**Image processing:** is any form of signal processing for which the input is an image and the output of image processing can be either an image or a set of characteristics or parameters related to the image.

Some of the image processing operations are:

* Euclidean geometry transformations such as enlargement, reduction, and rotation
* Color corrections such as brightness and contrast adjustments, quantization, or color translation to a different color space
* Digital compositing or optical compositing
* Interpolation, [demosaicing](http://en.wikipedia.org/wiki/Demosaicing), and recovery of a full image from a [raw image format](http://en.wikipedia.org/wiki/Raw_image_format) using a Bayer filter pattern
* Image registration, the alignment of two or more images
* Image differencing and [morphing](http://en.wikipedia.org/wiki/Morphing)
* [Image recognition](http://en.wikipedia.org/wiki/Image_recognition), for example, extract the text from the image by using optical character recognition
* Image segmentation
* High dynamic range imaging by combining multiple images
* Geometric hashing for 2-D object recognition with affine invariance

**Euclidean geometry transformations:** In geometry, an **affine transformation** or **affine map** or an **affinity** (from the Latin, *affinis*, "connected with") between two vector spaces consists of a linear transformation followed by a translation: 

In the finite-dimensional case each affine transformation is given by a matrix A and a vector *b*, satisfying certain properties described below.

Geometrically, an affine transformation in Euclidean space is one that preserves

* The [collinearity](http://en.wikipedia.org/wiki/Collinearity) relation between points; i.e., three points which lie on a line continue to be collinear after the transformation
* Ratios of distances along a line; i.e., for distinct collinear points *p*1, *p*2, *p*3, the ratio | *p*2 − *p*1 | / | *p*3 − *p*2 | is preserved

**Enlargement:** In Euclidean geometry, **uniform scaling** is a linear transformation that enlarges or increases or diminishes objects; the scale factor is the same in all directions; it is also called a homothety. The result of uniform scaling is similar (in the geometric sense) to the original. A scale factor of 1 is normally allowed, so that congruent shapes are also classed as similar.

More general is scaling with a separate scale factor for each axis direction. Non-uniform or anisotropic scaling is obtained when at least one of the scaling factors is different from the others; a special case is directional scaling (in one direction). Non-uniform scaling changes the shape of the object.



**Rotation:** is a transformation in a plane or in space that describes the motion of a rigid body around a fixed point. A rotation and the above-mentioned transformations are isometries; they leave the distance between any two points unchanged after the transformation.

Rotating a vector or coordinate pair counterclockwise about the origin is the same as rotating the plane or axes clockwise about the origin.



**Color correction:**  by using color gels, or filters, is a process used in stage lighting, photography, television, typically the light color is measured on a scale known as color temperature, as well as along a green–magenta axis orthogonal to the color temperature axis.

**Brightness:** is an attribute of visual perception in which a source appears to be radiating or reflecting light. In other words, brightness is the perception elicited by the luminance of a visual target. In the RGB color space, brightness can be thought of as the arithmetic mean *μ* of the red, green, and blue color coordinates. 

**Contrast:** is the difference in visual properties that makes an distinguishable from other objects and the background. In visual perception of the real world, contrast is determined by the difference in the color and brightness of the object and other objects within the same field of view.

**Quantization:** is the process of approximating ("mapping") a continuous range of values (or a very large set of possible discrete values) by a relatively small ("finite") set of ("values which can still take on continuous range") discrete symbols or integer values.

Quantization is referred to as scalar quantization, since it operates on scalar input data. In general, a scalar quantization operator can be represented as: 

Where an integer results that is sometimes referred to as the *quantization index*.

**Color Translation:** or color space conversion, is the translation of the representation of a color from one color space to another. This calculation is required whenever data is exchanged inside a color-managed chain. Transforming profiled color information to different output devices is achieved by referencing the profile data into a standard color space.

**Digital compositing:** is the process of digitally assembling multiple images to make a final image, typically for print, motion pictures or screen display. It is the evolution into the digital realm of optical film compositing.

The basic operation used is known as 'alpha blending', where an opacity value, 'α' is used to control the proportions of two input pixel values that end up a single output pixel.

**Image registration:** sets of data acquired by sampling the same scene or object at different times, or from different perspectives, will be in different coordinate systems. It is the process of transforming the different sets of data into one coordinate system. Registration is necessary in order to be able to compare or integrate the data obtained from different measurements.

**Image differencing:** is an image processing technique used to determine changes between images. The difference between two images is calculated by finding the difference between each pixel in each image, and generating an image based on the result. For this technique to work, the two images must first be aligned so that corresponding points coincide. The complexity of the pre-processing needed before differencing varies with the type of image.

**Image 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.

* **Recognition:** 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:** An individual instance of an object is recognized.
* **Detection:** the image data is scanned for a specific condition.

**Image segmentation:** refers to the process of partitioning a digital image into multiple segments. The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. More precisely, image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain visual characteristics.

**High dynamic range imaging:** is a set of techniques that allow a greater dynamic range of luminance’s between light and dark areas of a scene than normal digital imaging techniques or photographic prints. This wider dynamic range allows HDR images to represent more accurately the wide range of intensity levels found in real scenes ranging from direct sunlight to faint starlight.

**Geometric hashing:** is a method for efficiently finding two-dimensional objects represented by discrete points that have undergone an affine transformation. In an off-line step, the objects are encoded by treating each triple of points as a geometric basis. The remaining points can be represented in an invariant fashion with respect to this basis using two parameters. All such quadruples of object points are stored in a two-dimensional table that represents a discretization of these parameters. In the on-line step, randomly selected triples of data points (for example, from an image) are considered as candidate bases. For each candidate basis, the remaining data points are encoded according the basis and possible correspondences from the object are found in the previously constructed table. The candidate basis is accepted if a sufficiently large number of the data points index a consistent object basis.

**Pattern Recognition:** the act of taking in raw data and taking an action based on the “category”.

The process of recognizing in some real world problem of recognition involves these procedure steps:

First the process needs features for the objects to be recognized in order to be classified.

**Model:** different descriptions, which are typically mathematical in form. The overarching goal and approach in pattern classification is to hypothesize the class of these models, process the sensed data to eliminate noise (not due to the models), and for any sensed pattern choose the model that corresponds best.

**Pre-processing & Segmentation:** signals are preprocessed to simplify subsequent operations without losing relevant processing information. In particular, we might use a segmentation operation in which the images of different objects are somehow isolated from one another and from the background.

**Feature Extraction:** purpose is to reduce the data by measuring certain “features” or “properties.” These features (or, more precisely, the values of these features) are then passed to a *classifier* that evaluates the evidence presented and makes a final decision as to the species.

**Training Samples:** Choose an obvious feature, and attempt to classify the objects merely by seeing whether or not the feature *l* of an object exceeds some critical value *l∗*. To choose *l∗* we could obtain some *design* or training of the different types of fish, (somehow) make length measurements, samples and inspect the results.

**Cost:** Although sometimes we can have same cost of decision on different object types, sometimes this isn’t desirable so in these cases we need to have different measures of cost for different object types.

**Decision Theory:** considerations suggest that there is an overall single cost associated with our decision, and our true task is to make a decision rule (i.e., set a decision boundary) so as to minimize such a cost. This is the central task of *decision theory* of which decision pattern classification is perhaps the most important subfield.

**Decision Boundary:** This the problem where we have to partition the feature space into the number of classes existing for different objects, where for all patterns in each region we will call the objects with the class name of that region. So with this process is done the classification of objects based on a decision boundary on the definition of the classes that tell which pattern matches their features.

**Generalization:** In a situation where we have a wide feature space and a complex decision boundary, and in many cases many of the features doesn’t have any effect in the classification, also another implication is that there are cases where too many features doesn’t allow to recognize unknown patterns of objects, as a solution to this problem is the generalization approach where we remove features that don’t have any effect in the classification of the objects and those that make impossible to classify unknown patterns.

Rather, then, we might seek to “simplify” the recognizer, motivated by a belief that the underlying models will not require a decision boundary that is as complex. Indeed, we might be satisfied with the slightly poorer performance on the training samples if it means that our classifier will have better performance on novel patterns.

**Analysis by synthesis:** A central technique, when we have insufficient training data, is to incorporate knowledge of the problem domain. Indeed the less the training data the more important is such knowledge, for instance how the patterns themselves were produced. One analysis method that takes this notion to its logical extreme is that of *analysis by synthesis*, by synthesis where in the ideal case one has a model of how each pattern is generated.

**Recognition of 2D patterns:** The use of moment invariants as features for identification and inspection of 2D shape has received much attention. The geometrical moments of an image are integrals of the image function over space, and the image can be uniquely determined by its geometrical moments of all orders. Low-order moments are relatively more stable than high-order moments. However, low-order moments can be used to differentiate between images of real objects only if their shapes are significantly different. Since these moments are designed to capture global information about the image, they are not suitable for classifying similar objects when corrupted by a significant amount of random noise. In fact, the geometrical moments are sensitive to digitization error, minor shape deformations, camera non-linearity, and non-ideal position of camera. An important step in 2D recognition is the process of feature extraction which has a major impact in the results of recognition on later steps.

In order to reduce the problem of feature extraction from a 2D image object to that from a 1D sequence: 

where **. Note that *Sq*(*r*) is now a 1D sequence of variable *r*. It is important to note from Eq. (2) that if *gp*(*r*) is defined on the whole domain of variable *r*, then *Fpq* is a *global feature*, on the other hand, if the function of *gp*(*r*) is locally defined, then *Fpq* may be seen as a *local feature*.

**Intensity & Color:** is equal to the amount of light reflected by the corresponding point on the object in the direction of the viewer, multiplied by some constant factor that depends on the parameters of the image-forming system. To be precise, we have to think of intensity as light flux per unit area and correspondingly also have to consider the reflected light per unit area as seen by the viewer. Light entering the human visual system originates either from a self-luminous source or from light reflected from some object or from light transmitted through some translucent object. Where E is the spectral energy distribution, of light emitted from some primary light source, t and r denote the wavelength-dependent transmissivity and reflectivity, respectively, of an object. For a *transmissive object,* the observed light spectral energy distribution is: and for *reflective object* is: 

The attribute of light that distinguishes a red light from a green light or a yellow light, for example, is called the *hue* of the light. If two light sources with the same spectral energy distribution are observed under identical conditions, they will appear to possess the same hue. However, it is possible to have two light sources with different spectral energy distributions that are perceived identically. Such lights are called *metameric pairs.* The third perceptual descriptor of a colored light is its *saturation*, the attribute that distinguishes a spectral light from a pastel light of the same hue. In effect, saturation describes the whiteness of a light source.

**Types of implementation applications/tasks:** The area of speech recognition is a multidisciplinary field involving many other techniques as for e.g.:

* Signal processing
* Physics(acoustics)
* Pattern recognition
* Communication and information theory
* Linguistics
* Physiology
* Computer science, and
* Psychology

Examples of speech recognition systems that can handle tasks like Voice Control and Dialog. It consists of many parts that do specific jobs that are (speech recognizer, language analyzer, expert system, physical system being controlled by voice commands, and a text to speech synthesizer).

***Speech recognizer:*** converts spoken input into grammatically correct text as considered by the recognizer vocabulary and grammar model.

***Language analyzer:*** can be implemented as a semantic concept spotter using statistical methods or as a rule based system.

***Expert system:*** the decoded input is sent to the expert system which first selects a desired action then it issues a command to the physical system under voice control then receives data from the command results and after it, it converts the text into a speech using the text to speech synthesizer and transmits the voice results.

It must be taken into account also when we should implement speech recognition systems; here are some of the characteristics to follow:

* It must provide real benefit to the user in form of increased productivity, ease of use, better machine human communication or more a natural mode of communication.
* It must be user friendly in a way that the user should be comfortable with the voice dialogue, and provide helpful voice prompts.
* It must be accurate, in a way that it must achieve at least a specified performance on the task associated with the recognition decision, although this is a non-linear decision platform in which if the system reaches an acceptable result in total than it is considered successful.

The area where this system applies is in many diverse fields as for e.g.:

* Office or business systems, such as data entry onto forms, database management and control, etc.
* Manufacturing. It is used to provide eyes-free, hands-free monitoring process for manufacturing processes.
* Telephone or telecommunications.
* Medical. Voice creation and editing of specialized medical reports.
* Others. Including voice controlled and operated games and toys, voice recognition aid for handicapped people, and voice control for nonstrategic functions in a moving vehicle (such as air conditioning, audio system, etc).

**Document analysis:** it is used to encode documents for computer processing; it also has an effect in data compression since scanned documents saved as bitmaps use much more space than text documents. Here blocks containing text can be analyzed with a pattern-matching program to group similar symbols from the document and creates a prototype pattern to represent each group. The prototype patterns are then identified either manually using an interactive display or by automatic recognition logic if this is available for the fonts used in the document. Another important issue when considering document analysis is layout analysis, because we need to take different approaches when the text is in multiple columns, if the document is needed to be reproduced exactly then it’s needed also to extract line spaces, etc.

Most of the methods for document analysis consist of two main steps: first, segmentation procedure of dividing the document in sub-regions (text, images, etc). Next some basic features of these blocks are calculated, in this way a linear classifier is created that adapts to different heights of fonts and also that discriminates between text and images.

**Biometric recognition:** A biometric system is essentially a pattern-recognition system that recognizes a person based on a feature vector derived from a specific physiological or behavioral characteristic that the person possesses. A biometric system based on physiological characteristics is generally more reliable than one which adopts behavioral characteristics, even if the latter may be easier to integrate within certain specific applications. Biometric system can than operate in two modes: verification or identification. While identification involves comparing the acquired biometric information against templates corresponding to all users in the database, verification involves comparison with only those templates corresponding to the claimed identity.

A simple biometric system consists of four basic components:

1. *Sensor module* which acquires the biometric data;
2. *Feature extraction module* where the acquired data is processed to extract feature vectors;
3. *Matching module* where feature vectors are compared against those in the template;
4. *Decision-making module* in which the user's identity is established or a claimed identity is accepted or rejected.

There are number of biometric methods in use:

**Infrared thermogram** (facial, hand or hand vein): It is possible to capture the pattern of heat radiated by the human body with an infrared camera. That pattern is considered to be unique for each person.

**Gait**: Basically, gait is the peculiar way one walks and it is a complex spatio-temporal biometrics. It is not supposed to be very distinctive but can be used in some low-security applications.

**Keystroke**: It is believed that each person types on a keyboard in a characteristic way. This is also not very distinctive but it offers sufficient discriminatory information to permit identity verification.

**Odor**: Each object spreads around an odor that is characteristic of its chemical composition and this could be used for distinguishing various objects. This would be done with an array of chemical sensors, each sensitive to a certain group of compounds.

**Ear**: It has been suggested that the shape of the ear and the structure of the cartilaginous tissue of the pinna are distinctive. Matching the distance of salient points on the pinna from a landmark location of the ear is the suggested method of recognition in this case. This method is not believed to be very distinctive.

**Hand geometry**: The essence of hand geometry is the comparative dimensions of fingers and the location of joints, shape and size of palm. Dry weather or individual anomalies such as dry skin do not appear to have any negative effects on the verification accuracy. Since hand geometry is not very distinctive it cannot be used for identification of an individual from a large population, but rather in a verification mode. Further, hand geometry information may not be invariant during the growth period of children. Limitations in dexterity (arthritis) or even jewelry may influence extracting the correct hand geometry information.

**Fingerprint**: A fingerprint is a pattern of ridges and furrows located on the tip of each finger. Fingerprint recognition for identification acquires the initial image through live scan of the finger by direct contact with a reader device that can also check for validating attributes such as temperature and pulse. Since the finger actually touches the scanning device, the surface can become oily and cloudy after repeated use and reduce the sensitivity and reliability of optical scanners. The matching process involves comparing the two-dimensional minutiae patterns extracted from the user's print with those in the template.

**Face:** Facial images are the most common biometric characteristic used by humans to make a personal recognition, hence the idea to use this biometric in technology. This is a nonintrusive method and is suitable for covert recognition applications. The applications of facial recognition range from static ("mug shots") to dynamic, uncontrolled face identification in a cluttered background (subway, airport). Face verification involves extracting a feature set from a two-dimensional image of the user's face and matching it with the template stored in a database. The most popular approaches to face recognition are based on either: 1) the location and shape of facial attributes such as eyes, eyebrows, nose, lips and chin, and their spatial relationships, or 2) the overall (global) analysis of the face image that represents a face as a weighted combination of a number of canonical faces.

**Retina**: Retinal recognition creates an "eye signature" from the vascular configuration of the retina which is supposed to be a characteristic of each individual and each eye, respectively. Since it is protected in an eye itself, and since it is not easy to change or replicate the retinal vasculature, this is one of the most secure biometric. Image acquisition requires a person to look through a lens at an alignment target; therefore it implies cooperation of the subject. Also retinal scan can reveal some medical conditions and as such public acceptance is questionable.

**Iris**: Its complex pattern can contain many distinctive features such as arching ligaments, furrows, ridges, crypts, rings, corona, freckles and a zigzag collarette. Iris scanning is less intrusive than retinal because the iris is easily visible from several meters away. Responses of the iris to changes in light can provide an important secondary verification that the iris presented belongs to a live subject. Irises of identical twins are different, which is another advantage

**Palm-print:** Palms of the human hands contain unique pattern of ridges and valleys. Since palm is larger than a finger, palm-print is expected to be even more reliable than fingerprint. Palm-print scanners need to capture larger area with similar quality as fingerprint scanners, so they are more expensive. A highly accurate biometric system could be combined by using a high-resolution palm-print scanner that would collect all the features of the palm such as hand geometry, ridge and valley features, principal lines, and wrinkles.

**Voice**: The features of an individual's voice are based on physical characteristics such as vocal tracts, mouth, nasal cavities and lips that are used in creating a sound. These characteristics of human speech are invariant for an individual, but the behavioral part changes over time due to age, medical conditions and emotional state. Voice recognition techniques are generally categorized according to two approaches: 1) Automatic Speaker Verification (ASV) and 2) Automatic Speaker Identification (ASI). Speaker verification uses voice as the authenticating attribute in a two-factor scenario. Speaker identification attempts to use voice to identify who an individual actually is. Voice recognition distinguishes an individual by matching particular voice traits against templates stored in a database. Feature extraction typically measures formants or sound characteristics unique to each person's vocal tract.

**Signature**: Signature is a simple, concrete expression of the unique variations in human hand geometry. The way a person signs his or her name is known to be characteristic of that individual. Collecting samples for this biometric includes subject cooperation and requires the writing instrument. Signatures are a behavioral biometric that change over a period of time and are influenced by physical and emotional conditions of a subject. In addition to the general shape of the signed name, a signature recognition system can also measure pressure and velocity of the point of the stylus across the sensor pad.

**DNA**: Is probably the most reliable biometrics. It is in fact a one-dimensional code unique for each person. Exceptions are identical twins. This method, however, has some drawbacks: 1) contamination and sensitivity, since it is easy to steal a piece of DNA from an individual and use it for an ulterior purpose, 2) no real-time application is possible because DNA matching requires complex chemical methods involving expert's skills, 3) privacy issues since DNA sample taken from an individual is likely to show susceptibility of a person to some diseases. All this limits the use of DNA matching to forensic applications. It is obvious that no single biometric is the "ultimate" recognition tool and the choice depends on the application.