5].

6. Computer Vision: Algorithms

Algorithms in computer vision are techniques used to perform various operations on the image, to bring the desired output or get close to it. There are various algorithms which perform different kinds of operations on an image, leading to either a different output or a similar output. In short, the same output can be acquired performing different operations, which leads to the thought of which algorithm is more efficient, in terms of space and time complexity. Table 2, mentions some of the major algorithms used for computer vision.

Difference between SIFT, SURF, ORB is given in [4].

Table 2
Major algorithms used for computer vision

Algorithm	Usage
SIFT (Scale-invariant feature transform)	Detect and describe local features in an image
SURF(Speeded up robust features)	Local feature detector and descriptor
ORB (Oriented FAST and rotated BRIEF)	Object recognition , image rotation detection
Deep Fashion	Recognize clothing items
Colorful image colorization	Converts a black and white image into RGB image
Car Make and Model Recognition	Identify make, model, body style and model year of any care.
OCR (Optical Character Recognition)	Detect characters from an image
Crowd Counter	Count number of people in an image

7. Functions utilized in every computer visions system

A. Image acquisition

A digital image is generated by one or more image sensors, which include range sensors, tomography devices, radar, ultrasonic cameras, etc. in addition to multiple kinds of light-sensitive cameras. The resulting image data are an ordinary 2D image, a 3D volume, or an image sequence, depending on the type of sensor. Typically, the pixel values correspond to light intensity in one or more spectral bands (gray pictures or color pictures), but can also be associated with multiple physical measures such as depth, absorption or reflection of sonic or electromagnetic waves, or nuclear magnetic resonance.

B. Pre-processing

In order to extract some specific piece of information, it is usually necessary to process the data before a computer vision method can be applied to image data to ensure that it fulfills certain assumptions implied by the method. Examples are:

- Re-sampling to assure that the image coordinate system is correct.
- Noise reduction to ensure that noise from the sensor does not provide false information.
- Enhancement of contrast to ensure the detection of relevant information.
- Scale space representation to improve image structures at locally suitable scales.

C. Feature extraction

The image data extracts object features at various levels of complexity. Typical examples of these features are:

- Lines, edges and ridges.
- Localized interest points such as corners, blobs or points.
- More complex characteristics may have to do with texture, form and movement.

D. Detection/segmentation

A choice on which image points or areas of the image are important for further processing is produced at some stage in the processing. Examples are:

- Selection of a specific set of interest points.
- Segmentation of one or more regions of image containing a particular 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), while the visual salience is often implemented as spatial and temporal attention.

International Journal of Research in Engineering, Science and Management
Volume-2, Issue-11, November-2019
www.ijresm.com | ISSN (Online): 2581-5792

- Segmentation or co-segmentation of one or multiple videos into a series of per-frame foreground masks, while maintaining its temporal semantic continuity.

E. High-level processing

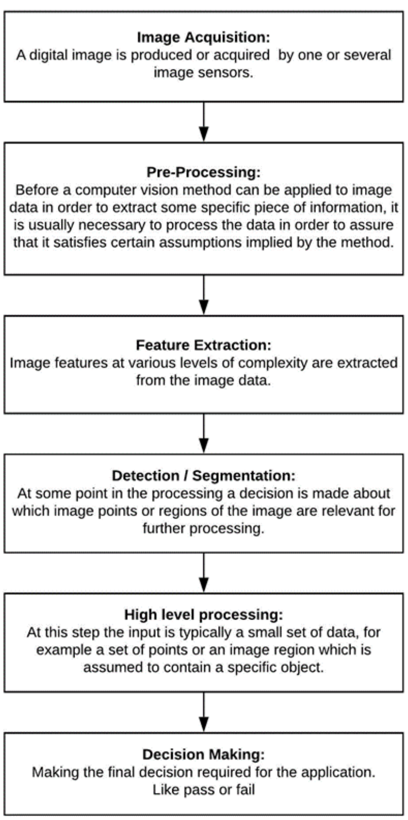
In this phase, typically, the input is a tiny collection of information, such as a set of points or an image region that is supposed to contain a particular object. The rest of the processing is handled, for example:

- Verify the data meet model-based and application-specific hypotheses.
- Application-specific parameter estimation, such as pose of objects or object length.
- Image recognition — sorting into different categories of a perceived object.
- Image registration — two different views of the same object are compared and merged.

F. Decision making

Making the final decision required for the application, 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.



Functions of a Computer Vision System

8. Applications

Applications range from duties such as industrial machine vision systems that, say, check bottles speeding on a production line, to research into artificial intelligence and computers or robots capable of understanding the world around them. There is an important overlap between computer vision and machine vision areas. Computer vision includes the key technology used in many areas of automated image analysis. Machine vision generally relates to a method of mixing automated image analysis with other techniques and techniques in industrial apps to provide automated inspection and robot instruction. The computers are pre-programmed in many computer-vision apps to address a specific job, but learning-based techniques are becoming increasingly prevalent. Major fields applications of computer visions are mentioned in Table 3.

Table 3
Major fields applications of computer visions

Application	Usage
Automatic Inspection	Manufacturing applications to check for any unusualities
Assisting humans in identification tasks.	Object recognition and differentiation system
Controlling processes	An industrial robot
Detecting events	Visual surveillance
Modeling objects or environments	Humidity maps
Navigation	Autonomous Vehicle
Organizing information	Indexing databases of images and image sequences
Medical image processing	Detection of tumor in a patient's body
Military applications	Remote control tank guidance
Space Station applications	Image acquisition and analysis through autonomous rovers
Education	Spreading the teachings of computer visions

9. Conclusion

This paper covers various information like what is Computer Vision, systems which use it, components used to make such systems, different algorithms used to process images and professional fields which make use of its abilities. Within 15–20 years computer vision will become a core component of the day to day lifestyle since its applications will be implemented in almost every device with cameras, especially for analytics and infrastructure. Evaluation of facial expression, body language, emotions and intentions, as well as audio evaluation of the tone and rhythm of spoken words for latent intentions and assumptions will all be possible and will have a huge impact on the way things work. Thus there is a wide variety of scope in terms of easing and speeding up many processes which relate to the daily lives of people.

References

[1] Gregory Robert, "How to plan your PC-Based Machine Vision system" April 1998.
[2] Athanasios Voulodimos, Nikolaos Doulamis, Anastasios Doulamis and Eftychios Protopapadakis, Deep Learning for Computer Vision: A brief review, IEEE Transactions On Education, Vol. 46, No. 1, February 2003.

[3] Niall O' Mahony, Sean Campbell, Lenka Krpalkova, Daniel Riordan, Joseph Walsh, Aidan Murphy, Conor Ryan Computer Vision for 3D Perception, Intelligent Systems Conference, September 2018.
[4] Anu Priya George, X. Felix Joseph Object Recognition Algorithms for Computer Vision System: A Survey, International Journal of Pure and Applied Mathematics, Volume 117, No. 21, 2017.
[5] Kevin Chiu, Ramesh Raskar Computer vision on Tap, IEEE, August 2015.
[6] Keiron O'Shea, Ryan Nash, "An Introduction to Convolutional Neural Networks," ArXiv e-prints, November 2015.