International Engineering Journal For Research & Development

Vol.2
Issue 6

E-ISSN NO:-2349-0721

Impact factor : 2.25

ANALYSIS OF VARIOUS SEGMENTATION TECHNIQUES FOR OBJECT DETECTION AND RECOGNITION

H. G. Watgule[1] , P. A. Khodke[2]
[1] P.G. Department: Computer Science and Engineering Amravati, India, [2] P.G. Department: Computer Science and Engineering Amravati, India.
[1]harshalwatgule1@gmail.com, [2]priti.khodke@gmail.com

ABSTRACT—

Segmentation technique is used to divide the image into meaningful sub parts. It is also very useful in object recognition and extraction for images or logos. There are various methods of Segmentation among which Threshold method is considered as simplest one. Some techniques are suitable for noisy images. The strongest method of noise cancellation in images is Markov Random Field (MRF). In recent years, a new approach proposed for extraction and recognition of logos in image archives named as Context Dependent Similarity (CDS) kernel. Object Recognition provides domain independent technique for data analysis. Segmentation is very useful in many image applications because it is the first step in image analysis and recognition. The novel variational frameworks are available to match and recognize multiple instances of multiple reference logos in image archives. Test images and Reference logos are seen like constellations of local features and matched by minimizing an energy function mixing like fidelity term and so on. Local features which consider for matching are regions, interest points, etc.

Index terms—Object Extraction, Object Recognition and Segmentation.

I. INTRODUCTION

The meaningful structured division of image is known as segmentation. It is the one of the most important part in image processing. The entire image is divided into several parts which are more meaningful and easier for further use. Segmentation is classified into various types like Region Based, Edge Based, Threshold, Feature Based Clustering and Model Based. To reduce the information for easy analysis purpose is the main motto of segmentation. Segmentation is also useful in Image Analysis, Image Compression, Image Extraction and Image Recognition.

The companies, institutions and individuals have the expanding and massive production of visual data. The popularity of social sites like Flickr, YouTube are increasing day-by-day for diffusion and sharing of images and video. So they have urgent need of research in effective solutions for object detection and recognition to support automatic annotation of video and images and content-based retrieval of visual data [1]–[3]. The special class of visual objects are Graphic logos and they are extremely important for assessing the identity of something. Logos are nothing but the graphic productions that either simply displays some abstract signs or recall some real world objects or a name that have strong perceptual appeal to attract the customers attention and recall the expectations associated with a particular product or service [see Fig. 1]. Such economical relevance has motivates the companies to actively participates in developing smart image analysis solutions to scan logo archives to find evidence of already existing similar logos, malicious tempering in logos to deceive customers like small variations in the originals, unauthorized use of their logo.

Some logo has color relevance but the distinctiveness of logos is more often given by a few details carefully studied by graphic designers. To attract the attention of the customer graphic layout is very much important and to convey the message and to give surety permanently that they buying the best product.

Fig.1: Examples of popular logos

It is possible that with slightly different spatial disposition of the graphic elements different logos may have similar graphic layout such as having some difference in the shape, size and orientation or differ by some masks or traits [see Fig. 2].

Fig.2: Pairs of logos with small malicious changes

Logos are often appears in images of any real world indoor or outdoor scenes. They are superimposed on objects of any geometry. Logos are printed on shirts of persons or jerseys of players. Other than that they are seen on boards of shops or billboards and posters in sports playfields. In most of the cases they are subjected to perspective transformations and deformations, often corrupted by lighting effects or noise or occluded partially. Such images and logos have often relatively low resolution and quality. Regions that include logos might be small and contain little information [see Fig. 3]. Logo detection and recognition in these scenarios has become important for a number of applications like detection of near duplicate logos andunauthorized uses, verification of the visibility of advertising logos in sports events [3], [4].

Fig.3: Examples of logos displayed in real world images

I. Related Work

Providing some automatic support to the logo registration process was concerned with early work on logo detection and recognition. In order to ensure that logo is sufficiently distinctive and avoid confusion, the system must check in archives of millions that whether other registered logos have similar appearance to the newcoming logo image or not [5].

Kato's system [6] was among the earliest ones. In this a normalized logo image mapped to a 64 pixel grid and calculated a global feature vector from the frequency distributions of edge pixels. More recently, Wei [7] proposed a different solution, where logos were described by global Zernike moments, local curvature and distance to centroid. Wei proposed a novel content-based trademark image retrieval(TIR) system with a set of feature descriptors, which performs extraction of edges using Canny edge detector, performs a shape normalization procedure, then extract the global and local features of the trademarks. A strategy of two-component feature matching is used to measure the similarity between the database images and query [7].

Other methods have used different global descriptors of the full logo image either accounting for logo contours or exploiting shape descriptors such as *shape context* [8], [9]. A key characteristic of Shape Matching and Object Recognition approach is the estimation of shape similarity and correspondences based on a novel descriptor which is shape context. All these methods assume that a logo picture is fully visible in the image, is not subjected to transformations and is not corrupted by noise. According to this, they cannot be applicable to real world images.

Several authors proposed the use of global descriptor for real world images. Phan [10], considered pairs of color pixels in the edge neighborhoods and accumulated differences between pixels at different spatial distances into a Color-Edge Co-occurence Histogram [3]. This global descriptor permits to perform fast approximate detection of logos, but it is not suited for incomplete information or transformed versions of the original logo. Some authors used interest point and local descriptor to detect and recognise logos in real world image. Local visual descriptors like MSER [11], SIFT [12], SURF [13], have been proved to be able to capture sufficiently discriminative local elements with some invariant properties to geometric or photometric transformations and are robust to occlusions. The bag of visual words approach exploited to represent affine covariant local regions from a codebook of SIFT descriptors proposed by Sivic and Zisserman [14]. They showed good capability to discriminate between objects and gave also examples of logo matching in unconstrained environments. In Sivic and Zisserman approach, they did not account for relationships between near keypoints but simply defined a spatial proximity criterion, by checking the local context of the 15 nearest neighbors of each feature match.

Zhang [2] proposed a new and general pre-locating algorithm based on "Spatial Connected Component Descriptor" (SCCD) for rapid logo detection in unconstrained images. By employing SCCD for connected component (CC) matching, author get as few and exact LPRs as possible for logo detection. Lazebnik [15] proposed a technique works by partitioning the image into increasingly fine sub-regions and computing histograms of local features found inside each sub-region. The resulting "spatial pyramid" is a simple and computationally efficient extension of an orderless bag-of-features image representation. But Contextual information at the image level, such as in the spatial pyramid approach for whole image categorization is clearly not appropriate.

Jim [16] introduced a multiresolution spatial pyramid mining technique which treating each brand as a class of objects in order to identify frequent spatial configurations of local features of logos in real world scenes. Earlier work in logo detection has primarily focused on the problem of locating design marks in documents [16] or similarity matching for trademark retrieval [17]. Shape and contour based features are employed for analysis of well-registered data sets. Other work attempts detection in real world scenes. Rigid planar backgrounds and utilize line-based intensity profiles assume by Den Hollander and Hanjalic [18] for logos in sports videos. Quantized local SIFT descriptors frequent spatial configurations are capture by logo detection using data mining [19] association rules. Bagdanov [20], using a bag of local features and cloud localization employ normalized matching in sports videos. Quack [21], introduced association rules employing idea for object detection to select features. Jim extend Quack's bag-of-features method and apply it to logo detection in his work.

II. SEGMENTATION

Segmentation is the one of the most important part in image processing. The entire image is divided into several parts which are more meaningful and easier for further use. By rejoined these several parts, the entire image will cover. Segmentation may also depend on various features contained in the image. It may be either color or texture. An image is segmented before denoising to recover the original image. Segmentation is classified into various types as follows:

1. Edge Based
2. Region Based
3. Threshold
4. Feature Based Clustering
5. Model Based

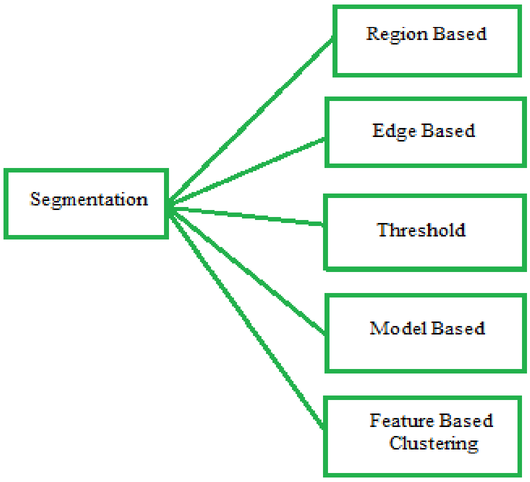

Fig.4: Various Types of Segmentation

1. *EDGE BASED SEGMENTATION*

By using edge detection techniques Segmentation can also be done. There are various techniques described in Fig 5. In this technique the boundary is identified to segment. To identify the discontinuities in the image edges are detected. Edges on the region are traced by identifying the pixel value and it is compared with the neighbouring pixels. They use both adaptive and fixed feature of Support Vector Machine **(SVM)** [7] for this classification. There is no need for the detected edges to be closed in this edge based segmentation. To segment the image various edge detectors are used [22].

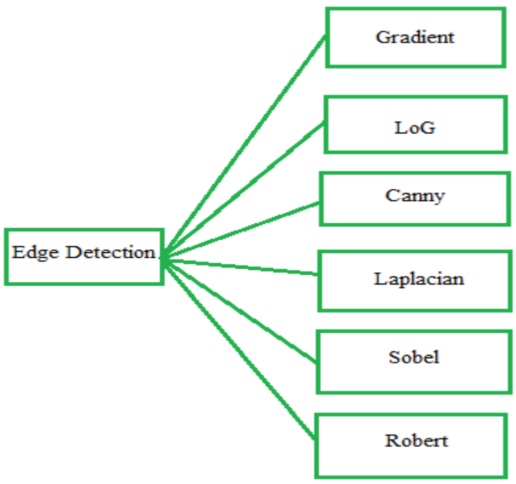

Fig.5: Type of Edge Detection

- *Algorithm: Edge-Based Segmentation*

Given an image *f*,

1. Compute an edgeness image ∇f from *f*. Any preferred gradient operator can be used for this.

2. Threshold ∇f to an image $(\nabla f)t$, so we have a binary image showing edge pixels.

3. Compute a Laplacian image Δf from *f*. Any preferred discrete or continuous

Laplacian operator may be used.

4. Compute the image $g = (\nabla f)t \cdot \mathrm{sgn}(\Delta f)$.

The sgn operator is used to returns the sign of its argument. Therefore the result image *g* will only contain three values: 0 at non-edge pixels of *f*, 1 at edge pixels on the bright side of an edge, and −1 at edge pixels on the dark side of an edge. The boundaries of the objects to be segmented are contains by the image *g*. For facilitate the final steps of the algorithm the Laplacianis used. Turning the boundary image into a segmented image *h* containing the solid objects. If the image g will traverse from left to right, two adjacent pixels with values 1 and −1 means we move out of one, and the values −1 and 1 means we move into an object. Therefore by setting all pixel values to zero, the image *h* can be created, except for those pixels that are between the transitions$-1 \rightarrow 1$ and $1 \rightarrow -1$ in each line of *g*, which are set to 1. A labelling algorithm can be run on *h*, if unique values are desired for each separate segment [23].

2. *REGION BASED SEGMENTATION*

This approach finds the object *region* instead of its edges. In theory, finding an object by establishing the region it covers and finding it by locating its boundary will give exactly the same object. The region and boundary are nothing but different representations of the same object. Region based segmentation methods uses two basic operations i.e. *Merging* and *Splitting*, many methods even used only one of these [23].

The basic approach using merging is:

1. Obtain an initial (over)segmentation of the image,

2. Merge those adjacent segments that are similar – in some respect– to form single segments,

3. Go to step 2 until no segments that should be merged remain.

The initial segmentation each pixel is a segment by itself, *i.e.* simply be all pixels. The similarity criterion is the heart of the merging approach. It used to decide whether two segments should be merged or not. This criterion may be based on the texture of the segments, grey value similarity, the edge strength of the boundary between the segments or one of many other possibilities [23].

The basic form using splitting is:

1. Obtain an initial (under)segmentation of the image,

2. Split each segment that is inhomogeneous in some respect (*i.e.*, each segment that is unlikely to *really* be a single segment).

3. Go to step 2 until all segments are homogeneous.

No segmentation at all may be the initial segmentation, *i.e.* there is only a single segment.

The inhomogeneity criterion of a segment may be the variance of its texture, the occurrence of strong internal edges, and the variance of its grey values or various other criteria.

The basic splitting and merging methods seem to be the same as bottom-up and top-down approach of segmentation but there is an intrinsic difference. The merging of two segments is straightforward but the splitting of a segment requires suitable sub-segments of the splited segments. Still there is a segmentation problem, but now it defined on a more local level.

For avoiding this problem, the basic splitting approach is often enhanced to a combined into *split and merge* approach. In which inhomogeneous segments are split in simple geometric forms (usually into four squares) recursively. This creates arbitrary boundaries of segment and to remove incorrect boundaries merge steps are included [23].

3. *THRESHOLD BASED SEGMENTATION*

The easiest way of segmentation is Thresholding. From the histogram of edges of the original image threshold values are obtained. From the edge detected image, the threshold values are obtained. Therefor if the edge detections are accurate then the threshold value is also accurate. As compared to other techniques segmentation using Thresholding has fewer computations. On "his ton" Segmentation is based [22]. "his ton" is defined as there may be set of pixels for a particular segment. For image segmentation a thresholding method followed Roughness measure. Adaptive thresholding is used for segmentation. The gradient is high in the gray level points. After that for segmentation gray level point's gradient is then added to thresholding surface. It is not suitable for complex images is the drawback of Threshold segmentation technique. The most frequent technique used to segment an image is Thresholding. The thresholding operation is nothing but a grey value remapping operation *g* defined by:

$$g(v) = \begin{cases} 0 \; if \; v < t \\ 1 \; if \; v \geq t, \end{cases}$$

Where,

v represents a grey value and

t is the threshold value

In thresholding there is mapping between a grey-valued image to a binary image. The image has been segmented into two segments after the thresholding operation. Pixels value is identified by the values 0 and 1. Thresholding can be used to segment the image if an image contains bright objects on a dark background. To segment an image into objects and background thresholding is often a well-suited method. If there are not objects overlapping then from each object create a separate segment by running a labelling algorithm on the thresholded binary image, thus unique pixel value assigning to each object.

- *Algorithm: Iterative Thresholding*

1. Assume that the four corner points of the image are background pixels (part of segment 0), and set $\mu 0$ to the average grey value of these four pixels. Assume all of the other pixels are object pixels, and set $\mu 1$ to their average grey value.

2. Set the threshold t to $t = 1/2\ (\mu 0 + \mu 1)$, and segment the image.

3. Recompute $\mu 0$ and $\mu 1$, the mean of the original grey values of the two segments.

4. Go to step 2 and iterate until the threshold value no longer changes (or no longer changes significantly).

4: Feature Based Clustering Techniques

Feature Based Clustering is also used in Segmentation. The collective name for methods that attempt to group together measurements points is clustering techniques. For data of arbitrary dimension clustering techniques are often formulated. But on two or three-dimensional images many clustering methods can be applied. Images similar in appearance to the rabbit plot are

best suited for applying clustering techniques [23]. K-means is a basic clustering algorithm which is used for segmentation in textured images.

- *Algorithm: K-means Clustering*

This algorithm minimizes the total distance of data points to the cluster center of the cluster they are assigned to. Note that the actual computation of distances it does not require.

1. Select the number of desired clusters *k* arbitrarily (or better: intelligently) place the *k* cluster centers at different initial locations in the image.

2. Assign each data point to the cluster whose center is closest.

3. Recompute the cluster centers; the cluster center should be at the average coordinates (center of gravity) of the data points that make up the cluster.

4. Go to step 2 until no more changes occur or a maximum number of iterations is reached.

5: Model Based Segmentation

Model based segmentation is known as Markov Random Field (MRF) based segmentation. An inbuilt region smoothness constraint which is used for color segmentation is presented in MRF. For further processing, components of the color pixel tuples are considered as independent random variables. To identify the edges accurately MRF is combined with edge detection. There are correlations among the color components and MRF has spatial region smoothness constraint [22].

Expectation-Maximization (EM) algorithm values the parameter is based on unsupervised operation. “Narrow Band” are called as Multiresolution based segmented technique. This technique is faster than the traditional approach. Coarse resolution is used to performed the initial segmentation and then at finer resolution. In an iterative fashion the process moves on. It is fast because only to the part of the image the resolution based segmentation is done. For the process where the spatial dependencies between pixels are considered in that the segmentation may also be done by using Gaussian Markov Random Field **(**GMRF**).** For region growing Gaussian Markov Model (GMM) based segmentation is used. By using this technique feature space is also detected [23].

COMPARISON TABLE: Comparison Between Various Segmentation Techniques

Sr. No.	Technique Name	Parameters			
		Approach used	*Color*	*Computation based on*	*Algorithm*
1.	Region Based	Splitting and Merging	Gray Level	Similarity	Region Growing
2.	Edge Based	Edge Image	Gray Level and Colour	Difference	Edge-Based Segmentation
3.	Threshold	Adaptive Thresholding	Gray Level	Threshold	Iterative Thresholding
4.	Feature Based Clustering	Clustering	Colour	Similarity	K-Means Clustering
5.	Model Based	Markov Random Field (MRF) Based	Colour	Colour Components	Expectation-Maximizatio (EM)

The above table describes the comparison between various segmentation techniques. The table contains parameters like approach used, colour, computation based on and algorithm used by the technique. It having techniques name such as Region Based, Edge Based, etc. Other parameters are also described in brief in above section named as segmentation.

I. Conclusion

The special class of visual objects are Graphic logos and they are extremely important for assessing the identity of something. Logo detection and localization approach is based on segmentation. On any type of image segmentation it can be applied. Depending on the variance of the application technique, thresholding is the simplest and computationally fast as compare to other methods.

On the basis of segmentation, future work includes use of segmentation techniques for the refinement of the definition of context in order to handle rigid and non-rigid logo transformations and also logo retrieval in videos.

References

[1] H. Sahbi, L. Ballan, G. Serra and A. D. Bimbo, “Context-Dependent logo matching and recognition,” IEEE Transactions on image processing, vol. 22, no. 3, March 2013.

[2] Y. Zhang, S. Zhang, W. Liang, H. Wang “Spatial connected component pre-locating algorithm for rapid logo detection,” 978-1-4673-0046-9/12 IEEE, 2012.

[3] N.Krishnan, P. Vijayalakshmi ,R. Nishanthini, D. A. Dev, “Automatic detection of colour objects using spatially weighted GVF,” International Journal of Advanced Engineering Applications, Vol.4, Iss.4, pp.62-66, 2011.

[4] J.-L. Shih and L.-H. Chen, “A new system for trademark segmentation and retrieval,” Image Vis. Comput., vol. 19, no. 13, pp. 1011–1018, 2001.

[5] A. Joly, O. Buisson, C. Frélicot, “Content-based copy retrieval using distortion-based probabilistic similarity search” IEEE Transactions On Multimedia, Vol. 9, No. 2, February 2007.

[6] T. Kato, “Database architecture for content-based image retrieval,” Proc. SPIE, Image Storage Retr. Syst., vol. 1662, no. 1, pp. 112–123, Jan. 1992.

[7] C.-H. Wei, Y. Li, W.-Y. Chau, and C.-T. Li, “Trademark image retrieval using synthetic features for describing global shape and interior structure,” Pattern

Recognit., vol. 42, no. 3, pp. 386–394, 2009.

[8] J. Rodriguez, P. Aguiar, and J. Xavier, “ANSIG—an analytic signature for permutation-invariant two-dimensional shape representation,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Anchorage, AK, pp. 1–8, Jun. 2008.

[9] S. Belongie, J. Malik, J. Puzicha, “Shape matching and object recognition using shape contexts,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 4, pp. 509–522, Apr. 2002.

[10] R. Phan and D. Androutsos, “Content-based retrieval of logo and trademarks in unconstrained color image databases using color edge gradient co-occurrence histograms,” Comput. Vis. Image Understand., vol. 114, no. 1, pp. 66–84, 2010.

[11] D. Crandall, J. Luo, “Robust color object detection using spatial-color joint probability functions,” Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition.

[12] D. Lowe, “Distinctive image features from scale-invariant keypoints,” Int. J. Comput. Vis., vol. 60, no. 2, pp. 91–110, 2004.

[13] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, “Speeded-up Robust Features (SURF),” Comput. Vis. Image Understand., vol. 110, no. 3, pp. 346–359, 2008.

[14] J. Sivic and A. Zisserman, “Efficient visual search of videos cast as text retrieval,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 31, no. 4, pp. 591–606, Apr. 2009.

[15] S. Lazebnik, C. Schmid, and J. Ponce, “Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories,” in Proc. Conf. Comput. Vis. Pattern Recognit., pp. 2169–2178, vol. 2. 2006.

[16] F. Cesarini, E. Francesconi, M. Gori, S. Marinai, J. Q. Sheng, and G. Soda, “A neural-based architecture for spot-noisy logo recognition,” in ICDAR, 1997.

[17] Jain and A. Vailaya, “Shape-based retrieval: A case study with trademark image databases,” in Pattern Recognition, 31(9):1369–13990, 1998.

[18] R.J.M. den Hollander and Alan Hanjalic, “Logo recognition in video stills by string matching,” in ICIP, 2003.

[19] Jim Kleban, Xing Xie, Wei-Ying Ma, “Spatial pyramid mining for logo detection in natural scene”, IEEE 978-1-4244-2571-6/08, 2008.

[20] A. D. Bagdanov, L. Ballan, M. Bertini, and A. Del Bimbo, “Trademark matching and retrieval in sports video databases,” in MIR, 2007.

[21] T. Quack, V. Ferrari, B. Liebe, and L. V. Gool, “Efficient mining of frequent and distinctive feature configurations,” in ICCV, 1997.

[22] R.Yogamangalam, B.Karthikeyan, “Segmentation techniques comparison in image processing”, International Journal of Engineering and Technology (IJET), Vol 5 No 1 Feb-Mar 2013.

[23] http://www.cs.uu.nl/docs/vakken/ibv/reader/chapter10.pdf