

# Shape based Object Recognition in Images: A Review

Garima Sharma  
Assistant Professor  
Geeta Engineering College  
Panipat, Haryana

Abhishek Gupta  
M.Tech Scholar  
BRCM College  
Bahal, Haryana

Ravi Malik  
Assistant Professor  
Geeta Engineering College  
Panipat, Haryana

## ABSTRACT

Object recognition is gaining its importance in many fields especially in the field of surveillance. Recognizing objects is much more than just detecting lines, colors or shapes. Rather it requires a large amount of background knowledge. Recognition by components theory by Irving Biederman says that object can be recognized with the help of geons. Object recognition mostly focuses on the viewpoint variations.

## General Terms

Object Recognition, Recognition by Components, Geons, perception, recognize.

## 1. INTRODUCTION

In modern vision system it is necessary to be able to navigate in real world. To enable better understanding of the scenes it is necessary that objects can be recognized easily. Objects differ from each other in many respects like form, structure, color, texture etc. One can recognize an object from any angle whether it is front, side view or back [1]. The appearance of the object depends on different conditions like changes in the lighting or color, the direction from which we are observing the object, the size and the shape of the object and humans have the capability to identify object under such crucial conditions also [4].

The perception of objects is defined as the way humans recognize familiar objects in the scene [13]. In our daily life we come across many objects, some of the objects are known to us and some are unknown. The main thing is to determine which of the given set of objects appear in a given image or image sequence.

## 2. OBJECT RECOGNITION

Object recognition deals with identifying the object from an image [4]. Object recognition is the field in computer vision which aims at recognizing objects from image data and to estimate the position and orientation of the recognized object. Object recognition aims at representing the objects in the surrounding and creating their relations with the 3D structures of the object [8]. Shapes can differ in appearance especially in the way of perception. Objects can be viewed from different angles and positions which make an object look different from all the angles.

Two trees are also different but we label them as one object. Many theories have been developed on how humans perform object recognition. Humans have the capability to easily identify objects in our surroundings, regardless of their circumstances, whether they are upside down, of different color or occluded or have different. Object recognition is one of the most fascinating abilities that humans easily possess since childhood. By just looking at the object, humans are able to tell its identity or category despite of the appearance variation due to change in pose, illumination, texture, deformation, and under occlusion [4] [11]. Object recognition utilizes certain other features also for the better recognition

like orientation, distance, size of object etc. Humans can uniquely identify individual objects not just face but also other objects like cars, houses, animals etc. Just the presence of a hand can indicate that there's a person present or a set of wheels can give reason to believe a car might be in the scene.

But shape is the main feature of identity. For identification, discrimination between physically similar objects is made but for the characterization of objects their generalization is necessary [8]. To recognize, identify or do categorization of an object involves comparing its visual representation with some stored information [9] [12].

Object recognition sometimes becomes difficult as objects change themselves with respect to the scene for e.g. in case of varying lightning conditions or the viewpoint of the object, presence of background clutter etc. [10].

Different steps are involved in the object recognition with certain assumptions in each case. In identification stage object shape is considered fixed. In the generalization task recognition of an object is done despite the changes in its appearances due to the transformations. Categorization involves the assignment of the object to a similar class of objects [11].

Some issues that are needed to be considered in the object recognition system are [7]:

1. *Object representation*: How an object is represented is a matter of concern as it determines the perception by a person to that particular object. The primary attribute of any object is the shape of the object which is a matter of concern whereas the secondary attributes like color, size, texture are also needed but they only aid in getting to know more information about any object. The representation of any object should be such that it gives all the information about it at a single instance.
2. *Feature extraction*: This step deals with the features that are needed to be extracted.
3. *Feature model matching*: This step deals with how the features that were extracted are matched with the model already in the database. This matching helps in the recognition process.
4. *Hypothesis formation*: This deals with how the probabilities are assigned to the objects. The presence of an object is verified by using their models. This step aims at reducing the size of search space.
5. *Object verification*: The presence of objects is verified from their models.



Figure 1: Image showing clean and dirty room

Figure 1 shows two images. In the first image a clean room is shown from which the things needed are easily visible like the green color book on the rack. Whereas in the second image there is large amount of clutter and recognition of object is a difficult task.

Figure 2 consists of a scene which can be categorized as outdoor or a city. Then the image is further looked into the features it contains for e.g. the image consists of people, cars, and buildings. The features are extracted by locating their exact position [5]. After the extraction of features image parsing is done. Parsing involves detection of non accidental properties such as symmetry, parallelism which may vary due to difference in the viewpoint angle [14].



Figure 2: Random scene

In order to recognize objects, a database is first created against which matching takes place. Once the appropriate geon has been extracted from the image, the object is recognized from the matches of the previous objects [14].

The complex objects are divided into simpler components. Then the known or unknown objects are recognized by parsing it in terms of its components and then comparing the smaller components with those of known objects [5] [10].

### 3. APPLICATIONS OF OBJECT RECOGNITION

Object recognition finds its application in many areas like the following:

1. Security and surveillance systems
2. Finding objects in digital images, photos or video
3. In robotics.
4. Inspection of different factors including quality involved in production.
5. Security in banks and airports.
6. Detecting moving objects in a scene or image.
7. Keeping check of the assets to keep away from fraud
8. In medical field to detect unknown areas.

9. Finding people, objects during the time of natural calamity like flood, earthquake.
10. Counting and measuring objects (cars, people, animals etc.)

### 4. RECOGNITION BY COMPONENTS (RBC)

Irving Biederman gave a theory named Recognition by Components to explain the concept of object recognition. According to this theory objects are recognized through geons obtained by segmentation of objects [15] [17]. Objects are generally represented by how a human perceives and interprets the object. According to RBC how a person represents an object is generally stored in brain as structural description. Visual input is matched against structural representations of objects in the brain [16]. Structural representation consists of geons and their inter relations (for e.g. an ice cream cone is broken into a sphere and a cone) [6] [16] [18]. The inter relations include aspects like the location and size. The geons and their inter relations of the perceived objects are matched against the structural descriptors if a successful match is found. If a good match is found then successful object recognition occurs [16] [18].

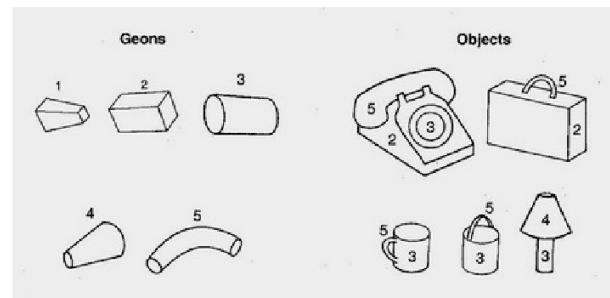


Figure 3: Biederman Geons and objects created from them

In figure 3 the left side shows the geons according to the theory of Biederman whereas the right side shows the objects created from the geons.

The objects according to RBC theory are not only described by the geons but also according to the way the objects are connected to each other [6] [9].

While viewing any object two main things are considered [17] [18]:

- (a) *Edges*: this property helps in recognizing the objects even if their orientation varies.
- (b) *Concavities*: it is defined as the area where two edges meet and this helps to determine the separation between two or more geons.

Recognition means *the act or state of recognizing*, means what we are seeing is stored in our mind as something we already know and are familiar with [13].

According to Biederman theory when people try to recognize an object, they first create a rough classification of the objects based on their 3D shapes. Then according to their classification the objects are classified into different categories [13].

Object recognition depends on [18]

1. Detecting edges
2. Detecting regions of concavity

3. Detecting non accidental properties
4. Putting these together to form object components (geons)

Recognition by components theory says that humans recognize object as they see it by assembling the 3D components called geons in a way so as to form the related object [12]. For example as shown in figure 4.



Figure 4: Perceived Object and Constituent Geons [13]

Geons can be combined to form a large number of objects. Whenever we see an object, we can recognize the geons the object is made up of if the angle of viewing is kept the same every time [2]. Geons can be varied in many ways according to its orientation and also according to the ratio between its length and width [3] [16].

Geons are generally different from each other and because of this they are easily discriminated [12]. They can also be recognized even if they have some part of the missing so they are robust to noise.

Geons are defined through a 4 digit code that describes the object with respect to its edges, symmetry and its relation with other geons [15]. The first digit of the code indicates whether the edges are straight or curved; the second digit encodes whether the object has reflection symmetry, rotation symmetry, or both; the third digit tells how the shape of the object changes with distance along its central axis, and the fourth digit encodes whether that central axis is straight or curved [16].

For e.g. 1321: straight edged, reflection & rotation symmetric, contracting sweep, curved central axis [15].

First digit- Edges: 0: Straight 1: Curved	0XXX 	1XXX 		
Second digit- Symmetry: 0: None 1: Reflection 2: Rotation 3: Both	X0XX 	X1XX 	X2XX 	X3XX 
Third digit- Sweep: 0: Constant Size 1: Expanding 2: Contracting 3: Both	XX0X 	XX1X 	XX2X 	XX3X 
Fourth digit- Axis: 0: Straight 1: Curved	XXX0 	XXX1 		
Object Relations: <0>: Smaller Than <1>: Bigger Than <2>: Above <3>: Beside <4>: Below <5>: Join End to Side <6>: Join Side to End <7>: Join both ends to side	Examples: Brick:  0300 Cylinder:  1300 Teapot:  1301<037>1310<136>1321			

Figure 5: Table for 4 digit code of geon [15]

According to the theory by Biederman, the number of Geons that can be combined to see the objects in our daily life are less than 36. For e.g. if a hand fan is broken into two parts then we are left with its handle and a semicircle shaped structure. Similarly a complex object can be made of many geons.

36 Geons leads to 30,000 different objects [14].

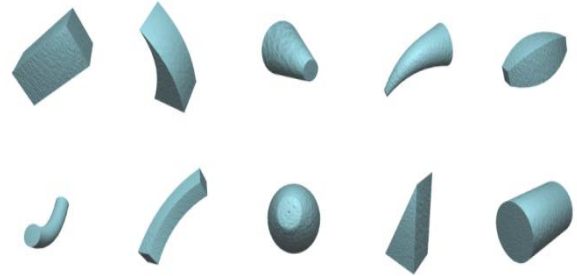


Figure 6: 10 out of 30 geons of the Biederman theory

The number of objects that would yield from 36 geons depends upon two assumptions [14]:

1. The numbers of objects need to specify each object.
2. The number of readily discriminable relation among the geons.

According to Biederman, edges and vertices are sufficient to recognize the objects and the color and texture of the object doesn't matter much [3] [13]. Although color and texture properties are not fully ignored by Biederman as they are necessary for the recognition of mass nouns. Mass nouns indicate those groups of objects whose boundaries are not clearly defined for e.g. objects like sand, water etc [10] [14].

The disadvantage of Recognition by Components theory is that it differentiates between the object types but do not compare between the same types of objects for example one orange versus another.

According to the RBC theory, a relation between two objects occur only if they exhibit some similarities in their parts for e.g. if an object is rotated without distorting the parts of objects, the object remains recognizable and the rotated and unrotated objects are same [14]. Geons theory proposes a four layer structure for object recognition. First layer consists of decomposing the information into the edges. Then the component axes, blobs and vertices are determined. At layer 3 geons (cones, cylinders and boxes) are identified. The layer 4 determines how the geon components interconnect with each other. Finally object recognition is achieved [10] [15].

Inter relation between the neighboring objects is also determined. When geons are connected in a number of ways then a large number of objects are created. The primary attribute of the geons is its shape. The secondary attributes consist of the color and texture [10].

Biederman geons share two properties [13]:

1. They can be distinguished from one another in any object.
2. When some randomly chosen part of geon is erased then the rest of the object remains almost the same and is easily recognizable. For e.g. if middle section of the cylinder is removed and the remaining part is rejoined again then the left over part looks same as the previous.

## 5. CONCLUSION

Object recognition provides a way to find specific object in an image. Humans have the ability to recognize any object even if they see it for a fraction of second in any condition like lightning, clutter, distortion, shape deviations. This review considers the RBC theory given by Irving Biederman which considers the way of classifying objects by identifying their components and its related properties. Biederman concluded that recognition is based on the decomposition of objects in geons. It is based on the shape information whereas color and texture do not contribute to the object recognition.

## 6. REFERENCES

- [1] Ferrari V, Jurie F, Schmid C, "From images to shape models for object detection" *International Journal of Computer Vision*, 87(3), 284-303, 2010.
- [2] Toshev A, Taskar B, Daniilidis K, "Object detection via boundary structure segmentation" *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* 2010.
- [3] Gu C, Lim J, Arbelaez P, Malik J, "Recognition using regions" *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2009.
- [4] Ming-Hsuan Yang "Object Recognition" *Encyclopedia of Database Systems* (eds. Ling Liu and M. Tamer Ozsu), pp. 1936-1939, 2009.
- [5] Peter M. Roth and Martin Winter, "survey of appearance-based methods for object recognition" *Technical Report ICG/TR{01/08* Graz, January 15, 2008.
- [6] Ferrari V, Jurie F, Schmid C, "Accurate object detection with deformable shape models learnt from images" *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* 2007.
- [7] M.Anji Reddy, Y.Hari Shankar "Textbook of Digital Image Processing", copyright © 2006, BS Publications.
- [8] Thomas J. Palmeri\* and Isabel Gauthier\* "Visual object understanding" *Volume 5 | April 2004 | © 2004 Nature Publishing Group*
- [9] Serge Belongie, Jitendra Malik, and Jan Puzicha. Shape matching and object recognition using shape contexts. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 24:509-522, 2002.
- [10] Pourang Irani, Colin Ware, "Diagrams Based on Structural Object Perception", Faculty of Computer Science, University of New Brunswick Published in proceedings of the working conference on Advanced visual interfaces ACM, New York, © 2000.
- [11] Shimon Edelman "Theories of representation and object recognition chapter 2- Representation and recognition in vision" © 1999 the MIT press
- [12] Shimon Edelman "Computational theories of object recognition" School of Cognitive and Computing Sciences University of Sussex June 1997. Revised September 1997
- [13] Robert Sekuler, Randolph Blake "Perception chapter 6- Object Perception: Recognizing the Things We See" Publication Date: January 1, 1994
- [14] I. Biederman. "Recognition by components: A theory of human image understanding." *Psychological Review*, 94(5):115-147, November 1987.
- [15] Biederman's Geon Theory <http://cns-alumni.bu.edu/~slehar/webstuff/pcave/biederman.html>.
- [16] Geon physiology from Wikipedia, the free encyclopedia: [http://en.wikipedia.org/wiki/Geon\\_\(psychology\)](http://en.wikipedia.org/wiki/Geon_(psychology)).
- [17] Recognition by Components theory from Wikipedia, the free encyclopedia: [http://en.wikipedia.org/wiki/Recognition\\_by\\_Components\\_Theory](http://en.wikipedia.org/wiki/Recognition_by_Components_Theory).
- [18] Kimberly Kirkpatrick "Object Recognition", Department of Psychology, University of York <http://www.pigeon.psy.tufts.edu/avc/kirkpatrick/default.htm>.