

G06V

IMAGE OR VIDEO RECOGNITION OR UNDERSTANDING

Definition statement

This place covers:

Higher-level interpretation and cognition of images or videos, which includes pattern recognition, pattern learning and semantic interpretation as fundamental aspects. These aspects involve the detection, categorisation, identification, authentication of image or video patterns by acquiring, preprocessing, extracting distinctive features or matching, supervised or unsupervised clustering or classification of these features or representations derived from them leading to one or several decisions, related confidence values (e.g. probabilities) or classification/clustering labels for explanatory purposes or to derive a certain meaning.

Pattern recognition or pattern learning in a specific, image or video-related context that includes:

- scene-related patterns and scene-specific elements - [G06V 20/00](#)
- character recognition or recognising digital ink; document-oriented image-based pattern recognition – [G06V 30/00](#)
- human-related, animal-related or biometric patterns in in image or video data [G06V 40/00](#).

Typical image or video understanding systems that include one or more of the following steps:

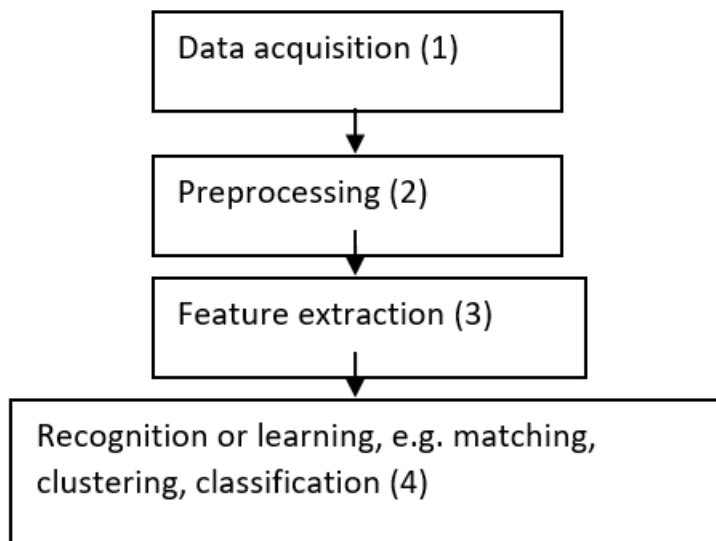


Figure 1. Processing steps involved in a pattern recognition system

These steps are provided as follows: image acquisition ([G06V 10/10](#)), image preprocessing ([G06V 10/20](#)), feature extraction ([G06V 10/40](#)), pattern recognition or pattern learning, e.g. matching, clustering or classifying techniques ([G06V 10/70](#)).

Image or video recognition can be carried out by using electronic means ([G06V 10/70](#)) or by using optical means ([G06V 10/88](#)).

Typically, a pattern recognition system involves one or more of the following techniques:

	Individual data entity	Groups of data entities (classes)
One data sample	Authentication	Categorisation
Several data samples	Identification	Clustering

When a collection of data samples is provided, identification means selecting a particular sample having a (predefined) characteristic which distinguishes it from the others. Several data samples are generally matched against the one to be identified in a many-to-one process.

Authentication involves verifying the identity of a sample using a test of genuineness and it involves a one-to-one comparison with the genuine (authentic) sample.

Categorisation means assigning a data sample to a class according to certain distinguishing properties (or characteristics) of that class and it generally involves a one-to-many test in which one data sample is compared with the characteristics of several classes.

Clustering means grouping data samples in groups or classes based on their properties (or characteristics) and it generally involves a many-to-many (dis)similarity test.

Relationships with other classification places

Data mining and pattern recognition techniques in general are classified in [G06F 18/00](#).

Some techniques of image or video understanding performed in the preprocessing step (Fig. 1, elem. 2) — which start with a bitmap image as an input and derive a non-bitmap representation from it — can also be encountered in general image analysis. If these techniques do not involve one of the functions of image or video pattern authentication, identification, categorisation or clustering, classification should be made only in the appropriate subgroups of subclass [G06T](#). Some examples of these techniques are: general methods for image segmentation, e.g. obtaining contiguous image regions with similar pixels, for position and size determination of an object without establishing its identity, for calculating the motion of an image region corresponding to an object irrespective of the identity of the object, for camera calibration, etc.

Biometric recognition provided under [G06V 40/00](#) does not cover techniques for medical image inspection or medical diagnosis.

Techniques based on coding, decoding, compressing or decompressing digital video signals using video object coding are provided under [H04N 19/20](#).

Velocity or trajectory determination systems or sense-of-movement determination systems using radar, sonar or lidar are provided for under [G01S 13/58](#), [G01S 15/58](#), [G01S 17/58](#), respectively.

General purpose image data processing, in particular image watermarking is provided under [G06T 1/00](#), while selective content distribution, such as generation or processing of protective or descriptive data associated with content involving watermarking is covered by [H04N 21/8358](#).

General purpose image data acquisition and related pre-processing using digital cameras and processing used to control digital cameras is provided under [H04N 5/00](#).

Play-back, editing, or synchronising of a music score, including interpretation therefor, as well as transmission of a music score between systems of musical instruments for play-back, editing or synchronising is provided under [G10H](#).

References

Application-oriented references

Examples of places where the subject matter of this place is covered when specially adapted, used for a particular purpose, or incorporated in a larger system:

Detecting, measuring and recording for medical diagnostic purposes	A61B 5/00
Identifications of persons in medical applications	A61B 5/117
Postal sorting	B07C 3/10
Input arrangements for interaction between user and computer	G06F 3/01

Application-oriented references

Testing to determine the identity or genuineness of paper currency or similar valuable papers	G07D 7/00
---	---------------------------

Informative references

Attention is drawn to the following places, which may be of interest for search:

Programme-controlled manipulators	B25J 9/00
Optical viewing arrangements in vehicles	B60R 1/00
Testing machines or structures	G01M
Investigating or analysing materials by determining their chemical or physical properties	G01N
Radio direction-finding; Radio navigation; Determining distance or velocity by use of radio waves; Locating or presence-detecting by use of the reflection or reradiation of radio waves; Analogous arrangements using other waves	G01S
Geophysics	G01V
Optical elements, systems, or apparatus	G02B
Photomechanical production of textured or patterned surfaces, e.g. for printing, for processing of semiconductor devices	G03F
Control or regulating systems in general	G05B
Input arrangements for transferring data to be processed into a form capable of being handled by the computer; Output arrangements for transferring data from processing unit to output unit, e.g. interface arrangements	G06F 3/00
Content-based image retrieval	G06F 16/50
Fourier, Walsh or analogous domain transformations	G06F 17/14
Security arrangements for protecting computer systems against unauthorised activity	G06F 21/00
Authentication of user input in security arrangements for computers	G06F 21/30
Computer-aided design	G06F 30/00
Handling natural language data	G06F 40/00
Methods or arrangements for sensing record carriers	G06K 7/00
Record carriers for use with machines and with at least a part designed to carry digital markings	G06K 19/00
Computer systems based on specific computational models	G06N
Data processing for business purposes, logistics, stock management	G06Q
General purpose image data processing, e.g. general purpose image acquisition, image watermarking or specific image analysis processor architectures or configurations	G06T 1/00
Geometric image transformation in the plane of the image, e.g. rotation of a whole image or part thereof	G06T 3/00
Image enhancement or restoration	G06T 5/00
Image analysis in general	G06T 7/00
Image analysis, in particular analysis of motion using feature-based methods	G06T 7/246

Image analysis, in particular determination of transform parameters for the alignment of images using feature-based methods	G06T 7/33
Image analysis of texture	G06T 7/40
Image analysis, in particular depth or shape recovery	G06T 7/50
Image analysis, in particular determining position and orientation of objects using feature-based methods	G06T 7/73
Image analysis, in particular determination of colour characteristics	G06T 7/90
Image coding, e.g. from bit-mapped to non bit-mapped	G06T 9/00
Contour coding, e.g. using detection of edges	G06T 9/20
Two-dimensional image generation	G06T 11/00
Three-dimensional image rendering	G06T 15/00
Lighting effects in 3D image rendering	G06T 15/50
Three-dimensional modelling, e.g. data description of 3D objects	G06T 17/00
Manipulating 3D models or images for computer graphics	G06T 19/00
Individual entry and exit registers	G07C 9/00
Intruder alarms using image scanning and comparing means	G08B 13/194
Burglar, theft or intruder alarm	G08B 13/196
Traffic control systems for road vehicles	G08G 1/00
Labels, tag tickets or similar identification or indication means	G09F 3/00
Speech recognition	G10L 15/00
Speaker recognition	G10L 17/00
Bioinformatics	G16B
Chemoinformatics and computational material science	G16C
Healthcare informatics	G16H
Semiconductor devices	H01L
Secret or secure communication	H04L 9/00
Scanning, transmission or reproduction of documents, e.g. facsimile transmission	H04N 1/00
Studio circuitry for television systems	H04N 5/222
Closed circuit television systems	H04N 7/18
Methods or arrangements for coding, decoding, compressing or decompressing digital video signals using video object coding	H04N 19/20
Methods or arrangements for coding, decoding, compressing or decompressing digital video signals, region motion estimation for predictive coding	H04N 19/543

Special rules of classification

The subject-matter classified in the application-oriented groups [G06V 20/00](#) - [G06V 40/00](#) is also classified in the groups [G06V 10/10](#) (Image acquisition) or [G06V 10/20](#) (Image preprocessing) when specific adaptations of the application-related context occur at the stages of image acquisition or, respectively, image preprocessing.

Pattern recognition or pattern learning techniques for images or video understanding involving feature extraction (Fig. 1, elem. 3) or matching, clustering or classification (Fig. 1, elem. 4) should be classified in [G06V 10/40](#) or [G06V 10/70](#) in general, thus irrespective whether an application-related context provided by the groups [G06V 20/00](#) - [G06V 40/00](#) exists.

Glossary of terms

In this place, the following terms or expressions are used with the meaning indicated:

Pattern	Data with inherent variability, or a representation derived from it, having some explanatory characteristic or a meaning, e.g. an object depicted in an image
Image and video understanding	Techniques for semantic interpretation, pattern recognition or pattern learning specifically applied to images and videos
Feature extraction	Deriving descriptive or quantitative measures from data
Clustering	Grouping or separating patterns according to their (dis)similarity or closeness
Classification	Assigning labels to patterns