Image Features Based on a Mixed Fractal Model and Evaluation of Their Effectiveness in Image Retrieval

Toshinori HAYASHI

Computing Center, Hokkaido University
N11W5, Sapporo, 060 Japan
E-mail: hayashi@cc.hokudai.ac.jp

Yuzuru TANAKA

Graduate School of Engineering, Hokkaido University
N13W8, Sapporo, 060 Japan
E-mail: tanaka@huee.hokudai.ac.jp

Abstract

Many research works on content-based image retrieval have made use of image features. On the other hand, many kinds of image features have also been developed in the research field of computer vision and image analysis. The fractal signature is an example, which is based on the assumption that an image consists of a single fractal. For complex natural objects like grass and trees, however, a mixed fractal model is more suitable, which assumes an image consists of multiple different fractals. The naive use of the fractal signature for a mixed fractal images causes the problem of missing information. In this paper we propose a new image feature called the average fractal signature and its variants to solve this problem. We also demonstrate the effectiveness of the features by experiments classifying real photograph images into four object types, grass, trees, clouds, and waves. To perform the experiments in the context of image retrieval we used a simple box decomposition strategy for image partitioning without a sophisticated segmentation technique. We obtained good classification results for grass and trees by using new image features.

1. Introduction

An increase of image data in multimedia applications demands image retrieval. As textual databases made it possible to retrieve textual information by words, image retrieval systems capable of retrieving by object types like trees and grass are expected to be developed. Many research works on image retrieval systems have made use of image features[5, 6, 7]. Since their classification ability plays an important role in retrieval systems, development of image features is one of crucial issues for its realization.

Many kinds of image features have also been developed in the research field of computer vision and image analysis. Image features based on fractal models including the fractal dimension[2] and the fractal signature[4] are examples that have been used effectively for classifying and segmenting images of natural objects. But, its classification ability is not necessarily enough for

---

grass and trees as you see in out experiments in section 3.

Since a natural object is created by various kinds of physical or biological processes, a fractal model might be too simple to model the object. Hence, we adopt a mixed fractal model[3], which assumes an image consists of multiple different fractals. Application of the fractal dimension and the fractal signature to a mixed fractal image, however, causes the problem of missing information, which will be explained in detail in section 2. To solve the problem, in this paper we propose a new image feature called the average fractal signature, as well as its variants, based on a mixed fractal model.

We also evaluate their effectiveness through classification experiments. Evaluation should be performed in the context of image retrieval. Many image retrieval systems have assigned one entry for each decomposed square image instead of one for each object in an image to avoid the difficulty of segmentation. Hence, we adopted a simple box decomposition strategy for image partitioning in the experiments. We used various kinds of images in a commercial CD-ROM as experimental data.

The outline of this paper is as follows. After we propose new image features in section 2, they are evaluated through classification experiments in section 3. We obtained good classification results for the experimental data that conventional image features does not work well for. We conclude the paper in section 4.

2. Image Features

A fractal model and a mixed fractal model are suitable tools for analyzing various natural objects and their photographs. In this section, after brief introduction of conventional fractal-based image features, we propose a new image feature and its variants based on mixed fractal models.

2.1 Image Features for Fractals and Mixed Fractals

The r-dilation of a surface is called the Minkowski sausage (see the 2-D illustration in Fig. 1-a). When we plot the volume V of the Minkowski sausage to r in a logarithmic diagram (Fig. 1-b), the fractal dimension is defined by $3 - g$, where g is the gradient of the plot[1]. That is, the fractal dimension is

$3 - \text{grad} \log V$,

where grad is a gradient to $\log r$. The fractal dimension of the intensity surface of an fractal image can be used as
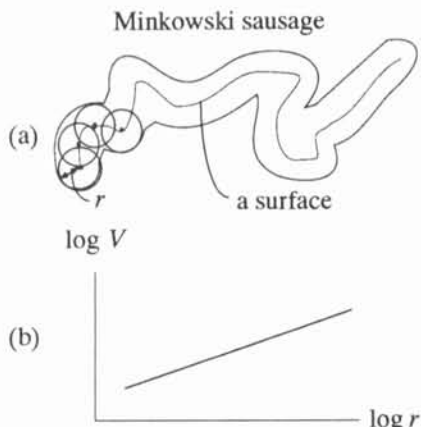
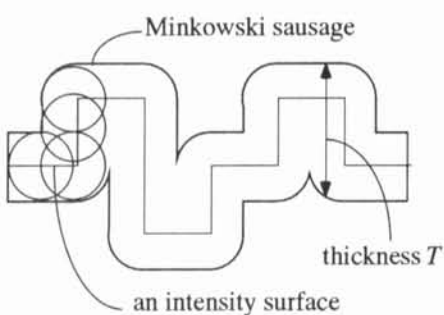Fig. 1. (a) The 2-D illustration of the Minkowski sausage. (b) The log-log plot of its volume V.

Fig. 2. The cross-section of a Minkowski sausage.

an image feature. This feature represents visual 'roughness' of the image[2].

By using a thickness T of the Minkowski sausage (Fig. 2), the fractal dimension of the image is expressed by

$3 - \text{grad} \log \sum T,$

where $\sum$ is summation over the image plane. Or, by using the averaging operation E, it is equivalently expressed by

$$3 - \text{grad} \log E(T). \tag{1}$$

To represent complex natural objects, it might be more appropriate to use the mixed fractal model[3]. To define it precisely, first we introduce the homogeneity of a fractal, then define a mixed fractal using it. Roughly speaking, a homogeneous fractal image has the same fractal dimension everywhere in the image. More strictly, it is defined as follows.

Definition 1[3]: A fractal image is said to be homogeneous, if any subimages of the image are fractals that have the same common fractal dimension.

A mixed fractal is simply a mixture of different homogeneous fractals as follows.

Definition 2[3]: An image that consists of multiple homogeneous fractal images is called a mixed fractal image.

The use of fractal dimension as an image feature for a mixed fractal image causes the problem of missing information as explained below. To make our arguments easier, we only treat mixed fractal images with only two homogeneous regions, say R_1 and R_2 (Fig. 3). The log-log
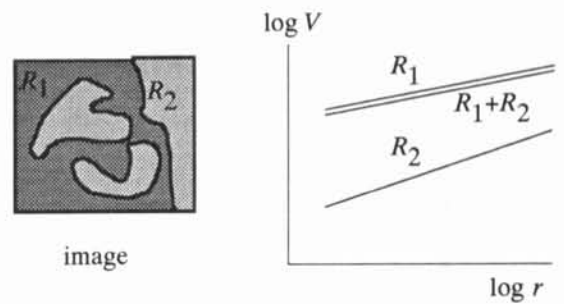

Fig. 3. A log-log plot of a mixed fractal.

plot of a mixed fractal image indicated by R_1+R_2 in Fig. 3 is mostly determined by the plot for one of the two regions, say R_1 in this case, because of the logarithmic nature of the definition of the fractal dimension. Therefore, the information about one of the two regions, say R_2 in this case, is lost in the fractal dimension.

2.2 The Average Fractal Dimension

To solve this problem, we propose a new image feature called the average fractal dimension. The basic idea of the new feature is the use of the average value instead of the effective use of the single value of the fractal dimensions. The average fractal dimension of an image is defined from the average plot indicated by R_1+R_2 in Fig. 4 of the two plots indicated by R_1 and R_2. That is, we average two plots in the logarithmic domain instead of in the non-logarithmic domain as in defining expression (1) of the fractal dimension.
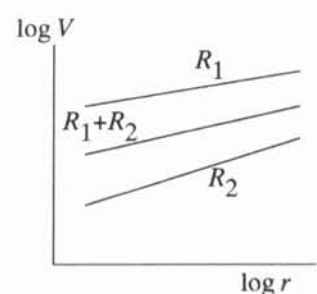

Fig. 4. A log-log plot for the average fractal dimension.

To see it more precisely, we rewrite (1) using the additive property of an averaging operation to

$$3 - \text{grad} \log \{w_1 E_1(T) + w_2 E_2(T)\}, \tag{2}$$

where w_1 and w_2 are the ratio of the area of R_1 and R_2 to the total image area, and E_1 and E_2 are averaging over regions R_1 and R_2, respectively. We can confirm by (2) that averaging in the form of a linear combination is really performed in the non-logarithmic domain.

In order to average in the logarithmic domain, we use the linear combination of logs instead of the log of a linear combination as in the following definition.

Definition 3: The average fractal dimension for a mixed fractal image that consists of two regions is defined by

$$3 - \text{grad} \{w_1 \log E_1(T) + w_2 \log E_2(T)\}. \tag{3}$$

Although it seems from this definition that we need to

partition the image into two regions to calculate the average fractal dimension, there is indeed an alternative way. That is, it can be calculated by the easier expression,

$$3 - \mathrm{grad}\, E(\log T). \tag{4}$$

We explain below why this calculation method works. By using the additive property of averaging, (4) can be rewritten to

$$3 - \mathrm{grad}\, \{w_1 E_1(\log T) + w_2 E_2(\log T)\}. \tag{5}$$

By comparing (3) with (5), we know that their equality or validity of the calculation method is established if we can prove

$$\mathrm{grad}\, E_i(\log T) = \mathrm{grad}\, \log E_i(T) \ (i = 1, 2). \tag{6}$$

As you see in the following lemma, the homogeneous property of the regions R_1 and R_2 derives (6).

Lemma 1: For any homogeneous fractal image,

$$\mathrm{grad}\, E(\log T) = \mathrm{grad}\, \log E(T). \tag{7}$$

Proof: Omitted.

Since (4) doesn't include any regional operations, e.g. E_1 and E_2, it is easy to extend the arguments to the general mixed fractals. But, we also omit the details for shortage of the space.

2.3 The Variants of the Average Fractal Dimension

Since the log-log plot of fractals in nature is linear in some limited regions of abscissa, Peleg et al. measured the fractal dimension in multiple narrow regions of abscissa[4] (Fig. 5). The tuple consisting of the fractal dimensions in each region of abscissa is called the fractal signature. In addition to this ordinary fractal signature, there are two variants, the upper and lower fractal signatures[4]. They are defined using the upper and lower parts of the Minkowski sausage cut by an intensity surface. In the same way of deriving the fractal signature from the fractal dimension, we introduce the average fractal signature from the average fractal dimension. We also introduce the upper and lower average fractal signatures in the same way.
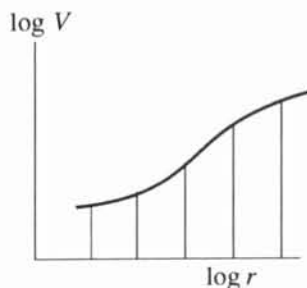

Fig. 5. An illustration of the fractal signature.

3. Evaluation

The classification ability of image features plays an important role in image retrieval systems. We evaluate the effectiveness of the proposed features comparing with conventional image features by classification experiments of real photographs. There have been several proposals of image retrieval methods[5, 7] that share the two common

strategies below.
1) Use of image features instead of direct analysis of images at retrieval time.
2) Partitioning of an image into small square images.
We adopted these strategies in the experiments.

The main purpose of the experiments is the evaluation of the maximum performance of the features. To achieve this purpose, we chose sophisticated methods for statistical analysis e.g. the k nearest neighbor method and the leave-one-out method[8]. We used 2,000 photograph images in a commercial CD-ROM[†]. The difficulty of the experiments using real images lies in the evaluation phase. Since most retrieval methods have used small square images, evaluation should be done to these images. However, it is difficult to decide which type a square image belongs to, because many square images lies in boundaries between different objects. On the other hand, it is easier to decide whether an image as a whole contains an object or not. Our solution to the above problem is that we perform the automatic classification to individual small square images in a photograph and then we check whether the photograph really contains the object.

We used three image features described below in the experiments.
1) The feature based on cooccurence matrix[9], which is well known in texture analysis: 60 floating point numbers.
2) The fractal signature: the ordinary, upper and lower fractal signatures calculated by variation method[10] at 5 points ($r = 2, 4, 8, 16, 32$) to provide 15 floating point numbers in total.
3) The proposed average fractal signature: 15 floating point numbers are provided in the same way as in 2).

We partitioned each image in the CD-ROM into square images of 64×64 pixels. Then we manually classified all the classifiable square images into one of the four object types to form training data. The numbers of classified training images of grass, trees, clouds, and waves are 954, 1,283, 4,512, and 2,873, respectively. We also used 4,935 images unclassified. We manually check whether each photograph contains each object type. The numbers of photographs containing each type are 135, 222, 361, and 264, respectively.

We performed automatic classification to all the square images in the CD-ROM using the training data. Then, we calculated likelihood a photograph contains each object. More specifically, we regard the maximum likelihood over all square images contained in a photograph as likelihood of that photograph.

Fig. 6 shows the top 10 photograph images judged by automatic classification using the average fractal signature to have each of the four object types. Fig. 7. shows the recall-precision relationship graph for each object type. This graph shows the trade-off relationship between the recall ratio and the precision ratio.

It is well known that fractal models are suitable to represent trees and clouds[1], and a cooccurence matrix

---

[†] PHOTOGRAPH 1996 by Photo Library Myojyo.

model is suitable to represent waves[9]. However, our experiments show the conventional image features based on these models do not have enough ability to classify grass and trees, but the average fractal signature have clear effects on classifying them.

4. Conclusions

In this paper we discussed image features for image retrieval. Mixed fractal models are suitable tools for representing natural objects. However, application of conventional fractal-based image features to mixed fractal images causes a problem. To solve the problem, we proposed new image features, and then confirmed their effectiveness. Good classification results for grass and trees were obtained by the experiments performed in the context of image retrieval.

References

[1] B. B. Mandelbrot, "The fractal geometry of nature", Freeman and Company, 1982.

[2] A. P. Pentland, "Fractal-based description of natural scenes", IEEE Transactions of Pattern Analysis and Machine Intelligence, Vol. PAMI-6, No. 6, pp. 661-674, Nov. 1984.

[3] J. C. Russ, "Fractal surfaces", Plenum Press, 1994.

[4] S. Peleg, J. Naor, R. Hartley, D. Avnir, "Multiple resolution texture analysis and classification", IEEE Transactions of Pattern Analysis and Machine Intelligence, Vol. PAMI-6, No. 4, pp. 518-523, July. 1984.

[5] R. W. Picard, T. P. Minka, "Vision texture for annotation", Multimedia Systems, Vol. 3, pp. 3-14, 1995.

[6] W. Niblack, R. Barber, W. Equitz, M. Flickner, E. Glasman, D. Petkovic, P. Yanker, C. Faloutsos, G. Taubin, "The QBIC project: querying images by content using color, texture, and shape", SPIE, Vol. 1908, pp. 173-187, 1993.

[7] J. R. Smith, S. F. Chang, "Quad-tree segmentation for texture-based image query", Proc. of ACM Multimedia 94, pp. 279-286, 1994.

[8] K. Fukunaga, "Introduction to statistical pattern recognition", 2nd ed., Academic Press, 1990.

[9] J. S. Weszka, C. R. Dyer, A. Rosenfeld, "A comparative study of texture measures for terrain classification", IEEE System, Man, and Cybernetics, Vol. 6, No. 4, pp. 269-285, Apr. 1976.

[10] B. Dubuc, S. W. Zucker, C. Tricot, J. F. Quiniou, D. Wehbi, "Evaluating the fractal dimension of surfaces", Proc. R. Soc. Lond. Vol. A-425, pp. 113-127, 1989.
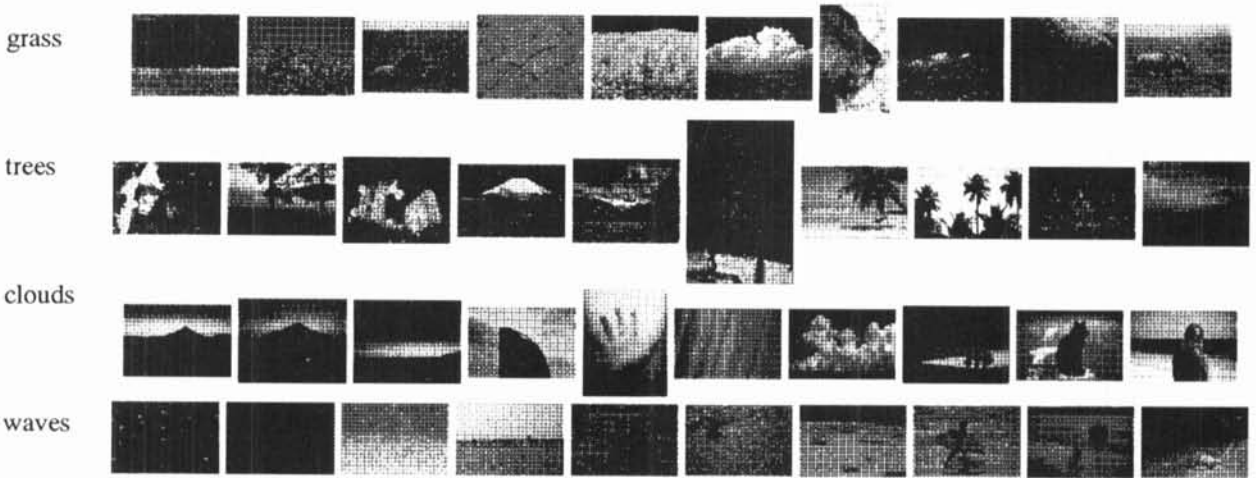
Fig. 6. The top 10 images judged by automatic classification using average fractal signature to have one of the four object types.
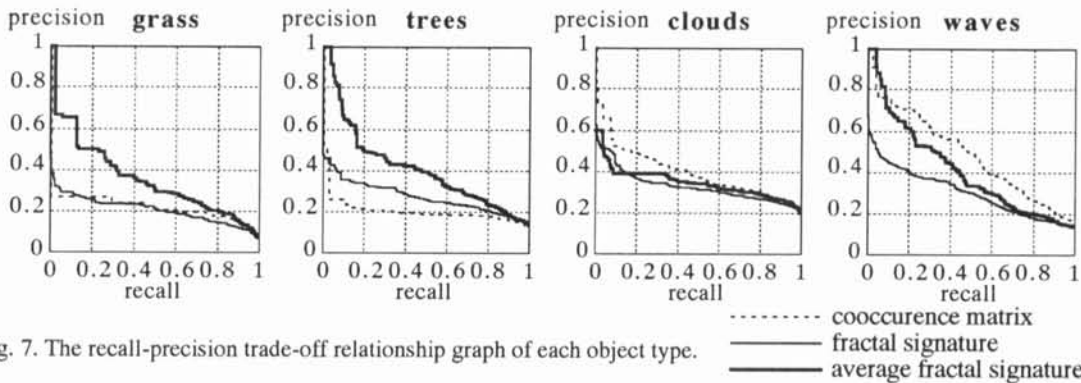

Fig. 7. The recall-precision trade-off relationship graph of each object type.

-------- cooccurence matrix
———— fractal signature
▬▬▬ average fractal signature