Video content analysis

From Wikipedia, the free encyclopedia

**Video content analysis** (also **Video content analytics**, **VCA**) is the capability of automatically analyzing [video](https://en.wikipedia.org/wiki/Video) to detect and determine temporal and spatial events.

This technical capability is used in a wide range of domains including entertainment,[[1]](https://en.wikipedia.org/wiki/Video_content_analysis#cite_note-KINECT-1) health-care, retail, automotive, transport, [home automation](https://en.wikipedia.org/wiki/Home_automation), flame and smoke detection, safety and security.[[2]](https://en.wikipedia.org/wiki/Video_content_analysis#cite_note-VCA_in_Security-2) The [algorithms](https://en.wikipedia.org/wiki/Algorithm) can be implemented as software on general purpose machines, or as hardware in specialized video processing units.

Many different functionalities can be implemented in VCA. Video Motion Detection is one of the simpler forms where motion is detected with regard to a fixed background scene. More advanced functionalities include [video tracking](https://en.wikipedia.org/wiki/Video_tracking) and [egomotion](https://en.wikipedia.org/wiki/Egomotion) estimation.

Based on the internal representation that VCA generates in the machine, it is possible to build other functionalities, such as [identification](https://en.wikipedia.org/wiki/Identification_of_human_individuals), [behavior](https://en.wikipedia.org/wiki/Behavior) analysis or other forms of [situation awareness](https://en.wikipedia.org/wiki/Situation_awareness).

VCA relies on good input video, so it is often combined with video enhancement technologies such as [video denoising](https://en.wikipedia.org/wiki/Video_denoising), [image stabilization](https://en.wikipedia.org/wiki/Image_stabilization), [unsharp masking](https://en.wikipedia.org/wiki/Unsharp_masking) and [super-resolution](https://en.wikipedia.org/wiki/Super-resolution).

**Contents**

  [hide]

* [1Functionalities](https://en.wikipedia.org/wiki/Video_content_analysis#Functionalities)
* [2Commercial applications](https://en.wikipedia.org/wiki/Video_content_analysis#Commercial_applications)
* [3Law Enforcement](https://en.wikipedia.org/wiki/Video_content_analysis#Law_Enforcement)
* [4Academic research](https://en.wikipedia.org/wiki/Video_content_analysis#Academic_research)
* [5References](https://en.wikipedia.org/wiki/Video_content_analysis#References)

Functionalities[[edit](https://en.wikipedia.org/w/index.php?title=Video_content_analysis&action=edit&section=1)]

Several articles provide an overview of the modules involved in the development of video analytic applications.[[3]](https://en.wikipedia.org/wiki/Video_content_analysis#cite_note-Nik_Gagvani-3)[[4]](https://en.wikipedia.org/wiki/Video_content_analysis#cite_note-Cheng_Peng-4) This is a list of known functionalities and a short description.

|  |  |
| --- | --- |
| **Function** | **Description** |
| Dynamic masking | Blocking a part of the video signal based on the signal itself, for example because of privacy concerns. |
| Flame and smoke detection | IP cameras with intelligent video surveillance technology can be used to detect flame and smoke in 15–20 seconds or even less because of the built-in DSP chip. The chip processes algorithms that analyzes the videos captured for flame and smoke characteristics such as color chrominance, flickering ratio, shape, pattern and moving direction. |
| [Egomotion estimation](https://en.wikipedia.org/wiki/Egomotion) | Egomotion estimation is used to determine the location of a camera by analyzing its output signal. |
| Motion detection | Motion detection is used to determine the presence of relevant motion in the observed scene. |
| Shape recognition | Shape recognition is used to recognize shapes in the input video, for example circles or squares. This functionality is typically used in more advanced functionalities such as object detection. |
| Object detection | Object detection is used to determine the presence of a type of object or entity, for example a person or car. Other examples include fire and smoke detection. |
| Recognition | [Face recognition](https://en.wikipedia.org/wiki/Facial_recognition_system) and [Automatic Number Plate Recognition](https://en.wikipedia.org/wiki/Automatic_Number_Plate_Recognition) are used to recognize, and therefore possibly identify, persons or cars. |
| Style detection | Style detection is used in settings where the video signal has been [produced](https://en.wikipedia.org/wiki/Filmmaking#Production), for example for television broadcast. Style detection detects the style of the production process.[[5]](https://en.wikipedia.org/wiki/Video_content_analysis#cite_note-Style_detection-5) |
| Tamper detection | Tamper detection is used to determine whether the camera or output signal is tampered with. |
| [Video tracking](https://en.wikipedia.org/wiki/Video_tracking) | Video tracking is used to determine the location of persons or objects in the video signal, possibly with regard to an external reference grid. |
| [Video Error Level Analysis](http://www.forensicprotection.com/VELA__Peer_review__by_Doug_Carner.pdf) | Video scene content tamper analysis using free software. Video [Error level analysis](https://en.wikipedia.org/wiki/Error_level_analysis) (VELA) concept by [Dr. Neal Krawetz](https://www.blackhat.com/html/bh-usa-07/bh-usa-07-speakers.html#Krawetz). |

Commercial applications[[edit](https://en.wikipedia.org/w/index.php?title=Video_content_analysis&action=edit&section=2)]

VCA is a new technology. New applications are frequently found, however the track record of different types of VCA differs widely. Functionalities such as motion detection and people counting are believed to be available as [commercial off-the-shelf](https://en.wikipedia.org/wiki/Commercial_off-the-shelf) products with a decent track-record, even freeware such as dsprobotics Flowstone can handle movement and color analysis.

In many domains VCA is implemented on [CCTV](https://en.wikipedia.org/wiki/Closed-circuit_television) systems, either distributed on the cameras (at-the-edge) or centralized on dedicated processing systems. Video Analytics and Smart CCTV are commercial terms for VCA in the security domain. In the UK the [BSIA](https://en.wikipedia.org/w/index.php?title=BSIA&action=edit&redlink=1) has developed an introduction guide for VCA in the security domain.[[6]](https://en.wikipedia.org/wiki/Video_content_analysis#cite_note-BSIA-6) In addition to video analytics and to complement it, audio analytics can also be used.[[7]](https://en.wikipedia.org/wiki/Video_content_analysis#cite_note-audioanalytics-7)

Independent video management software manufacturers are constantly expanding the range of video analytics modules available. With the new suspect tracking technology, it is then possible to track all of this subject's movements easily: where they came from, and when, where, and how they moved. Within a particular surveillance system, the indexing technology is able to locate people with similar features who were within the cameras’ viewpoints during or within a specific period of time. Usually, the system finds a lot of different people with similar features and presents them in the form of snapshots. The operator only needs to click on those images and subjects which need to be tracked. Within a minute or so, it’s possible to track all the movements of a particular person, and even to create a step-by-step video of the movements.

[Kinect](https://en.wikipedia.org/wiki/Kinect) is an add-on peripheral for the Xbox 360 gaming console that uses VCA for part of the user input.[[8]](https://en.wikipedia.org/wiki/Video_content_analysis#cite_note-Natal_101-8)

In retail industry, VCA is used to track shoppers inside the store.[[9]](https://en.wikipedia.org/wiki/Video_content_analysis#cite_note-9) By this way, a heatmap of the store can be obtained, which is beneficial for store design and marketing optimisations.

The quality of VCA in the commercial setting is difficult to determine. It depends on many variables such as [use case](https://en.wikipedia.org/wiki/Use_case), [implementation](https://en.wikipedia.org/wiki/Implementation), [system configuration](https://en.wikipedia.org/wiki/Computer_configuration) and [computing platform](https://en.wikipedia.org/wiki/Computing_platform). Typical methods to get an objective idea of the quality in commercial settings include independent [benchmarking](https://en.wikipedia.org/wiki/Benchmarking)[[10]](https://en.wikipedia.org/wiki/Video_content_analysis#cite_note-i-Lids-10) and designated test locations.

VCA has been used for [crowd management](https://en.wikipedia.org/wiki/Crowd_control) purposes, notably at [The O2 Arena](https://en.wikipedia.org/wiki/The_O2_Arena) in London and [The London Eye](https://en.wikipedia.org/wiki/The_London_Eye).[[11]](https://en.wikipedia.org/wiki/Video_content_analysis#cite_note-11)

Law Enforcement[[edit](https://en.wikipedia.org/w/index.php?title=Video_content_analysis&action=edit&section=3)]

Police and forensic scientists analyse CCTV video when investigating criminal activity. Police use software which performs video content analysis to search for key events in video and find suspects. Surveys have shown that up to 75% of cases involve CCTV. Police use video content analysis software to search long videos for important events.[[12]](https://en.wikipedia.org/wiki/Video_content_analysis#cite_note-12)[[13]](https://en.wikipedia.org/wiki/Video_content_analysis#cite_note-13)

Academic research[[edit](https://en.wikipedia.org/w/index.php?title=Video_content_analysis&action=edit&section=4)]

Video content analysis is a subset of [computer vision](https://en.wikipedia.org/wiki/Computer_vision) and thereby of [artificial intelligence](https://en.wikipedia.org/wiki/Artificial_intelligence). Two major academic benchmark initiatives are [TRECVID](https://en.wikipedia.org/wiki/TRECVID),[[14]](https://en.wikipedia.org/wiki/Video_content_analysis#cite_note-TRECVID-14) which uses a small portion of i-LIDS video footage, and the PETS Benchmark Data.[[15]](https://en.wikipedia.org/wiki/Video_content_analysis#cite_note-PETS_Benchmark_Data-15) They focus on functionalities such as tracking, left luggage detection and virtual fencing.

The [EU](https://en.wikipedia.org/wiki/EU) is funding a [FP7](https://en.wikipedia.org/wiki/Seventh_Framework_Programme#Framework_Programme_7) project called P-REACT [[16]](https://en.wikipedia.org/wiki/Video_content_analysis#cite_note-P-REACT-16) to integrate video content analytics on embedded systems with police and transport security databases.[[17]](https://en.wikipedia.org/wiki/Video_content_analysis#cite_note-17) [[18]](https://en.wikipedia.org/wiki/Video_content_analysis#cite_note-18)