INTERNATIONAL JOURNAL OF RESEARCH IN COMPUTER APPLICATIONS AND ROBOTICS
ISSN 2320-7345

Review paper on Overview of Image Processing and Image Segmentation

[1]Kavita, [2]Ritika Saroha, [3]Rajani Bala, [4]Ms. Sunita Siwach

[1-4] School of engineering and sciences
Bhagat Phool Singh Mahila Vishwavidyalaya, khanpur kalan
Sonepat, India
[1]yadavkavita42@gmail.com, [2]saroharitika@gmail.com

Abstract— Image **processing** is a method to convert an image into digital form and perform some operations on it, in order to get an enhanced image or to extract some useful information from it. Image processing usually refers to digital image processing, but optical and analog image processing also are possible. This article is about general techniques that apply to all of them. The acquisition of images (producing the input image in the first place) is referred to as *imaging*. ***Image segmentation*** is a process of partitioning an image into sets of segments to change the representation of an image into something that is more meaningful and easier to analyze.

Keywords: **Image processing; form of signal processing; image segmentation**

1. Introduction

An image is an array, or a matrix, of square pixels (picture elements) arranged in columns and rows.

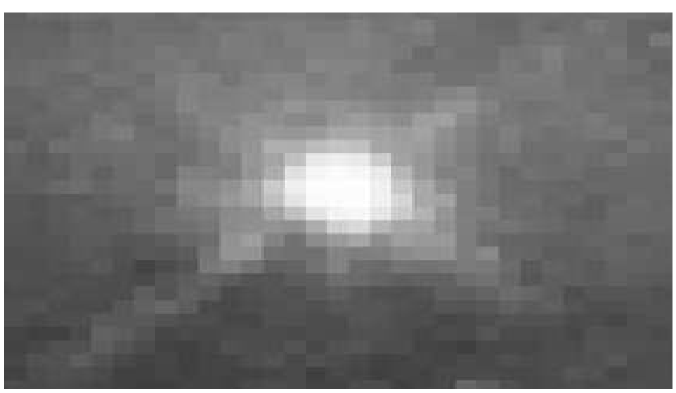


Fig. an image-array of pixels

Image processing is a method to convert an image into digital form and perform some operations on it, in order to get an enhanced image or to extract some useful information from it. It is a type of signal dispensation in which input is image, like video frame or photograph and output may be image or characteristics associated with that image. Usually **Image Processing** system includes treating images as two dimensional signals while applying already set signal processing methods to them.

It is among rapidly growing technologies today, with its applications in various aspects of a business. Image Processing forms core research area within engineering and computer science disciplines too.

In imaging science, image processing is any form of signal processing for which the input is an image, such as a photograph or video frame; the output of image processing may be either an image or a set of characteristics or parameters related to the image. Most image-processing techniques involve treating the image as a two-dimensional signal and applying standard signal-processing techniques to it.

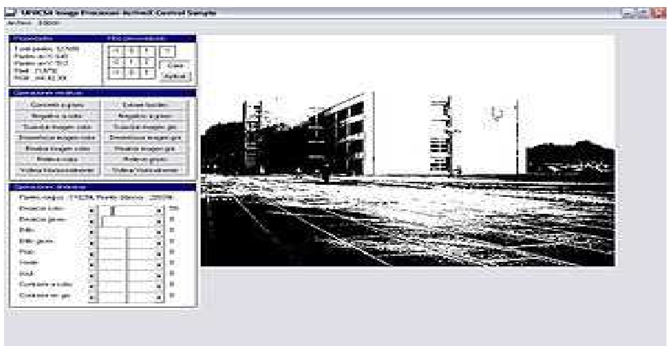


Fig. ***Monochrome black/white image***

Image processing usually refers to digital image processing, but optical and analog image processing also are possible. This article is about general techniques that apply to all of them. The acquisition of images (producing the input image in the first place) is referred to as imaging.

Image processing basically includes the following three steps.

- Importing the image with optical scanner or by digital photography.
- Analyzing and manipulating the image which includes data compression and image enhancement and spotting patterns that are not to human eyes like satellite photographs.
- Output is the last stage in which result can be altered image or report that is based on image analysis.

I. IMAGE PROCESSING AND ITS CATEGORIES

Image processing refers to processing of a 2D picture by a computer.

Basic definitions:

An image is defined in the "real world" is considered to be a function of two real variables, for example, a(x, y) with a as the amplitude (e.g. brightness) of the image at the real coordinate position (x,y).

Modern digital technology has made it possible to manipulate multi-dimensional signals with systems that range from simple digital circuits to advanced parallel computers.

The goal of this manipulation can be divided into three categories:

I. Image Processing (image in -> image out)
II. Image Analysis (image in -> measurements out)
III. Image Understanding (image in -> high-level description out)

Purpose of Image processing

The purpose of image processing is divided into 5 groups. They are:

1. Visualization - Observe the objects that are not visible.
2. Image sharpening and restoration - To create a better image.
3. Image retrieval - Seek for the image of interest.
4. Measurement of pattern – Measures various objects in an image.
5. Image Recognition – Distinguish the objects in an image.

Types of method used for image processing

The two types of methods used for Image Processing are

- Analog and
- Digital Image Processing

Analog or visual techniques of image processing can be used for the hard copies like printouts and photographs. Image analysts use various fundamentals of interpretation while using these visual techniques. The image processing is not just confined to area that has to be studied but on knowledge of analyst. Association is another important tool in image processing through visual techniques. So analysts apply a combination of personal knowledge and collateral data to image processing.

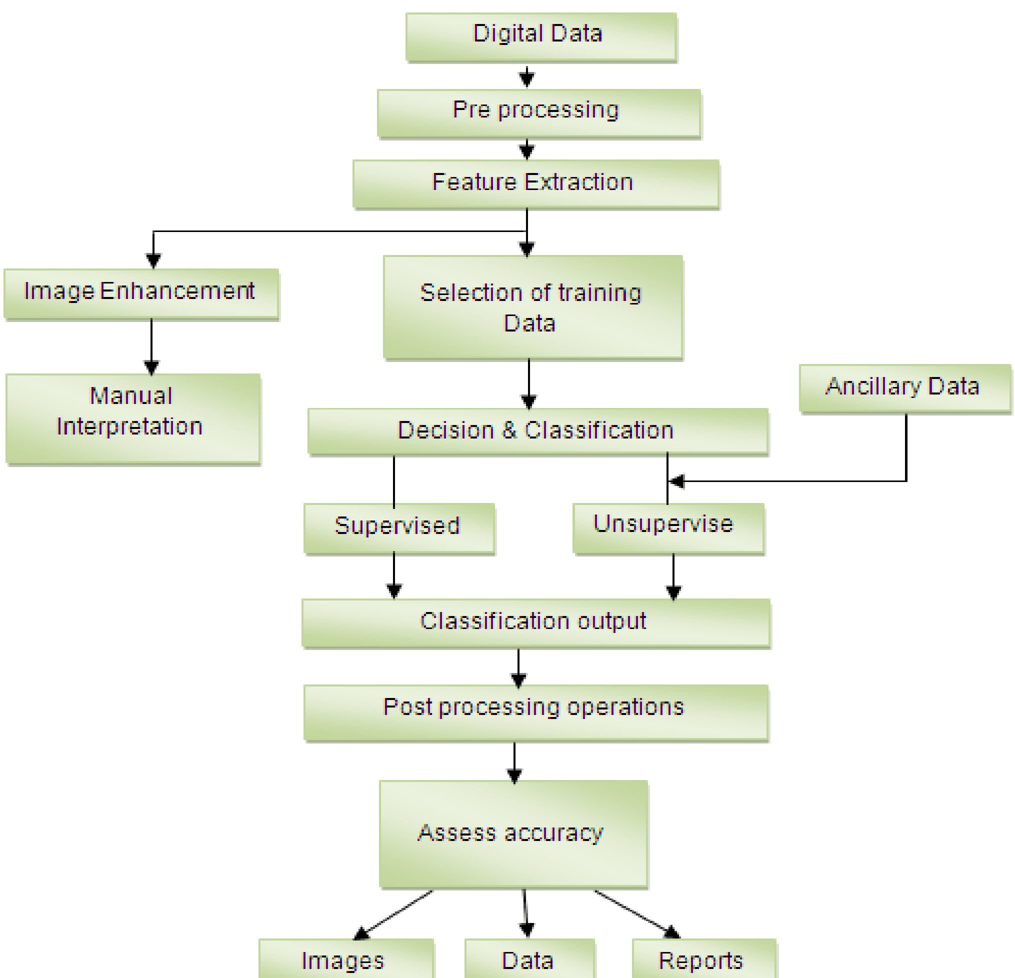
Digital Processing techniques help in manipulation of the digital images by using computers. As raw data from imaging sensors from satellite platform contains deficiencies. To get over such flaws and to get originality of information, it has to undergo various phases of processing. The three general phases that all types of data have to undergo while using digital technique are Pre- processing, enhancement and display, information extraction.

Closely related to image processing are:

Computer graphics and computer vision-

In computer graphics, images are manually made from physical models of objects, environments, and lighting, instead of being acquired (via imaging devices such as cameras) from natural scenes, as in most animated movies.

Computer vision, on the other hand, is often considered high-level image processing out of which a machine/computer/software intends to decipher the physical contents of an image or a sequence of images (e.g., videos or 3D full-body magnetic resonance scans).



Before going to processing an image, it is converted into a digital form. Digitization includes sampling of image and quantization of sampled values. After converting the image into bit information, processing is performed.

This processing technique may be Image enhancement, Image restoration, and Image compression.

Image enhancement:

It refers to accentuation, or sharpening, of image features such as boundaries, or contrast to make a graphic display more useful for display & analysis. This process does not increase the inherent information content in data. It includes gray level & contrast manipulation, noise reduction, edge crispening and sharpening, filtering, interpolation and magnification, pseudo coloring, and so on.

Image restoration:

It is concerned with filtering the observed image to minimize the effect of degradations. Effectiveness of image restoration depends on the extent and accuracy of the knowledge of degradation process as well as on filter design. Image restoration differs from image enhancement in that the latter is concerned with more extraction or accentuation of image features.

Image compression:

It is concerned with minimizing the number of bits required to represent an image. Application of compression are in broadcast TV, remote sensing via satellite, military communication via aircraft, radar, teleconferencing, facsimile transmission, for educational & business documents, medical images that arise in computer tomography, magnetic resonance imaging and digital radiology, motion, pictures, satellite images, weather maps, geological surveys and so on.

II. IMAGE SEGMENTATION:

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

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. Each of the pixels in a region is 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). 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.

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 :

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.

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 **Clustering methods**

Source image

Image after running *k*-means with $k = 16$.

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

The K-means algorithm is an iterative technique that is used to partition an image into *K* clusters. 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*.

Compression-based methods:

Compression based methods postulate that the optimal segmentation is the one that minimizes, over all possible segmentations, the coding length of the data. 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:

Histogram-based methods are very efficient when 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. 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 is repeated with smaller and smaller clusters until no more clusters are formed.

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 occur over 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 information result 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:

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.

Segmentation methods can also be applied to edges obtained from edge detectors. Lindeberg and Li developed an integrated method that segments edges into straight and curved edge segments for parts-based object recognition, based on a minimum description length (MDL) 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.

Region-growing methods:

The first region-growing method was 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 comparing 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 this way is allocated to the respective region. This process continues until all pixels are allocated to a region.

Seeded region growing requires seeds as additional input. The segmentation results are dependent on the choice of seeds. Noise in the image can cause the seeds to be poorly placed. Unseeded region growing is a modified algorithm that doesn't require explicit seeds. It starts off with a single region A_1 – the pixel chosen here does not significantly influence 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 significantly 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) is based on pixel intensities. The mean and scatter of the region and the intensity of the candidate pixel is 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) will be 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.

Split-and-merge methods:

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 so forth. Conversely, if four son-squares are homogeneous, they can be 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. 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.

Partial differential equation-based methods:

Using a partial differential equation (PDE)-based method and solving the PDE equation by a numerical scheme, one can segment the image. 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:

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 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:

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 nineties. 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, where its zero level 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 encodes numerous advantages: it is implicit, parameter free, provides a direct way to estimate the geometric properties of the evolving structure, can change the topology and is intrinsic. Furthermore, they can be used to define an optimization framework as proposed by Zhao, Merriman and Osher in 1996. Therefore, one can conclude that it is a very convenient framework to address numerous applications of computer vision and medical image analysis. Furthermore, research into various level set data structures has led to very efficient implementations of this method.

Fast marching methods:

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

Graph partitioning methods:

Graph partitioning methods can effectively be used for image segmentation. 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 dissimilarity 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, random walker, minimum cut, isoperimetric partitioning and minimum spanning tree-based segmentation.

Watershed transformation:

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:

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.

Multi-scale segmentation:

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:

Witkin's seminal work 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:

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 proposed to study how iso-intensity contours evolve over scales and this approach was investigated in more detail by Lifshitz and Pizer. 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 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 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 and been carried over to clinical use by Dam Vincken et al. Proposed a hyper stack for defining probabilistic relations between image structures at different scales. The use of stable image structures over scales has been furthered by Ahuja 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 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. Bijaoui and Rué 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:

In this 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.

Trainable segmentation:

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 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.

REFERENCES

a) Linda G. Shapiro and George C. Stockman (2001): "Computer Vision", pp 279-325, New Jersey, Prentice-Hall, ISBN 0-13-030796-3

b) 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.

c) K J. Batenburg, and J. Sijbers, "Adaptive thresholding of tomograms by projection distance minimization", Pattern Recognition, vol. 42, no. 10, pp. 2297-2305, April, 2009http://dx.doi.org/10.1016/j.patcog.2008.11.027

d) 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 Download paper

e) 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.

f) 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.

g) 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.

h) T. Lindeberg and M.-X. Li "Segmentation and classification of edges using minimum description length approximation and complementary junction cues", Computer Vision and Image Understanding, vol. 67, no. 1, pp. 88--98, 1997.

i) 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.

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

k) S.L. Horowitz and T. Pavlidis, Picture Segmentation by a Tree Traversa