- A*: A search technique that performs best-first searching based on an evaluation function that combines the cost so far and the estimated cost to the goal. [WP:A*_search_algorithm]
- *a posteriori* probability: Literally, "after" probability. It is the probability p(s|e) that some situation *s* holds after some evidence *e* has been observed. This contrasts with the <u>*a priori*</u> probability, p(s), the probability of *s* before any evidence is observed. Bayes' rule is often used to compute the *a posteriori* probability from the *a priori* probability and the evidence. See also posterior distribution. [JKS95:15.5]
- *a priori* probability: A probability distribution that encodes an agent's beliefs about some uncertain quantity before some evidence or data is taken into account. See also prior distribution. [Bis06:1.2.3]
- aberration: Problem exhibited by a lens or a mirror whereby unexpected results are obtained. Two types of aberration are commonly encountered: <u>chromatic aberration</u>, where different frequencies of light focus at different positions:



and spherical aberration, where light passing through the edges of a lens (or mirror) focuses at slightly different positions. [FP03:1.2.3]

- **absolute conic**: The conic in 3D projective space that is the intersection of the unit (or any) sphere with the plane at infinity. It consists only of complex points. Its importance in computer vision is because of its role in the problem of <u>autocalibration</u>: the image of the absolute conic (IAC), a 2D conic, is represented by a 3×3 matrix ω that is the inverse of the matrix KK^T, where K is the matrix of the internal parameters for <u>camera</u> <u>calibration</u>. Subsequently, identifying ω allows the camera calibration to be computed. [FP03:13.6]
- **absolute coordinates**: Generally used in contrast to *local* or *relative* coordinates. A coordinate system that is referenced to some external datum. For example, a pixel in a satellite image might be at (100,200) in image coordinates, but at (51:48:05N, 8:17:54W) in georeferenced absolute coordinates. [JKS95:1.4.2]
- **absolute orientation**: In photogrammetry, the problem of <u>registration</u> of two corresponding sets of 3D points. Used to register a photogrammetric reconstruction to some <u>absolute</u> coordinate system. Often expressed as the problem of determining the rotation **R**, translation \vec{t} and scale *s* that best transforms a set of *model* points $\{\vec{m}_1, \ldots, \vec{m}_n\}$ to corresponding data points $\{\vec{d}_1, \ldots, \vec{d}_n\}$ by minimizing the least-squares error

$$\epsilon(R, \vec{t}, s) = \sum_{i=1}^{n} \|\vec{d}_i - s(\mathbf{R}\vec{m}_i + \vec{t})\|^2$$

Dictionary of Computer Vision and Image Processing, Second Edition.

R. B. Fisher, T. P. Breckon, K. Dawson-Howe, A. Fitzgibbon, C. Robertson, E. Trucco and C. K. I. Williams. © 2014 John Wiley & Sons, Ltd. Published 2014 by John Wiley & Sons, Ltd.

to which a solution may be found by using singular value decomposition. [JKS95:1.4.2]

absolute point: A 3D point defining the origin of a coordinate system. [WP: Cartesian_coordinate_system]

absolute quadric: The symmetric 4×4

rank 3 matrix $\Omega = \begin{pmatrix} \mathbf{I}_3 & \mathbf{\vec{0}}_3 \\ \mathbf{\vec{0}}_3^{\top} & \mathbf{0} \end{pmatrix}$. Like the

absolute conic, it is defined to be invariant under Euclidean transformations, is rescaled under similarities, takes the

form $\mathbf{\Omega} = \begin{pmatrix} \mathbf{A}^{\mathsf{T}}\mathbf{A} & \vec{0}_3 \\ \vec{0}_3^{\mathsf{T}} & 0 \end{pmatrix}$ under affine

transforms and becomes an arbitrary 4×4 rank 3 matrix under projective transforms. [FP03:13.6]

- **absorption**: Attenuation of light caused by passing through an optical system or being incident on an object surface. [Hec87:3.5]
- accumulation method: A method of accumulating evidence in <u>histogram</u> form, then searching for peaks, which correspond to hypotheses. See also <u>Hough transform</u> and <u>generalized</u> Hough transform. [Low91:9.3]
- accumulative difference: A means of detecting motion in image sequences. Each frame in the sequence is compared to a reference frame (after registration if necessary) to produce a difference image. Thresholding the difference image gives a binary motion mask. A counter for each pixel location in the accumulative image is incremented every time the difference between the reference image and the current image exceeds some threshold. Used for change detection. [JKS95:14.1.1]
- accuracy: The error of a value away from the true value. Contrast this with <u>precision</u>. [WP:Accuracy_and_ precision]
- acoustic sonar: Sound Navigation And Ranging. A device that is used primarily for the detection and location of objects (e.g., underwater or in air, as in mobile robotics, or internal to a human body, as in medical <u>ultrasound</u>) by reflecting and <u>intercepting</u>

acoustic waves. It operates with acoustic waves in a way analogous to that of <u>radar</u>, using both the time of flight and Doppler effects, giving the radial component of relative position and velocity. [WP:Sonar]

- ACRONYM: A vision system developed by Brooks that attempted to recognize three-dimensional objects from twodimensional images, using <u>generalized</u> <u>cylinder</u> primitives to represent both stored model and objects extracted from the image. [Nev82:10.2]
- action cuboid: The 3D <u>spatio-temporal</u> <u>space in which an action detection</u> may be localized in a video sequence:



Analogous to a window (or region of interest) in which an object detection may be localized within a 2D image. [GX11:6.4]

- action detection: An approach to the automated <u>detection</u> of a given human, vehicle or animal <u>activity</u> (action) from imagery. Most commonly carried out as a <u>video analysis</u> task due to the temporal nature of actions. [Sze10:12.6.4]
- action localization: An approach to in-image or in-scene positional <u>localization</u> of a given human, vehicle or animal activity. See also action <u>detection</u>. [Sze10:12.6.4]
- action model: A pre-defined or learned model of a given human action which is matched against a given unseen action instance to perform action recognition or action detection. Akin to the use of models in model-based object recognition. [NWF08]
- action recognition: Similar to action detection but further considering the

classification of actions (e.g., walking, running, kicking, lifting, stretching). See also activity recognition and behavior classification, of which action recognition is often a sub-task, i.e., an activity or behavior is considered as a series of actions:



Commonly the terms action, activity and behavior are used inter-changeably in the literature. [Sze10:12.6.4]

- action representation: A model-based approach whereby an action is represented as a spatio-temporal <u>feature</u> vector over a given video sequence. [GX11:Ch. 6]
- action unit: The smallest atom or measurement of action within an action sequence or <u>action representation</u> removed from the raw measurement of pixel movement itself (e.g., <u>optical</u> flow). [LJ11:18.2.2]
- active appearance model: A generalization of the widely used active shape model approach that includes all of the information in the image region covered by the target object, rather than just that near modeled edges. The active appearance model has a statistical model of the shape and graylevel appearance of the object of interest. This statistical model generalizes to cover most valid examples. Matching to an image involves finding model parameters that minimize the difference between the image and a synthesized model example, projected into the image. [NA05:6.5]
- active blob: A region-based approach to the tracking of non-rigid motion in which an active shape model is used. The model is based on an initial region that is divided using Delaunay

triangulation and then each patch is tracked from frame to frame (note that the patches can deform). [SI98]

- active calibration: An approach to <u>camera calibration</u> that uses naturally occurring <u>features</u> within the scene with active motion of the camera to perform calibration. By contrast, traditional approaches assume a static camera and a predefined <u>calibration object</u> with fixed features. [Bas95]
- active contour model: A technique used in <u>model-based vision</u> where object boundaries are detected using a deformable curve representation such as a <u>snake</u>. The term "active" refers to the ability of the snake to deform shape to better match the image data. See also active shape model. [SQ04:8.5]
- active contour tracking: A technique used in <u>model-based vision</u> for tracking object boundaries in a <u>video sequence</u> using active contour models. [LL93]
- active illumination: A system of lighting where intensity, orientation or pattern may be continuously controlled and altered. This kind of system may be used to generate <u>structured light</u>. [CS09:1.2]
- active learning: A <u>machine-learning</u> approach in which the learning agent can actively query the environment for data examples. For example, a classification approach may recognize that it is less reliable over a certain sub-region of the input example space and thus request more training examples that characterize inputs for that sub-region. Considered to be a <u>supervised learning</u> approach. [Bar12:13.1.5]
- active net: An <u>active shape model</u> that parameterizes a <u>triangulated mesh</u>. [TY89]
- active recognition: An approach to <u>object recognition</u> or <u>scene</u> <u>classification</u> in which the recognition agent or algorithm collects further evidence samples (e.g., more images after moving) until a sufficient level of confidence is obtained to make a decision on identification. See also active learning. [RSB04]

- active sensing: 1) A sensing activity carried out in an active or purposive way, e.g., where a camera is moved in space to acquire multiple or optimal views of an object (see also <u>active vision</u>, <u>purposive vision</u> and <u>sensor planning</u>).
 2) A sensing activity implying the projection of a pattern of energy, e.g., a laser line, onto the scene (see also <u>laser stripe triangulation</u> and <u>structured light triangulation</u>). [FP03:21.1]
- active shape model: Statistical model of the shape of an object that can deform to fit a new example of the object. The shapes are constrained by a <u>statistical shape model</u> so that they may vary only in ways seen in a training set. The models are usually formed using <u>principal component analysis</u> to identify the dominant modes of shape variation in observed examples of the shape. Model shapes are formed by linear combinations of the dominant modes. [WP:Active_shape_model]
- active stereo: An alternative approach to traditional <u>binocular stereo</u>. One of the cameras is replaced with a <u>structured</u> <u>light</u> projector, which projects light onto the object of interest. If the camera calibration is known, the <u>triangulation</u> for computing the 3D coordinates of object points simply involves finding the intersection of a ray and known structures in the light field. [CS09:1.2]
- active structure from X: The recovery of scene <u>depth</u> (i.e., 3D structure) via an active sensing technique, such as shape from X techniques plus motion.



[Sze10:12.2] The figure shows the shape from structured light method, with the light plane being swept along the object:

- active surface: 1) A surface determined using a range sensor; 2) an active shape model that deforms to fit a surface. [WP:Active_surface]
- active triangulation: Determination of surface depth by <u>triangulation</u> between a <u>light source</u> at a known position and a camera that observes the effects of the illuminant on the scene. <u>Light stripe ranging</u> is one form of active triangulation. A variant is to use a single scanning laser beam to illuminate the scene and a stereo pair of cameras to compute depth. [WP: 3D_scanner#Triangulation]
- active vision: An approach to computer vision in which the camera or sensor is moved in a controlled manner, so as to simplify the nature of a problem. For example, rotating a camera with constant angular velocity while maintaining fixation at a point allows absolute calculation of scene point depth, instead of relative depth that depends on the camera speed. See also <u>kinetic</u> depth. [Nal93:10]
- active volume: The volume of interest in a machine vision application. [SZH+10:Ch. 1]
- activity: A temporal sequence of actions performed by an entity (e.g., a person, animal or vehicle) indicative of a given task, behavior or intended goal. See activity classification. [Sze10:12.6.4]
- activity analysis: Analyzing the behavior of people or objects in a video sequence, for the purpose of identifying the immediate actions occurring or the long-term sequence of actions, e.g., detecting potential intruders in a restricted area. [WP:Occupational_ therapy#Activity_analysis]
- $\begin{array}{c} \textbf{activity} \quad \textbf{classification:} & \text{The} \\ \underline{\textbf{classification}} & \text{of a given temporal} \\ \overline{\textbf{sequence of}} & \text{actions forming a given} \\ \underline{\textbf{activity}} & \text{into a discrete set of labels.} \\ \overline{[\text{Sze10:}12.6.4]} \end{array}$
- activity graph: A graph encoding the activity transition matrix where each node in the graph corresponds to an

activity (or stage of an activity) and the arcs among nodes represent the allowable next activities (or stages): [GX11:7.3.2]



- activity model: A representation of a given <u>activity</u> used for <u>activity</u> <u>classification</u> via an approach akin to that of model-based object recognition.
- activity recognition: See <u>activity</u> classification.
- activity representation: See <u>activity</u> model.
- activity segmentation: The task of segmenting a video sequence into a series of sub-sequences based on variations in <u>activity</u> performed along that sequence. [GX11:7.2]
- activity transition matrix: An NxN matrix, for a set of N different activities, where each entry corresponds to the transition probability between two states and each state is itself an activity being performed within the scene. See also <u>state</u> transition probability: [GX11:7.2]

Activity	Walk	Browse	Meet	Fight
Walk	0.9	0.84	0.63	0.4
Browse	0.3	0.78	0.73	0.2
Meet	0.74	0.79	0.68	0.28
Fight	0.32	0.45	0.23	0.60

acuity: The ability of a vision system to discriminate (or resolve) between closely arranged visual stimuli. This can be measure using a grating, i.e., a pattern of parallel black and white stripes of equal widths. Once the bars become too close, the grating becomes indistinguishable from a uniform image of the same average intensity as the bars. Under optimal lighting, the minimum spacing that a person can resolve is 0.5 min of arc. [Umb98:7.6]

AdaBoost: An Adaptive Boosting approach for <u>ensemble learning</u> whereby the (weak) classifiers are trained in sequence such that the *n*th classifier is trained over a <u>training set</u> re-weighted to give greater emphasis to training examples upon which the previous (n - 1) classifiers performed poorly. See boosting. [Bis06:14.3]

adaptation: See adaptive.

- adaptive: The property of an algorithm to adjust its parameters to the data at hand in order to optimize performance. Examples include <u>adaptive</u> <u>contrast enhancement</u>, <u>adaptive</u> <u>filtering and adaptive smoothing</u>. [WP: Adaptive_algorithm]
- adaptive behavior model: A behaviormodel exhibiting adaptive propertiesthat facilitate the online updating of themodel used for behavior analysis. Generally follows a three-stage process:model initialization, online anomalousbehavior detectionupdating via unsupervised learning.See also unsupervised behaviormodeling. [GX11:8.3]
- adaptive bilateral filter: A variant on <u>bilateral filtering</u> used as an <u>image-sharpening operator</u> with simultaneous <u>noise removal</u>. Performs image sharpening by increasing the overall "slope" (i.e., the gradient range) of the edges without producing overshoot or undershoot associated with the unsharp operator. [ZA08]
- adaptive coding: A scheme for the transmission of signals over unreliable channels, e.g., wireless links. Adaptive coding varies the parameters of the encoding to respond to changes in the channel, e.g., "fading", where the signal-to-noise ratio degrades. [WP: Adaptive_coding]

- adaptive contrast enhancement: An image processing operation that applies <u>histogram</u> equalization locally across an image. [WP: Adaptive_histogram_equalization]
- adaptive edge detection: Edge detection with adaptive thresholding of the gradient magnitude image. [Nal93:3.1.2]
- adaptive filtering: In signal processing, any filtering process in which the parameters of the filter change over time or where the parameters are different at different parts of the signal or image. [WP:Adaptive_filter]
- adaptive histogram equalization: A localized method of improving image contrast. A histogram is constructed of the gray levels present. These gray levels are re-mapped so that the histogram is approximately flat. It can be made perfectly flat by dithering: [WP: Adaptive_histogram_equalization]



original

after adaptive histogram equalization

- adaptive Hough transform: A Hough transform method that iteratively increases the resolution of the parameter space quantization. It is particularly useful for dealing with highdimensional parameter spaces. Its disadvantage is that sharp peaks in the histogram can be missed. [NA05:5.6]
- **adaptive meshing**: Methods for creating simplified meshes where elements are made smaller in regions of high detail (rapid changes in surface orientation) and larger in regions of low

detail, such as planes. [WP:Adaptive_ mesh_refinement]

- **adaptive pyramid**: A method of multiscale processing where small areas of image having some feature in common (e.g., color) are first extracted into a graph representation. This graph is then manipulated, e.g., by pruning or merging, until the level of desired scale is reached. [JM92]
- adaptive reconstruction: Data-driven methods for creating statistically significant data in areas of a 3D data cloud where data may be missing because of sampling problems. [YGK95]
- **adaptive smoothing**: An iterative smoothing algorithm that avoids smoothing over edges. Given an image I(x, y), one iteration of adaptive smoothing proceeds as follows:
 - 1. Compute gradient magnitude image $G(x, y) = |\nabla I(x, y)|.$
 - 2. Make weights image $W(x, y) = e^{-\lambda G(x, y)}$.
 - 3. Smooth the image: $S(x, y) = \frac{\sum_{i=1}^{l} \sum_{j=1}^{l} A_{xyij}}{\sum_{i=1}^{l} \sum_{j=1}^{l} B_{xyij}}$ where $A_{xyij} = I(x+i, y+j)W(x+i, y+j)$ $B_{xyij} = W(x+i, y+j)$ [WP:Additive_smoothing]
- adaptive thresholding: An improved
- image thresholding technique where the threshold value varies at each pixel. A common technique is to use the average intensity in a neighbourhood to set the threshold: [Dav90:4.4]



- adaptive triangulation: See <u>adaptive</u> meshing.
- adaptive visual servoing: See visual servoing. [WP:Visual_Servoing]

- adaptive weighting: A scheme for weighting elements in a summation, voting or other formulation such that the relative influence of each element is representative (i.e., adapted to some underlying structure). For example this may be the similarity of pixels within a neighborhood (e.g., an adaptive bilateral filter) or a property changing over time. See also <u>adaptive</u>. [YK06]
- additive color: The way in which multiple wavelengths of light can be combined to allow other colors to be perceived (e.g., if equal amounts of green and red light are shone onto a sheet of white paper, the paper will appear to be illuminated with a yellow light source (see below). Contrast this with subtractive color: [Gal90:3.7]



additive noise: Generally, imageindependent noise that is added to it by some external process. The recorded image *I* at pixel (*i*, *j*) is then the sum of the true signal *S* and the noise *N*.

$$I_{i,j} = S_{i,j} + N_{i,j}$$

The noise added at each pixel (i, j) could be different. [Umb98:3.2]

adjacency: See adjacent.

adjacency graph: A graph that shows the adjacency between structures, such as segmented image regions. The nodes of the graph are the structures and an arc implies adjacency of the two structures connected by the arc. The figure shows the graph associated with the segmented image on the left: [AFF85]



Regions

Adjacency graph

- **adjacent**: Commonly meaning "next to each other", whether in a physical sense of <u>pixel connectivity</u> in an image, image regions sharing some common boundary, nodes in a <u>graph</u> connected by an arc or components in a <u>geometric model</u> sharing some common bounding component. Formally defining "adjacent" can be somewhat heuristic because you may need a way to specify closeness (e.g., on a quantized grid of pixels) or to consider how much shared "boundary" is required before two structures are adjacent. [Nev82:2.1.1]
- **affective body gesture**: See <u>affective</u> gesture.
- **affective gesture**: A gesture made by the body (human or animal) which is indicative of emotional feeling or response. Used in gesture analysis to indicate social interaction. [GX11:5.4]
- affective state: The emotional state on an entity (human or animal) relating to emotional feeling or response. Often measured via gesture analysis or facial expression analysis. See affective gesture.
- **affine**: A term first used by Euler. Affine geometry is a study of properties of geometric objects that remain invariant under <u>affine transformations</u> (mappings), including parallelness, cross ratio and adjacency. [WP:Affine_ geometry]
- **affine arc length**: For a parametric equation of a curve $\vec{f}(u) = (x(u), y(u))$, arc

length is not preserved under an affine transformation. The affine length

$$\tau(u) = \int_0^u (\dot{x}\ddot{y} - \ddot{x}\dot{y})^{\frac{1}{3}}$$

is invariant under affine transformations. [SQ04:8.4]

- affine camera: A special case of the projective camera that is obtained by constraining the 3×4 camera parameter matrix **T** such that $T_{3,1} = T_{3,2} = T_{3,3} =$ 0 and reducing the camera parameter vector from 11 degrees of freedom to 8. [FP03:2.3.1]
- affine curvature: A measure of curvature based on the <u>affine arc length</u>, τ . For a parametric equation of a curve $\vec{f}(u) = (x(u), y(u))$, its affine curvature, μ , is

$$\mu(\tau) = x''(\tau)y'''(\tau) - x'''(\tau)y''(\tau)$$

[WP:Affine_curvature]

- **affine flow**: A method of finding the movement of a surface patch by estimating the <u>affine transformation</u> parameters required to transform the patch from its position in one view to another. [Cal05]
- affine fundamental matrix: The fundamental matrix which is obtained from a pair of cameras under affine viewing conditions. It is a 3×3 matrix whose upper left 2×2 submatrix is all zero. [HZ00:13.2.1]
- **affine invariant**: An object or shape property that is not changed by (i.e., is <u>invariant</u> under) the application of an affine transformation. [FP03:18.4.1]
- affine length: See <u>affine arc length</u>. [WP: Affine_curvature]
- **affine moment**: Four shape measures derived from second and third order <u>moments</u> that remain invariant under <u>affine transformations</u>. They are given by the following equations, where each μ is the associated <u>central</u> moment: [NA05:7.3]

$$I_1 = \frac{\mu_{20}\mu_{02} - \mu_{11}^2}{\mu_{00}^4}$$

$$I_2 = (\mu_{30}^2 \mu_{03}^2 - 6\mu_{30}\mu_{21}\mu_{12}\mu_{03})$$

$$+ 4\mu_{30}\mu_{12}^{3} + 4\mu_{21}^{3}\mu_{03}
- 3\mu_{21}^{2}\mu_{12}^{2})/\mu_{00}^{10}
I_{3} = (\mu_{20}(\mu_{21}\mu_{03} - \mu_{12}^{2})
- \mu_{11}(\mu_{30}\mu_{03} - \mu_{21}\mu_{12})
+ \mu_{02}(\mu_{30}\mu_{12} - \mu_{21}^{2}))/\mu_{00}^{7}
I_{4} = (\mu_{20}^{3}\mu_{03}^{2} - 6\mu_{20}^{2}\mu_{11}\mu_{12}\mu_{03}
- 6\mu_{20}^{2}\mu_{02}\mu_{21}\mu_{03} + 9\mu_{20}^{2}\mu_{02}\mu_{12}^{2}
+ 12\mu_{20}\mu_{11}^{2}\mu_{21}\mu_{03}
+ 6\mu_{20}\mu_{11}\mu_{02}\mu_{30}\mu_{03}
- 18\mu_{20}\mu_{11}\mu_{02}\mu_{21}\mu_{12}
- 8\mu_{11}^{3}\mu_{30}\mu_{03} - 6\mu_{20}\mu_{02}^{2}\mu_{30}\mu_{12}
+ 9\mu_{20}\mu^{2}\mu^{2} + 12\mu^{2}\mu_{20}\mu_{10}\mu_{10}\mu_{10}$$

- $+9\mu_{20}\mu_{02}^{-}\mu_{21}^{-}+12\mu_{11}^{-}\mu_{02}\mu_{30}\mu_{12}^{-}\\-6\mu_{11}\mu_{02}^{2}\mu_{30}\mu_{21}^{-}+\mu_{02}^{3}\mu_{30}^{2})/\mu_{00}^{11}$
- **affine quadrifocal tensor**: The form taken by the <u>quadrifocal tensor</u> when specialized to the viewing conditions modeled by the <u>affine camera</u>. [HTM99]
- affine reconstruction: A threedimensional reconstruction where the ambiguity in the choice of basis is affine only. Planes that are parallel in the Euclidean basis are parallel in the affine reconstruction. A projective reconstruction can be upgraded to an affine reconstruction by identification of the plane at infinity, often by locating the absolute conic in the reconstruction. [HZ00:9.4.1]
- affine registration: The registration of two or more images, <u>surface</u> <u>meshes or point clouds</u> using an <u>affine</u> transformation. [JV05]
- affine stereo: A method of scene reconstruction using two calibrated views of a scene from known viewpoints. It is a simple but very robust approximation to the geometry of stereo vision, to estimate positions, shapes and surface orientations. It can be calibrated very easily by observing just four reference points. Any two views of the same planar surface will be related by an affine transformation that maps one image to the other. This consists of a translation and a tensor, known as the disparity gradient tensor, representing the distortion in image shape. If the standard unit vectors X and Y in one

JWST364-c01 JWST364-Fisher Printer: Yet to Come October 7, 2013 9:57 Trim: 229mm × 152mm

image are the projections of some vectors on the object surface and the linear mapping between images is represented by a 2×3 matrix **A**, then the first two columns of **A** will be the corresponding vectors in the other image. Since the centroid of the plane will map to both image centroids, it can be used to find the surface orientation. [Oua93]

- **affine transformation**: A special set of transformations in Euclidean geometry that preserve some properties of the construct being transformed:
 - Points remain collinear: if three points belong to the same straight line, their images under affine transformations also belong to the same line and the middle point remains between the other two points.
 - Parallel lines remain parallel and concurrent lines remain concurrent (images of intersecting lines intersect).
 - The ratio of lengths of the segments of a given line remains constant.
 - The ratio of areas of two triangles remains constant.
 - Ellipses remain ellipses; parabolas remain parabolas and hyperbolas remain hyperbolas.
 - Barycenters of triangles (and other shapes) map into the corresponding barycenters.



Analytically, affine transformations are represented in the matrix form

$f(x) = \mathbf{A}x + b$

where the determinant det(**A**) of the square matrix **A** is not 0. In 2D, the matrix is 2×2 ; in 3D it is 3×3 . [FP03:2.2]

- **affine trifocal tensor**: The form taken by the <u>trifocal tensor</u> when specialized to the viewing conditions modeled by the affine camera. [HTM99]
- **affinely invariant region**: Image patches that automatically deform with changing <u>viewpoint</u> in such a way that they cover identical physical parts of a scene. Since such regions are describable by a set of <u>invariant</u> features they are relatively easy to match between views under changing <u>illumination</u>. [TG00]
- **affinity matrix**: A matrix capturing the similarity of two entities or their relative attraction in a force- or flow-based model. Often referred to in graph <u>cut</u> formulations. See <u>affinity metric</u>. [Sze10:5.4]
- **affinity metric**: A measurement of the similarity between two entities (e.g., features, nodes or images). See <u>similarity metric</u>.
- affordance and action recognition: An affordance is an opportunity for an entity to take an action. The recognition of such occurrences thus identifies such action opportunities. See action recognition. [Gib86]
- **age progression**: Refers to work considering the change in visual appearance because of the human aging process. Generally considered in tasks such as <u>face recognition</u>, <u>face detection</u> and <u>face modeling</u>. Recent work considers artificial aging of a sample facial image to produce an aged interpretation of the same: [GZSM07]



agglomerative clustering: A class of iterative <u>clustering</u> algorithms that begin with a large number of clusters and, at each iteration, merge

pairs (or tuples) of clusters. Stopping the process at a certain number of iterations gives the final set of clusters. The process can be run until only one cluster remains and the progress of the algorithm can be represented as a <u>dendrogram</u>. [WP: Hierarchical_clustering]

- AIC: See <u>Akaike Information Criterion</u> (AIC).
- AkaikeInformationCriterion(AIC):A method for statisticalmodelselectionwherethe best-fitlog-likelihood is penalized by the number of adjustable parameters in themodel, so as to counterover-fitting.Compare withBayesianBayesianinformationcriterion.[Bis06:1.3]
- **albedo**: Whiteness. Originally a term used in astronomy to describe reflecting power.



If a body reflects 50% of the light falling on it, it is said to have albedo of 0.5. [FP03:4.3.3]

algebraic curve: A simple parameterized <u>curve</u> representation using <u>Euclidean geometry</u> for an object that cannot be represented by <u>linear</u> properties (see figure). Parameterized in \mathbb{R}^n <u>Euclidean space</u>, in the form of $\{\vec{x}: f(\vec{x}) = 0\}$: [Gib98:Ch. 1]



algebraic distance: A linear distance <u>metric</u> commonly used in <u>computer</u> <u>vision</u> applications because of its simple form and standard matrix-based <u>least mean square estimation</u> operations. If a curve or surface is defined implicitly by $f(\vec{x}, \vec{a}) = 0$ (e.g., $\vec{x} \cdot \vec{a} =$ 0 for a hyperplane), the algebraic distance of a point \vec{x}_i to the surface is simply $f(\vec{x}_i, \vec{a})$. [FP03:10.1.5]

- algebraic point set surfaces: A smooth <u>surface model</u> defined from a <u>point</u> <u>cloud</u> representation using localized <u>moving least squares fitting</u> of an <u>algebraic surface</u> (namely an algebraic <u>sphere</u>). [GG07]
- algebraic surface: A parameterized <u>surface</u> representation using <u>Euclidean</u> <u>geometry</u> defined in \mathbb{R}^n <u>Euclidean</u> <u>space</u>. Regular 3D surfaces such as <u>planes</u>, <u>spheres</u>, tori and <u>generalized</u> <u>cylinders</u> occur in \mathbb{R}^3 , in the form of $\overline{\{\vec{x}: f(\vec{x}) = 0\}}$. [Zar71]
- aliasing: The erroneous replacement of high spatial frequency (HF) components by low-frequency ones when a signal is sampled. The affected HF components are those that are higher than the Nyquist frequency, or half the sampling frequency. Examples include the slowing of periodic signals by strobe lighting and corruption of areas of detail in image resizing. If the source signal has no HF components, the effects of aliasing are avoided, so the low-pass filtering of a signal to remove HF components prior to sampling is one form of anti-aliasing. Consider the perspective projection of a checkerboard. The image is obtained by sampling the scene at a set of integer locations. The spatial frequency increases as the plane recedes, producing aliasing artifacts (jagged lines in the foreground, moiré patterns in the background):



Removing high-frequency components (i.e., <u>smoothing</u>) before downsampling mitigates the effect: [FP03:7.4]



- alignment: An approach to geometric model matching by registration of a geometric model to the image data. [FP03:18.2]
- ALVINN: Autonomous Land Vehicle In a Neural Network; an early attempt, at Carnegie Mellon University, to learn a complex behavior (maneuvering a vehicle) by observing humans. [Pom89]
- **ambient light**: Illumination by diffuse reflections from all surfaces within a scene (including the sky, which acts as an external distant surface). In other words, light that comes from all directions, such as the sky on a cloudy day. Ambient light ensures that all surfaces are illuminated, including those not directly facing light sources. [FP03:5.3.3]
- **ambient space**: Refers to the dimensional space surrounding a given mathematical object in general terms. For example, a line can be studied in isolation or within a 2D space in which case, the ambient space is a plane. Similarly a sphere may be studied in 3D ambient space. This is of particular relevance if the ambient space is nonlinear or skewed (e.g., a magnetic field). [SMC05]

- AMBLER: An autonomous active vision system using both structured light and sonar, developed by NASA and Carnegie Mellon University. It is supported by a 12-legged robot and is intended for planetary exploration. [BHK+89]
- **amplifier noise**: Spurious <u>additive noise</u> signal generated by the electronics in a sampling device. The standard model for this type of noise is Gaussian. It is independent of the signal. In color cameras, where more amplification is used in the blue color channel than in the green or red channels, there tends to be more noise in the blue channel. In well-designed electronics, amplifier noise is generally negligible. [WP: Image_noise#Amplifier_noise_.28 Gaussian_noise.29]
- analog/mixed analog-digital image processing: The processing of images as analog signals (e.g., by <u>optical</u> <u>image processing</u>) prior to or without any form of <u>image digitization</u>. Largely superseded by <u>digital image</u> processing. [RK82]
- analytic curve finding: A method of detecting parametric curves by transforming data into a feature space that is then searched for the hypothesized curve parameters. An example is line finding using the Hough transform. [XOK90]
- anamorphic lens: A lens having one or more cylindrical surfaces. Anamorphic lenses are used in photography to produce images that are compressed in one dimension. Images can later be restored to true form using a reversing anamorphic lens set. This form of lens is used in wide-screen movie photography. [WP:Anamorphic_lens]
- **anatomical map**: A biological model usable for <u>alignment</u> with, or <u>region</u> <u>labeling</u> of, a corresponding image dataset. For example, one could use a model of the brain's functional regions to assist in the identification of brain structures in an <u>NMR</u> dataset. [GHC+00]

AND operator: A Boolean logic operator that combines two input binary images:

р	q	p&q
0	0	0
0	1	0
1	0	0
1	1	1

This approach is used to select image regions by applying the AND logic at each pair of corresponding pixels. The rightmost image below is the result of ANDing the two leftmost images: [SB11:3.2.2]



- angiography: A method for imaging blood vessels by introducing a dye that is opaque when photographed by Xray. Also the study of images obtained in this way. [WP:Angiography]
- angularity ratio: Given two figures, X and Y, $\alpha_i(X)$ and $\beta_i(Y)$ are angles subtending convex parts of the contour of the figure X and $\gamma_k(X)$ are angles subtending plane parts of the contour of figure *X*; the angularity ratios are:

360°

$$\sum_{i} \frac{\alpha_i(X)}{360^\circ}$$

and

$$\frac{\sum_i \beta_j(X)}{\sum_k \gamma_k(X)}$$

[Lee64]

- diffusion: An edgeanisotropic preserving smoothing filter commonly used for noise removal. Also called Perona-Malik diffusion. See also bilateral filtering. [Sze10:3.3]
- anisotropic filtering: Any filtering technique where the filter parameters vary

over the image or signal being filtered. [WP:Anisotropic_filtering]

- anisotropic structure tensor: A matrix representing non-uniform, secondorder (i.e., gradient/edge) information of an image or function neighborhood. Commonly used in corner detection and anisotropic filtering approaches. [Sze10:3.3]
- annotation: A general term referring to the labeling of imagery either with regard to manual ground truth labeling or automatic image labeling of the output of a scene understanding, semantic scene segmentation or augmented reality approach: [Sze10:14.6, 13.1.2]



- anomalous behavior detection: Special case of surveillance where human movement is analyzed. Used in particular to detect intruders or behavior likely to precede or indicate crime. [WP:Anomaly_detection]
- anomaly detection: The automated detection of an unexpected event, behavior or object within a given environment based on comparison with a model of what is normally expected within the same. Often considered as an unsupervised learning task and commonly applied in visual industrial inspection and automated visual surveillance. [Bar12:13.1.3]
- antimode: The minimum between two maxima. One method of threshold selection is done by determining the antimode in a bimodal histogram. [WP: Bimodal_distribution#Terminology]



aperture: Opening in the lens diaphragm of a camera through which light is admitted. This device is often arranged so that the amount of light can be controlled accurately. A small aperture reduces the amount of light available, but increases the <u>depth of field</u>. The figure shows nearly closed (left) and nearly open (right) aperture positions: [TV98:2.2.2]



- **aperture control**: Mechanism for varying the size of a camera's <u>aperture</u>. [WP:Aperture#Aperture_control]
- **aperture problem**: If a motion sensor has a finite receptive field, it perceives the world through something resembling an aperture, making the motion of a homogeneous contour seem locally ambiguous. Within that aperture, different physical motions are therefore indistinguishable. For example, the two motions of the square below are identical in the circled receptive fields: [Nal93:8.1.1]



- **apparent contour**: The apparent contour of a surface *S* in 3D is the set of critical values of the projection of *S* onto a plane, in other words, the silhouette. If the surface is transparent, the apparent contour can be decomposed into a collection of closed curves with double points and cusps. The convex envelope of an apparent contour is also the boundary of its <u>convex hull</u>. [Nal93:Ch. 4]
- **apparent motion**: The 3D motion suggested by the image <u>motion field</u>, but not necessarily matching the real 3D motion. The reason for this mismatch is that motion fields may be ambiguous; that is, they may be generated by different 3D motions or light source movement. Mathematically, there may be multiple solutions to the problem of reconstructing 3D motion from the image motion field. See also <u>visual</u> <u>illusion</u> and <u>motion estimation</u>. [WP: <u>Apparent_motion]</u>
- **appearance**: The way an object looks from a particular viewpoint under particular lighting conditions. [FP03:25.1.3]
- appearance-based recognition: Object recognition where the object model encodes the possible appearances of the object (as contrasted with a geometric model that encodes the shape, in model-based recognition). In principle, it is impossible to encode all appearances when occlusions are considered; however, small numbers of appearances can often be adequate, especially if there are not many models in the model base. There are many approaches appearance-based recognition, to such as using a principal component model to encode all appearances in a compressed framework, using color histograms to summarize the appearance or a set of local appearance descriptors such as Gabor filters extracted at interest points. A common feature of these approaches is learning the models from examples. [TV98:10.4]
- appearance-based tracking: Methods for object or target recognition in real

time, based on image pixel values in each frame rather than derived features. Temporal filtering, such as the Kalman filter, is often used. [BJ98]

- **appearance change**: Changes in an image that are not easily accounted for by motion, such as an object actually changing form. [BFY98]
- **appearance enhancement transform**: Generic term for operations applied to images to change, or enhance, some aspect of them, such as brightness adjustment, contrast adjustment, edge sharpening, histogram equalization, saturation adjustment or magnification. [Hum77]
- appearance feature: An object or scene feature relating to visual appearance, as opposed to features derived from shape, motion or behavior analysis. [HFR06]
- **appearance flow**: Robust methods for real-time object recognition from a sequence of images depicting a moving object. Changes in the images are used rather than the images themselves. It is analogous to processing using <u>optical</u> flow. [DTS06]
- appearance model: A representation used for interpreting images that is based on the appearance of the object. These models are usually learned by using multiple views of the objects. See also active appearance model and appearance-based recognition. [WP: Active_appearance_model]
- **appearance prediction**: Part of the science of appearance engineering, where an object texture is changed so that the viewer experience is predictable. [Kan97]
- **appearance singularity**: An image position where a small change in viewer position can cause a dramatic change in the appearance of the observed scene, such as the appearance or disappearance of image features. This is contrasted with changes occurring when in a generic viewpoint. For example, when viewing the corner of a cube from a distance, a small change in viewpoint still leaves the three surfaces at the corner visible. However, when the

viewpoint moves into the infinite plane containing one of the cube faces (a singularity), one or more of the planes disappears. [MR98]

arc length: If f is a function such that its derivative f' is continuous on some closed interval [a, b] then the arc length of f from x = a to x = b is the integral: [FP03:19.1]

$$\int_{a}^{b} \sqrt{1 + [f'(x)]^2} dx$$

f(x)



arc of graph: Two <u>nodes</u> in a graph can be connected by an arc (also called an edge). The edges can be either <u>directed</u> or <u>undirected</u>. The dashed lines here are the arcs: [WP: Graph_(mathematics)]



- architectural model reconstruction: A generic term for reverse engineering buildings based on collected 3D data as well as libraries of building constraints. [WZ02]
- **area**: The measure of a region or surface's extension in some given units. The units could be image units, such as square pixels, or scene units, such as square centimeters. [JKS95:2.2.1]

- **area-based**: Operation applied to a region of an image; as opposed to pixel-based. [CS09:6.6]
- ARMA: See <u>autoregressive moving</u> average model.
- **array processor**: A group of timesynchronized processing elements that perform computations on data distributed across them. Some array processors have elements that communicate only with their immediate neighbors, as in the topology shown below. See also <u>single instruction multiple</u> data. [WP:Vector_processor]



- arterial tree segmentation: Generic term for methods used in finding internal pipe-like structures in medical images, such as <u>NMR</u> images, <u>angiograms</u> and <u>X-rays</u>. Example trees include bronchial systems and veins. [BWB+05]
- **articulated object**: An object composed of a number of (usually) rigid subparts or components connected by joints, which can be arranged in a number of different configurations. The human body is a typical example. [BM02:1.9]
- articulated object model: A representation of an <u>articulated object</u> that includes its separate parts and their range of movement (typically joint angles) relative to each other. [RK95]
- articulated object segmentation: Methods for acquiring an <u>articulated</u> object from 2D or 3D data. [YP06]

- articulated object tracking: Tracking an <u>articulated object</u> in an image sequence. This includes both the <u>pose</u> of the object and also its shape parameters, such as joint angles. [WP: Finger_tracking]
- aspect graph: A graph of the set of views (aspects) of an object, where the <u>arcs</u> of the graph are transitions between two neighboring views (the <u>nodes</u>) and a change between aspects is called a visual event. See also <u>characteristic</u> <u>view</u>. This graph shows some of the aspects of a hippopotamus: [FP03:20]



aspect ratio: 1) The ratio of the sides of the <u>bounding box</u> of an object, where the orientation of the box is chosen to maximize this ratio. Since this measure is scale invariant, it is a useful metric for <u>object recognition</u>.

2) In a camera, the ratio of the horizontal to vertical pixel sizes.

- 3) In an image, the ratio of the image width to height – an image of 640 by 480 pixels has an aspect ratio of 4:3. [Low91:2.2]
- **aspect**: See <u>characteristic view</u> and aspect graph.
- **asperity scattering**: A light scattering effect, common to the modeling or photography of human skin, caused by sparsely distributed point scatters over the surface. In the case of human skin, these point scatters are vellus (short, fine, light-colored and barely noticeable) hairs present on the surface. [INN07:3.3]
- association graph: A graph used in structure matching, such as <u>matching</u> a <u>geometric model</u> to a data description. In this graph, each <u>node</u> corresponds

to a pairing between a model and a data feature (with the implicit assumption that they are compatible). <u>Arcs</u> in the graph mean that the two connected nodes are pairwise compatible. Finding <u>maximal cliques</u> is one technique for finding good matches. The graph below shows a set of pairings of model features A, B and C with image features a, b, c and d. The maximal clique consisting of A:a, B:b and C:c is one match hypothesis: [BB82:11.2.1]



astigmatism: A refractive error with where the light is focused within an optical system. It occurs when a lens has irregular curvature causing light rays to focus at an area, rather than at a point:



- It may be corrected with a toric lens, which has a greater refractive index on one axis than the others. In human eyes, astigmatism often occurs with nearsightedness and farsightedness. [FP03:1.2.3]
- asymmetric SVM: A variant on traditional <u>support vector machine</u> <u>classification</u> where the <u>false positives</u> are modeled in the training objective of finding a maximal margin of classification. An asymmetric SVM maximizes the margin between the negative class

and the core (i.e., high confidence subset) of the positive class by introducing a secondary core margin in addition to the traditional inter-class margin. They are jointly optimized within the training optimization cycle. [WLCC08]

- atlas-based segmentation: A segmentation technique used in medical image processing, especially with brain images. Automatic tissue segmentation is achieved using a <u>model</u> of the brain structure and imagery (see <u>atlas</u> <u>registration</u>) compiled with the assistance of human experts. See also <u>image</u> segmentation. [VYCL03]
- atlas registration: An image registration technique used in medical image processing, especially to register brain images. An atlas is a <u>model</u> (perhaps statistical) of the characteristics of multiple brains, providing examples of normal and pathological structures. This makes it possible to take into account anomalies that single-image registration could not. See also <u>medical image</u> <u>registration</u>. [HSS+08]
- atomic action: In gesture analysis and <u>action recognition</u>, a short sequence of basic limb movements that form the pattern of movement associated with a higher level action. For example, "lift right leg in front of left leg" for a running action or "swing right hand" and "rotate upper body" for taking a badminton shot. [GX11:Ch. 1]
- ATR: See automatic target recognition. [WP:Automatic_target_recognition]
- attached shadow: A shadow caused by an object on itself by self-occlusion. See also cast shadow. [FP03:5.3.1]
- attention: See <u>visual attention</u>. [WP: Attention]
- attenuation: The reduction of a particular phenomenon, e.g., noise attenuation is the reduction of image noise. [WP:Attenuation]
- attributed graph: A graph useful for representing different properties of an image. Its nodes are attributed pairs of image segments, their color or shape e.g. The relations between them, such

as relative texture or brightness are encoded as arcs. [BM02:4.5.2]

- atypical co-occurrence: An unusual joint occurrence of two or more events or observations against *a priori* expectation. [GX11:Ch. 9]
- **augmented reality**: Primarily a projection method that adds, e.g., graphics or sound as an overlay to original image or audio. For example, a fire-fighter's helmet display could show exit routes registered to his or her view of the building. [WP:Augmented_reality]
- autocalibration: The recovery of a <u>camera's calibration</u> using only point (or other feature) correspondences from multiple uncalibrated images and geometric consistency constraints (e.g., that the camera settings are the same for all images in a sequence). [Low91:13.7]
- **autocorrelation**: The extent to which a signal is similar to shifted copies of itself. For an infinitely long 1D signal $f(t) : \mathbb{R} \mapsto \mathbb{R}$, the autocorrelation at a shift Δt is

$$R_f(\Delta t) = \int_{-\infty}^{\infty} f(t)f(t + \Delta t)dt$$

The autocorrelation function R_f always has a maximum at 0. A *peaked* autocorrelation function decays quickly away from $\Delta t = 0$. The sample autocorrelation function of a finite set of values $f_{1..n}$ is $\{r_f(d)|d=1,\ldots,n-1\}$ where

$$r_f(d) = \frac{\sum_{i=1}^{n-d} (f_i - \bar{f})(f_{i+d} - \bar{f})}{\sum_{i=1}^{n} (f_i - \bar{f})^2}$$

and $\bar{f} = \frac{1}{n} \sum_{i=1}^{n} f_i$ is the sample mean. [WP:Autocorrelation]

autofocus: Automatic determination and control of image sharpness in an optical or vision system. There are two major variations in this control system: active focusing and passive focusing. Active autofocus is performed using a sonar or infrared signal to determine the object distance. Passive autofocus is performed by analyzing the image itself to optimize differences between adjacent pixels in the CCD array. [WP:Autofocus] automated visual surveillance: The generalized use of automatic scene understanding approaches commonly including object detection, tracking and, more recently, behavior classification: [Dav90:Ch. 22]



- **automatic**: Performed by a machine without human intervention. The opposite of "manual". [WP: Automation]
- automatic target recognition (ATR): The detection of hostile objects in a scene using sensors and algorithms. Sensors are of many different types, sampling in <u>infrared</u> and visible light and using <u>acoustic sonar</u> and <u>radar</u>. [WP:Automatic_target_recognition]
- autonomous vehicle: A mobile robot controlled by computer, with human input operating only at a very high level, e.g., stating the ultimate destination or task. Autonomous navigation requires the visual tasks of route detection, self-localization, landmark location and obstacle detection, as well as robotic tasks such as route planning and motor control. [WP:Autonomous car]
- **autoregressive model**: A model that uses statistical properties of the past behavior of some variable to predict future behavior of that variable. For example, a signal x_t at time t satisfies an autoregressive model of order p if $x_t = \sum_{n=1}^{p} \alpha_n x_{t-n} + \omega_t$, where ω_t is noise. The model could also be nonlinear. [WP:Autoregressive_ model]
- autoregressive moving average (ARMA) model: Combines an autoregressive model with a moving average for the statistical analysis and future value prediction of time series information. [BJ71]
- autostereogram: An image similar to a random dot stereogram in which

the corresponding features are combined into a single image. <u>Stereo fusion</u> allows the perception of a 3D shape in the 2D image. [WP:Autostereogram]

average smoothing: See <u>mean</u> smoothing.

- AVI: Microsoft format for audio and video files ("audio video interleaved"). Unlike <u>MPEG</u>, it is not a standard, so that compatibility of AVI video files and AVI players is not always guaranteed. [WP:Audio_Video_Interleave]
- axial representation: A region representation that uses a curve to describe the image region. The axis may be a <u>skeleton</u> derived from the region by a thinning process. [RM91]
- axiomaticcomputervision:Approachesrelating to the coreprinciples (or axioms) of computervision. Commonly associated with theinterpretation of Marr's theory for thebasic building blocks of how visualinterpretationshould ideallybe performed. Mostrecently associated with low-levelfeature detection (e.g., edge detectionand corner detection). [KZM05]
- **axis of elongation**: 1) The line that minimizes the second moment of the data points. If $\{\vec{x}_i\}$ are the data points, and $d(\vec{x}, L)$ is the distance from point \vec{x} to line *L*, then the axis of elongation *A* minimizes $\sum_i d(\vec{x}_i, A)^2$.

Let $\vec{\mu}$ be the mean of $\{\vec{x}_i\}$. Define the scatter matrix $\mathbf{S} = \sum_i (\vec{x}_i - \vec{\mu})(\vec{x}_i - \vec{\mu})^T$. Then the axis of elongation is the <u>eigenvector</u> of **S** with the largest <u>eigenvalue</u>. See also <u>principal</u> <u>component analysis</u>. The figure shows a possible axis of elongation for a set of points:



2) The longer midline of the <u>bounding</u> box with largest length-to-width ratio. [JKS95:2.2.3]

- **axis of rotation**: A line about which a rotation is performed. Equivalently, the line whose points are fixed under the action of a rotation. Given a 3D rotation matrix \mathbf{R} , the axis is the eigenvector of \mathbf{R} corresponding to the eigenvalue 1. [JKS95:12.2.2]
- **axis–angle curve representation**: A rotation representation based on the amount of twist θ about the axis of rotation, here a unit vector \vec{a} . The <u>quaternion</u> rotation representation is similar. [WP:Axisangle_representation]