From Cognitive Psychology to Image Segmentation: A Change of Perspective

Anju Mishra1, Priya Ranjan1, Sanjay Kumar3 and Amit Ujlayan2

[1]Amity University Uttar Pradesh, Noida, India
{amishra1,pranjan}@amity.edu
[2]Gautam Buddha University, Greater Noida, India
iit.amit@gmail.com
[3]Oxford Brookes University, Oxford, UK
skumar@brookes.ac.uk

Abstract. Image segmentation is a complex and essential task used in many computer vision applications. The problem of image segmentation can essentially be formulated as a grouping problem which in its simplest form tries to group the pixels of image into distinguished regions of interest so that further processing of the extracted regions can be achieved. This work proposes an image segmentation model which is inspired by the findings in cognitive psychology theories to divide the image into separate coherent regions. The proposed work tries to correlate between human and machine cognition by studying the segmentation process under the light of psychology of human vision.

Keywords: Image Segmentation·Foreground Extraction·Cognitive Psychology·Machine vision

1 Introduction

Image segmentation is the task of decomposing an image into its constituent regions of interest i.e., objects and the background present in that image. A lot of work [1-4, 10, 13] has already been done in this direction and various types of approaches have been formulated and proposed by researchers across the globe to cope up with this simple task. For many applications separating figure and ground is a necessary and inherent step for further processing e.g., developing intelligent machines and cameras that can automatically separate the object of interest based upon the context of application. Even though, for a human the task of separating figure from its ground is very simple and inherent, from computational or machine point of view, the same task is considered to be a difficult one and as the nature or the characteristics of the environment (e.g., noise and illumination variations and variations in camera angles and perspective etc.) changes, the task is transformed into a complex problem. From a psychological perspective the problem can be defined as a grouping problem in which the visual system tries to group the visual stimuli based on different properties of these stimuli e.g., intensity, contrast, distance, similarity of color, texture etc.
The Gestalts' school of Psychology [6-8, 11-12] follows that perceptual grouping plays a vital role in human visual perception and based on these thoughts, many researchers have formulated various algorithms which follow closely the line of perceptual grouping. Since, perceptual grouping can be linked most naturally to the problem of figure-ground separation (i.e., segmentation process), all these proposed algorithms tried to group image pixels based on different pixel properties- be it the color, texture, or contrast between adjacent pixels. An efficient image segmentation algorithm must have two important properties:

(a) It must capture perceptually important regions within an image.
(b) It must be computationally faster in terms of running time so as to be applicable in real-time scenario.

2 Related Work

The latest computer vision algorithms follow one or more of the following mathematical models [2, 4, 9, 13, 14, 15, 17]:

Graph-based Approaches: Models the set of image pixels as a graph and segmentation process to be a graph partitioning problem.

Clustering Methods: Models image pixels to belong to different clusters or groups and segmentation process to be the problem of finding non-overlapping coherent groups of pixels.

Appearance-based Models: Models the image as having different appearances for figure and ground and segmentation process to be the problem of finding key appearance features e.g., key feature points etc.

Classification Techniques: Attach a class label with each image pixel to find different areas/regions/patterns inside the image.

Thresholding and Histogram Methods: Focus on forming the bins of pixels based on intensity value and try to find a threshold value suitable to divide the image into segments.

Edge Detection Methods: Segment the image into edges and background.

Region-based Methods: Partition the image into separate regions.

Template Matching: In this method segmentation is achieved by creating templates (representation of objects) and then matching them against the input image to find the presence of object in question.

Many psychologist have proposed different viewpoints supported by experiments to prove that visual perception process is either a top-down process or bottom-up process or a combination of these processes [5]. According to Gestalt theory, visual perception is a global phenomenon which starts from local characteristics of visual stimuli. This leads to the perception of those structures which are constituted from elementary perceptual stimuli but are more stable and strong in terms of global properties. The winning visual structure is constituted as 'Figure' and the loser as 'Ground'. The Gestaltist have proposed the laws of visual perception which are classified into two categories:

(i) *Local or Elementary grouping laws*
(ii) *Global grouping laws*

The process of visual perception starts by considering local characteristics such as similarity (of color, shape, texture etc.), constant width, vicinity etc. [6-8]. Whenever the points in a neighborhood have common characteristics, they get grouped and form visually extended objects. Since, all the elementary grouping laws work simultaneously, many overlapping visually extended (global) perceivable structures are formed simultaneously. Finally, global grouping laws governing these extended (global) structures compete with each other and the winning law which governs the most stable structure contributes to the perception of that structure as an object.

The Gestalt view suggests a bottom-up approach for figure-ground perception, where elementary grouping laws works first and form higher-level groups and in the next step of refinement process these higher-level groups are synthesized by application and collaboration of global gestalt laws, as a result different interpretations for figure and ground in the same image are possible.

By the efforts made by the psychologist it was proved experimentally that visual cognition and perception is not only a bottom up process but sometimes it is a top-down process or a mixture of both. Apparently, prior information and cues or previously learned facts play an important role in the visual cognition process and the domain knowledge and past experiences also speeds up the whole process of visually perceiving objects from a scene [5,11-12].

3 Perceptual Analysis of Existing Computer based Segmentation Methods

By making use of these psychological findings, computer-based algorithms can be formulated to perform image segmentation. [6,8] has identified that the gap between human and computer vision is due to the fact that human vision primarily depends on the physiology as well as the psychology of brain system while all computer based methods and systems primarily focus on the physiological aspects of human brain and try to imitate the functioning of the brain system. The most commonly available computer based segmentation techniques can be broadly classified into three different categories inspired by psychological findings:

(i) *Top-Down Approaches*
(ii) *Bottom-Up Approaches*
(iii) *Knowledge based or hybrid approaches*

Top-Down Approach	**Bottom-up Approach**	**Learning(Knowledge) based/ Hybrid Approach**
Appearance models	Graph based	Template Matching
Classification	Clustering	Interactive Techniques
Region Splitting	Thresholding	
	Edge based	
	Region growing	

Table 1. Classification of Computer based Segmentation Methods according to psychological findings.

The working principal of popular automated image segmentation techniques are explained below which clearly shows that all these techniques got their intuition from perceptual grouping principles proposed by Gestalt's which says that process of visual

perception starts by considering local characteristics such as similarity (of color, shape, texture etc.), constant width, vicinity etc. Whenever the points in a neighborhood have common characteristics, they get grouped and form visually extended objects.

1. Graphs: Graph-based approaches to image segmentation tries to create a graph of image pixels wherein pixels represent the vertices of graph and these are linked by edges between them based on some similarity criteria.

2. Clustering: These methods group the pixels in different clusters of interest based on pixel properties and the neighborhood around them.

3. Appearance-based models: In these methods, the appearance of object region is created (e.g., by extracting key feature points of object region) to model the object shapes.

4. Classification: The methods that use a classifier fall in this category. The classifier learns to divide image pixels into coherent segments by utilizing supervised or unsupervised learning methods.

5. Threshold: In thresholding technique one or more intensity values (generally called 'the Threshold' value(s)) are identified and all the intensity values greater or equal to threshold value are grouped together (generally by setting them e.g. 1) and all others are set to a different intensity, say 0. In this manner all the image pixels are divided into two or more groups which corresponds to object(s) and ground.

6. Edge: The edge –based methods use the contrast information and finds the edges and boundaries of object regions. Normally, closed boundaries indicates the location of objects in image and rest of the information is treated as background.

7. Region: Region based approaches start by choosing initial seed points from image space and then growing or accumulating more similar and neighborhood points into the same region.

8. Template Matching: In this method initial cues of actual objects are created and saved, these are called the 'Templates' and then by using some sort of matching function the similarity between the object in question and the template is decided.

4 Proposed Model

By studying human visual psychology, it has been concluded that visual cognition is not only about seeing or sensing the environment from visual organ but the perceiving of the objects is a complex process which involves previous knowledge and cues to

make the overall process faster and effective. Based on these facts, a generalized automated image segmentation model must have the following components:

(a) An elementary feature extraction and grouping module
(b) A global feature extraction module
(c) A System memory module to hold past knowledge
(d) A decision maker module to select the most stable candidate segmentation.

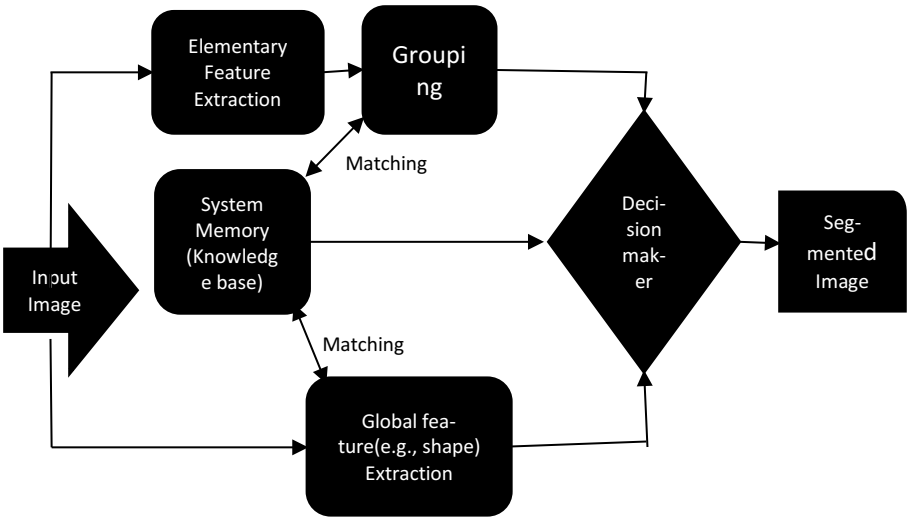

Fig. 1 Proposed Model for Image Segmentation

5 Conclusion and Future Scope

The proposed work is an attempt to link psychological findings with computer vision. The proposed model makes use of past knowledge to select between the possible interpretations of object which may help in achieving higher segmentation accuracy. The future work will be to incorporate this model for developing an automated image segmentation system.

References

1. Andreas Richtsfeld, Thomas Mörwald, Johann Prankl, Michael Zillich, Markus Vincze, “Learning of perceptual grouping for object segmentation on RGB-D data”, Journal of Visual Communication and Image Representation, Volume 25, Issue 1, January 2014, Pages 64-73, ISSN 1047-3203, http://dx.doi.org/10.1016/j.jvcir.2013.04.006.
2. M. M. Cheng, N. J. Mitra, X. Huang, P. H. S. Torr and S. M. Hu, "Global Contrast Based Salient Region Detection," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 37, no. 3, pp. 569-582, March 1 2015.
3. Laura Walker Renninger, Jitendra Malik, When is scene identification just texture recognition?, Vision Research, Volume 44, Issue 19, September 2004, Pages 2301-2311, ISSN 0042-6989.
4. Felzenszwalb, P.F. & Huttenlocher, D.P., “Efficient Graph-Based Image Segmentation”, International Journal of Computer Vision (2004) 59: 167. doi:10.1023/B:VISI.0000022288.19776.77
5. Vecera, S. P. & Farah, M. J.,”Is visual image segmentation a bottom-up or an interactive process?”, Perception and Psychophysics (1997), 59:1280–96.[PDG]
6. M. Wertheimer, Untersuchungen zur Lehre von der Gestalt. II, Psychological Research 4 (1) (1923) 301–350.
7. M. Wertheimer, Principles of perceptual organization, in: D.C. Beardslee, M. Wertheimer (Eds.), A Source Book of Gestalt Psychology, Van Nostrand, Inc., 1958, pp. 115–135.
8. Explorations in Cognitive Neuropsychology by Parkin, A. 2016, Taylor & Francis, ISBN 9781317715795.
9. A. Borji, M. M. Cheng, H. Jiang and J. Li, "Salient Object Detection: A Benchmark," in IEEE Transactions on Image Processing, vol. 24, no. 12, pp. 5706-5722, Dec. 2015. doi: 10.1109/TIP.2015.2487833
10. T. Liu, Z. Yuan, J. Sun, J. Wang, N. Zheng, T. X., and S. H.Y., “Learning to detect a salient object,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 33, no. 2, pp. 353–367, Feb. 2011.
11. Advances in Computer Vision, Volume 1 by C. Brown Psychology Press, 2014, Taylor & Francis, ISBN 1317767667, 9781317767664.
12. Jon Driver, Greg Davis, Charlotte Russell, Massimo Turatto, Elliot Freeman, Segmentation, attention and phenomenal visual objects, Cognition, Volume 80, Issues 1–2, June 2001, Pages 61-95, ISSN 0010-0277, http://dx.doi.org/10.1016/S0010-0277(00)00151-7.
13. Boykov, Y., Veksler, O., and Zabih, R. 2001. Fast Approximate Energy Minimization via Graph Cuts. IEEE Trans. Pattern Anal. Mach. Intell. 23, 11 (November 2001), 1222-1239.
14. Xianbin Gu, Jeremiah D. Deng, Martin K. Purvis: Image segmentation with superpixel-based covariance descriptors in low-rank representation. CoRR abs/1605.05466 (2016).
15. Nameirakpam Dhanachandra, Khumanthem Manglem, Yambem Jina Chanu, Image Segmentation Using K -means Clustering Algorithm and Subtractive Clustering Algorithm, Procedia Computer Science, Volume 54, 2015, Pages 764-771, ISSN 1877-0509, http://dx.doi.org/10.1016/j.procs.2015.06.090.
16. K. Koffka, Principles of Gestalt Psychology, International Library of Psychology, Philosophy, and Scientific Method, vol. 20, Harcourt, Brace and, World, 1935.
17. Zhang, Hui, Jason E. Fritts, and Sally A. Goldman. "Image segmentation evaluation: A survey of unsupervised methods." computer vision and image understanding 110.2 (2008): 260-280.