Unsupervised Categorization of Objects into Artificial and Natural Superordinate Classes Using Features from Low-Level Vision

Zahra Sadeghi *zahra.sadeghi@ut.ac.ir*
Cognitive Robotics Lab
School of Electrical and Computer Engineering
College of Engineering, University of Tehran, Iran
School of Cognitive Sciences
Institute for Research in Fundamental Sciences (IPM), Tehran, Iran

Majid Nili Ahmadabadi *mnili@ut.ac.ir*
Cognitive Robotics Lab
School of Electrical and Computer Engineering
College of Engineering, University of Tehran, Iran
School of Cognitive Sciences
Institute for Research in Fundamental Sciences (IPM), Tehran, Iran

Babak Nadjar Araabi *araabi@ut.ac.ir*
Cognitive Robotics Lab
School of Electrical and Computer Engineering
College of Engineering, University of Tehran, Iran
School of Cognitive Sciences
Institute for Research in Fundamental Sciences (IPM), Tehran, Iran

Abstract

Object recognition problem has mainly focused on classification of specific object classes and not much work is devoted to the problem of automatic recognition of general object classes. The aim of this paper is to distinguish between the highest levels of conceptual object classes (i.e. artificial vs. natural objects) by defining features extracted from energy of low level visual characteristics of color, orientation and frequency. We have examined two modes of global and local feature extraction. In local strategy, only features from a limited number of random small windows are extracted, while in global strategy, features are taken from the whole image.

Unlike many other object recognition approaches, we used unsupervised learning technique for distinguishing between two classes of artificial and natural objects based on experimental results which show that distinction of visual object super-classes is not based on long term memory. Therein, a clustering task is performed to divide the feature space into two parts without supervision. Comparison of clustering results using different sets of defined low level visual features show that frequency features obtained by applying Fourier transfer could provide the highest distinction between artificial and natural objects.

Keywords: Objects' Super-class Categorization, Low Level Visual Features, Categorization of Objects to Artificial and Natural, Local and Global Features, Color, Orientation, Frequency.

1. INTRODUCTION

Object recognition is a prominent problem in many fields of study such as computer vision, robotics, and cognitive sciences and has been studied for four decades [1]. The ultimate goal of this problem is to find a proper visual representation to identify each object effectively. However, the emphasis has been mainly laid on classifying very specific groups of objects in which the members of each class have many similarities in shapes and textures. In contrast, not much work has been devoted to the problem of recognizing objects of more general classes due to vagueness in identifying the common properties for all the members of a high level conceptual group.

Moreover, object recognition problem has been mostly studied as a similarity measurement problem in a supervised environment and so it needs many labeled training examples from some predefined and known classes before making prediction about the label of unknown and unseen examples. This process is tedious and time-consuming and relies heavily on the goodness of the selected training examples. Also, as stated in [2] there is no comparison between the different methods of object classification and each paper took an ad hoc approach on a special dataset.

In this paper we intend to categorize members of classes in the highest generality level. We assume that artificial/natural objects are positioned in the highest level of abstraction and hence investigate the problem of finding an efficient representation for each object to show this difference. Although the problem of indoor vs. outdoor scene or city versus landscape image classification is studied by many authors [3],[4],[5],[6],[7] little attention has been given to the problem of artificial vs. natural object distinction which is the subject of this paper.

2. RELATED WORK

The topic of artificial/natural discrimination is studied in both computer vision and neuroscience papers. In [8] a method for identification of artificial objects in natural scenes based on applying Zipf's law is proposed. Their idea is based on the observation that man-made objects are much simpler in their texture and generally contain much more uniform and flatter regions compared to natural ones. The idea of uniformity is also used by Caron et. al [9]. They measured the amount of uniformity by computing gradient or derivate of images. Fractal patterns are another approach applied to measure the amount of uniformity in images [10],[11]. In a specific problem, artificial and natural fruits are tried to be classified using color and texture features [12].

In a separate study the energy of Gabor orientation filter maps for Natural/man-made object classification is used as feature representatives [13].They showed that Gabor orientation energies of man-made objects have more variations than natural objects and their associated Gabor diagrams have sharper points. KNN is used to classify feature vectors obtained from computing orientation energy of each image. Line and edge properties are also used to categorize images into natural and artificial classes [14].

Yet in another attempt, a model based approach is used to detect artificial objects in which basic shapes are applied as templates to represent for artificial shapes [15].

In addition, there are many significant studies on the animate/inanimate recognition in primate's brain in the field of neuroscience. A detailed review on two different theories related to animate/inanimate object categorization is studied in [16]. It is shown that the representation of animate and inanimate objects in brain is separated both functionally and anatomically by recording the BOLD activity in visual cortex [17]. In [18] an experience is conducted to evaluate the effect of different conceptual levels of animate and inanimate objects based on the response time. Bell et al. studied the relation between three different hypotheses about the organization of IT cortex for object representation based on animacy, semantic category, and visual features using FMRI [19].

3. SUPERORDINATE OBJECT CLASSES

Our purpose is to make the most general distinction between object categories. While in some studies, animate/inanimate categories are assumed to be at the highest level of abstraction [20],[21], in our view, artificial/natural categories encompass more general visual characteristics and are located in the highest level of inclusiveness. In animate/inanimate categorization the discerning characterization is life. However, in artificial/ natural classification, the question is whether the object is a man-made product or not. Our proposed object categorization diagram is depicted in Figure 1 in which objects are first divided into artificial and natural classes. The natural objects can further be subdivided into animate and inanimate objects. However, artificial group include only inanimate objects. For example, shells and rocks belong to inanimate natural objects.

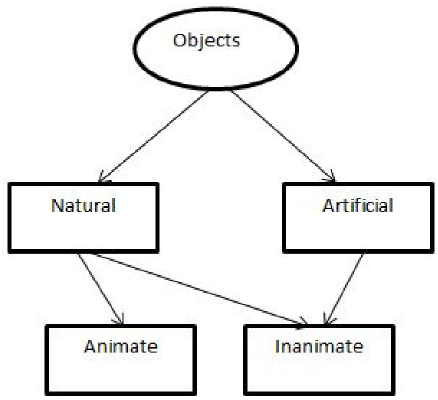


FIGURE 1: The Highest Levels of Object Categories.

3.1 Rationale Behind Unsupervised Learning of Superordinate Object Classes

Object recognition is mostly considered as a supervised problem in which the learning system needs to see many training samples to determine the label of test data with respect to the stored class representatives (eg. [22],[23],[24],[25].) One problem with supervised learning methods is that they are devoid of the power of generalization and they are dependent on the pre-computed features which are selected subjectively by their designers.

In this paper, we aim to perform the categorization of artificial/natural entities without the intervention of long term memory. In other words, instead of object classification, we try to group or cluster image data according to their similarities. We conjecture that the visual differences between these two categories can be found without any recall from memory, i.e., the difference between man-made and natural objects is visually observable and it doesn't depend on prior information. This hypothesis is supported with studies that have shown infants under 20-month year old can distinguish between superordinate level classes but cannot discriminate between basic level categories [26],[27],[28],[29].Other studies have shown that children at the earliest age of living can categorize more inclusive classes much better than other levels [30],[31],[32],[33]. Recently, Rostad et al. have shown that the starting age at which children are capable of animate/inanimate categorization is found to be around 14 month year old [34]. These findings encourage the idea that the task of superordinate class discrimination is not relied on prior knowledge. In this direction, we first define feature vectors for each image and then perform a clustering task using k-means algorithm in order to divide feature space into separate regions.

4. DESCRIPTION OF PROPOSED FEATURES FROM LOW LEVEL VISION

In order to distinguish between artificial and natural classes we need to find general visual characteristic of objects. As Rosch and Lloyd pointed out superordinate categories are low in category resemblance [35], i.e., the superordinate level contains a very wide range of basic and subordinate object categories with different appearances and shapes, and so finding an effective representation with a high capability of discrimination is a very challenging task because it is not visually clear what is the common set of features which is shared among all the members of a superordinate category.

In the present study, we intend to show the power of several complex features derived from basic visual features, i.e., color, orientation, and frequency feature sets.

Note that we considered low level attributes of visual information i.e., orientation, frequency and color as basic features and the resulted features after particular computation on them as complex features. This is similar to Huble and Wisel Findings [36] that there exist simple and complex sets of neurons in visual system for processing visual information. In addition, according to psychologists, there are special visual receptors that respond to particular visual features [37]. We now explain how these features are computed.

For extracting orientation features we used Gabor wavelet functions:

$$g(x, y) = \frac{1}{\pi(\sigma s)^2} e^{-\frac{x^2+y^2}{2\sigma^2}} (\cos(x) + i\sin(x)) = G\cos(x) + G\sin(x) \quad (1)$$

$$x = (x_0 \cos(\alpha) + y_0 \sin(\alpha)) / s$$

$$x = (y_0 \cos(\alpha) - x_0 \sin(\alpha)) / s$$

$$-(2s + .5) \le x_0 \le 2s + .5,$$

$$-(2s + .5) \le y_0 \le 2s + .5$$

$$s \in \{.5, 1, 1.5, 2\}$$

$$\alpha \in \{0, \pi/6, \pi/3, \pi/2, 2\pi/3, 5\pi/6\}$$

We then removed the DC component from the cosine part.

To obtain complex features of orientation, we took the same approach as described in [13] by computing the sum of absolute values of pixel of images convolved with a bank of 24 Gabor or log Gabor filters.

Frequency features are computed using Fourier Transform functions of input images and then the sum of squared absolute values are computed on both phase and amplitude. We noticed that the shape of magnitude and phase of images in frequency domain is different for artificial and natural groups. The effect of using phase and amplitude spectrum of Fourier transform in man-made and natural scene images is discussed in [38].

The entropy of each input image is computed as well to measure the amount of randomness in object images. For this, we have computed the entropy of both RGB and gray input images.

We also defined four more complex features based on statistical analysis of histogram of edge and color attributes motivated by the fact that in general, artificial objects are much simpler in their texture properties. These features can be grouped into two main attributes, namely diversity and variability characteristics which represent two basic characteristics of artificial and natural images. Diversity feature demonstrates the number of distinct orientations and colors in each input image.

For computing the diversity of orientations, we convolved the input image with Gabor filters at 360 different orientations. We then applied a max pooling operator by selecting the max orientation that causes the highest output in convolution, i.e., in each pixel we look for an orientation corresponding to the maximum magnitude of the result of convolution. For Gabor filter computation we followed the same approach proposed by Riesenhuber and Poggio [39] and its parameter values are represented in equations (2-6) and Table 1 correspondingly.

$g(x,y)=\exp\{-(x^2+G^2.y^2)/(2\sigma^2)\}\cos(2\pi x/\lambda)$ (2)
$x=x_0\cos(\theta)-y_0\sin(\theta+\rho)$ (3)
$y=x_0\sin(\theta)+y_0\cos(\theta+\rho)$ (4)
$-2\le x_0\le 2, -2\le y_0\le 2$ (5)
$\lambda=(RF\_size)^2/div$ (6)
$\sigma=\lambda*.8$ (7)

G	θ	ρ	RF_size	div
.3	0,1,2,…,359	0	5	4

TABLE 1: Gabor Filter Parameter's Value.

For diversity of color, RGB space is transformed into HSV color space and then a 360 bin histogram of hue values is computed. While RGB space is so sensitive to intensity variation, in HSV color space, intensity is separated from other components. It has also been found that this space is more likely to human visual system [40]. Having computed the histogram of edge and color, we finally count the number of different orientations which ranges from 0 to 359 and the number of hue colors which differs from 1 to 360.

In contrast to diversity, the variability attribute checks the number of peaks in each histogram which represents the number of dominant orientations and colors in each input image. This property shows the amount of change and variation and measures the coarseness of input image, i.e. how often we have a noticeable alteration in the orientations and colors. We implemented this feature by counting the number of peaks in the histogram of orientations and colors of input images. In a nutshell, having computed the histogram of orientation/color, the defined features can be explained as follows:

Diversity= the number of filled bins of histogram

Variability=the number of peaks of histogram

Generally, in comparison to artificial objects, natural objects group are characterized by higher values in Gabor energy, and much more amount of values in entropy, variability and diversity of pixels. In other words, large regions of artificial objects are of the same color and orientation, but in contrast, it is unlikely to find a region in natural objects with exactly identical color and orientation. However, sometimes, the opposite property can also be observed in some artificial and natural objects. For instance, artificial objects with complicated patterns in texture tend to have the attribute of naturalness due to its high variation in color and orientation.

5. LOCAL PROCESSING AND RANDOM EYE MOVEMENT

As was mentioned earlier, we have applied two different modes for feature extraction which are local and global strategies. In the local computation, complex features are extracted from random patches and are averaged and then the overall decision is made with respect to the average normalized feature values obtained from randomly selected regions from each object. However, in the global strategy, the defined features are extracted from the whole image. Thus based on the global or local method, we select the whole image or patches of image as the input data. The whole procedures of these strategies are described in Algorithm 1 and 2 respectively.

```
for im = 1 to size(ImageDataset) do
    InputImage = ImageDataset(im)
    BF = BasicF eatures(InputImage)
   CF(im; :) = ComplexF eatures(BF)
end for
Cluster(CF)
```

ALGORITHM 1: Global strategy

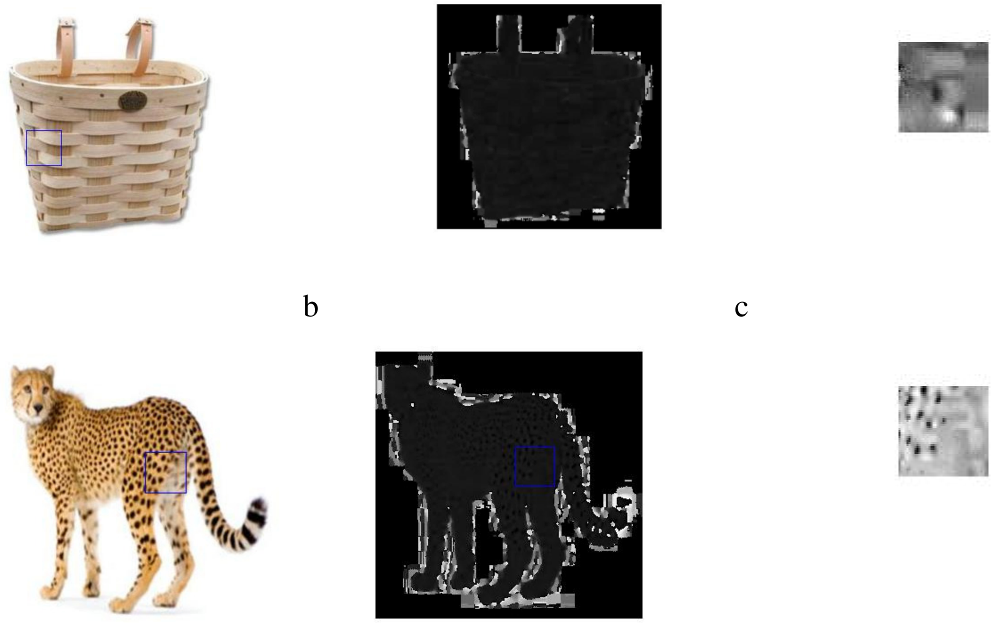
```
for im = 1 to size(ImageDataset) do
    image = ImageDataset(im)
    for i = 1 to exploringWindowsNum do
      InputImage = randomPatch(image)
      BF(i; :) = BasicF eatures(InputImage)
     CFi(i; :) = ComplexF eatures(BF; inputImage)
   end for
CF(im; :) = Average(CFi)
end for
Cluster(CF)
```

ALGORITHM 2: Local strategy

Our local approach is based on random area selection, i.e. we explore different regions of objects randomly based on the assumption that the visual system has no prior knowledge for discerning between artificial and natural objects. In other words, we hypothesize that for a neutral subjective viewer with no previous knowledge, fovea attention wanders around the central middle point of an image which we call it gaze point.

Moreover, we didn't take the saliency detection approaches for selecting the local windows, because the salient regions (which are basically defined as regions with great amount of edges and colors) are not the only source of information to help us make a decision about the objects' type as being artificial or natural. Rather, the steady and monotonous regions are informative as well, and therefore, all areas of objects are important in order to decide the artificial/natural group of each object.

Computing the features locally have this advantage that the extracted features can represent small variation in features' differences much more accurately. In addition, since each image includes only one object, a considerable amount of each image contains plain background and it can reduce the accuracy of performance in the global strategy. Note that in local hue computation, the minimum and maximum of each block are computed for each local patch instead of the whole image. And in orientation and frequency computation, the Gabor filters and Fourier transform are applied only on a single patch. Figure 2 compares the local and global approach in hue computation for two sample images.



a b c

a b c

FIGURE 2: Global vs. Local Hue Computation
a: original image with a random selected patch,
b: Hue component computed globally for the whole image,
c: Hue component computed locally for the selected patch

6. EXPERIMENTAL RESULTS

One of the major issues in object recognition evaluation is the lack of proper dataset [41]. Most of the available datasets are composed of very limited range of classes. For example, UIUC database [42] contains only car images, CORE database [43] contains only vehicle and animals, and Pascal VOC challenges [44] are made of motorbikes, bicycles, cars, and people classes. Other datasets which cover a wider range of classes (ex. SUN dataset [45]) are specified to scene understanding purposes.

For the task of artificial/natural object discrimination we decided to use the images from two available datasets, i.e. Caltech-101 [46] and Coil-100 [47] object collections. Our dataset contains two main categories of artificial objects (selected from COIL-100) and natural objects (selected from Caltech-101 object libraries). The selected subclasses for artificial and natural groups are listed in Table 2.

In addition, we created another dataset by selecting 454 images from Hemera object collection database [48] and divided them into two groups of artificial and natural objects. In addition, as a preprocessing all images are converted to Jpeg format.

Natural Objects			
Ant	Dalmatian	Leopard	Rhino
Bass	dolphin	kangaroo	Rooster
Beaver	dragonfly	llama	scorpion
Bonsai	elephant	Lobster	sea-horse
brontosaurus	emu	Lotus	starfish
butterfly	flamingo	nautilus	stegosaurus
cougar-body	flamingo-head	Octopus	strawberry
cougar-face	Gerenuk	Okapi	sunflower
Crab	hawksbill	Panda	Tick
crayfish	hedgehog	Pigeon	water-lilly
crocodile	Ibis	Pizza	wild-cat
crocodile-head	joshua-tree	platypus	
Artificial Objects			
Obj1,obj3			
Obj5 to obj62			
obj76 to obj82			
Obj64 to obj74			

TABLE 2: Artificial and Natural object classes selected from Caltech and Coil datasets.

As was mentioned before, we have defined complex feature vectors derived from different basic features of frequency, orientation and color.

The frequency features is a 3-dimensional feature vector obtained from:

$$FI = F(input\_image) \quad (8)$$

$$magnitude(x,y) = \sqrt{\mathrm{Re}(FI(x,y))^2 + \mathrm{Im}(FI(x,y))^2} \quad (9)$$

$$phase(x,y) = \tan^{-1}(\mathrm{Im}(FI(x,y) / \mathrm{Re}(FI(x,y)) \quad (10)$$

$$FreqFeat(1) = \sum_{x,y \in input\_image} |\, magnitude(x,y)\,| \quad (11)$$

$$FreqFeat(2) = \sum_{x,y \in input\_image} \log(1 + magnitude(x,y)) \quad (12)$$

$$FreqFeat(3) = \sum_{x,y \in input\_image} |\, phase(x,y)\,| \quad (13)$$

Where FI is the result of Fourier Transform of input images.

For orientation feature, 24 dimensional feature vectors are obtained from sum of the absolute energy of convolution of images with Gabor filters which are computed for scale values of .5 to 2 with step sizes of .5 and orientations from 0 to 150 with step size of 30:

$$GI(x,y,s,o) = G(input\_image, s, \alpha) \quad (14)$$

$$gaborFeat(s,\alpha) = \sum_{x,y \in input\_image} |\, GI(x,y,s,\alpha)\,| \quad (15)$$

$$s \in \{.5, 1, 1.5, 2\}$$

$$\alpha \in \{0, \pi/6, \pi/3, \pi/2, 2\pi/3, 5\pi/6\}$$

Where GI is obtained by convolving the input image with Gabor filters with a specific scale s and orientation α .

The entropy feature vector is a 2-dimensioanl feature vector including entropy of orientation and color which is obtained by computing the entropy of both RGB and gray input images using the following equation:

$$EntropyFeat(1) = entropy(Input\_image\_RGB) \quad (16)$$
$$EntropyFeat(2) = entropy(Input\_image\_gray) \quad (17)$$
$$entropy = \sum -H(I).\log(H(I)) \quad (18)$$

Where H(I) stands for 256-bin histogram counts of the input image I.

The histogram of orientation and color feature vectors are a two dimensional vector composed of diversity and variablity attributes which were explained in section 4 and can be obtained by:

orientHFeat(1) = diversity(GrayInputImage) (19)
orientHFeat(2) = variablity(GrayInputImage) (20)
colorHFeat(1) = diversity(RGBInputImage) (21)
colorHFeat(2) = variablity(RGBInputImage) (22)

All the local strategy results are averaged for 10 independent runs of the whole procedure and the number of selected patches in each run is selected as 20 and 30 for the first and second datasets respectively. We used more patches for the second dataset (Hemera objects) because the size of images is larger than the images of the first dataset (Coil-Caltech). Note that, all the images of the first dataset are resized to 128*128 pixels, but the size of images of the second dataset is more than 200 pixels in each dimension.

As was mentioned before, for grouping the images, K-means clustering technique is used with value of k=2. To evaluate the quality of the generated clusters we used R-index, F-measure, precision, and recall of the obtained clusters [49] which are defined by:

$$RI = \frac{TP+TN}{TP+FP+FN+TN} \quad (23)$$
$$P = \frac{TP}{TP+FP} \quad (24)$$
$$R = \frac{TP}{TP+Fn} \quad (25)$$

$$F = \frac{(b^2+1).P.R}{b^2.P+R}, b=2 \quad (26)$$

Where TP, TN, FP, and FN stand for true positive, true negative, false positive, and false negative respectively.

The performance results of clustering with the local and global methods for both datasets are listed in Tables 3 to 6. All the results are rounded to two decimal points. Each Table shows the evaluation for the corresponding feature dimensions (explained in equations (8-22). In bold are represented the best results obtained from each feature set.

It can be inferred from the results that frequency features showed dominance in making distinction between artificial and natural images. While local strategy is applied on a sequence of image patches instead of the whole image, it has generated superior or equal results in comparison to global strategy due to high amount of similarity and cohesion between pixels of each patch. Note that we only have considered the patches which are located mostly on foreground (i.e. the objects) and the patches that fall on background are automatically removed. It may be concluded that in artificial/natural distinction, texture and local information plays more important role than shape and global information.

Feature name	**Feature dimension**	**RI**	**F**	**P**	**R**
FreqFeat	1	0.89	**0.93**	0.88	**0.94**
FreqFeat	2	0.88	0.88	0.93	0.86
FreqFeat	3	0.51	0.60	0.57	0.61
FreqFeat	1:3	**0.90**	0.89	**0.94**	0.89
GaborFeat	1:24	0.82	0.90	0.79	0.93
EntropyFeat	1	**0.61**	**0.62**	**0.75**	**0.60**
EntropyFeat	2	0.55	0.58	0.65	0.58
EntropyFeat	1:2	0.58	0.60	0.73	0.58
colorHFeat	1	**0.56**	**0.57**	**0.64**	**0.56**
colorHFeat	2	0.54	0.55	0.62	0.53
colorHFeat	1:2	0.54	0.55	0.62	0.54
orientHFeat	1	**0.59**	**0.60**	**0.67**	**0.59**
orientHFeat	2	0.50	0.52	0.58	0.51
orientHFeat	1:2	0.56	0.59	0.63	0.58

TABLE 3: Performance of Local methods for Caltech-Coil Dataset.

Feature name	**Feature dimension**	**RI**	**F**	**P**	**R**
FreqFeat	1	0.79	0.87	0.78	0.89
FreqFeat	2	0.87	0.88	0.90	0.88
FreqFeat	3	0.87	0.86	0.92	0.84
FreqFeat	1:3	**0.90**	**0.90**	**0.94**	**0.89**
GaborFeat	1:24	0.69	0.82	0.69	0.86
EntropyFeat	1	0.56	0.57	0.65	0.55
EntropyFeat	2	**0.64**	**0.65**	**0.72**	**0.63**
EntropyFeat	1:2	0.61	0.62	0.69	0.60
colorHFeat	1	**0.53**	**0.57**	**0.60**	**0.56**
colorHFeat	2	0.50	0.52	0.58	0.51
colorHFeat	1:2	0.52	0.55	0.59	0.54
orientHFeat	1	0.58	0.60	0.66	0.58
orientHFeat	2	0.64	0.64	0.72	0.62
orientHFeat	1:2	**0.65**	**0.64**	**0.73**	**0.63**

TABLE 4: Performance of global methods for Clatech-Coil dataset.

Feature name	**Feature dimension**	**RI**	**F**	**P**	**R**
FreqFeat	1	0.70	0.72	0.69	**0.73**
FreqFeat	2	0.71	0.71	0.72	0.71
FreqFeat	3	0.64	0.66	0.64	0.67
FreqFeat	1:3	**0.72**	**0.72**	**0.73**	0.72
GaborFeat	1:24	0.54	0.70	0.53	0.76
EntropyFeat	1	0.72	0.72	0.73	0.72
EntropyFeat	2	0.73	0.73	0.73	0.73
EntropyFeat	1:2	**0.73**	**0.73**	**0.73**	**0.73**
colorHFeat	1	0.70	0.70	0.69	0.70
colorHFeat	2	**0.73**	**0.73**	**0.73**	**0.73**
colorHFeat	1:2	0.73	0.73	0.73	0.73
orientHFeat	1	**0.69**	**0.69**	**0.69**	**0.69**
orientHFeat	2	0.63	0.64	0.63	0.64
orientHFeat	1:2	0.68	0.68	0.68	0.68

TABLE 5: Performance of Local methods for Hemera dataset.

Feature name	**Feature dimension**	**RI**	**F**	**P**	**R**
FreqFeat	1	0.82	0.75	0.83	0.82
FreqFeat	2	0.80	0.74	0.80	0.80
FreqFeat	3	0.78	0.74	0.78	0.78
FreqFeat	1:3	**0.84**	**0.76**	**0.85**	**0.84**
GaborFetat	1:24	0.56	0.60	0.56	0.61
EntropyFeat	1	0.51	0.56	0.51	0.58
EntropyFeat	2	0.51	0.56	0.51	0.58
EntropyFeat	1:2	**0.51**	**0.56**	**0.51**	**0.58**
colorHFeat	1	0.50	0.50	0.51	0.50
colorHFeat	2	0.51	0.51	0.51	0.51
colorHFeat	1:2	**0.51**	**0.51**	**0.51**	**0.51**
orientHFeat	1	0.56	0.73	0.55	0.79
orientHFeat	2	0.59	0.72	0.57	0.78
orientHFeat	1:2	**0.56**	**0.73**	**0.55**	**0.79**

TABLE 6: Performance of Global methods for Hemera dataset.

7. CONCLUSION & DISCUSSION

One possible approach for solving object categorization problem is a top-down view. In this direction, different levels of categories need to be recognized subsequently in which the inclusiveness of recognition levels decreases in a descending order. In this paper, the first conceptual level of abstraction is associated with artificial and natural categories. Based on neuroscientific views, artificial and natural objects are represented differently in brain. However, the processing and encoding of visual features is under debate. Experimental studies on children can support the theory that human may distinguish between these two categories without referring to their long term memory and hence our feature definition mechanism is an unsupervised learning algorithm which doesn't use pre-learned parameter sets for dividing the feature space into two general categories. However, automatic unsupervised artificial/natural grouping is a complicated task. First, the artificial/natural category is located in the highest level of abstraction. Thus, finding appropriate generic properties is not an easy task. Second, in contrast to classification problems in which there exists a set of labeled data that helps the categorization problem, in clustering problem, there is no access to any prior information in advance. Taking into account objects' characteristics we derived different high level features from basic low level features which can make distinction between artificial and natural categories of objects. We compared the discriminating effect of different features obtained by using Fourier transform, Gabor filter, entropy, and histogram of color and orientation for artificial/natural object distinction. Feature extraction is applied by using two different strategies of local and global processing and then a clustering task is performed to group the similar features with regard to Euclidean distance between them. The obtained simulation results showed that frequency features derived from Fourier Transform achieved the first highest efficiency tier in distinguishing between artificial and natural objects. Also, local strategy which is based on random patch selection corresponding to random eye movement resulted in comparable performance in term of accuracy and with regard to the lower amount of information processing.

8. REFERENCES

[1] D. Marr. (1982). Vision. A computational investigation into the human representation and processing of visual information.New York. W.H. Freeman.

[2] T. Tuytelaars, C. H.Lampert, M. B. Blaschko, W.Buntine. "Unsupervised Object Discovery: A Comparison."

[3] M. Szummer, R. W. Picard. (1998). "Indoor-outdoor image classification, Proceeding of International Workshop on Content-Based Access of Image and Video Databases, Bombay".

[4] E.C. Yiu, (1996). "Image classiffication using color cues and texture orientation."

[5] M.J. Kane, A.E. Savakis. (2004). "Bayesian Network StructureLearning and Inference in Indoor vs. Outdoor Image Classification."

[6] N. Serrano, A.E. Savakis, L. Luo. (2002). "A Computationally Efficient Approach to Indoor/Outdoor Scene Classification."

[7] M. Szummer, R.W. Picard. (1998). "Indoor-Outdoor Image Classification."

[8] Y.CARON, P. MAKRIS, N. VINCENT. (2002). "A method for detecting artificial objects in natural environments." pp.600-603.

[9] Y. Caron, P. Makris, N. Vincent. (2002). "A method for detecting artificial objects in natural environments." IPCR, pp. 600 603.

[10] D. Chenoweth, B. Cooper, J. Selvage. (1995). "Aerial Image Analysis Using FractalBased Models." IEEE Aerospace Applications Conference Proceedings, pp. 277-285.

[11] G. Cao, X. Yang, and Z. Mao. (2005). "A two-stage level set evolution scheme for man-made objects detection in aerial images." Proc. IEEE Conf. Comput. Vis. Pattern Recog., San Diego, CA, pp. 474-479.

[12] B. S., Anami, V. Burkapalli, V. Handur, H.K.: Bhargav. "Discrimination of Artificial Objects from Natural Objects."

[13] M., Kim, C., Park K. Koo. "Natural / Man-Made Object Classification Based on Gabor Characteristics, Image and Video Retrieval, LNCS 3568, pp. 550-559

[14] J. Tajima, H. Kono. (2008). "Natural Object/Artifact Image Classification Based on Line Features." IEICE Transactions. 91-D(8), pp. 2207-2211.

[15] Z.Wang, J. Ben Arie. (1999). "Generic Object Detection using Model Based Segmentation." IEEE Computer Society Conference on Computer Vision and Pattern Recognition.

[16] A. Caramazza, J. R. Shelton. (1998). "Domain-specific knowledge systems in the brain - the animate-inanimate distinction." Neuroscience, Vol. 10, pp. 1-34.

[17] T. Naselaris, D.E. Stansbury, J. L. Gallant. "Cortical representation of animate and inanimate objects in complex natural scenes."

[18] A.J., Wiggett, , I.C., Pritchard P.E. Downing. (2009). "Animate and inanimate objects in human visual cortex: Evidence for task-independent category."

[19] A.H., Bell, F. Hadj-Bouziane, J.B., Frihauf, R.B., Tootell, L.G. Ungerleider. (2009). Object representations in the temporal cortex of monkeys and humans as revealed by functional magnetic resonance imaging. J Neurophysiol 101, pp. 688-700.

[20] D. Poulin-Dubois, S. Graham, L. Sippola. (1995). Early lexical development: The contribution of parental labeling and infants categorization abilities. Journal of Child Language: 22, pp. 325-343.

[21] D. H. Rakison. Parts, motion, and the development of the animate inanimate distinction in infancy. In D. H. Rakison, L. M. Oakes (Eds.) (2003). Early category and concept development: Making sense of the blooming, buzzing confusion, pp.159-192.

[22] P.F. Felzenszwalb, D.P Huttenlocher. (2000). “Efficient matching of pictorial structures.” IEEE Conference on Computer Vision and Pattern Recognition, pp.66-73.

[23] B. Heisele. (2003). “Visual Object Recognition with Supervised Learning.” IEEE Intelligent Systems - AI's Second Century, pp. 38-42.

[24] D.J., Crandall, P.F., Felzenszwalb, D.P. Huttenlocher, (2005). “Spatial priors for part- based recognition using statistical models”. In IEEE Conference on Computer Vision and Pattern Recognition, pp.10-17.

[25] T. Deselaersa, G. Heigoldb, H. Ney. “Object Classification by Fusing SVMs and Gaussian Mixtures.”

[26] J. M. Mandler, P. J. Bauer. (1988). “The cradle of categorization: Is the basic level basic? Cognitive Development.” 3(3), pp. 247-264.

[27] J. M. Mandler, P. J. Bauer, L. McDonough. (1991). “Separating the sheep from the goats: Di_erentiating global categories.” Cognitive Psychology, 23, pp.263-298.

[28] M. H. Bornstein, M. E. Arterberry. (1991). “The development of object categorization in young children: Hierarchical inclusiveness, age, perceptual attribute and group versus individual analyses.” Developmental Psychology: 46, pp. 350-365.

[29] J. M. Mandler, L. McDonough. (1993). “Concept formation in infancy. Cognitive Development.” 8, pp. 291-318.

[30] B.A., Younger, D.D. (2000). “Fearing, A global-to-basic trend inearly categorization: Evidence from a dualcategory habituation task.” Infancy. 1, pp. 47-58.

[31] P.C. Quinn, M.H. Johnson, D. Mareschal, D.H. Rakison, B.A. Younger. (2000.). “Understanding early categorization: One process or two?.” Infancy. 1, pp.111-122.

[32] S. Pauen. (2002). “The global-to-basic level shift in infants categorical thinking: First evidence from a longitudinal study. International Journal of Behavioral Development.” 26, pp. 492-499.

[33] S. Pauen. (2002). “Evidence for knowledge-based category discrimination in infancy.” Child Development. 73,pp. 1016-1033.

[34] K. Rostad, J. Yott, D. Poulin-Dubois. (2012). “Development of categorization in infancy: Advancing forward to the animate/inanimate level.” Infant Behav Dev.

[35] E. Rosch, B.B. Lloyd. (1978). Cognition and categorization. Pp.27-48.

[36] D.H. Hubel, T.N. Wiesel. (2005). “Brain and visual perception: the story of a 25-year collaboration.” Oxford University.

[37] S.M. Zeki. (1976). “The functional organization of projections from Striate to prestriate visual cortex in the rhesus monkey. Cold Spring Harbor Symposia on Quantitative Biology.” 5, pp. 591-600.

[38] A. Oliva, A. Torralba. (2001). Modeling the shape of the scene: a holistic representation of the spatial envelope. International Journal of Computer Vision, 42(3), pp. 145-175.

[39] M. Riesenhuber, T. Poggio. (1999). "Hierarchical models of object recognition in cortex." Nature Neuroscience.

[40] M. Khosrowpour. Encyclopedia of Information Science and Technology, 1-5

[41] J. Ponce, T.L. Berg, M. Everingham, D.A. Forsyth, M. Hebert, S. Lazebnik, M. Marszalek, C. Schmid, , Russell B.C., A. Torralba, C.K.I, Williams, J. Zhang, A Zisserman. "Dataset Issues in Object Recognition."

[42] Agarwal, S., Roth, D. (2002). "Learning a sparse representation for object detection. In: Proc. European Conf. Comp. Vision." LNCS 23-53., Copenhagen, Denmark. pp.113-127.

[43] A. Farhadi, I. Endres, D. Hoiem. (2010). "Attribute-Centric Recognition for Cross-category Generalization."

[44] R. Fergus, P. Perona, A. Zisserman. (2003.). Object class recognition by unsupervised scale-invariant learning. In: Proc. IEEE Conf. Comp. Vision Patt. Recog. 2, pp. 264-271.

[45] J. Xiao, J. Hays, K. Ehinger, A. Oliva, A. Torralba. (2010). "SUN Database: Large-scale Scene Recognition from Abbey to Zoo." IEEE Conference on Computer Vision and Pattern Recognition.

[46] L. Fei-Fei, R. Fergus, P. Perona. (2004). "Learning generative visual models from few training examples: an incremental Bayesian approach tested on 101 object categories." CVPR, Workshop on Generative-Model Based Vision.

[47] S. A., Nene, S. K. Nayar, H. Murase. (1996). "Columbia Object Image Library (COIL-100)" Technical Report CUCS-006 96.

[48] http://www.hemera.com

[49] C.J., Van Rijsbergen. (1979). Information retrieval, Butterworths, London, second edition.