

# IMAGE INDEXING AND RETRIEVAL USING VISUAL PERCEPTUAL LEARNING IN HUMAN OBJECT RECOGNITION

<sup>1</sup>FIRAS ABDALLAH, <sup>2</sup>SHADI KHAWANDI, <sup>3</sup>ANIS ISMAIL

Lebanese University, Faculty of Technology, Saida, Lebanon

E-mail: <sup>1</sup>firas.abdallah@gmail.com, <sup>2</sup>ckh@ul.edu.lb, <sup>3</sup>anismaail@yahoo.com

## ABSTRACT

Image indexing and retrieval became an interesting field of research nowadays due to the lack of advanced methodologies to index and retrieve images and to the existence of huge quantities of images available everywhere; especially on the web. The available solutions are able to find similar items having the exact shape but not the same item if it has a different shape. In this paper, we propose cloning the human being visual perception and cognitive learning to recognize, index and retrieve images. In Computer Vision, Mesh data structure is used for its capacity of neighboring, therefore Mesh data structure will be used for the fact that human being groups objects in categories where neighboring is essential for indexing and retrieval. The proposed methodology is intelligent in the way that retrieving images lead into an insertion of the searched image. In addition to that the proposed solution will be able to recognize items having different shape from their features.

**Keywords:** *Image, Indexing And Retrieval, Visual Perception, Cognitive Learning, Shape.*

## 1. INTRODUCTION

The number of pictures has been growing enormously with the existence of digital cameras. Storing those huge amounts of images is not a disability with the existence of new hard disks that can stock up to Terra's of Bytes. Picking a photo of a bulb for instance from those set of images stockpiled on a hard disk is a handicap. Subsequently retrieving images in a quick and relative way is a must. To retrieve an image you are looking among between plenty of different images each of which containing items unrelated (Figure 1) is a very hard task if images were not indexed prior to storing them on the hard disk or even prior to searching for any of them. Image indexing and retrieval is a very important task in many various fields like Crime prevention, Military, Architectural and Engineering design, Medical and diagnosis, remote Sensing systems and many others.

Image indexing is a mandatory preliminary step in the process of the retrieval of an image. This will lead to the process called "image indexing and retrieval" that will be comprehensively covered in this paper with all

its available techniques. Their advantages and disadvantages over other techniques and relatively to our need of having not only an image with high similarity of the one searched for but sometimes the exact image in the procedure of face recognition.

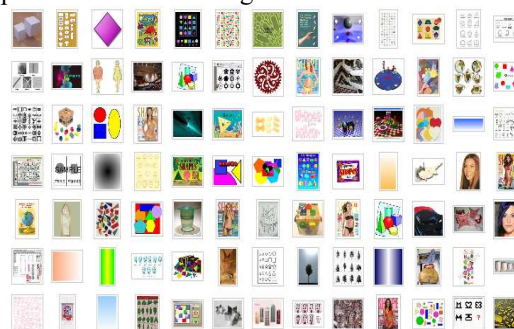


Figure 1: Collection of Images [32]

Many Techniques to index and/or retrieve images exist and are to be introduced in the following paragraph. Some of those techniques are very efficient to identify a specific item that is taken in a picture alone [1-9]; some other techniques are capable of segmenting and extracting the variety of items that exist in a

certain picture; then processing each item for identification [10-13].

Items can be identified using variant methods and techniques that vary from invariant feature illumination [7] to content based retrieval [3, 8-10, 14, 20] passing through Description based retrieval [6, 15], shape recognition [5, 16], wavelet or color correlogram [17, 18, 24], Object based retrieval [3, 9, 11], intermediate features retrieval [12], geometric shape recognition [7, 16], Expressive Fuzzy Description Logics [13, 23], Mapping low level features to semantic concepts [19], feature extraction [5, 7, 11, 12, 20, 21, 22] and last but not least ontology-based information retrieval [15]. Various procedures are concerned not only with shape but also with texture and color to retrieve an image. Algorithms' threshold is set on each function to reduce the error of non-matching identifications where some of the algorithms [1, 3, 6, 7, 9] are good in matching the same item when it is translated, scaled, or rotated.

Image features can be textual like the tags manually set all through the web or visual significant to low level feature like colors or shapes such as red or triangle. Human beings search for images using a semantic search like "light bulb". Whereas computers have to interpret this information and a gap is found between low-level features extracted by computers and semantics used by human beings which is "the semantic gap".

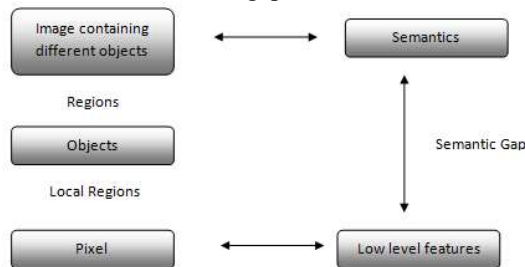


Figure 2: Semantic Gap

Moreover, if one uses any of the available methods to search for a bulb (Figure 2), he/she will be find nothing, even with an important threshold because any object could be redesigned with a new shape even within the same company. Subsequently, with a high tolerance the unidentified item will not be found. No matter how the created algorithm is, it will always be looking for similarity between images because it is not interpreting the image content. Whether using Content Based Image Retrieval (CBIR) [25] or Description Based Image

Retrieval (DBIR) [6] [15], none of the existing methods can identify any similarity between the



light bulbs in Figure 4.

Figure 3 : Light Bulb



Figure 4: Different Light Bulb Shapes [33]

The second problem is that most techniques use Tree and advanced Tree Data Structures like VP-TREE [26] to index images. A Tree can in an efficient way group images containing different objects with common attributes whereas neighboring is limited to two nodes that is not so supportive for indexing images.

## 2. RELATED WORKS

The most important techniques are to be exploited. The first procedure that comes to mind is Content Based Image retrieval (CBIR). Many years of researching have been made on this

topic using this methodology. This paragraph is divided into two major sections. The first section of this paragraph explains the functionality of this technique and shows all work done in CBIR. The second part discusses the Description Based Image Retrieval (DBIR) also called technique of the web, the second widely used technique. The first set of subsections summarizes the job done in general feature extraction techniques, then more precisely, shape extraction and color extraction techniques are explained.

Extended Image retrieval with a thesaurus of shapes was discussed using CBIR method [4]. While indexing an image, not only features extraction is needed, but also text annotation to be stored in a shape thesaurus. This thesaurus will contain a precompiled list of shapes for major objects, statistics about those shapes, textual and semantic annotation that describe those shapes, and finally for each shape its set of shapes that are similar to it. Then he talks about relationships between objects and shapes that can be Hierarchical or Related for the first and Part-of or variant-of for the shapes.

Mikolajczyk, Zisserman and Schmid have provided a method [5], to recognize an object by detecting its edges, that is invariant to similarity transformations and scaling. They work on neighborhood descriptors which were generalized from The SIFT method (Scale Invariant Feature Transform). The edge based descriptors can identify well drawn edges in objects like a bicycle.

Md., Farooque [30] has well defined in his article image attributes that were identified as three categories perceptual, interpretive and reactive. Perceptual is mainly literal objects, people, color, location, visual elements and description. Interpretive goes more into people qualities, abstract concepts and content story. The reaction is based on external relationships by comparing the image to a set of images like in the thesaurus. The use of this approach defines perceptual as low level features, interpretive as semantics and meanings, and reactive for relations with other related objects. This method is one of our interests in this paper as it complements the thesaurus to be used.

Feature extraction can lead into object extraction as shown by Zlattof, Ryder, Tellez and Baskurt when they grouped extracted features in order to get assembled object [9]. Those are two similar and different techniques the first in image retrieval and the second for image indexing. Segmenting an image using Expressive Fuzzy

Description Logics by Simou, Athanasiadis, Stoilos and Kollias was really efficient using Semantic Recursive Shortest Spanning Tree and fuzzy Logic [13]. First they divide the image into regions by using low level feature and color extraction, second they extract available visual descriptors and spatial relations between regions and then the different parts of the images are classified based on different extracted regions.

Images could contain more than one object, the fact that led into working on image segmentation and categorization. Object categorization is important to build up the thesaurus and index images, therefore some of the selected chosen article describe how to segment and categorize image. The initial tree containing the possible categories is explicitly shown by Djeraba [20], classifies a small set of objects into a tree having a one level of root and four sub levels of first child that goes from Naturals like water, flowers, mountains and snow to people, industries and transports. As for Chen and Wang they went further with categories to show 20 different detailed categories as beaches, flowers, sunsets, waterfalls, cars, buses, antiques, dogs, elephants, lizards and many more. A small set of categories could go to 20 different categories (Figure 2.3), but what is needed is a lot more than that, which could go to millions of categories that are a must in image indexing and retrieval. Some went further also by combining segmentation and categorization for fuzzy knowledge based semantic annotation [23].

A variety of techniques have been discussed in the last decade, starting by shape descriptor using polar plot [6], illumination of invariant features [7], geometric shape recognition [17], salient points reduction [8], noise tolerant approaches, comparing dissimilarity measures [14], object retrieval based on rough set theory [3], real valued boundary point [2], new algorithms using wavelet correlogram [18], ontology-based information retrieval using expressive resource descriptions [15], local scale invariant feature extraction [11], global and region features extraction [22] and last but not least, scalable color image using vector wavelets [24].

In all researching approach, the main issue was to fill the semantic gap produced between description based and content based. Description based image indexing and retrieval count on human interactive, which is not objective, to annotate images. Content based

uses feature extraction to annotate images that can extract shape, color and texture. CBIR is more efficient than DBIR and is used while researching the image indexing and retrieval topic but CBIR cannot give a semantic meaning to what it extracts. Some techniques focus on minimizing the gap between low level feature and semantics. Obeid, Jedynek and Daoudi have introduced what is called “Intermediate features” [12] to fill the semantics gap. Stan and Sethi mapped low-level image features to semantic concepts [19].

### 3. PROPOSED SOLUTION

In paragraph, we have showed most of the techniques used in image indexing and retrieval and all their advantages and disadvantages. Part of those techniques will be used in our solution like segmentation and feature extraction. In the previous paragraph, we have talked in details about human object recognition using visual perception. We have described how human being categorizes objects and assemble features for comparison and recognition or recognition and indexing. This depends on whether we are introducing a new item or extracting an already existing item. As we have seen human being are capable of extracting features and indexing or more accurately storing items in the brain using categories without labeling the items.

Human Cognitive learning is still being a dense material for researchers. Effectively, when the human brain is explored, amazing functionalities are found. It is important to mention that human neural network has the Mesh shape that store short and long term memory, does computational work, dreams, and imagines. Visual perception that is interpreted by our same brain that does all above is also something amazing for one simple fact that is possible to see in the most difficult circumstances and identify objects in the same conditions. Machine Learning is a very interesting topic in computer science in general and in robotics in particular. Making a system auto-learn is something all researchers are working hard to achieve.

#### 3.1 Indexing using Human Cognitive Theory

Our proposed solution is based on Human cognitive theory and visual learning, combined with an indexing in a Mesh Data

Structure [27] [28] [29] instead of a Tree or a relational database. Humans are able to identify items in tough circumstances, therefore in this paper; the aim of this study is to understand how humans recognize objects and try to apply their method to image indexing and retrieval. Storing and indexing images will be in a Mesh Data Structure because of its ability to store same items and similar items all next to each others with the neighboring concept.

#### 3.1.1 Indexing using a VP – TREE

An accurate and fast retrieval search result must be preceded by a reliable and efficient indexing. Indexing images is a crucial and critical task that facilitates the act of retrieval; the more precise the indexing is, the better the retrieval is. In the previous paragraph we have seen that human beings see objects as a collection of Geons. Geons help identify items while reducing errors in identification with a kind of zero tolerance on comparison. Comparing items for retrieval will be done by comparing Geons and not the entire unidentified item. Geons are specific types of geometric structures that form an item or object when grouped together.

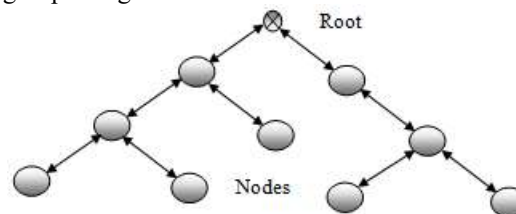


Figure 5: VP-TREE Connections

Tree Data structure is used in several methodologies to store images; indexing in a tree structure is efficient but not enough for having a fast and accurate retrieval with the huge amount of image that is available. Indexing images is not like indexing text that could be very efficient for retrieval with a Tree data structure. All the methods that uses Tree data structure for indexing, do index images as a whole with all what contains in terms of objects or items. The Vantage Point Tree or VP-Tree [26] (Figure 5) is the tree structure that is used in image indexing as it is a Binary Space Partitioning Tree (BSP Tree) with a neighboring concept that gets the nearest indexed items with a threshold. As known, a Binary Tree is fast in indexing and efficient for text and the best used for now for images. Connections in a VP-Tree are upper

lower, left and right which make region indexing faster from a normal binary tree.

In our solution, indexing will be constructed using a Mesh data structure that will be reviewed in the last section of this paragraph after going through all the aspect of indexing and retrieval in order to show how efficient a Mesh data structure could give value to the method. The Mesh data structure represents a set of nodes grouped together with a high access to the neighborhood. Nodes will not store only images as other methods do but also the set of Geons that represent the objects included in the images.

Categorization, the fact of grouping similar objects together, will take place in the Mesh data structure will be discussed in one of the sections of this paragraph. As our method will work similar to object recognition in human beings, Categorization is one of the tasks that are produced by our brain while indexing items together.

### 3.1.2 Two Nested Levels of Segmentation

There are two levels of segmentation that needs to be processed on the images to extract features or Geons out of it. The first segmentation needs to be developed on the image itself to separate the different objects inside an image, then the features extraction of each object. Two levels of segmentation must be processed over the images in order to extract the different features for them to be used in storing and indexing the image in the Mesh Data structure.

An image usually contains at least one object, if pictures of any specific object are not taken with a unique background color and not texture color or view. Even with one object in the image, we need to segment the image to extract the exact shape of the item to be identified. After the first step of Segmentation, we need to process the highlighted objects into another mechanism that needs to do some feature extraction on the existing objects. By that we will have an image segmented into objects each of which partitioned into a set of features or Geons.

#### 3.1.2.1 Image Segmentation

The study is made on images having raster format only. Images have textures or/and shapes with different colors where objects and textures might be rotated, scaled, or translated in space. The segmentation process has to be able to identify the different part of the image

discarding all kind of transformation on the objects and textures.



Figure 6: Original Image

N. Simou was able to segment an image containing sand on the beach the water and the waves and a man standing in the middle of the picture [13]. The method was to index and retrieve images based on fuzzy logic descriptor. The process of the method can be defined by initial segmentation followed by an extraction of spatial relations between regions then extraction of visual descriptors and region based classification in a Semantic Recursive Shortest Spanning Tree (S-RSST). The S-RSST combined with a Fuzzy Reasoning Engine (FiRE) [13]. This method was proved to be very efficient in segmenting images having a texture that is usually considered as a hard task in image indexing and retrieval. For those purposes this method will be used in our proposed solution in order to segment the images into different objects and textures. Textures can be considered as object and stored accordingly.



Figure 7: Segmented Image into Objects

### 3.1.2.2 Object Feature Extraction

After segmenting the image, a set of unidentified items are available for recognition, here comes the task of feature extraction as human beings do with the partitioning of the object into a set of Geons. Nicolas Zlatof has showed an efficient way of segmenting the objects into regions, where each region has its own characteristics. Those regions can be considered as Geons. Taking the hammer as a research model, Zlatof tried to extract different region of the hammer that are the handle and the head. Zlatof succeeded in dividing the object into two regions after adjusting the algorithm to the necessary parts for extraction (Figure 8). This algorithm is very essential to our work done on image indexing as we need to extract subdivide the unidentified items into different set of regions that are needed in the global definition of the object [9].

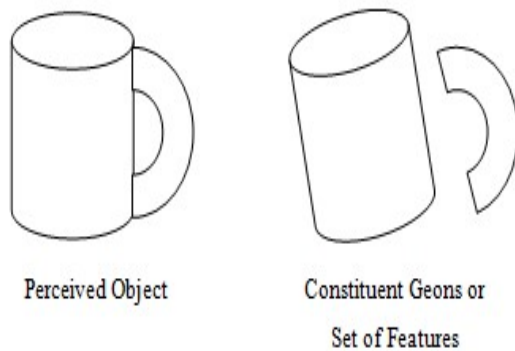


Figure 8: Features Extraction

### 3.1.3 Categorization

Human beings group items by characteristics, those groups go from basic level to the real object category. A natural grouping of attributes occurs by basic level such as apple and fruit, fruit being the basic level and apple the object level. As performed extraction on the objects in the previous section, we can now group those objects by their attributes or features into categories. We need to remember that a human being does not need to have a name or label for any object in order to categorize it or extract its features.

#### 3.1.3.1 Features Grouping

After segmenting the image in regions, then segmenting the objects found in images into regions too, we need to group those images by their features and not as a whole image or objects. For example, having a mug and a bucket that are both formed using the same set of features the cylinder and an arc. Assembling the cylinder and the arc in two different ways will

lead us to two different objects. Therefore, by putting the arc on the side of the cylinder, we will have the mug, and while putting the arc on top of the cylinder the bucket is formed. On a lower level, categorization will be done by grouping features together rather than grouping images or objects together. At the end we will have same objects grouped together and similar objects grouped together too as displayed in Figure 9.

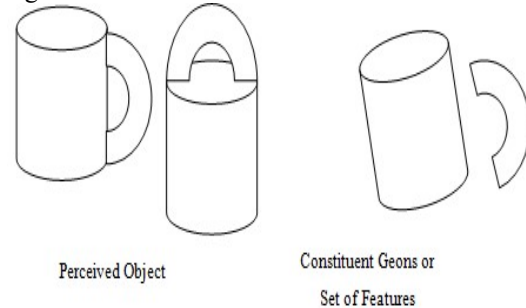


Figure 9: Same Geometric Icons Form Different Objects

Categorizing the set of features instead of categorizing objects themselves is done for faster retrieval. When we group features together, we will have the possibility while retrieving to get all possible items having the same set of features. Taking again the light bulb and comparing each bulb using the currently used techniques, it will be hard to match any of them. The currently used methods are very efficient if we want to retrieve exactly the same item. As if we want to have the collection of objects called blub light, we need to extract the features of those bulbs and index them all next to each others for later accurate retrieval. The same industry could produce different shapes of light bulbs. If someone tries to search for a light bulb in a database, the system must give him/her all the bulbs and not only bulbs similar to the one that he/she provided if the search is done by example, or even give him/her some of the bulbs if the search is done by tags. All bulbs with no matter what shape they have must be returned to the user.

#### 3.1.3.2 Indexing Objects in a Mesh Data Structure

Image indexing and retrieval is not complete without an efficient storage for a fast and reliable retrieval. Categorization and similarities are better implemented when having a concept that is very implicated in region neighboring. A fast and reliable retrieval should be based on an efficient, fast and reliable indexing [29]. Mesh data structure has the ability

to store items on region based next to each others. As our method consists of extracting features and grouping features together by categorization then we need a data structure that can do the job. All other methods use Tree data structure and some use the advanced form of the BSP Tree the VP-Tree. The tree data structure is very useful for indexing on region based but the tree will remain efficient on a 2D level on indexing and not 3D [27]. 3D indexing means to group objects having similar features next to each other. Tree data structure can only group similar objects together and not objects with similar features.

Mesh data structure is a very efficient structure for computer vision as it has the ability to produce an object in the space and in 3D. We will try to benefit from this capacity in order to produce our model. The Mesh Data Structure is not a simple structure that we can easily navigate through its nodes and it is very hard to search inside it as it is huge and contains millions of items many of them connected to each others. Connections in trees are limited to three, each node can have only three connections in a binary tree, as for a mesh we can have plenty of nodes. Having plenty of nodes is what we need to construct our solution and have a region based structure. The problem remains on how to index images in a fast and efficient way for us to have a fast and reliable retrieval.

The study of Biederman [31], Recognition by components, has showed that 36 Geons can build millions of objects. Using those 36 Geons as entry to the Mesh will help us a lot in our indexing and later retrieval, though we have 36 nodes as being the entry to our Mesh data structure that will smooth the progress of indexing through the structure. Each of those nodes can be in connection with up to 36 nodes to form all possible objects with those combinations. The same Feature can be used more than once in the same object; will lead us into a huge number of objects to be formed from those 36 Geons.

### 3.1.4 The Method

Our proposed indexing technique will start by segmenting the images into regions called objects where each object will be disassembled into regions called features. Those features are to be indexed in a Mesh data structure (Figure 10), as meshes have a 3D region grouping capacity. Each object is defined by a set of features is indexing accordingly in the same region with other similar objects having the

same set of features or with some other objects having the same features and forming another object.

The example of the bucket and the mug are a tangible example of how with the same set of features we can build objects depending on the relation between those features. The relationships between components are very important to differentiate between objects having the same set of facets.

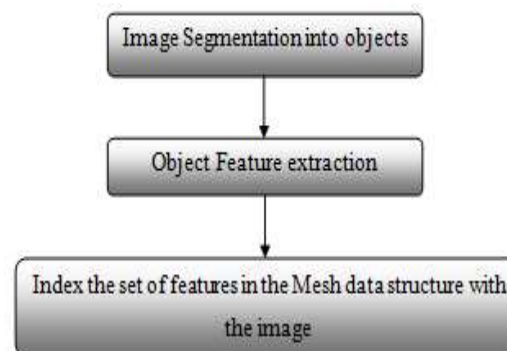


Figure 10: Indexing Algorithm

### 3.2 Retrieval Based on Features of Similarities and Categorization

Due to the fact that retrieval by itself is a very hard task, we tried to make a competent indexing to reduce the complexity of retrieval. The algorithm for retrieving images takes a big part of the indexing algorithm plus some more functionality to extract the found images and the insertion of the new image. Searching is done in two ways either by labeling, names, tags and annotation or by query by example; we will only discuss the second approach in this paper. Taking the second means for searching, we have to do all what we have done as a primitive step in indexing from segmenting the given image into regions considering each region as an object. Objects are then disassembled into set of features; this task is called features extraction. After extracting the set of features of all available objects, a search through the nodes of the mesh is to be performed in parallel by comparing all existing features extracted with the set of 36 Geons that were predefined to be the root nodes of the indexed Mesh.

For example, taking the mug as the inquiry item, we need to extract the two features out of it. As the first characteristic is the cylinder, then we need to access the Mesh from the cylinder node that was predefined in advance. From this first node we need to find all nodes that have both the cylinder and the second

feature that is the arc. If we had more than two characteristics for the queried item, we must perform the same task on the entire set of attributes, until we find all the nodes containing the whole set. After that we must look at the relations that join each of the couples of attributes together to form the object.

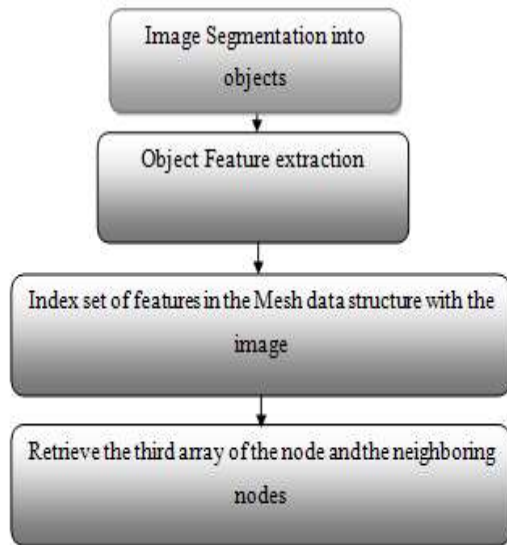


Figure 11: Retrieval Algorithm

We must do that to make sure we are having the same object as the one we are seeking. If all items are found and all relations are satisfied between the sample and the objects in the mesh, we extract all the images that are stored in the array of images within the node. We realize that the same indexing algorithm is applied on the sample image we are looking for (Figure 11).

### 3.3 Smart System

Artificial intelligence is normally interfered when the system is capable to learn by itself, the user set a list of rules that will make the system learn and continue learning without human intervention. The system is called a smart system when trying to recreate the functionalities of the human brain. Applying human recognition and visual perception to recognize identified items is helping to implement a kind of artificial intelligence. Not only using the human cognitive theory is making from this methodology artificial intelligence but also letting the system auto learn, is the second task performed in an artificial intelligence approach. Actually, when we give the system a sample image to look for, we will do the same task of indexing an image and in addition to that we will insert the image in the list of array of the found node. This way each time the system is used to find a certain object,

the Mesh data structure is fed by a new image. This approach will help the system improve its search result by learning all new objects we are looking for. With time we will have a robust system able to give the user a collection of images coming from the users themselves.

#### 3.3.1 Initialization Process

As previously defined there will be a set of about 36 features to be used as the root nodes. We do not need while searching by example to have any name to notation, as the system will extract the features and categorize them. Objects will be grouped in the same array having the same set of features and the same relationships between features. Not using query by example, the system at first use must be configured and for only one time to know the labels and annotation of each object. Items must have tags for only once, if user is searching by tag and names. This task is performed once and after that the system will learn alone. This task is only performed if the search by tag and name is used, if there is no need to learn the names because human beings do not have to know the name of the item in order to recognize it. As the object is disassembled into a set of features and those features are compared with the existing features, we disregard all naming that is to be used in the process of indexing and retrieving objects. Nevertheless, if we want to have the possibility to search by labels, a big task is to be done prior the first use of the system to label and tag at least one element of each existing object.

#### 3.3.2 The Learning Process

Artificial intelligence, and more precisely machine learning, will take a part in image indexing and retrieving. The search using query by example extracts attributes of the given item and compares it to the available attributes' items until we find a maximum number of identical attributes common to the searched item and a set of items already indexed. When finding the group of items, the new item will be indexed next to the similar images in the Mesh. With the Mesh data structure and the artificial intelligence using machine learning, an important auto expand of the indexed images will be achieved very fast. This database will help automatic indexing be constructed when using a mesh data structure.



#### 4. DESIGN SPECIFICATIONS

In this paragraph, we show the model by describing each step of the algorithm, the data structure of the Mesh with detailed descriptions of the nodes and giving some sample pseudo-code used.

##### 4.1 Programming Language and Framework

The proposed solution is accomplished based on some open source libraries in computer vision that are available online. Aforge.Net is the library the most used, it is a done in C# and .NET framework. It provides several functionalities used in computer vision. Image segmentation does not exist in the Aforge.Net Library and had to be written down according to watershed algorithm learned in the computer vision class. All code is written in C# 3.0 under the .NET 3.5 Framework.

##### 4.2 Mesh Data Structure

The final form of the Mesh (figure 12) will be like the neural network millions of nodes connected together to form all possible combinations that means all possible objects. In all this procedure naming and labeling will never interfere to index and item whereas naming could come up while retrieving an object if the user have chose to search by tag.

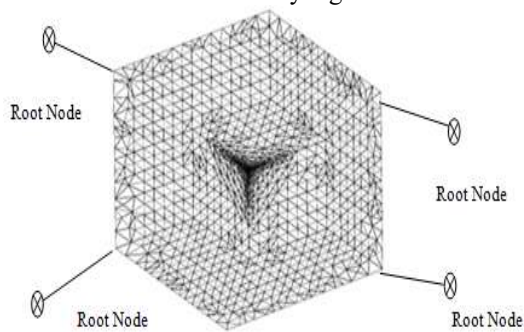


Figure 12: Mesh Data Structure Presentation with Many Root Nodes

At the end, we want to retrieve images containing objects and not a set of features, as features are only used for indexing. Users can search for images either by query-by- example or by annotations and tags. Therefore, each node we are creating will contain the set of features that constitute each object, the possible names for this object to be identified by the user, and the collection of images containing this particular object. Each node of the Mesh data structure will contain three arrays (Figure 13) the first one the set of features that contains a set of relations

between each feature for example the arc is on top of the cylinder to form a bucket or the arc is on the side of the bucket to form the mug. The second array contains the collection of images that were already indexed in the Mesh and the third one contains the list of names, tags and/or annotations. The representation of the arrays in different shape sizes is meant to show that storing images in a database is very consuming, in addition to that, a disk space is to be reserved for the set of Geons, which is used only for comparison and not visual for users. About a double storage is paid in this algorithm for a better image indexing and retrieval. On the other hand, One Label could have many images or set of Geons, that explains the small size of the second array containing labels.

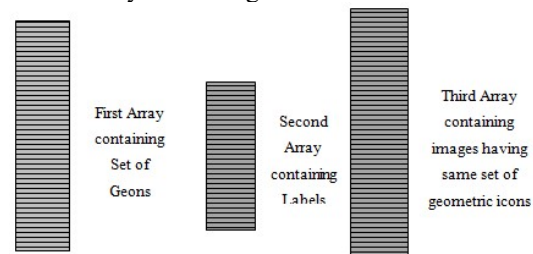


Figure 13: Node Representation

##### 4.3 Pseudo Code

After segmenting the image, there is an algorithm used for item identification through feature extraction. Noise removal must be applied on the image before extracting feature. If Noise removal is not applied on the image, it will be very hard to extract the feature, because there will be no defined shape. The noise removal is done using the salt and pepper algorithm learned in the computer vision class. The black points in the image are called pepper need to be deleted. As for the white dots are salt and needs to be filled. The algorithm can be found as an open source in the community to be used. Figure 14 shows the core algorithm in pseudo code that apply the whole process from reading the image to segmenting it and extracting features of each item then inserting the features into the first array of the Mesh Data Structure, inserting the item itself into the second array of the Mesh Data Structure, and finally inserting the image itself in the third array of the Mesh data structure. After that all images in the third array of the neighboring nodes will be extracted as they contain the same item.

**Enter Program****Read the Image****Segment the image into items****While unidentified items exists****Extract features from current item****Insert extracted features from items into the Mesh Data Structure in the first array****Insert the item into the Mesh Data Structure in the second array****Insert the image itself in to the Mesh Data structure in the third array****Retrieve all neighboring images****End While****Exit Program**

Figure 14: Pseudo Code

Taking a first at the algorithm we can see being  $O(n)$  as there is only one loop. Looking closely, extracting features will require a loop on the current unidentified item which would pass the efficiency class to  $O(n^2)$ . In addition to this retrieving the neighboring images require a loop through all neighboring images which confirm the  $O(n^2)$  as the two loops are not nested and not  $O(n^3)$ .  $O(n^2)$  can be passed to the  $O(n)$  by parallelization of the two inner loops. Doing the extraction and the retrieval in parallel will set the notation to  $O(n)$ .

#### 4.4 The Flowchart

The model consists of using existing solutions of image indexing and retrieval based on the human recognition theory as algorithm. The following subsection illustrates the algorithm used in the proposed solution. The algorithm is made of a preliminary step and three main parts that are Indexing, Retrieval and Learning each of which contains several sub-parts. (Figure 15) The algorithm will start by taking an input image as the image we want to index or retrieve similar to it.

The image is segmented into unidentified items each of which has its features extracted. So the algorithm will loop on the unidentified items and extract their features until there are no more unidentified items to be identified. The features, the items, and the images will be inserted in the Mesh data structure for indexing. All images and items in the neighborhood of the items that have been inserted will be extracted or retrieved. The retrieved images represents image contains similar items of the input image. The proposed

solution works on features extracted from unidentified items and not the image itself.

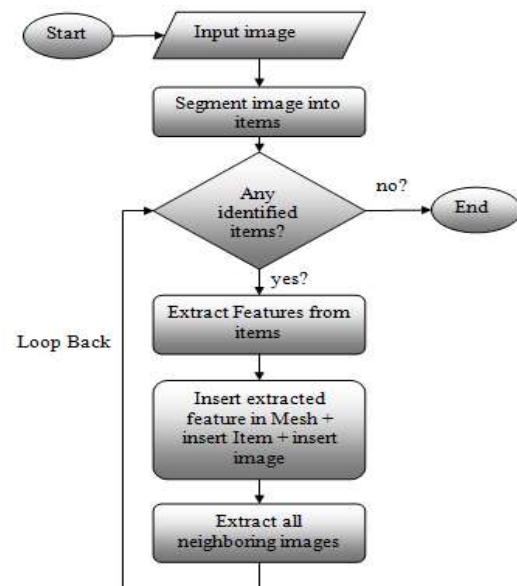


Figure 15: The Whole Process Algorithm

Getting back to the preliminary step that has to be done only once at the beginning of the procedure. This step consists of scanning all possible types of objects, running the feature extractor algorithm on those images, and finally labeling each feature then labeling the items themselves. It is a long process and mandatory process that should be worked on in the future solutions. This process is the fact of learning by a human being intervention but has nothing to do

with the artificial intelligence that is used to let the system learn by itself.

## 5. TESTING AND VALIDATION

In this paragraph, we show a sample test that is set up and the results are discussed.

### 5.1 Tests

The tests were conducted on the application part by part and not as the whole process, this technique was used in order to pre-identify any section that is not working properly. The project is divided into two sections: image segmentation which includes item feature extraction and the second section is comparing items with existing features.



Figure 16: Loading the Initial Image

As seen in Figure 16 the image is loaded to be used in the process, this image is segmented by detecting the edges first. Figure 17 shows the same Figure 16 segmented by edge detection technique. As the light bulb is not yet introduced in the system, in other words as the system did not learn to recognize the new shape “Light bulb”. A Learning step is needed for the system to know what shape it is dealing with (Figure 18).



Figure 17: Light bulb in Figure 16 Segmented

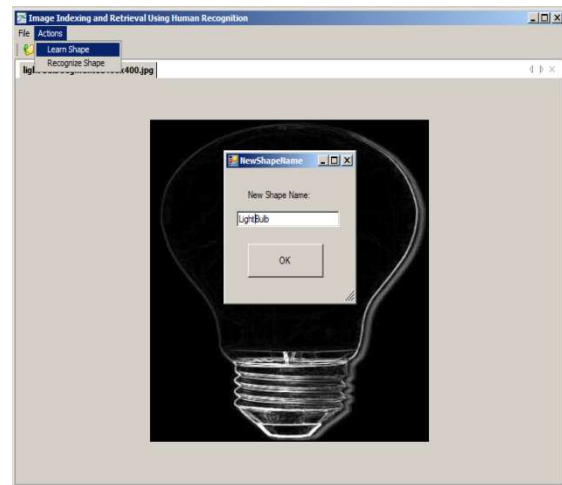


Figure 18: Introducing the New Shape to the System



Figure 19: Recognizing 10 Sample Items as a Light Bulb

After this step and as the system knows the new shape “Light bulb”, Figure 19 shows 10 sample tests done for light bulbs introduced in 10 separate images. A percentage of similarities for known shapes is shown below each image. There is a clear variation in the recognition percentage among the images.

## 5.2 Comparative Study

Table 1 below represents the percentages of successful retrieval from different methods in image indexing and retrieval. Many techniques have not provided any experimental efficiency results for the work done.

TABLE 1: *Retrieval Efficiency for Different Retrieval Algorithms*

Retrieval Methods	Average Percentage
A Shape Recognition Scheme Based on Real-Valued Boundary Points [2]	97 %
Scalable Color Image Indexing and Retrieval using Vector Wavelets [17]	30 %
A New Algorithm for Image Indexing and Retrieval Using Wavelet Correlogram [18]	71 %
Mapping Low-Level Image Features To Semantic Concepts [19]	68 %
Image Indexing & Retrieval Using Intermediate Features [12]	40 %
Image Indexing and Retrieval using Visual Perceptual Learning in Human Object Recognition	76.97 %

As seen in Table 1, some methods are not more than 40 % accurate on retrieval. The

proposed solution shows a mean of 76.97 %  $\pm$  11.13 % that is way far from the 97 % showed by the first algorithm that is worked on grey scale images only and low resolution 200 \* 256. Comparing methods with the same resolutions and number of colors show that our proposed solution is way in front of other methods. The proposed solution is more into building a base to construct on more than having some accurate results. The method used provide a new way for image indexing and retrieval by disassembling objects into a set of well defined Geons as the human being percepts objects.

## 5.3 Results

As seen, the system must go through an initialization phase in order to accomplish the task of retrieving images. This is initializing phase is also a learning step that helps the system recognize the images or items inside the images. After the system has learned the image, it is able to recognize any image similar to the first one introduced. The new items introduced are learned by the system and retrieve in later search results. The system passes through different phase in order to recognize the image. At the end the system is not recognizing the image itself but features of the containing items that exists in the image. Then the system extracts all images containing similar items in the image we are looking for. The threshold of error in recognition is set to 75 %, therefore all recognitions having more than 75 % of similarity are considered as similar items. Finally, the item is recognized and the similar images are retrieved.

## 6. CONCLUSIONS AND FUTURE WORK

In this paper, the problem of identification of a unique object with different designs as the light bulb is solved. Working with object features and not complete objects to index and retrieve objects features using human cognitive learning was a new concept in image indexing and retrieval; other methods extract features but store images, whereas we extracted features and indexed images according to those features. The second advantage is the usage of the Mesh data structure which is exceptionally efficient in space and region grouping that helped us build a virtual grouping of items next to each others. The concept using human object recognition had give in new path in this researching field. As the human brain is very efficient in object recognition, we tried to clone

the functionality of the human capabilities to recognize objects. We are still far away to clone exactly all the functionalities with the smallest details. Another advantage is the ability of this method to learn by indexing the searched sample in the mesh for future use; the more we search using this algorithm the more the database grow, which lead us to return more accurate results with time.

Main disadvantages for this solution is that it is not applicable in recognizing people by comparing faces; this solution will be able to get you all faces available but not the exact face of the person. The Solution cannot either get the exact imagery in a hospital. For instance it can be used to get all faces in a database then apply another method to compare exact features of the existing faces with the face we are looking for. The other disadvantage is the need to construct at least once the Mesh data structure to contain all available items for later indexing and retrieval and this part requires human intervention for once. After this human intervention, we run the system on several huge databases and we will have the Mesh data structure filled with all the items scanned. As the Mesh data structure grows, we will face another problem in iterating the mesh to get the available images. Even with the new generation of processors, iterating a mesh containing billions of images will consume a lot of time and resources. However this problematic will be solved with future generations of processors as we have an active research field for processing. In addition to that another type of identification was not supported by this method which is related to patterns. This method can classify patterns in a category alone but is not able to recognize patterns that exist in images. Finally, Mesh data structure needs to be implemented to be used as a data structure that is not the case in this study. The study was made using a Mesh data structure but the implementation is using index file to store data as the implementation is only a prototype.

Image indexing and retrieval is wide field of research, we were able to provide a solution that can solve a part of the problematic in retrieving images from large databases. Our conclusion consisted of segmenting an image into regions; each region can be either a pattern or an object. Patterns were not reviewed in this paper however, objects are well identified. Disassembling objects was the next step in the process of extracting features, those features were used to index the images in a Mesh Data

Structure and not a Tree as almost all other methodologies have used. Each node of the Mesh contains three Arrays one for the set of features and the relations between them, the second is for the list of images in which assembling the set features we will get any of those images and finally an array containing names, tags and labels for the current object node. Mesh has the ability to store items grouped in regions, which helps in categorizing the objects after feature extraction. The act of retrieval is exactly the same as indexing with two more additions, one the insertion of the sample image in the Mesh data structure and the extraction of the images found in the same node. Inserting the sample will give the possibility to the system to auto learn and improve it database of content. Our method consists of comparing features of objects and not the whole object with a similar.

No human intervention in the whole process is needed, on the other side, there is a big implication of the user to configure the system before it is used. Configuration consists of defining the set of Geometric Icons to be the root nodes of the Mesh and the entry point for scanning and to insert all possible items that does exist with their set of features. This task is simple but long and mandatory for the system to be able to do the job of indexing and retrieval. After that configuration the user must give a huge database and run the algorithm to insert all available images in the Mesh data structure for future use. After that phase the system will be ready to learn, index, and retrieve images.

Building a perfect system to index and retrieve images is impossible, on the other side we were able to propose a solution for the project that is very efficient, even though there are still some points not considered but overall the solution is an innovation by itself first because it is similar to the human brain object recognition and second that is different in vision than all other available methodologies. We were able to build a system by comparing set of features related to objects and store those features or attributes into a Mesh Data Structure. The indexing happens on the set of features and not on the object itself, as for the retrieval is done on image basis. Relying on the Mesh Data Structure was a plus in this solution as it has the ability to group items by regions. We did not use the VP-TREE as the storage data structure as it is limited in space to two nodes in forming the grouping. The system administrator must interfere for a

single time to adjust the system for first use and some more work has to be done on the solution for better efficiency and use.

The algorithm has provided a solution for image indexing and retrieval using query by example, the other method of retrieving images is to use text and labeling. Searching an image by label was not implemented however the platform was prepared for such a solution. The node of the Mesh Data Structure contains a list of names and labels for the items index at the node level. To improve the system the whole Mesh data structure must be indexed to retrieve images on naming conventions.

Face recognition is not implemented, therefore adding an algorithm that do face recognition to the system at the region that contains faces in the Mesh Data Structure will be a plus for the image indexing and retrieval, moreover if we add for each specific region or category a specific algorithm for more accurate identification the work can pass to a higher level. Nevertheless the projected solution has prepared the fundamental level for image indexing and retrieval and all coming methods should build over this one.

## REFERENCES

- [1] Cole L., Austin D., Cole L. "Visual Object Recognition using Template Matching." Australasian Conference on Robotics and Automation, 2004.
- [2] D. J. Buehrer, C. S. Lu and C. C. Chang. "A Shape Recognition Scheme Based on Real-Valued Boundary Points." Tamkang Journal of Science and Engineering, 2000: 55-77.
- [3] Gaber A. Sharawy, Neveen I. Ghali, Wafaa A. Ghoneim. "Object-Based Image Retrieval System Based On Rough Set Theory." International Journal of Computer Science and Network Security (IJCSNS International Journal of Computer Science and Network Security) 9 (2009): 160-165.
- [4] Hove, L.J. "extending image retrieval systems with a thesaurus for shapes." The Norwegian Information Technology Conference, 2004.
- [5] Mikolajczyk, K. and Zisserman, A. and Schmid, C. "Shape recognition with edge-based features." Proceedings of the British Machine Vision Conference, 2003. 779-788.
- [6] Pillai, Brijesh. "shape descriptor using polar plot for shape recognition." Clemson University, 2008.
- [7] Ronald Alferez, Yuan-Fang Wang. "Image Indexing and Retrieval Using Image-Derived, Geometrically and Illumination Invariant Features." IEEE International Conference on Multimedia Computing and Systems. IEEE Computer Society, 1999.
- [8] Tsai, Yao-Hong. "Salient Points Reduction for Content-Based Image Retrieval." Pwaset.World Academy of Science, Engineering And Technology, 2009. 656-659.
- [9] Zlatoff N., Ryder G., Tellez, B., Baskurt A. "Content-Based Image Retrieval: on the Way to Object Features." 18th International Conference on Pattern Recognition. ICPR 2006, 2006. 153-156.
- [10] D., Dimov. "Rapid and Reliable Content Based Image Retrieval." Proceedings of NATO ASI on Multisensor Data and Information Processing for Rapid and Robust Situation and Threat Assessment, 2005.
- [11] Lowe, David G. "Object Recognition from Local Scale-Invariant Features." The Proceedings of the Seventh IEEE International Conference. IEEE International Conference on In Computer Vision, 1999. 1150-1157.
- [12] Mohamad Obeid, Bruno Jedynek, Mohamed Daoudi. "Image indexing & retrieval using intermediate features." Proceedings of the ninth ACM international conference on Multimedia . ACM, 2001. 531-533.
- [13] N. Simou, Th. Athanasiadis , G. Stoilos, S. Kollias. "Image indexing and retrieval using expressive fuzzy description logics ." Signal, Image and Video Processing, 2008.
- [14] Haiming Liu, Dawei Song, Stefan Rüger, Rui Hu, Victoria Uren. Comparing Dissimilarity Measures for Content-Based Image Retrieval. Asia Information Retrieval Symposium , 2008.
- [15] Tran Thanh, Bloehdrn Stephan, Cimiano Philipp, Haase Peter. Expressive Resource Descriptions For Ontology-Based Information Retrieval. 2007. CiteSeerX - Scientific Literature Digital Library and Search Engine (United States) .
- [16] Gershon Elber, Myung-Soo Kim. "Geometric shape recognition of freeform curves and surfaces." Graphical Model and Image Processing , 1997. 417-433.
- [17] Elif Albuz, Erturk Kocalar , Ashfaq A. Khokhar. "Scalable Color Image Indexing and Retrieval Using Vector Wavelets." IEEE

- Transactions on Knowledge and Data Engineering (IEEE Educational Activities Department) 13, no. 5 (2001): 851-861.
- [18] H. Abrishami Moghaddama, T. Taghizadeh Khajoieb, A.H. Rouhib, M. Saadatmand Tarzjana. "Wavelet correlogram: A new approach for image indexing and retrieval." *Pattern Recognition* (Elsevier B.V.) 38, no. 12 (2005): 2506-2518.
- [19] Stan Daniela, Sethi Ishwar K. "Mapping low-level image features to semantic concepts." *Storage and Retrieval for Media Databases*, 2001: 172-179.
- [20] Djeraba, C. "When image indexing meets knowledge discovery." *ACM MDM/KDD*, 2000. 73-81.
- [21] Yixin Chen, James Z. Wang. "Image Categorization by Learning." *Journal of Machine Learning Research*, 2004: 913-939.
- [22] Suresh Pabboju, Dr. A.Venu Gopal Reddy. "A Novel Approach for Content-Based Image Indexing and Retrieval System using Global and Region Features." *IJCSNS International Journal of Computer Science and Network Security* 9, no. 2 (2009): 119-130.
- [23] G. Papadopoulos, T. Athanasiadis, N. Simou, R. Benmokhtar, K. Chandramouli, V. Tzouvaras, V. Mezaris, M. Phiniketos, Y. Avrithis, Y. Kompatsiaris, B. Huet, E. Izquierdo. "Combining Segmentation and Classification Techniques for Fuzzy Knowledge-based Semantic Image Annotation." 3rd international conference on Semantics And digital Media Technologies . Koblenz: (SAMT), 2008.
- [24] Elif Albuz, Erturk Kocalar, Ashfaq A. Khokhar. "Scalable Image Indexing and Retrieval using Wavelets." 1998.
- [25] Remco C. Veltkamp, Mirela Tanase. *Content-Based Image Retrieval Systems: A Survey*. Technical Report UU-CS, Department of Computing Science, Utrecht University, 2002.
- [26] Markov I, "VP-tree: Content-Based Image Indexing," in 4th Spring Young Researcher's Colloquium on Database and Information Systems (SYRCoDIS'2007), Moscow, Russia, 2007.
- [27] Daniel K. Blandford, Guy E. Blelloch, David E. Cardoze, Clemens Kadow. "Compact Representations of Simplicial Meshes in Two and Three Dimensions." *International Journal of Computational Geometry and Applications*, 2003: 135--146.
- [28] Tyler J. Alumbaugh, Xiangmin Jiao. "Compact Array-Based Mesh Data Structures." 14th International Meshing Roundtable. *IMR 2005*, 2005. 485-503.
- [29] Mount, F. Betul Atalay and David M. "Pointerless Implementation of Hierarchical Simplicial Meshes and Efficient Neighbor Finding in Arbitrary Dimensions." In *Proc. International Meshing Roundtable*. (IMR 2004), 2004. 15--26.
- [30] Md., Farooque. "Image Indexing and Retrieval." *DRTC Workshop on Digital Libraries: Theory & Practice*, 2003.
- [31] Biederman, I. (1987). "Recognition-by-components: A theory of human image understanding". *Psychological Review*, 94, 115-147
- [32] Live.com image search, June 2009.
- [33] <http://www.peterborough.gov.uk>, June 2009.