The scene classification based on topological properties

LIU SHUOYAN, XU DE, YANG XU
Institute of Computer Science, Beijing Jiaotong University, Beijing, 100044, China

Abstract: -In this paper, we propose a novel approach of scene classification using topological properties. The scene is represented by low dimensional features, that we term topological properties. There is a set of topological properties, including Euler number. The experimental results show the performance of the scene classification based on topological properties is satisfying.

Key-Words: - scene classification; topological property; natural images

1 Introduction

In last few years, many different approaches on scene classification have been proposed. Most of those approaches are based on Marr's theory using local features. In contrast, some experimental studies have suggested that recognition of real world scenes may be initiated from the encoding of the global configuration, ignoring the details and object information [1] .The global topological perception is prior to the perception of local properties [2].

There are some defects in traditional approaches based on local features, since low-level features may be different among the images in the same scene [5]. For example, here are three images of a coast scene, respectively colored in red, blue and white. Fig. 1 shows that the color feature can not be used as the criterion to classify the three coast images as the same category.

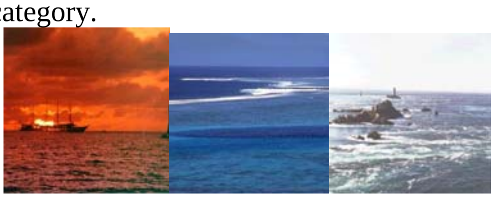


Fig. 1 Images with different low-level features for the same scene.

According to the theory of topological perception, the more one feature is invariant under general geometric transformation, the more stable this feature's perception is [2]. Particularly, since topological properties are the most invariant under various geometric transformations, such as rotation, shift and scaling, they may be more stable and global for scene classification than other local features. So this paper exploits topological properties to improve the performance of scene classification.

There is a set of well known topological properties, such as holes, Euler number, connectivity, and the inside/outside relationship, etc. In our experiment, we extract Euler number as a topological property for scene classification through image segmentation. It is obtained by spectral graph partitioning, namely Normalize cut [4]. The topological property extracted from an image provides a meaningful description of its scene category.

The paper is structured as follows: Section 2 introduces the theory of the topological perception. Section 3 describes the procedure how to compute the Euler Number. Section 4 gives the evaluation of this approach, along with the discussion of the results .A summary and conclusions form this work end this paper.

2 The theory of topological perception

In recent research, there are two different ideas in the study of perception, feature analysis and early listic registration. The former holds that perceptual processing is from local to global. The latter claims that the perceptual processing is from global to local, which has two typical theories, namely Gestalt and topological perception.

The core contribution of the Gestalt idea goes far beyond the notion that "Whole is more than the simple sum of it parts"; rather it is that "Holistic registration is prior to local analysis". [2]

The topological perception is a formal and systematic of early holistic registration. A topological transformation is a one-to-one and continuous transformation. The properties preserved under an arbitrary topological transformation are called topological properties. Euler Number, connectivity, the number of holes and the inside/outside relationship are some of typical topological properties, while the local geometric properties such as size, parallelism and collinearity are not topological properties. The Klein's hierarchy of geometries includes, in a descending order of stability, topological, projective, affine, and

Euclidean geometrical invariants [7]. A property is considered more global (or stable) the more general the transformation group is. The topological transformation is the most general and hence topological properties are the most global .And the time of perceiving form topological properties is less than other features.

Thorpe et al. find that human is able to categories complex natural scenes containing animals or vehicles very quickly. Fei-Fei et al. later shows that a little or no attention is required for such rapid natural scene categorization [3]. The topological properties are obtained by human at the earliest stage of the analysis of an images. So scene recognition by the topological properties is reasonable and effective.

3 The Euler number: A Topological Property

Euler number is one of the topological properties [6]. Euler Number equals to the number of objects in the region minus the number of holes in those objects.

$$E = C - H \tag{1}$$

E: Euler Number

C: the number of objects in the region

H: the number of holes.

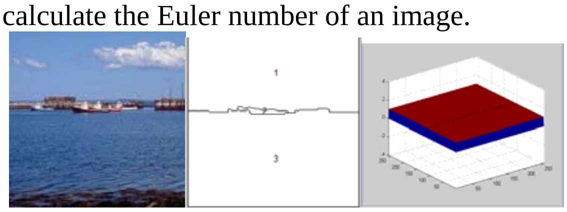
The Euler Number reflects the complexity of images. The complexity means that the object included in the image is complex with more holes. The larger the Euler Number is, the more complex the image is. Euler Number of different images in the same scene is similar, expressed by dissimilar degree (Eq.2) in the table1. But the distribution of the Euler Number on the different scenes is varied, as showed in the Fig. 2.

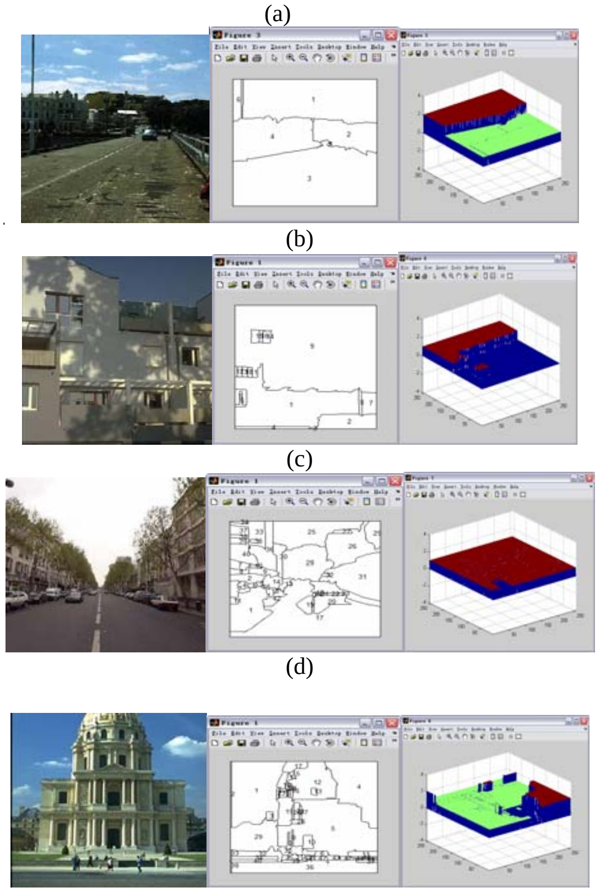
Table1the dissimilar degree of images for same scene

	Inside _city	Street	Tall building
dissimilar degree	0.6	0.67	0.8

$$\text{dissimilar degree} = \frac{\text{standard deviation}}{\text{average}} \tag{2}$$

In this paper, we segment the real-world scene images into regions by the algorithm of Normalized cut [4]. And then we compute the Euler number for each region using formula 1. Finally, by adding up all of the Euler numbers from each region in one image, we calculate the Euler number of an image.





(e)

Fig.2.Scenes with different Euler number, where the height level corresponds to the Euler Number at each pixel: a) coast) highway)inside_city,d)street)tall building

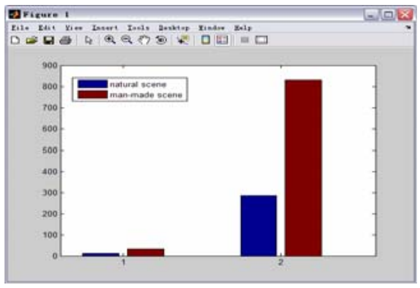
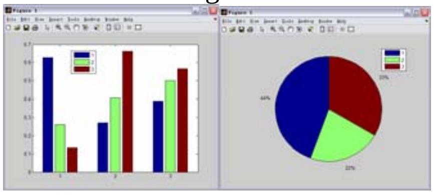


Fig.3 the means and sigma of Euler number in the natural and the man-made scene

Fig.2 shows that Euler number can be intuitively distinctive among different scene categories. The left column is the original images that come from each scene category. The middle column represents the images segmented by Normalized cut. The last column shows the Euler number distribution in each regions. The difference on the average and variance of Euler Number between the natural scene and the man-made scene can be also seen obviously in the Fig. 3. The average Euler Number of the man-made scene is 32.167, and the average of the natural scene is 12.3.

The topological properties have more advantages than other traditional features, for instance color.

Firstly, the topological features can also deal with gray images where the color property fails to classify scenes. Secondly, the images in the same scene category has different colors as shown in Fig. 1, but the topological properties among these images are similar as illustrated in Fig. 4.



(a) (b)

Fig.4 the contrast among the images showed in the Fig.1 between color and the Euler number (one of the topological properties) a: the distribution of the color b: the distribution of the Euler Number

The topological properties are invariably, even though the images are processed. But the local features will change unpredictably. So the topological features are usually more useful than local features.

4 Experiment Results

In our experiment, the database includes 2503 images, consisting of eight categories: 360 coasts, 165 forests, 260 highway, 306 inside_city, 374 mountains, 410 open countries, and 272 street and 356 tall buildings. They come from the Corel stock photo library [1]. The pictures are taken from a digital camera and downloaded from the web. The size of the images is 256×256. The natural scene is composed of 1144 images (e.g., coast, mountains and open country), while the man-made scene contains about 934 images (e.g., inside_city, street and tall buildings).

4.1 Classification results

We use 150 randomly selected training images from every scene and the rest for testing. Feature used is the Euler number through the segmentation. The final feature vector is 1-dimensioanl. The correct rate of recognition the man-made (inside_city, street and tall building) and natural scene (coast, mountain and open country) is 71%. The confusion table is shown on the table 2. Table 3 shows the result for coast and forest, while the high way and the inside_city are showed in the table 4.

Table 2. Confusion matrix between natural and man-made scene(N=1832)

	Natural	Man-made
natural	849	165
Man-made	366	443

Table 3. Confusion matrix between coast and forest scene(N=389)

	coast	forest
coast	296	17
forest	50	71

Table 4. Confusion matrix between high way and inside city scene(N=473)

	High way	Inside_city
High way	207	6
Inside_city	112	148

Classification is performed by a Bayes classification. The Bayes uses a labeled scenes database with the corresponding category (training set) in order to classify new unlabeled images: in the training stage, the mean and sigma, which are the necessary parameters in the Bayes classification, are computed through the training images. In the test stage, the Gauss model (Eq.3) [8] is computed using the two parameters and the topological properties form the testing images.

$$g_i(x) = -\frac{1}{2}(x-\mu_i)^T \Sigma_i^{-1}(x-\mu_i) - \frac{d}{2}\ln 2\pi - \frac{1}{2}\ln|\Sigma_i| + \ln P(\omega_i) \quad (3)$$

Then, the new pictures are assigned to the scene category with the largest $g_i(x)$.

The classification task assigns each testing image to the category. The performance is measured by a confusion table, and overall performance rates are measured by the average value of the diagonal entries of the confusion table [3].Generally, the performance of scenes classification is evaluated by the recall and precision.

From the table 1, the recall of natural scene is 70%, while the precision is 84%; the recall of man-made scene is 72%, while the precision is 55%. From the table 2, the correct rate between coast and forest is 85%, at the same way; the recall and precision of coast are 85.6% and 95% respectively, while the recall and precision of forest are 81% and 59% respectively. From the table 3, the recall and precision of high way are 65% and 97% respectively, while the recall and precision of forest are 96% and 57% respectively.

4.2 Discussion

The performance of scene classification is different by the number of training images. The result based on the number of training images is 15, 30, 150, 300 and 450 are showed on the table 5.

Table 5 the error rate on the different number of training images

The number of images	15	30	150	300	450
Error rate	27%	28%	29%	29.3%	29.6%

The results show that the larger number of training images leads to a worse result of scene classification.

Here are the reasons. First, using one-dimension features and too many training images result in over-fit.Thus, 15 training images work well than 450 ones. Secondly, the single-Gauss model classified in our experiment weakens the performance. Since the man-made scene includes many different scenes such as inside-city, street and tall building, so is the natural scene. The more training images are used, the less performance is attained for this model. So, the performance of classification is also not good enough. As far as I know, the mix-Gauss model may be improved the performance.

When the training images are 150, the rate of error is 28.9%. Fig.5 shows some misclassified images.

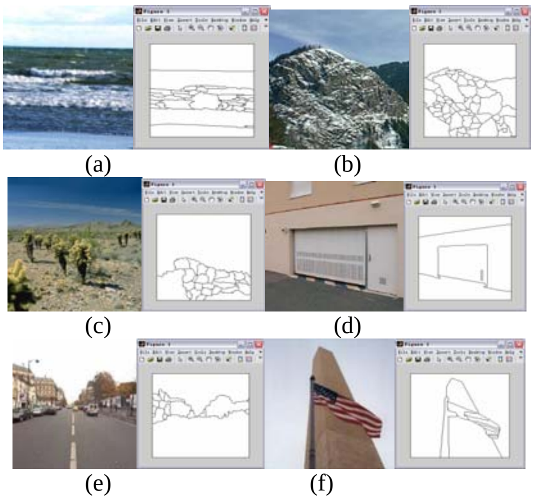


Fig. 5 the misclassified scenes using the Euler number representation a) a coast, b) a mountain ,c)open country classified as a man-made d) a high way, e)tall building ,f)street classified as natural

The local feature used in mostly scene classfication is color. We used the same training images and testing images to classify the scene. And the result compared between the local feature like color and the topologcial properties like Euler number showed in table 6.

Table 6. Confusion matrix between natural and man-made scene compared by color and Euler number (N=60)

	natural		Man-made	
	color	Euler	color	Euler
natural	22	27	8	3
Man-made	13	12	17	18

The color feature computed by 12 bins,with each 4 bins for color components r, g and b. In the top experiments, the number of training images and testing images are 30. The rate of the correct classification is 75% based on the topological properties, while the local feature bring 65% correct rate.

The dimension of the local features is 12, while the topological property is only one. And, less training images can also bring better performance，so the topological features are more effective. In my opinion, by increasing the number of topological properties, the performance can be improved continually.

Now there are many topological features, and the next job is to recognize the real-world scene using more topological features.

6 Conclusions

In this paper, we present a novel approach for scene classification based on topological properties. The topological properties of an environmental scene can be described as Euler number. These properties are meaningful to human observers by many psychological experiments. The topological properties provide a holistic description of the scene and they can't be replaced by other local features.

Although the experimental results show the encouraging performance, the research also shows that there is plenty rooms for improvement. Future work will be focused on other topological features for scene recognition, which is important to high-level semantic content analysis of images. Also, the topological properties can be extracted directly that avoids image segmentation.

References:

[1]AUDE OLIVA, *Modeling the Shape of the Scene: A Holistic Representation of the Spatial Envelope*, International Journal of Computer Vision 42(3), 145-175, 2001

[2] Lin Chen, *The topological approach to perceptual organization*, Visual Cognition 12(4), 553-637, 2005

[3] Anna Bosch, Xavier Munoz, *A review: Which is the best way to organize/classify images by content?* Image and Vision Computing,2006.

[4] Jianbo Shi and Jitendra Malik, *Normalied Cuts and Image Segmentation*, IEEE Trans. Pattern Anal. Mach. Intell, Vol(22), No8 ,August 2000.

[5] Anna Bosch, *Scene classification via plsa*, ECCV,2006

[6] Kang jinwang. Intuitive topology，Beijing Normal university Press,1995

[7] Zhu Ying, *Chen Lin's Theory of Topological Perception.*Psychological Science Vol(28) 1031-1034, 2005

[8] Qi zhaobian, *Pattern Recognition*, Tsinghua Univerisity Press, 2000