Image segmentation

From Wikipedia, the free encyclopedia
(Redirected from Segmentation (image processing))

In computer vision, **image segmentation** is the process of partitioning a digital image into multiple segments (sets of pixels, also known as superpixels). 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.[1][2] 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 characteristics.

The result of image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image (see edge detection). Each of the pixels in a region are similar with respect to some characteristic or computed property, such as color, intensity, or texture. Adjacent regions are significantly different with respect to the same characteristic(s).[1] When applied to a stack of images, typical in medical imaging, the resulting contours after image segmentation can be used to create 3D reconstructions with the help of interpolation algorithms like Marching cubes.
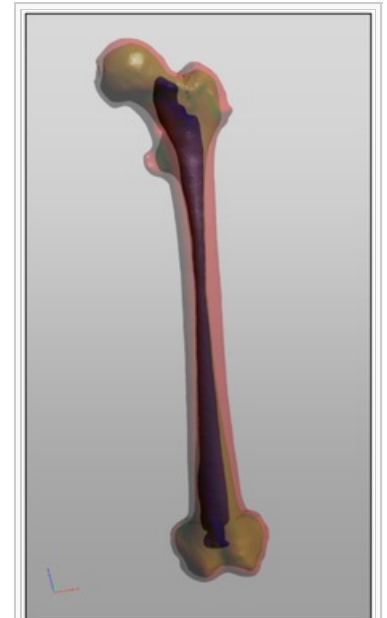
Model of a segmented femur. It shows the outer surface (red), the surface between compact bone and spongy bone (green) and the surface of the bone marrow (blue).

Contents [hide]

Applications   [edit]

Some of the practical applications of image segmentation are:

- Content-based image retrieval
- Machine vision
- Medical imaging[3][4]
  - Locate tumors and other pathologies[5][6]
  - Measure tissue volumes
  - Diagnosis, study of anatomical structure
  - Surgery planning
  - Virtual surgery simulation
  - Intra-surgery navigation
- Object detection[7]
  - Pedestrian detection
  - Face detection
  - Brake light detection
  - Locate objects in satellite images (roads, forests, crops, etc.)
- Recognition Tasks
  - Face recognition
  - Fingerprint recognition
  - Iris recognition
- Traffic control systems
- Video surveillance

Several general-purpose algorithms and techniques have been developed for image segmentation. To be useful, these techniques must typically be combined with a domain's specific knowledge in order to effectively solve the domain's segmentation problems.

Thresholding   [edit]

The simplest method of image segmentation is called the thresholding method. This method is based on a clip-level (or a threshold value) to turn a gray-scale image into a binary image. There is also a balanced histogram thresholding.

The key of this method is to select the threshold value (or values when multiple-levels are selected). Several popular methods are used in industry including the maximum entropy method, Otsu's method (maximum variance), and k-means clustering.

Recently, methods have been developed for thresholding computed tomography (CT) images. The key idea is that, unlike Otsu's method, the thresholds are derived from the radiographs instead of the (reconstructed) image[8] .[9]

New methods suggested the usage of multi-dimensional fuzzy rule-based non-linear thresholds. In these works decision over each pixel's membership to a segment is based on multi-dimensional rules derived from fuzzy logic and evolutionary algorithms based on image lighting environment and application.[10]

Clustering methods   [edit]

Main article: Data clustering

The K-means algorithm is an iterative technique that is used to partition an image into K clusters.[11] The basic algorithm is

1. Pick K cluster centers, either randomly or based on some heuristic
2. Assign each pixel in the image to the cluster that minimizes the distance between the pixel and the cluster center
3. Re-compute the cluster centers by averaging all of the pixels in the cluster

4. Repeat steps 2 and 3 until convergence is attained (i.e. no pixels change clusters)

In this case, [distance](#) is the squared or absolute difference between a pixel and a cluster center. The difference is typically based on pixel [color](#), [intensity](#), [texture](#), and location, or a weighted combination of these factors. *K* can be selected manually, [randomly](#), or by a [heuristic](#). This algorithm is guaranteed to converge, but it may not return the [optimal](#) solution. The quality of the solution depends on the initial set of clusters and the value of *K*.

Source image.

Image after running *k*-means with *k* = *16*. Note that a common technique to improve performance for large images is to downsample the image, compute the clusters, and then reassign the values to the larger image if necessary.

Compression-based methods [[edit](#)]

Compression based methods postulate that the optimal segmentation is the one that minimizes, over all possible segmentations, the coding length of the data.[12][13] The connection between these two concepts is that segmentation tries to find patterns in an image and any regularity in the image can be used to compress it. The method describes each segment by its texture and boundary shape. Each of these components is modeled by a probability distribution function and its coding length is computed as follows:

1. The boundary encoding leverages the fact that regions in natural images tend to have a smooth contour. This prior is used by [Huffman coding](#) to encode the difference [chain code](#) of the contours in an image. Thus, the smoother a boundary is, the shorter coding length it attains.
2. Texture is encoded by [lossy compression](#) in a way similar to [minimum description length](#) (MDL) principle, but here the length of the data given the model is approximated by the number of samples times the [entropy](#) of the model. The texture in each region is modeled by a [multivariate normal distribution](#) whose entropy has closed form expression. An interesting property of this model is that the estimated entropy bounds the true entropy of the data from above. This is because among all distributions with a given mean and covariance, normal distribution has the largest entropy. Thus, the true coding length cannot be more than what the algorithm tries to minimize.

For any given segmentation of an image, this scheme yields the number of bits required to encode that image based on the given segmentation. Thus, among all possible segmentations of an image, the goal is to find the segmentation which produces the shortest coding length. This can be achieved by a simple agglomerative clustering method. The distortion in the lossy compression determines the coarseness of the segmentation and its optimal value may differ for each image. This parameter can be estimated heuristically from the contrast of textures in an image. For example, when the textures in an image are similar, such as in camouflage images, stronger sensitivity and thus lower quantization is required.

Histogram-based methods [[edit](#)]

[Histogram](#)-based methods are very efficient compared to other image segmentation methods because they typically require only one pass through the [pixels](#). In this technique, a histogram is computed from all of the pixels in the image, and the peaks and valleys in the histogram are used to locate the [clusters](#) in the image.[1] [Color](#) or [intensity](#) can be used as the measure.

A refinement of this technique is to [recursively](#) apply the histogram-seeking method to clusters in the image in order to divide them into smaller clusters. This operation is repeated with smaller and smaller clusters until no more clusters are formed.[1][14]

One disadvantage of the histogram-seeking method is that it may be difficult to identify significant peaks and valleys in the image.

Histogram-based approaches can also be quickly adapted to apply to multiple frames, while maintaining their single pass efficiency. The histogram can be done in multiple fashions when multiple frames are considered. The same approach that is taken with one frame can be applied to multiple, and after the results are merged, peaks and valleys that were previously difficult to identify are more likely to be distinguishable. The histogram can also be applied on a per-pixel basis where the resulting information is used to determine the most frequent
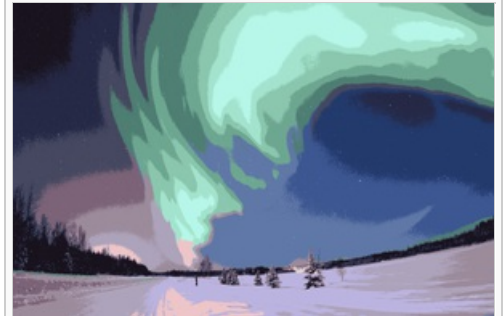
color for the pixel location. This approach segments based on active objects and a static environment, resulting in a different type of segmentation useful in Video tracking.

Edge detection [edit]

Edge detection is a well-developed field on its own within image processing. Region boundaries and edges are closely related, since there is often a sharp adjustment in intensity at the region boundaries. Edge detection techniques have therefore been used as the base of another segmentation technique.

The edges identified by edge detection are often disconnected. To segment an object from an image however, one needs closed region boundaries. The desired edges are the boundaries between such objects or spatial-taxons.[15] [16]

Spatial-taxons[17] are information granules.,[18] consisting of a crisp pixel region, stationed at abstraction levels within a hierarchical nested scene architecture. They are similar to the Gestalt psychological designation of figure-ground, but are extended to include foreground, object groups, objects and salient object parts. Edge detection methods can be applied to the spatial-taxon region, in the same manner they would be applied to a silhouette. This method is particularly useful when the disconnected edge is part of an illusory contour[19][20]

Segmentation methods can also be applied to edges obtained from edge detectors. Lindeberg and Li[21] developed an integrated method that segments edges into straight and curved edge segments for parts-based object recognition, based on a minimum description length (M_{DL}) criterion that was optimized by a split-and-merge-like method with candidate breakpoints obtained from complementary junction cues to obtain more likely points at which to consider partitions into different segments.

Dual clustering method [edit]

This method is a combination of three characteristics of the image: partition of the image based on histogram analysis is checked by high compactness of the clusters (objects), and high gradients of their borders. For that purpose two spaces has to be introduced: one space is the one-dimensional histogram of brightness H = H(B), the second space – the dual 3-dimensional space of the original image itself B = B(x, y). The first space allows to measure how compact is distributed the brightness of the image by calculating minimal clustering kmin. Threshold brightness T corresponding to kmin defines the binary (black-and-white) image – bitmap b = φ(x, y), where φ(x, y) = 0, if B(x, y) < T, and φ(x, y) = 1, if B(x, y) ≥ T. The bitmap b is an object in dual space. On that bitmap a measure has to be defined reflecting how compact distributed black (or white) pixels are. So, the goal is to find objects with good borders. For all T the measure M_{DC} =G/(k-L) has to be calculated (where k is difference in brightness between the object and the background, L is length of all borders, and G is mean gradient on the borders). Maximum of MDC defines the segmentation.[22]

Region-growing methods [edit]

Region-growing methods rely mainly on the assumption that the neighboring pixels within one region have similar values. The common procedure is to compare one pixel with its neighbors. If a similarity criterion is satisfied, the pixel can be set to belong to the cluster as one or more of its neighbors. The selection of the similarity criterion is significant and the results are influenced by noise in all instances.

The method of Statistical Region Merging[23] (SRM) starts by building the graph of pixels using 4-connectedness with edges weighted by the absolute value of the intensity difference. Initially each pixel forms a single pixel region. SRM then sorts those edges in a priority queue and decide whether or not to merge the current regions belonging to the edge pixels using a statistical predicate.

One region-growing method is the seeded region growing method. This method takes a set of seeds as input along with the image. The seeds mark each of the objects to be segmented. The regions are iteratively grown by comparison of all unallocated neighboring pixels to the regions. The difference between a pixel's intensity value and the region's mean, δ, is used as a measure of similarity. The pixel with the smallest difference measured in this way is assigned to the respective region. This process continues until all pixels are assigned to a region. Because seeded region growing requires seeds as additional input, the segmentation results are dependent on the choice of seeds, and noise in the image can cause the seeds to be poorly placed.

Another region-growing method is the unseeded region growing method. It is a modified algorithm that does not require explicit seeds. It starts with a single region A_1—the pixel chosen here does not markedly influence the final segmentation. At each iteration it considers the neighboring pixels in the same way as seeded region growing. It differs from seeded region growing in that if the minimum δ is less than a predefined threshold T then it is added to the respective region A_j. If not, then the pixel is considered different from all current regions

A_i and a new region A_{n+1} is created with this pixel.

One variant of this technique, proposed by Haralick and Shapiro (1985),[1] is based on pixel intensities. The mean and scatter of the region and the intensity of the candidate pixel are used to compute a test statistic. If the test statistic is sufficiently small, the pixel is added to the region, and the region's mean and scatter are recomputed. Otherwise, the pixel is rejected, and is used to form a new region.

A special region-growing method is called λ-connected segmentation (see also lambda-connectedness). It is based on pixel intensities and neighborhood-linking paths. A degree of connectivity (connectedness) is calculated based on a path that is formed by pixels. For a certain value of λ, two pixels are called λ-connected if there is a path linking those two pixels and the connectedness of this path is at least λ. λ-connectedness is an equivalence relation.[24]

Split-and-merge segmentation is based on a quadtree partition of an image. It is sometimes called quadtree segmentation.

This method starts at the root of the tree that represents the whole image. If it is found non-uniform (not homogeneous), then it is split into four son squares (the splitting process), and so on. If, in contrast, four son squares are homogeneous, they are merged as several connected components (the merging process). The node in the tree is a segmented node. This process continues recursively until no further splits or merges are possible.[25][26] When a special data structure is involved in the implementation of the algorithm of the method, its time complexity can reach $O(n \log n)$, an optimal algorithm of the method.[27]

Partial differential equation-based methods [edit]

Using a partial differential equation (PDE)-based method and solving the PDE equation by a numerical scheme, one can segment the image.[28] Curve propagation is a popular technique in this category, with numerous applications to object extraction, object tracking, stereo reconstruction, etc. The central idea is to evolve an initial curve towards the lowest potential of a cost function, where its definition reflects the task to be addressed. As for most inverse problems, the minimization of the cost functional is non-trivial and imposes certain smoothness constraints on the solution, which in the present case can be expressed as geometrical constraints on the evolving curve.

Parametric methods [edit]

Lagrangian techniques are based on parameterizing the contour according to some sampling strategy and then evolve each element according to image and internal terms. Such techniques are fast and efficient, however the original "purely parametric" formulation (due to Kass, Witkin and Terzopoulos in 1987 and known as "snakes"), is generally criticized for its limitations regarding the choice of sampling strategy, the internal geometric properties of the curve, topology changes (curve splitting and merging), addressing problems in higher dimensions, etc.. Nowadays, efficient "discretized" formulations have been developed to address these limitations while maintaining high efficiency. In both cases, energy minimization is generally conducted using a steepest-gradient descent, whereby derivatives are computed using, e.g., finite differences.

Level set methods [edit]

The level set method was initially proposed to track moving interfaces by Osher and Sethian in 1988 and has spread across various imaging domains in the late 90s. It can be used to efficiently address the problem of curve/surface/etc. propagation in an implicit manner. The central idea is to represent the evolving contour using a signed function whose zero corresponds to the actual contour. Then, according to the motion equation of the contour, one can easily derive a similar flow for the implicit surface that when applied to the zero level will reflect the propagation of the contour. The level set method affords numerous advantages: it is implicit, is parameter-free, provides a direct way to estimate the geometric properties of the evolving structure, allows for change of topology, and is intrinsic. It can be used to define an optimization framework, as proposed by Zhao, Merriman and Osher in 1996. One can conclude that it is a very convenient framework for addressing numerous applications of computer vision and medical image analysis.[29] Research into various level set data structures has led to very efficient implementations of this method.

Fast marching methods [edit]

The fast marching method has been used in image segmentation,[30] and this model has been improved (permitting a both positive and negative speed propagation speed) in an approach called the generalized fast marching method.[31]

Variational methods [edit]

The goal of variational methods is to find a segmentation which is optimal with respect to a specific energy functional. The functionals consist of a data fitting term and a regularizing terms. A classical representative is the Potts model defined for an image f by

$$\underset{u}{\operatorname{argmin}} \gamma \|\nabla u\|_0 + \int (u - f)^2 dx.$$

A minimizer u^* is a piecewise constant image which has an optimal tradeoff between the squared L2 distance to the given image f and the total length of its jump set. The jump set of u^* defines a segmentation. The relative weight of the energies is tuned by the parameter $\gamma > 0$. The binary variant of the Potts model, i.e., if the range of u is restricted to two values, is often called Chan-Vese model.[32] An important generalization is the Mumford-Shah model [33] given by

$$\underset{u,K}{\operatorname{argmin}} \gamma |K| + \mu \int_{K^C} |\nabla u|^2 dx + \int (u - f)^2 dx.$$

The functional value is the sum of the total length of the segmentation curve K, the smoothness of the approximation u, and its distance to the original image f. The weight of the smoothness penalty is adjusted by $\mu > 0$. The Potts model is often called piecewise constant Mumford-Shah model as it can be seen as the degenerate case $\mu \to \infty$. The optimization problems are known to be NP-hard in general but near-minimizing strategies work well in practice. Classical algorithms are graduated non-convexity and Ambrosio-Tortorelli approximation.
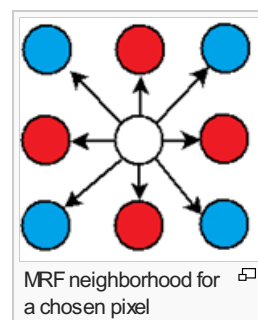
Graph partitioning methods [edit]

Graph partitioning methods are an effective tools for image segmentation since they model the impact of pixel neighborhoods on a given cluster of pixels or pixel, under the assumption of homogeneity in images. In these methods, the image is modeled as a weighted, undirected graph. Usually a pixel or a group of pixels are associated with nodes and edge weights define the (dis)similarity between the neighborhood pixels. The graph (image) is then partitioned according to a criterion designed to model "good" clusters. Each partition of the nodes (pixels) output from these algorithms are considered an object segment in the image. Some popular algorithms of this category are normalized cuts,[34] random walker,[35] minimum cut,[36] isoperimetric partitioning,[37] minimum spanning tree-based segmentation,[38] and segmentation-based object categorization.

Markov Random Fields [edit]

The application of Markov random fields (MRF) for images was suggested in early 1984 by Geman and Geman.[39] Their strong mathematical foundation and ability to provide a global optima even when defined on local features proved to be the foundation for novel research in the domain of image analysis, de-noising and segmentation. MRFs are completely characterized by their prior probability distributions, marginal probability distributions, cliques, smoothing constraint as well as criterion for updating values. The criterion for image segmentation using MRFs is restated as finding the labelling scheme which has maximum probability for a given set of features. The broad categories of image segmentation using MRFs are supervised and unsupervised segmentation.

Supervised Image Segmentation using MRF and MAP [edit]

In terms of image segmentation, the function that MRFs seek to maximize is the probability of identifying a labelling scheme given a particular set of features are detected in the image. This is a restatement of the Maximum a posteriori estimation method.

MRF neighborhood for a chosen pixel

The generic algorithm for image segmentation using MAP is given below:

Optimization algorithms   [edit]

Each optimization algorithm is an adaptation of models from a variety of fields and they are set apart by their unique cost functions. The common trait of cost functions is to penalize change in pixel value as well as difference in pixel label when compared to labels of neighboring pixels.

Iterated Conditional Modes/Gradient Descent   [edit]

The ICM algorithm tries to reconstruct the ideal labeling scheme by changing the values of each pixel over each iteration and evaluating the energy of the new labeling scheme using the cost function given below,

$$\alpha(1 - \delta(l_i - l_{\text{initial } i}) + \beta \Sigma_{q \epsilon N(i)}(1 - \delta(l_i, l_{q(i)}))$$

where α is the penalty for change in pixel label and β is the penalty for difference in label between neighboring pixels and chosen pixel. Here $N(i)$ is neighborhood of pixel i and δ is the Kronecker delta function. A major issue with ICM is that, similar to gradient descent, it has a tendency to rest over local maxima and thus not obtain a globally optimal labeling scheme.

Simulated Annealing(SA)   [edit]

Derived as an analogue of annealing in metallurgy, SA uses change in pixel label over iterations and estimates the difference in energy of each newly formed graph to the initial data. If the newly formed graph is more profitable, in terms of low energy cost, given by:

$$\Delta U = U^{\text{new}} - U^{\text{old}}$$

$$l_i = \begin{cases} l_i^{new}, & \text{if } \Delta U <= 0, \\ l_i^{new}, & \text{if } \Delta U > 0 \text{ and } \delta < e^{-\Delta U/T}, \\ l_i^{old} \end{cases}$$

the algorithm selects the newly formed graph. Simulated annealing requires the input of temperature schedules which directly affects the speed of convergence of the system, as well as energy threshold for minimization to occur.

Alternative Algorithms   [edit]

A range of other methods exist for solving simple as well as higher order MRFs. They include Maximization of Posterior Marginal, Multi-scale MAP estimation,[40] Multiple Resolution segmentation[41] and more. Apart from likelihood estimates, graph-cut using maximum flow[42] and other highly constrained graph based methods[43][44] exist for solving MRFs.

Unsupervised Image Segmentation using MRF and Expectation Maximization   [edit]

A subset of unsupervised machine learning, the Expectation–maximization algorithm is utilized to iteratively estimate the a posterior probabilities and distributions of labeling when no training data is available and no estimate of segmentation model can be formed. A general approach is to use histograms to represent the
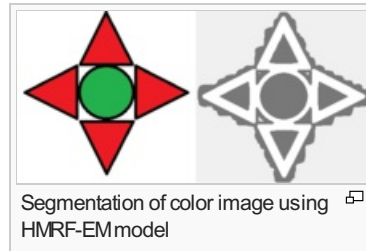
features of an image and proceed as outlined briefly in the 3-step algorithm mentioned below,

1. A random estimate of the model parameters (same as in supervised learning) is utilized.

2. E-Step: Estimate class statistics based on the random segmentation model defined. Using these, compute the conditional probability of belonging to a label given the feature set is calculated using naive Bayes' theorem.

$$P(\lambda|f_i) = \frac{P(f_i|\lambda)P(\lambda)}{\Sigma_{\lambda\epsilon\Lambda}P(f_i|\lambda)P(\lambda)}$$ Here $\lambda\epsilon\Lambda$ is the set of all possible labels.

3. M-Step: The established relevance of a given feature set to a labeling scheme is now used to compute the a priori estimate of a given label in the second part of the algorithm. Since the actual number of total labels is unknown (from a training data set), a hidden estimate of the number of labels given by the user is utilized in computations.

$$P(\lambda) = \frac{\Sigma_{\lambda\epsilon\Lambda}P(\lambda|f_i)}{|\Omega|}$$ where Ω is the set of all possible features.

Segmentation of color image using HMRF-EM model

Disadvantages of MAP and EM based image segmentation   [edit]

1. Exact MAP estimates cannot be easily computed.

2. Approximate MAP estimates are computationally expensive to calculate.

3. Extension to multi-class labeling degrades performance and increases storage required.

4. Reliable estimation of parameters for EM is required for global optima to be achieved.

5. Based on method of optimization, segmentation may cluster to local minima.

Implementations of MRF based Image segmentation   [edit]

1. HMRF-EM Implementation of EM algorithm based image segmentation.

2. Zoltan Kato's implementation of supervised image segmentation using MRFs.[45][46]

3. Purdue University's implementation of Discrete MRF and their application to segmentation.

4. Gaussian Mixture Model based HMRF segmentation in MATLAB.

5. CMU's implementation of multiple graph cut based segmentation algorithms.

Watershed transformation   [edit]

The watershed transformation considers the gradient magnitude of an image as a topographic surface. Pixels having the highest gradient magnitude intensities (GMIs) correspond to watershed lines, which represent the region boundaries. Water placed on any pixel enclosed by a common watershed line flows downhill to a common local intensity minimum (LIM). Pixels draining to a common minimum form a catch basin, which represents a segment.

Model based segmentation   [edit]

The central assumption of such an approach is that structures of interest/organs have a repetitive form of geometry. Therefore, one can seek for a probabilistic model towards explaining the variation of the shape of the organ and then when segmenting an image impose constraints using this model as prior. Such a task involves (i) registration of the training examples to a common pose, (ii) probabilistic representation of the variation of the registered samples, and (iii) statistical inference between the model and the image. State of the art methods in the literature for knowledge-based segmentation involve active shape and appearance models, active contours and deformable templates and level-set based methods.[citation needed]

Multi-scale segmentation   [edit]

Image segmentations are computed at multiple scales in scale space and sometimes propagated from coarse to fine scales; see scale-space segmentation.

Segmentation criteria can be arbitrarily complex and may take into account global as well as local criteria. A common requirement is that each region must be connected in some sense.

One-dimensional hierarchical signal segmentation   [edit]

Witkin's seminal work[47][48] in scale space included the notion that a one-dimensional signal could be unambiguously segmented into regions, with one scale parameter controlling the scale of segmentation.

A key observation is that the zero-crossings of the second derivatives (minima and maxima of the first derivative or slope) of multi-scale-smoothed versions of a signal form a nesting tree, which defines hierarchical relations between segments at different scales. Specifically, slope extrema at coarse scales can be traced back to corresponding features at fine scales. When a slope maximum and slope minimum annihilate each other at a larger scale, the three segments that they separated merge into one segment, thus defining the hierarchy of segments.

Image segmentation and primal sketch   [edit]

There have been numerous research works in this area, out of which a few have now reached a state where they can be applied either with interactive manual intervention (usually with application to medical imaging) or fully automatically. The following is a brief overview of some of the main research ideas that current approaches are based upon.

The nesting structure that Witkin described is, however, specific for one-dimensional signals and does not trivially transfer to higher-dimensional images. Nevertheless, this general idea has inspired several other authors to investigate coarse-to-fine schemes for image segmentation. Koenderink[49] proposed to study how iso-intensity contours evolve over scales and this approach was investigated in more detail by Lifshitz and Pizer.[50] Unfortunately, however, the intensity of image features changes over scales, which implies that it is hard to trace coarse-scale image features to finer scales using iso-intensity information.

Lindeberg[51][52] studied the problem of linking local extrema and saddle points over scales, and proposed an image representation called the scale-space primal sketch which makes explicit the relations between structures at different scales, and also makes explicit which image features are stable over large ranges of scale including locally appropriate scales for those. Bergholm proposed to detect edges at coarse scales in scale-space and then trace them back to finer scales with manual choice of both the coarse detection scale and the fine localization scale.

Gauch and Pizer[53] studied the complementary problem of ridges and valleys at multiple scales and developed a tool for interactive image segmentation based on multi-scale watersheds. The use of multi-scale watershed with application to the gradient map has also been investigated by Olsen and Nielsen[54] and been carried over to clinical use by Dam[55] Vincken et al.[56] proposed a hyperstack for defining probabilistic relations between image structures at different scales. The use of stable image structures over scales has been furthered by Ahuja[57][58] and his co-workers into a fully automated system. A fully automatic brain segmentation algorithm based on closely related ideas of multi-scale watersheds has been presented by Undeman and Lindeberg[59] and been extensively tested in brain databases.

These ideas for multi-scale image segmentation by linking image structures over scales have also been picked up by Florack and Kuijper.[60] Bijaoui and Rué[61] associate structures detected in scale-space above a minimum noise threshold into an object tree which spans multiple scales and corresponds to a kind of feature in the original signal. Extracted features are accurately reconstructed using an iterative conjugate gradient matrix method.

Semi-automatic segmentation   [edit]

In one kind of segmentation, the user outlines the region of interest with the mouse clicks and algorithms are applied so that the path that best fits the edge of the image is shown.

Techniques like SIOX, Livewire, Intelligent Scissors or IT-SNAPS are used in this kind of segmentation. In an alternative kind of semi-automatic segmentation, the algorithms return a spatial-taxon (i.e. foreground, object-group, object or object-part) selected by the user or designated via prior probabilities.[62][63]

Trainable segmentation   [edit]

Most segmentation methods are based only on color information of pixels in the image. Humans use much more knowledge than this when doing image segmentation, but implementing this knowledge would cost considerable computation time and would require a huge domain-knowledge database, which is currently not available. In

addition to traditional segmentation methods, there are trainable segmentation methods which can model some of this knowledge.

Neural Network segmentation relies on processing small areas of an image using an artificial neural network[64] or a set of neural networks. After such processing the decision-making mechanism marks the areas of an image accordingly to the category recognized by the neural network. A type of network designed especially for this is the Kohonen map.

Pulse-coupled neural networks (PCNNs) are neural models proposed by modeling a cat's visual cortex and developed for high-performance biomimetic image processing. In 1989, Eckhorn introduced a neural model to emulate the mechanism of a cat's visual cortex. The Eckhorn model provided a simple and effective tool for studying the visual cortex of small mammals, and was soon recognized as having significant application potential in image processing. In 1994, the Eckhorn model was adapted to be an image processing algorithm by Johnson, who termed this algorithm Pulse-Coupled Neural Network. Over the past decade, PCNNs have been utilized for a variety of image processing applications, including: image segmentation, feature generation, face extraction, motion detection, region growing, noise reduction, and so on. A PCNN is a two-dimensional neural network. Each neuron in the network corresponds to one pixel in an input image, receiving its corresponding pixel's color information (e.g. intensity) as an external stimulus. Each neuron also connects with its neighboring neurons, receiving local stimuli from them. The external and local stimuli are combined in an internal activation system, which accumulates the stimuli until it exceeds a dynamic threshold, resulting in a pulse output. Through iterative computation, PCNN neurons produce temporal series of pulse outputs. The temporal series of pulse outputs contain information of input images and can be utilized for various image processing applications, such as image segmentation and feature generation. Compared with conventional image processing means, PCNNs have several significant merits, including robustness against noise, independence of geometric variations in input patterns, capability of bridging minor intensity variations in input patterns, etc.

Open-source implementations of trainable segmentation:

- Trainable Segmentation (in Java)
- IMMI

Other methods [edit]

There are many other methods of segmentation like multispectral segmentation or connectivity-based segmentation based on DTI images.[65]

Segmentation benchmarking [edit]

Several segmentation benchmarks are available for comparing the performance of segmentation methods with the state-of-the-art segmentation methods on standardized sets:

- Prague On-line Texture Segmentation Benchmark[66]
- The Berkeley Segmentation Dataset and Benchmark[67]

See also [edit]

- Computer vision
- Image-based meshing
- Range image segmentation
- Vector quantization
- Image quantization
- Color quantization

Notes [edit]

1. ^ *a b c d e* Linda G. Shapiro and George C. Stockman (2001): "Computer Vision", pp 279-325, New Jersey, Prentice-Hall, ISBN 0-13-030796-3
2. ^ Barghout, Lauren, and Lawrence W. Lee. "Perceptual information processing system." Paravue Inc. U.S. Patent Application 10/618,543, filed July 11, 2003.
3. ^ Pham, Dzung L.; Xu, Chenyang; Prince, Jerry L. (2000). "Current Methods in Medical Image Segmentation". *Annual Review of Biomedical Engineering* 2: 315–337. doi:10.1146/annurev.bioeng.2.1.315. PMID 11701515.
4. ^ Forghani, M.; Forouzanfar, M.; Teshnehlab, M. (2010). "Parameter optimization of improved fuzzy c-means clustering algorithm for brain MR image segmentation". *Engineering Applications of Artificial Intelligence* 23 (2): 160–168.

5. ^ W. Wu, A. Y. C. Chen, L. Zhao and J. J. Corso (2014): "Brain Tumor detection and segmentation in a CRF framework with pixel-pairwise affinity and super pixel-level features", *International Journal of Computer Aided Radiology and Surgery*, pp. 241-253, Vol. 9.

6. ^ E. B. George and M. Karnan (2012): "MR Brain image segmentation using Bacteria Foraging Optimization Algorithm", *International Journal of Engineering and Technology*, Vol. 4.

7. ^ J. A. Delmerico, P. David and J. J. Corso (2011): "Building façade detection, segmentation and parameter estimation for mobile robot localization and guidance", International Conference on Intelligent Robots and Systems, pp. 1632-1639.

8. ^ Batenburg, K J.; Sijbers, J. "Adaptive thresholding of tomograms by projection distance minimization". *Pattern Recognition* **42** (10): 2297–2305. doi:10.1016/j.patcog.2008.11.027.

9. ^ K J. Batenburg, and J. Sijbers, "Optimal Threshold Selection for Tomogram Segmentation by Projection Distance Minimization", *IEEE Transactions on Medical Imaging*, vol. 28, no. 5, pp. 676-686, June, 2009 PDF

10. ^ A. Kashanipour, N Milani, A. Kashanipour, H. Eghrary," Robust Color Classification Using Fuzzy Rule-Based Particle Swarm Optimization", *IEEE Congress on Image and Signal Processing*, vol. 2, pp. 110-114 , May, 2008 [1]

11. ^ Barghout, Lauren, and Jacob Sheynin. "Real-world scene perception and perceptual organization: Lessons from Computer Vision." Journal of Vision 13.9 (2013): 709-709.

12. ^ Hossein Mobahi, Shankar Rao, Allen Yang, Shankar Sastry and Yi Ma. Segmentation of Natural Images by Texture and Boundary Compression, International Journal of Computer Vision (IJCV), 95 (1), pg. 86-98, Oct. 2011.

13. ^ Shankar Rao, Hossein Mobahi, Allen Yang, Shankar Sastry and Yi Ma Natural Image Segmentation with Adaptive Texture and Boundary Encoding, Proceedings of the Asian Conference on Computer Vision (ACCV) 2009, H. Zha, R.-i. Taniguchi, and S. Maybank (Eds.), Part I, LNCS 5994, pp. 135–146, Springer.

14. ^ Ohlander, Ron; Price, Keith; Reddy, D. Raj (1978). "Picture Segmentation Using a Recursive Region Splitting Method". *Computer Graphics and Image Processing* **8** (3): 313–333. doi:10.1016/0146-664X(78)90060-6.

15. ^ R. Kimmel and A.M. Bruckstein. http://www.cs.technion.ac.il/~ron/PAPERS/Paragios_chapter2003.pdf, *International Journal of Computer Vision* 2003; 53(3):225-243.

16. ^ R. Kimmel, http://www.cs.technion.ac.il/~ron/PAPERS/laplacian_ijcv2003.pdf, chapter in Geometric Level Set Methods in Imaging, Vision and Graphics, (S. Osher, N. Paragios, Eds.), Springer Verlag, 2003. ISBN 0387954880

17. ^ Barghout, Lauren. Visual Taxometric approach Image Segmentation using Fuzzy-Spatial Taxon Cut Yields Contextually Relevant Regions. Communications in Computer and Information Science (CCIS). Springer-Verlag. 2014

18. ^ Witold Pedrycz (Editor), Andrzej Skowron (Co-Editor), Vladik Kreinovich (Co-Editor). Handbook of Granular Computing. Wiley 2008

19. ^ Barghout, Lauren (2014). Vision. Global Conceptual Context Changes Local Contrast Processing (Ph.D. Dissertation 2003). Updated to include Computer Vision Techniques. Scholars' Press. ISBN 978-3-639-70962-9.

20. ^ Barghout, Lauren, and Lawrence Lee. "Perceptual information processing system." Google Patents

21. ^ Lindeberg, T.; Li, M.-X. (1997). "Segmentation and classification of edges using minimum description length approximation and complementary junction cues". *Computer Vision and Image Understanding* **67** (1): 88–98.

22. ^ Shelia Guberman, Vadim V. Maximov, Alex Pashintsev Gestalt and Image Understanding. GESTALT THEORY 2012, Vol. 34, No.2, 143-166.

23. ^ R. Nock and F. Nielsen, Statistical Region Merging, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 26, No 11, pp 1452-1458, 2004.

24. ^ L. Chen, H. D. Cheng, and J. Zhang, Fuzzy subfiber and its application to seismic lithology classification, Information Sciences: Applications, Vol 1, No 2, pp 77-95, 1994.

25. ^ S.L. Horowitz and T. Pavlidis, Picture Segmentation by a Directed Split and Merge Procedure, Proc. ICPR, 1974, Denmark, pp.424-433.

26. ^ S.L. Horowitz and T. Pavlidis, Picture Segmentation by a Tree Traversal Algorithm, Journal of the ACM, 23 (1976), pp. 368-388.

27. ^ L. Chen, The lambda-connected segmentation and the optimal algorithm for split-and-merge segmentation, Chinese J. Computers, 14(1991), pp 321-331

28. ^ V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. International Journal of Computer Vision, 22(1):61-79, 1997. http://www.cs.technion.ac.il/~ron/PAPERS/CasKimSap_IJCV1997.pdf

29. ^ S. Osher and N. Paragios. Geometric Level Set Methods in Imaging Vision and Graphics, Springer Verlag, ISBN 0-387-95488-0, 2003.

30. ^ James A. Sethian. "Segmentation in Medical Imaging". Retrieved 15 January 2012.

31. ^ Forcade, Nicolas; Le Guyader, Carole; Gout, Christian (July 2008), "Generalized fast marching method: applications to image segmentation", *Numerical Algorithms* **48** (1-3): 189–211, doi:10.1007/s11075-008-9183-x

32. ^ Chan, T.F.; Vese, L. (2001). "Active contours without edges". *IEEE Transactions on Image Processing* **10** (2): 266–277. doi:10.1109/83.902291.

33. ^ David Mumford and Jayant Shah (1989): Optimal approximations by piecewise smooth functions and associated variational problems, *Communications on Pure and Applied Mathematics*, pp 577-685, Vol. 42, No. 5

34. ^ Jianbo Shi and Jitendra Malik (2000): "Normalized Cuts and Image Segmentation" 📄, *IEEE Transactions on pattern analysis and machine intelligence*, pp 888-905, Vol. 22, No. 8
35. ^ Leo Grady (2006): "Random Walks for Image Segmentation" 📄, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1768–1783, Vol. 28, No. 11
36. ^ Z. Wu and R. Leahy (1993): "An optimal graph theoretic approach to data clustering: Theory and its application to image segmentation" 📄, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1101–1113, Vol. 15, No. 11
37. ^ Leo Grady and Eric L. Schwartz (2006): "Isoperimetric Graph Partitioning for Image Segmentation" 📄, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 469–475, Vol. 28, No. 3
38. ^ C. T. Zahn (1971): "Graph-theoretical methods for detecting and describing gestalt clusters" 📄, *IEEE Transactions on Computers*, pp. 68–86, Vol. 20, No. 1
39. ^ S. Geman and D. Geman (1984): "Stochastic relaxation, Gibbs Distributions and Bayesian Restoration of Images", IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 721-741, Vol. 6, No.6.
40. ^ A. Bouman and M. Shapiro (2002): "A multiscale Random field model for Bayesian image segmentation", IEEE Transactions on Image Processing, pp. 162-177, Vol. 3.
41. ^ J. Liu and Y. H. Yang (1994): "Multiresolution color image segmentation", IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 689-700, Vol. 16.
42. ^ S. Vicente, V. Kolmogrov and C. Rother (2008): "Graph cut based image segmentation with connectivity priors", CVPR
43. ^ Corso, Z. Tu, and A. Yuille (2008): "MRF Labelling with Graph-Shifts Algorithm", Proceedings of International workshop on combinatorial Image Analysis
44. ^ B. J. Frey and D. MacKayan (1997): "A Revolution: Belief propagation in Graphs with Cycles", Proceedings of Neural Information Processing Systems (NIPS)
45. ^ Z. Kato, J. Zerubia and M. Berthod (1992): "Satellite Image classification using modified metropolis dynamic", Proceedings of International Conference on Acoustics, Speech and Signal Processing, pp. 573-576, Vol. 3.
46. ^ Z. Kato, "Modelisations markoviennes multiresolutions en vision par ordinateur", Thesis 1994.
47. ^ Witkin, A. P. "Scale-space filtering", Proc. 8th Int. Joint Conf. Art. Intell., Karlsruhe, Germany,1019–1022, 1983.
48. ^ A. Witkin, "Scale-space filtering: A new approach to multi-scale description," in Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing (ICASSP), vol. 9, San Diego, CA, Mar. 1984, pp. 150–153.
49. ^ Koenderink, Jan "The structure of images", Biological Cybernetics, 50:363–370, 1984
50. ^ Lifshitz, L. and Pizer, S.: A multiresolution hierarchical approach to image segmentation based on intensity extrema, IEEE Transactions on Pattern Analysis and Machine Intelligence, 12:6, 529–540, 1990. 🔗
51. ^ Lindeberg, T.: Detecting salient blob-like image structures and their scales with a scale-space primal sketch: A method for focus-of-attention, International Journal of Computer Vision, 11(3), 283–318, 1993. 🔗
52. ^ Lindeberg, Tony, Scale-Space Theory in Computer Vision, Kluwer Academic Publishers, 1994 🔗, ISBN 0-7923-9418-6
53. ^ Gauch, J. and Pizer, S.: Multiresolution analysis of ridges and valleys in grey-scale images, IEEE Transactions on Pattern Analysis and Machine Intelligence, 15:6 (June 1993), pages: 635–646, 1993. 🔗
54. ^ Olsen, O. and Nielsen, M.: Multi-scale gradient magnitude watershed segmentation, Proc. of ICIAP 97, Florence, Italy, Lecture Notes in Computer Science, pages 6–13. Springer Verlag, September 1997.
55. ^ Dam, E., Johansen, P., Olsen, O. Thomsen,, A. Darvann, T. , Dobrzenieck, A., Hermann, N., Kitai, N., Kreiborg, S., Larsen, P., Nielsen, M.: "Interactive multi-scale segmentation in clinical use" in European Congress of Radiology 2000.
56. ^ Vincken, K., Koster, A. and Viergever, M.: Probabilistic multiscale image segmentation 🔗, IEEE Transactions on Pattern Analysis and Machine Intelligence, 19:2, pp. 109–120, 1997.]
57. ^ M. Tabb and N. Ahuja, Unsupervised multiscale image segmentation by integrated edge and region detection, IEEE Transactions on Image Processing, Vol. 6, No. 5, 642–655, 1997. 🔗
58. ^ E. Akbas and N. Ahuja, "From ramp discontinuities to segmentation tree" 🔗
59. ^ C. Undeman and T. Lindeberg (2003) "Fully Automatic Segmentation of MRI Brain Images using Probabilistic Anisotropic Diffusion and Multi-Scale Watersheds", Proc. Scale-Space'03, Isle of Skye, Scotland, Springer Lecture Notes in Computer Science, volume 2695, pages 641–656. 🔗
60. ^ Florack, L. and Kuijper, A.: The topological structure of scale-space images, Journal of Mathematical Imaging and Vision, 12:1, 65–79, 2000.
61. ^ Bijaoui, A., Rué, F.: 1995, A Multiscale Vision Model, *Signal Processing* **46**, 345 🔗
62. ^ Barghout, Lauren. Visual Taxometric Approach to Image Segmentation using Fuzzy-Spatial Taxon Cut Yields Contextually Relevant Regions. IPMU 2014, Part II. A. Laurent et al (Eds.) CCIS 443, pp 163-173. Springer International Publishing Switzerland
63. ^ Barghout, Lauren (2014). *Vision: How Global Perceptual Context Changes Local Contrast Processing (Ph.D. Dissertation 2003). Updated to include Computer Vision Techniques* 🔗. Scholars Press. ISBN 978-3-639-70962-9.
64. ^ Mahinda Pathegama & Ö Göl (2004): "Edge-end pixel extraction for edge-based image segmentation", *Transactions on Engineering, Computing and Technology,* vol. 2, pp 213–216, ISSN 1305-5313
65. ^ Menke, RA, Jbabdi, S, Miller, KL, Matthews, PM and Zarei, M.: Connectivity-based segmentation of the substantia nigra in human and its implications in Parkinson's disease 🔗, Neuroimage, 52:4, pp. 1175–80, 2010.]

66. **^** Haindl, M. – Mikeš, S. Texture Segmentation Benchmark 🔗, Proc. of the 19th Int. Conference on Pattern Recognition. IEEE Computer Society, 2008, pp. 1–4 ISBN 978-1-4244-2174-9 ISSN 1051-4651
67. **^** D.Martin; C. Fowlkes; D. Tal; J. Malik (July 2001). "A Database of Human Segmented Natural Images and its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics". *Proc. 8th Int'l Conf. Computer Vision* **2**. pp. 416–423.

References [edit]

- 3D Entropy Based Image Segmentation 🔗
- Frucci, Maria; Sanniti di Baja, Gabriella (2008). "From Segmentation to Binarization of Gray-level Images" 🔗. *Journal of Pattern Recognition Research* 🔗 **3** (1): 1–13. doi:10.13176/11.54 🔗.

External links [edit]

- Some sample code that performs basic segmentation 🔗, by Syed Zainudeen. University Technology of Malaysia.
- Generalized Fast Marching method 🔗 by Forcadel et al. [2008] for applications in image segmentation.
- Image Processing Research Group 🔗 An Online Open Image Processing Research Community.
- Segmentation methods in image processing and analysis 🔗

Categories: Image segmentation