J. Edu. & Sci., Vol. (22), No. (2) 2009

Recognizing Text Objects within Images By Uses Fractal Dimension

Hanaa F. Mahmood
Computer Science Dep.
University of Mosul

Eklass A. Sultan
Computer Center
University of Mosul

Omar A. Najim
Education College
University of Mosul

Received ١٧ / ٠٦ / 200٨

Accepted ٢٢ / 0٧ / 200٨

الخلاصة:

يقدم البحث خوارزمية جديدة لتمييز المقاطع النصية الصورية باستخدام الهندسة الكسورية . تم استخدام البعد الكسوري كصفة اساسية في تمييز المقاطع النصية الصورية . استخدمت طريقة عد الصناديق لحساب البعد الكسوري لمكونات الصورة . ولايجاد قيمه حد عتبة للمقاطع النصية الصورية استخدمت نماذج صورية نصية ذات تدرجات رمادية حيث تم حساب البعد الكسوري لكل نقطة صورية ومن ثم ايجاد المعدل لكافة نقاط الصورة. اعتمد على قيمة العتبة في تمييز المقاطع واسترجاعها.

طبقت الخوارزمية على ٧٥ نموذج صوري ، استخدم ٢٥ نموذج لغرض تدريب الخوارزمية المقترحة حيث يتم تحديد حد العتبة خلال هذه المرحلة؛ بينما استخدم ٥٠ نموذج لاختبار الخوارزمية. اظهرت الخوارزمية المقترحة قدرة عالية على تمييز المقاطع النصية ، حيث أعطت نسبة نجاح ٩١.٥% والتي تعتبر اداء جيدا .يمكن الاستفادة من هذه الخوارزمية في انظمة تمييز الوثائق النصية.

Abstract

This paper proposes a new algorithm for recognizing text object images by using fractal geometry. The fractal dimension was used as a main feature for recognizing text objects within images. Box-counting method was used to estimate the fractal dimension for image contents. In order to determine a threshold value for the textual objects within image, the fractal dimension was computed for a number of gray scale textual images. The fractal number of each pixel was calculated, then the mean value of all these fractal values were computed. The threshold value was used in recognizing and retrieving the textual objects within image.

This algorithm was applied on 75 image samples, 25 image samples were used in training phase, the threshold value was determined throughout this phase; whereas 50 image samples were used in testing the algorithm. The proposed algorithm has performed extremely well with recognition rates 91.5% which is considered good performance. It is a promising technique for optical character recognition system.

1. Introduction

Zhang and Chang proposed a system to detect and extract the overlay texts in digital video, the system used a multiple hypothesis testing approach: The region-of-interests (ROI) probably containing the overlay texts are decomposed into several hypothetical binary images using color space partitioning; A grouping algorithm then is conducted to group the identified character blocks into text lines in each binary image; If the layout of the grouped text lines conforms to the verification rules, the bounding boxes of these grouped blocks are output as the detected text regions[1].

Wang and Srihari used projection profiles of the pixel values to identify large “text” blocks by detecting valleys in these profiles. Wahl *et al* used constrained run lengths and connected component analysis to detect blocks of text. Fletcher and Kasturi used the fact that most text blocks lie in a straight line, and utilized Hough transform techniques to detect collinear elements. Taxt *et al* view the identification of print in document images as a two-category classification problem, where the categories are print and background. They use various classification methods to compute the segmentation including contextual classification and relaxation algorithms[2].

Many algorithms developed for document image content extraction, which find regions containing machine-printed text, handwriting, photographs, etc within images of documents, these algorithms cope with a rich diversity of document, image, and content types, vast and rapidly growing scale of document image collections has been compellingly documented. Information extraction and retrieval from document images is an increasingly important at the interface between document image analysis and information retrieval [3].

Individual pixels are classified, not regions, in order to avoid the arbitrariness and restrictiveness of limited families of region shapes, this policy has yielded, to date, modest per-pixel classification accuracies which already support usefully high recall and precision rates for queries on collections of documents. This exibility has another advantage: it allows greater accuracy in inventory statistics, by which we mean summaries of each page estimating, for each content class, the fraction of page area dominated by that class [4], in our approach we use the fractal geometry for extraction features.

Many structures in nature exhibit fractal properties. From trees, to frosted glass, to island coastlines, we are surrounded by shapes that upon closer inspection reveal more and more detail and yet retain the same general appearance no matter on what scale they are observed [5].

For example, a square contains no more detail than is readily apparent at a single glance. Yet the coastline of an island will display seemingly endlessly increasing detail when viewed from space, from a plane, from the ground or even portions of it through a magnifying glass. It is self-evident that increases in detail will translate into increased estimates of the length of such a fractal shape[5]. The fractal analysis of the large scale structure has been under consideration in the last twenty years in order to describe the distribution of galaxies. We consider, after mandelbrot, that it is fractal an object which [6]:

- has a very fragmented or irregular shape.
- presents self-similarity or self-affinity.
- presents scale-invariance.

Our interest in fractal dimension is primarily due to its ability to segment images into different textural regions. Textural analysis of subregions is important because it is connected to current models of the Human Visual System (*HVS*). Texture perception and, more specifically, texture roughness is a key cue feature used in recognition of objects. We began investigating the use of fractal dimension in segmenting images.[7]

2. Fractal Geometry and Fractal Dimension

In 1981, mathematician John Hutchinson used the theory of iterated function system to model collections of contractive transformations in a metric space as dynamical systems, which later provides the theoretical support of recognizing fractals in metric space. It was Michael Barnsley, eventually, who generated the fractal model using Iterated Function Systems (*IFS*),[8]

The fractal dimension is most commonly defined as the Hausdorff-Besicovitch (*HB*) dimension, *Dh(A)*, where A denotes the image/signal. In general, the *HB* dimension of *A* is defined in the following manner [7]

Let

$$R^n=\{X|X=(x1,\ldots\ldots\ldots.xn)\ xi\in R\} \qquad \ldots\ldots\ldots\ldots\ldots\ldots(1)$$

For some natural number n. Then, define the diameter of some cover C as

$$\text{diam}(C)=\sup\{de(x,y)|\ x,y\in C\} \qquad \ldots\ldots\ldots\ldots\ldots\ldots..(2)$$

where *de(x, y)* denotes the Euclidean distance function. An open cover of A is defined by covers *Ci* such that

$$A\subset\bigcup_{i=1}^{\infty}C_i \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(3)$$

Then let

$$h_{\varepsilon}^{s}(A)=\inf\{\sum_{i=0}^{\infty} diam(C_i)^s \,|_{ofAwithdiam(C_i)\le\varepsilon}^{\{C1,C2\},opencover}\} \quad \text{……………………..(4)}$$

So that the s-dimensional Hausdroff measure of *A* is

$$h^{s}(A)=\lim_{\varepsilon\to 0} h_{\varepsilon}^{s}(A) \quad \text{……………………..(5)}$$

Then, the Hausdorff-Besicovitch (*HB*) fractal dimension of *A* is defined as

$$D_h(A)=\inf\{S \mid h^{s}(A)=0\}=\sup\{S \mid h^{s}(A)=\infty\} \quad \text{……….(6)}$$

Due to the complexity and impracticality of finding the optimal cover defined by the *HB* dimension, it need a different bounding estimate, one that is easily calculated. One such upper bound estimation of the HB dimension is the box dimension definition[8].

A fractal is a set that has a non integer fractal dimension. The fractal dimension is formally defined as the (*HB*) dimension. However, there are a number of ways to estimate this fractal dimension, each of which uses a slightly different definition of the dimension [7].

3. Fractals and Texture Features

Texture analysis refers to a class of mathematical procedures and models that characterize the spatial variations within imagery as a means of extracting information.[9]

Image texture, defined as a function of the spatial variation in pixel intensities (gray values). One immediate application of image texture is the recognition of image regions using texture properties. Chen used fractal texture features to classify ultrasound images of livers, and used the fractal texture features to do edge enhancement in chest X-rays.[2]

Texture is represented as an index at each pixel, being the local fractal dimension within an window (11 × 11) pixel estimated according to the fractal Brownian motion model proposed by Chen. The texture feature is used in addition to a number of other traditional features, including the response to a Kirsch edge operator, the gray level, and the result of temporal operations [10].

4. Estimate the Fractal Dimension for Text Images

The box dimension, *Db(A)*, which is normally estimated using the box counting algorithm, is a commonly used upper-bound estimator for the *HB* dimension. In general, the box dimension can be defined as follows:

Let *Cd(A)* be the smallest number of closed covering elements of size *d*, see (2), that cover the set *A*. Then, *Db(A)* is defined as: [11]

$$D_b(A) = \lim_{d=0} \frac{\log C_d(A)}{\log \frac{1}{d}} \qquad \text{................(7)}$$

The calculation of the box dimension is often difficult. One must first find the optimal covering (smallest number of box shaped covering elements) of *A* for a given set of box covering elements with edges of length *d*. Placing the covering elements in an algorithmic way to minimize their total number for a given size is easily achieved for low-dimensional data sets. However, this process becomes much less tractable in the case of more general forms of data. This also requires algorithms to determine when and how to overlap covering elements when a whole number of elements does not “evenly” fit the space[12].

4.1 The Proposed Box Counting Dimension Algorithm

The box counting dimension *d(A)* is normally estimated by dividing a 1 dimensional object into N identical parts scaled from the ratio:

$r = 1/N$, $r = N^{(-1/1)} = 1/N$

For a 2 dimensional object see Figure (1):

$r = N^{(-1/2)} = \frac{1}{\sqrt{N}}$ and finally for a 3 dimensional object, $r = \frac{1}{\sqrt[3]{N}}$. In general for an object of a given dimension *D*;

$$r = N(r)^{(-1/D)} \qquad \text{........................(8)}$$

Figure (1): Sample 1 divided into 4 blokes

The dimension is then found by rearranging the above equation to the following form:

$$D_{box} = \frac{\log[AVR(\frac{sum(N)}{N}]}{\log(1/r)} \qquad \text{...................(9)}$$

When *D* has an integer value, the object in equation has Euclidean geometry. If, however, *D* takes a non integer value, the object in equation(4) is deemed to have fractal dimensions.

A square mesh of various sizes s is laid over the image (containing the object). The number of mesh boxes *N(r)* that contain part of the image are counted. The fractal (box) dimension *D* is given by the slope of the linear portion of a *log(AVR(Sum(N)/N)* via *log(1/r)* graph as shown Figure (2).

Because there is no preferred origin for the boxes with respect to the pixels in the image, multiple measures *N(r)* can be computed for different mesh origins. The graphed value of *N(r)* is usually the average of *N(r)* from the different mesh origins.

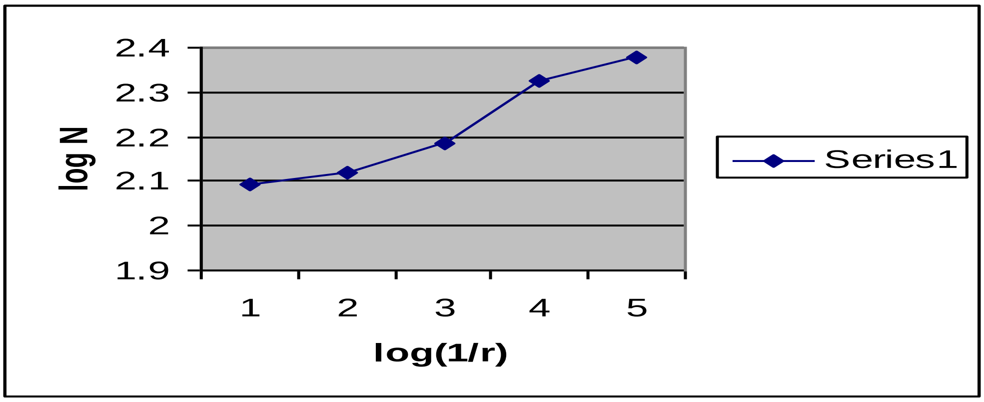


Figure (2): The slo for fractal dimension of sample 1

An important user option is the range of box sizes that are used. Some observations that can be used as a guide to box sizes: box sizes close to the object resolution are the most important. So for a detailed image the box sizes from 1,2,4... are important. For a crude or noisy image it may be appropriate to start the box sizes around [4, 6, 8 . . .16].

Figure (3): The sample 1 with 16 blokes

5. The Fractal Calculation Algorithm and Texture Segmentation

Text objects recognition pass through several stage, figure(4). This approach was applied on images of different sizes that contain an assorted objects, the image dimensions in figure (5) was [320×320] pixel.

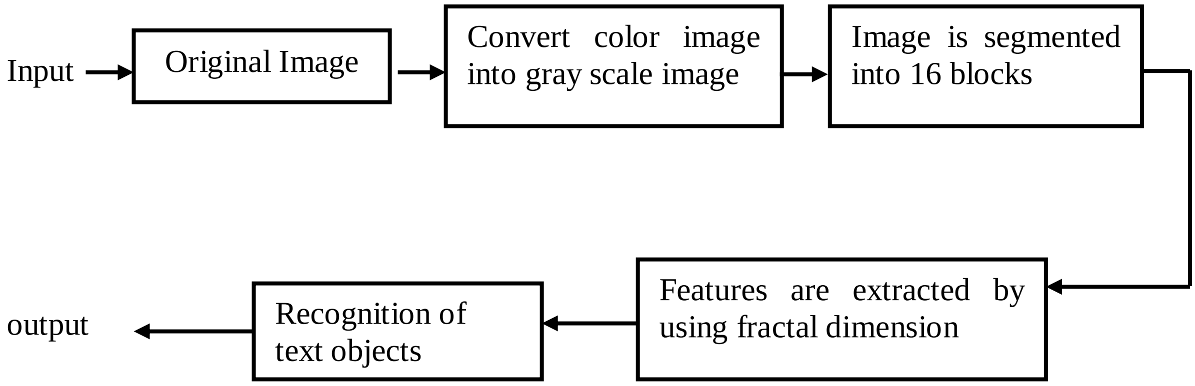


Figure (4): Text objects recognizing

For evaluating the fractal dimension for each pixel, the dimension of the image was change neither to accept the division on four. The size of the segmentation scalar was determined [2-16].

To calculate the fractal for each pixel in the image we made pixels the center of the box (the dimension of the box 2×2) and calculate the fractal using the equation(1). According to the box counting method the size of box was changed to [4×4] and repeat the operation. After applying the algorithm on sample (1) the fractal value saved in a matrices as shown in Figure (5).

Figure (5): Real image and fractal image for sample 1 with size [250×250] pixel

After calculates the fractal for each pixel in the image the next step is divide the fractal image into segments depending on F value for each pixel and save the result in a three dimensions matrix containing the slides and values of F as show in Figure (6) .

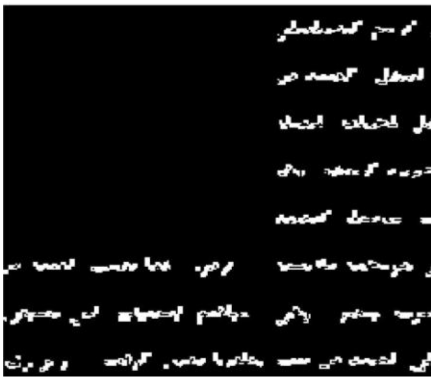


Figure (6): Sample of image slide

6. Threshold Value

In the segmentation scheme a region is labeled as being representative of an anomaly if the transformed domain with the lowest root mean square (*rms*) error value for that range is above the designated similarity threshold; cells with an *rms* error below the similarity threshold are labeled as being normal or expected. If the similarity threshold is too high, everything in the image is mapped and no segmentation of anomalous regions occurs; similarly, if the value is too low, nothing in the image is mapped and the entire image is seen as being representative of the anomaly. A similarity threshold is selected near the point at which partitions display a dramatic change in *rms* error level[13]. After applying the algorithm on the (25) image samples a threshold value of approximately 0.365 was obtained for recognize the text objects image. The finale step was retrieving the text objects image depending on the pixels positions in the slide shown in figure (6) and retrieve the pixel real value from the origin image as shown in Figure (7).

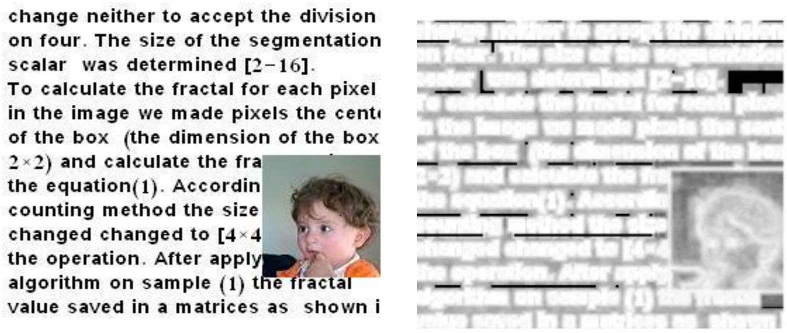
Figure (7): The retrieved image for sample 1

The algorithm has been applied on sample (2,3) as shown in Figure (8,9) the images size were [350×350] and [270×270] pixel and the text objects were in both Arabic and English font which does not effect on segmentation operation.

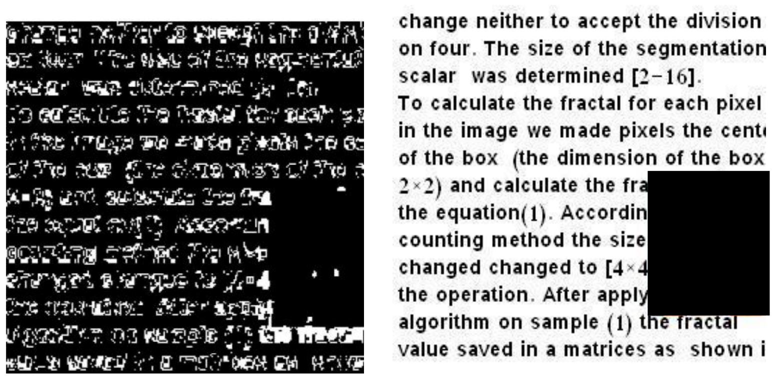
a **b**

c **d**

Figure (8) **a: The original image** **b: The image fractal**
c: Segmented part of the image **d: Retrieved Image**



a **b**



c **d**

Figure (8) **a: The original image** **b: The image fractal**
c: Segmented part of the image **d: Retrieved Image**

7. Conclusion and Future work :

This algorithm is very useful when it uses in image segmentation filed. The algorithm applied on a 75 image samples 25 image samples used for training and 50 image samples are used for testing, the text objects recognized ratio was 91.5%.

The proposed algorithm offers a new approach for the geometric characterization. With the advantage that there are no necessicity for any previous assumptions on the large scale homogeneity. It is provide a good result as the methods mention in section 1.

Segmenting a textual object within video film and textual object within color images are suggested as a future work.

References:

1. D. Zhang, S. Chang, 2002, “Accurate Overlay Text Extraction for Digital Video Analysis”, NY 10027.

2. M. Tuceryan, 1998, “Texture analysis” The Handbook of Pattern Recognition and Computer Vision (2nd Edition), pp. 207-248.

3. Y. Ishitani, 2001, “Model-based information extraction method tolerant of ocr errors for document images”, icdr 00, p. 908.

4. Casey M. R., 2006, “Fast approximate nearest neighbors”, M.S. thesis, Lehigh University, Bethjahem, Pennsylvania.

5. J. G. Zhou, Z. He and J. Guo, 1999, “Fractal growth modeling of electrochemical deposition in solid freeform fabrication", IEEE Transaction on System, Man and Cybernetics, vol.6, no.3, pp.119-125.

6. R. Murdzek, 2007, "The box-counting method in the large scale structure of the universe", Rom. Journ. Phys., Vol.52, No.1–2, P.149–154.

7. S. Weszka, C. R. Dyer and A. Rosenfeld, 1976, "A comparative study of texture measure for terrain classification", IEEE Transaction on System, Man and Cybernetics, vol.6, no.4, pp.269-285.

8. X. Henry., 2005., “Fractal audio coding”, Master's thesis, Queen's University, ON Canada

9. M. Tuceryan, J. K. Jain, 1990, “Texture segmentation using voronoi polygons”, IEEE Transaction on Pattern Analysis and machine intelligence 12, 211-216.

10. A. Pentland, 1984, "Fractal based description of natural scene", IEEE Transaction on Pattern Analysis and Machine Intelligence 6, 6, 661-674.

11. P. Blaudeck, S. Seeger, C. Schulzky, K. H. Hoffmann, T. Dutta and S. Tarafdar, 2006, "The coastline and lake shores on a fractal island", J. Phys. A: Math. Gen. 39:1609-1618.

12. T. Ojala, M. Pietikyinen and D. Harwood, 1996, "A comparative study of texture measure with classification based on feature distributive", Pattern Recognition 29:51-59.

13. H. Sari-Sarraf, S. S. Gleason K. T. Hudson, K. F. Hubner, 1996, “A novel approach to computer-aided diagnosis of mammographic images,” IEEE Workshop on Applications of Computer Vision, December.