Computer Vision

Theory and Applications

Haitham A. El-Ghareeb

April 27, 2019

Faculty of Computers and Information Sciences
Mansoura University
Egypt
helghareeb@mans.edu.eg

Contacts

- https://www.haitham.ws
- https://youtube.com/user/helghareeb
- https://www.github.com/helghareeb
- http://eg.linkedin.com/in/helghareeb
- helghareeb@mans.edu.eg

Computer Vision

Computer Vision

Definition

Computer Vision

- Interdisciplinary scientific field
- Deals with how computers can be made to gain high-level understanding from digital images or videos
- Seeks to automate tasks that the human visual system can do
- Automatic extraction, analysis and understanding of useful information from a single image or a sequence of images
- Development of a theoretical and algorithmic basis to achieve automatic visual understanding." As a scientific discipline, computer vision is concerned with the theory behind artificial systems that extract information from images

Image Data Forms

Image data can take many forms

- video sequences
- views from multiple cameras
- multi-dimensional data from a medical scanner

Related Fields

Related Fields

Artificial Intelligence

Artificial Intelligence

- Deal with autonomous planning or deliberation for robotical systems to navigate through an environment
- 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

Related Fields

Information Engineering

Information Engineering

- Computer vision is often considered to be part of information engineering
- Information engineering is the engineering discipline that deals with the generation, distribution, analysis, and use of information, data, and knowledge in systems

Related Fields

Solid-state Physics

Solid-state Physics

- Closely related to computer vision
- Computer vision systems rely on image sensors, which detect electromagnetic radiation, which is typically in the form of either visible or infra-red light
- Sensors are designed using quantum physics
- The process by which light interacts with surfaces is explained using physics

Related Fields

Neurobiology

Neurobiology

- Biological vision system
- Extensive study of eyes, neurons, and the brain structures devoted to processing of visual stimuli in both humans and various animals
- How "real" vision systems operate in order to solve certain vision related tasks
- Computer Vision studies and describes the processes implemented in software and hardware behind artificial vision systems

Related Fields

Signal Processing

Signal Processing

- 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
- Together with the multi-dimensionality of the signal, this defines a subfield in signal processing as a part of computer vision.

Related Fields

Other Fields

Other Fields

- Many of the related research topics can also be studied from a purely mathematical point of view
- Many methods in computer vision are based on statistics, optimization or geometry

Related Fields

Fields of Implementation

Fields of Implementation

- Significant part is devoted to the implementation aspect of computer vision
- How existing methods can be realized in various combinations of software and hardware
- 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
  - Beauty Industry

Distinctions

Distinctions

- Closely related fields to computer vision are
  - image processing
  - image analysis
  - machine vision
- Significant overlap in the range of techniques and applications that these cover
- Basic techniques are similar

Distinctions

Computer Graphics

Computer Graphics

- Computer Graphics produces image data from 3D models
- Computer Vision often produces 3D models from image data
- Combination of the two disciplines: augmented reality

Distinctions

Image Processing and Image Analysis

Image Processing and Image Analysis

- Both tend to focus on 2D images and how to transform one image to another
- By pixel-wise operations such as
  - contrast enhancement
  - local operations such as edge extraction or noise removal
  - geometrical transformations such as rotating the image
- Both neither require assumptions nor produce interpretations about the image content

Distinctions

2D Images

Computer Vision and 2D Images

- Computer vision includes 3D analysis from 2D images
- This analyzes the 3D scene projected onto one or several images
- How to reconstruct structure or other information about the 3D scene from one or several images

Distinctions

Machine Vision

Machine Vision

- Process of applying a range of technologies and methods to provide imaging-based automatic inspection, process control and robot guidance in industrial applications
- Tends to focus on applications, mainly in manufacturing
- Vision based robots and systems for vision based inspection, measurement, or picking (such as bin picking)
- Image sensor technologies and control theory often are integrated with the processing of image data to control a robot

Distinctions

Imaging

Imaging

- Focus on the process of producing images
- Sometimes deals with processing and analysis of images
- Medical imaging includes substantial work on the analysis of image data in medical applications

Distinctions

Pattern Recognition

Pattern Recognition

- Field which uses various methods to extract information from signals in general
- Mainly based on statistical approaches and artificial neural networks
- Significant part of this field is devoted to applying these methods to image data

Computer Vision Applications

Applications

- Tasks such as industrial machine vision systems which, say, inspect bottles speeding by on a production line
- Research into artificial intelligence and computers or robots that can comprehend the world around them
- 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

Computer Vision Applications

Learning 3D Shapes

Learning 3D Shapes - 01

- Challenging task
- 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
    - Automatic inspection, e.g., in manufacturing applications;
    - Assisting humans in identification tasks, e.g., a species identification system;
    - Controlling processes, e.g., an industrial robot;
    - Detecting events, e.g., for visual surveillance or people counting;

Learning 3D Shapes - 02

- 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.

Computer Vision Applications

Medical Computer Vision

Medical Computer Vision

- Medical Computer Vision / Medical Image Processing
- Characterized by the extraction of information from image data to diagnose a patient
  - Detection of tumours
  - Arteriosclerosis or other malign changes
  - Measurements of organ dimensions
  - Blood flow, etc.
- Supports medical research by providing new information: e.g., about the structure of the brain, or about the quality of medical treatments

Computer Vision Applications

Industry

Industry

- Information is extracted for the purpose of supporting a manufacturing process
- Quality Control where details or final products are being automatically inspected in order to find defects
- Measurement of position and orientation of details to be picked up by a robot arm
- Heavily used in agricultural process to remove undesirable food stuff from bulk material (Optical Sorting)

Computer Vision Applications

Other Application Areas

Other Application Areas

Other application areas include:

- Military
- Autonomous Vehicles
- Visual Effects creation for cinema and broadcast
- Surveillance
- Tracking and counting organisms in the biological sciences

Computer Vision Tasks

Computer Vision Tasks

- Tasks include methods for
  - Acquiring
  - Processing
  - Analyzing and understanding digital images
  - Extraction of high-dimensional data from the real world
- in order to produce numerical or symbolic information, e.g., in the forms of decisions

Computer Vision Tasks

Recognition

Recognition

- Classical problem in Computer Vision, Image Processing, and Machine Vision
- Determine whether or not the image data contains some specific object, feature, or activity
    - Object Recognition
    - Identification
    - Detection

Object Recognition

- a.k.a 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

Identification

- Individual instance of an object is recognized
- Identification of a specific person's face or fingerprint
- Identification of handwritten digits
- Identification of a specific vehicle

Detection

- Image data are scanned for a specific condition
- 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

Computer Vision Tasks

Specialized Tasks based on Recognition

Specialized Tasks based on Recognition

Several specialized tasks based on recognition exist, such as:

- Content-based Image Retrieval
- Pose Estimation
- Optical Character Recognition (OCR)
- 2D Code Reading
- Facial Recognition

Content-based Image Retrieval

- Finding all images in a larger set of images which have a specific content
- Content can be specified in different ways
  - In terms of similarity relative a target image
    - give me all images similar to image X
  - 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
- Assisting a robot arm in retrieving objects in an assembly line situation or picking parts

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

- Technology capable of identifying or verifying a person from a digital image or a video frame from a video source

Shape Recognition

- Shape Recognition Technology (SRT) in people counter systems differentiating human beings (head and shoulder patterns) from objects

Computer Vision Tasks

Motion analysis

Motion Analysis

- 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) 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.

Computer Vision Tasks

Scene Reconstruction

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

Computer Vision Tasks

Image Restoration

Image Restoration

- The aim 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 like, a model which distinguishes them from the 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

Computer Vision Typical Functions

Computer Vision Typical Functions

Typical functions which are found in many Computer Vision systems

1. Image Acquisition
2. Pre-Processing
3. Feature Extraction
4. Detection / Segmentation
5. High-Level Processing
6. Decision Making

Image Acquisition

- 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 color 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

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.

Pre-Processing Examples

- Re-sampling in order to assure that the image coordinate system is correct
- Noise reduction in order 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
- 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

Detection / Segmentation Examples

- Selection of a specific set of interest points
- Segmentation of one or multiple image regions which contain a specific object of interest
- Segmentation of image into nested scene architecture comprised 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
- Segmentation or co-segmentation of one or multiple videos into a series of per-frame foreground masks, while maintaining its temporal semantic continuity

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
- 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, 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 (IUS)

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
  - High level, includes: objects, scenes, or events

Summary

Summary

- Definition
- Distinctions
  - Computer Graphics
  - Image Processing and Image Analysis
- Applications
- Tasks
- Typical Functions
- Image Understanding Systems

https://www.haitham.ws