Survey of Image Object Recognition Techniques

Gurjinder Singh[1], Jaspreet Kaur[2]

[12]Guru Gobind Singh College of Science and Engineering Guru Kashi University Talwandi Sabo (Bathinda), India

---

Abstract- *Object recognition is a process of detecting the object present in an image or a video sequence, with the help of some recognition technique or methods. Object recognition is one of the techniques of digital image processing where we can process any image by applying some of the operation. The modern world is enclosed with gigantic masses of digital visual information. Increase in the images has urged for the development of robust and efficient object recognition techniques. Most work reported in the literature focuses on competent techniques for object recognition and its applications.*

Key words: Object recognition, Spatial, Temporal, Mining

1. INTRODUCTION

Object recognition is the task of finding a given object in an image or video sequence. For any object in an image, there are many 'features' which are interesting points on the object that can be extracted to provide a "feature" description of the object. This description extracted from a training image can then be used to identify the object when attempting to locate the object in a test image containing many other objects [1][2].

Goals of Object Recognition

- Goal is to retrieve information that is not apparent in the images we perceive. The name of things is one piece of information
- Animals recognize without words. Important information may be whether to ignore, eat, flee, etc.
- A robot could connect the objects it sees to the information it knows about them, and also connect new information about objects to what it already knows about them

1.2 IMAGE MINING PROCESS

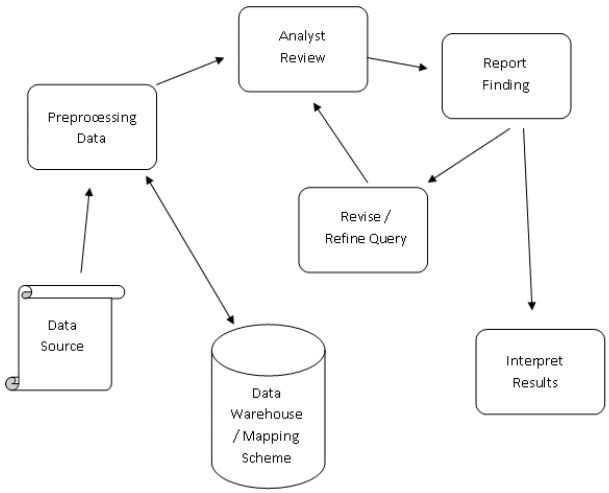

Fig-1: Data mining process interpreting mined results.

Data mining is the feature detection technique of image data and the use of software based techniques for finding image patterns and bitmap regularities in sets of image data. The computer program is responsible for observing the image bit patterns by elucidating the pre-defined rules and image features in the image data. Data collected from different sources like databases or flat files or any other resources like web is made to be preprocessed. Various standardization techniques of data cleaning or formatting

are to be used by data warehouse or mapping schemes. [4] Further preprocessed data is made to be reviewed by data analyst who is done nowadays by computers software. In earlier days it was manually done by data analysts. After reviewing preprocessed data analyst reports are to be made which is revised or refined if something is needed with hand on information to generate finally mined or interpreted results.

2. SOME OF THE IMPORTANT OBJECT RECOGNITION TECHNIQUES ARE LISTED AS BELOW:

2.1 Spatial Relations

One of the most mandatory features in many multimedia documents is management of spatial coherence. Management of spatial relation of components of image data is done by considering its rectangular Cartesian coordinates. The spatial axis position of a component in object in image data is represented by rectangular coordinates and the relations between components are calculated mathematically. [7, 3] A multimedia document which is composed of images, flowcharts, and other random graphics as well as image text is another example that proves the management of spatial coordinates for layout information in an image [5]. In other applications of image processing such as geographical information system (GIS), the image pattern representation and random indexing of abstract and finite spatial relations in some regions of image is studied. A 2D stream of bits [3], [7] is an image data indexing technique for representing a spatial pattern between different components of an image; 2D stream of bits represents few patterns in position of components in neighboring regions, which is composed of horizontal and vertical ordering of image components. Also, it gives several distinct levels of a sparse-strict spatial relation, where the stiffness of directional data differs from one level of hierarchy to other. In [6] a set of two dimensional bits such as '00' '11' '01' '10' '111' '100' '101' '001' and '011' is defined as primary relations for representing spatial region of image data coordinates in observed components of image. This method along with 2D stream of bits is suitable for sparse bit evaluation of spatial relation in image patterns. The advantage of above method of representation of bits of image is that it left subtle denominations of spatial relations that are not required to be evaluated. However, the spatial relation is not distance-dependent, non-interval-based bits in image data are out of their scope of findings.

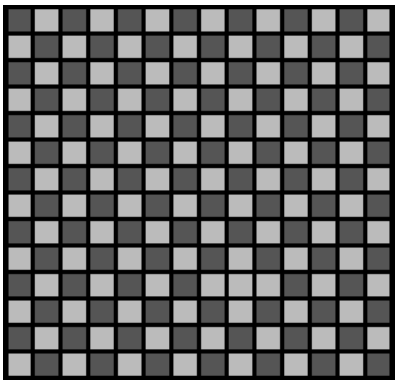

Fig- 2: Check-board with one square colored incorrectly

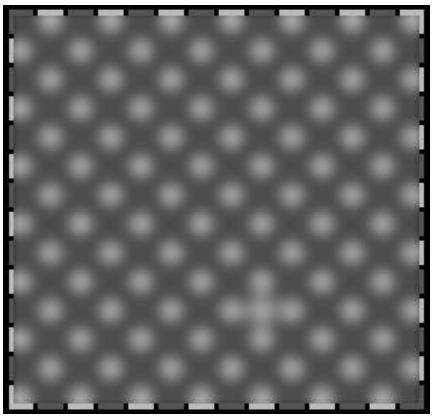

Fig-3: Blur image of check-board

2.2 Temporal Relations

A recent study depends upon bit representation and image data management of temporal relation considers video based applications of image processing such as video information databases (VID). There are two basic techniques for representing time based relationship between objects in multimedia data of a moving image: One is a point-based image data representation containing time lag of various colors in a multimedia document, and the other is an interval-based image data representation which contains intervals of data shifts from neighboring regions of data. [12, 14, 18] The point based representation contains the Cartesian position of objects by points in the image data on timeline, whereas the interval-based representation contains the cohesion of image objects by means of their intervals of the occurrences of similar intensities in neighboring regions of an image. Many studies shows

Interval based modeling technique of data representation. is used to manage the time lag in between various component videos objects that are used in multimedia document. Video elements of document (VID) are shown as if a stream of video bit patterns being defined by the time lag intervals, a video object comprises of different streams of video frames which are stored in contiguous manner. [23, 24] Video objects can be changed using two operations of merge and overlap. To be part of the definition of every video object the textual annotation of the contents of the video are merged or overlapped. Both the frame based and interval based specification are used to query a video object in OVID. Gibbs et.al [11] used a timed based representation of component objects. Sequential representation of component objects is used in this study. So this technique did not give a natural representation of a condition. Reverse and temporal relations were discussed and extended by little and Ghafoor. Allen's basic definition of temporal relations was followed by Hopner.

2.3 Data Retrieval In Conventional Databases

Earlier database systems were designed to manage and retrieve the textual data and their related keyword based retrieval techniques which are not suitable for retrieving the data which comprises of text, video and audio data. Adding to this, in earlier times lot of human effort was required for the manual annotation because as the data which was stored in binary form was of no meaning to the human. Similarity comparison techniques are considered to be the best techniques for the exact matching of the applications. To get the best solutions the mechanisms of filtering and ranking are used. The user can store and query various types of multimedia information with the help of multimedia database system. The process of retrieving images which are exactly same to a given image can be done by the typical multimedia type. [8, 11, 9, 10] So that is why it is known as the content based retrieval technique. This process can be carried out at three different levels of abstraction. First one is raw data in which the comparison is done by comparing each pixel with the other pixel. In this search is also done on the different features of the objects like its texture etc. The other is semantic in which the grouping of different objects is done based on its meaning. The configurationally similarity is based on the direction, topological address and the relation of image object with respect to distance. The preprocessing is required by almost all the techniques which are global or local.

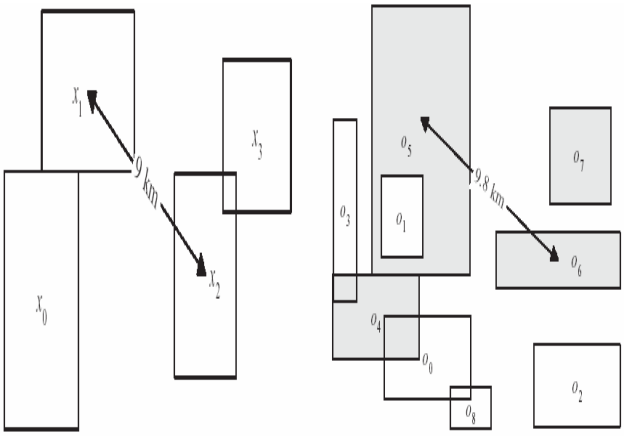

Fig-4: Nine regions of interests in a binary encoding of relations

2.4 Image Extraction Through Mining

The technique of handling the association of image data and its patterns which are not stored in its images is known as image mining. Different methods of image processing and retrieval are used by this technique. There are two approaches to mining, firstly to mine from a large database of images and the other one performed on different collections of images and the data. The co-ordination of structures and the human brain can be deteriorated by this technique. The actual meaning of image mining is the production of different patterns with no knowledge of the image content. There can be different kinds of patterns like description, temporal, spatial etc. [21, 23] All huge image databases are being handled but the image mining. This system includes different image retrieval as well as indexing technique for pattern matching. This method is considered as if user friendly for data patterns and can generate in depth knowledge of image representation. The different patterns are grouped together .The sample of different images is taken to extract the content. But during this process different type of distortions and disturbances are being met. The results of mining can only be obtained after the matching of the model description with the symbolic prescription.

2.5 Content Based Retrieval Technique

In this technique a complete image is segmented into small parts which further comprise to form the group of small images based on the index and layout feature. For the large scale data management (AIR)

Automated Image Retrieval System is used to develop the Image Science and machine Vision (ISMV) of ORNL. In the olden days textual description was handled manually. But there were different loopholes such as semantic structure can vary with respect to a retrieval task and it was also a very time consuming process with regard to the size of image databases. The solution to these problems was content descriptors. In these technique global dimensions of the image was used. [16, 18, 20] Due to this another concept came into being that was named as Image segmentation in which the internal content of image was sidelined .But there is no correct technique of segmentation. Although there are different techniques but we could not arrive at the best one. So a definition of segmentation came into being i.e. it is a process with the help of which we can label each and every point of image including the class label. Another feature of segmentation is the capability to work in the unplanned sequence. These techniques do not rely on continuous human efforts. This problem is due to the restriction applied on the size. Due to everyday increase in the size of image which leads to hindrance in the way of conversation. [12, 10, 13, 14] The gap between the low level and high level interpretation can be decreased with the help of image segmentation. The low level image content should be user friendly so that changes can be performed at any stage in a image in CBIR

system. So different steps must be taken to reduce this gap. The Semantic representation is based on color, spatial and texture information.

The deformities and ambiguities in different object recognition is due to spatial and spectral adjancies which are based on the type of shadow, object and their complex relations with the adjacent objects. [18] The object recognition can be improved with the use of a knowledge base. This process can be achieved by going through two stages, first one is to define the properties of each segmented region and the second one is to classify results based on the views of the classified objects. [15, 17, 19, 20] The scene contextual information is based on the height, sensory information, the analysis of the visibility mapping etc. the methodology of object recognition is performed at two databases i.e. multi view digital angular camera and multi angular worldview

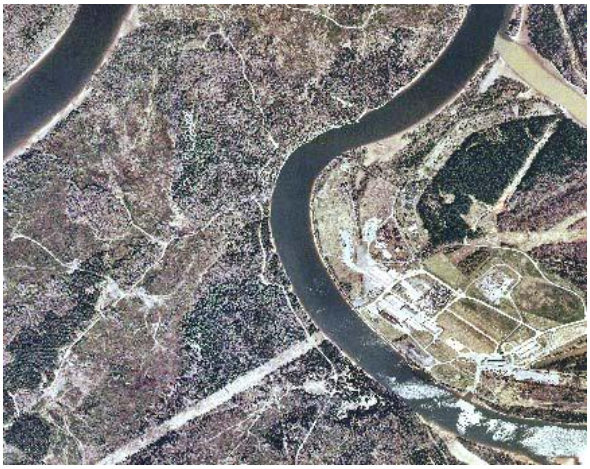

Fig-5: Oak Ridge image (2m×2m)

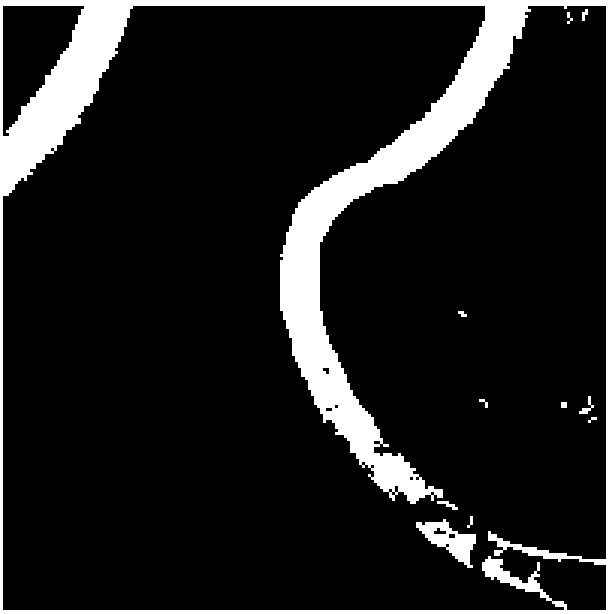

Fig-6: Color Histogram image of water area

2.6 Image Object Retrieval Using Context Aware Technique

The object recognition can be improved with the use of a knowledge base. This process can be achieved by going through two stages, first one is to define the properties of each segmented region and the second one is to classify results based on the views of the classified objects. The scene contextual information is based on the height, sensory information, the analysis of the visibility mapping etc. the methodology of object recognition is performed at two databases i.e. multi view digital angular camera and multi angular worldview. [22] The gap between the low level and high level interpretation can be decreased with the help of image segmentation. The low level image content should be user friendly so that changes can be performed at any stage in a image in CBIR system.

3. CONCLUSION

In image pre-processing loaded input image undergoes segmentation, finding bright intensity values of pixels and their belongingness to particular region is noted in the form of feature. Out of which extracted 2x1 matrix is maintained. This matrix shows presence of image object particles in an image. This algorithm is tested on light background images of courtesy standard benchmark images; this has found its moderate performance with these images as compared to other allied approaches like Convolution Based Detection,

REFERENCES

[1] V. Ferrari, T. Tuytelaars, and L.V. Gool(2004), "Simultaneous Object Recognition and Segmentation by Image Exploration," Proc. Eighth European Conf. Computer Vision, 2004.

[2] B. Leibe, A. Leonardis, and B. Schiele(2004), "Combined Object Categorization and Segmentation with an Implicit Shape Model," Proc. ECCV Workshop Statistical Learning in Computer Vision, 2004

[3] Burger and Lenddins, (2010), *Digital Image processing using Matlab*, Mcgraw Hill's New Delhi

[4] Dwayne and Philipps, (2000), *Analyzing and Enhancing Digital Images,* R & D Books, US

[5] Hanry and Ford, (2010), *Fundamentals of image Processing*, Dartmouth Inst Press, US

[6] Koprowsky, Wrobel and Ketty, (2011), *Image Processing in Optical Coherence Tomography,* University of Silesia, US

[7] Pettis and Jay, (2012), *Getting Started with Image Tools,* O'Reilly Media, New Delhi

[8] Strack and Murtagh, (2008), *Astronomical Image and Data Analysis,* Springer

[9] Bottou, Cortes, Denker, Drucker, Guyon, Jackel, LeCun and Muller,(1994) "Comparison of classifier methods: A case study in handwritten digit recognition," in Proc. 12th IAPR Int. Conf. Pattern Recognit., vol. 2, pp. 77–82.

[10] Phillips, Gleeson, (2001) "Volatile organic compounds in breath as markers of lung cancer:a cross-sectional study", Lancet , P.353:1930–33

[11] Yufeng Zeng, (2013), *Image Fusion and Its Applications,* Intech Press, Canada

[12] Edwads, Staples, Thei zNose, (2000) "A New Electronic Nose Using Acoustic Technology", Acoustical Society of America, P.242-250

[13] Berg, Olsson, Lindblad, and Chilo, (2008) "Automatic design of pulse coupled neurons for image segmentation," Neurocomputing, vol. 71, nos. 10–12, pp. 1980–1993.

[14] Cao, Luo, Kautz, and Huang, (2009) "Image annotation within the context of personal photo collections using hierarchical event and scene models," IEEE Trans. Multimedia, vol. 11, no. 2, pp. 208–219

[15] Wang, Hua, Tang, and Hong,(2009) "Beyond distance measurement: Constructing neighborhood similarity for video annotation," IEEE Trans. Multimedia, vol. 11, no. 3, pp. 465–476

[16] Afraz and Cavanagh, (2009)."The gender specific face after effect is based in retinotopic not spatiotopic

coordinates across several natural image transformations." J.Vis. 9,10.1–10.17.doi:10.1167/9.10.10

[17] Everingham, Gool, Williams, Winn, and Zisserman, (2010) "The Pascal visual object classes (VOC) challenge," Int. J. Comput. Vis., vol. 88, no. 2, pp. 303–338.

[18] Laeng, Okubo, Saneyoshi, Michimata, (2011), "Processing spatial relations with different apertures of attention", Cogn Sci35:297–329

[19] Arbelaez, Maire, Fowlkes, and Malik, (2011) "Contour Detection and Hierarchical Image Segmentation," PAMI, vol. 33, no. 5.

[20] Chen, Park, Ma, and Ala, (2011) "A new automatic parameter setting method of a simplified PCNN for image segmentation," IEEE Trans. Neural Netw., vol. 22, no. 6, pp. 880–892.

[21] Wei, Wang, and Lai, (2012) "Compact image representation model based on both nCRF and reverse control mechanisms," IEEE Trans. Neural Netw. Learn. Syst., vol. 23, no. 1, pp. 150–162.

[22] L.Tadic, (2012) "Color Space Basics", AMIA Tech Review. Issue 4. pno 123

[23] Westland S. (2012) "Computational Colour Science using MATLAB 2e".By WILEY ISBN: 978-0-470-66569-5 p:27-104.

[24] Wu, Chensheng, Christopher, Davis.(2014) "Modified plenoptic camera for phase and amplitude wavefront sensing." International Society for Optics and Photonics, pp. 88740I-88740I.

BIOGRAPHIES

Gurjinder Singh received BCA in 2012 and Msc (IT) in 2014 Under GNDU(Guru Nanak Dev University, Amritsar).Now His M.Tech(IT) perusing From Guru Kashi University ,Talwandi Sabo (Bathinda),India.