Image Digest II: A region sensitive Image Digest generation algorithm

[1]Sesha Kalyur, [2]G.S.Nagaraja

Research Scholar CSE Dept. R.V.C.E Bangalore, Professor CSE Dept. R.V.C.E Bangalore
Email: [1]sesha_kalyur@hotmail.com, [2]nagarajags@rvce.edu.in

Abstract— Automatic Image annotation and retrieval is an interesting area of research in the broader field of Image analysis and Image processing. A digest generated using the pixel properties of an image has interesting applications of its own. In this paper we present a methodology to generate a digest for a given image capturing the region information. The algorithm works on grayscale VGA images. So we convert color images to their corresponding monochrome counterparts. Images of resolution higher than VGA are converted to their VGA equivalents. A digital image can be segmented in several different ways prior to analysis. Here the given image is divided into equal sized segments also referred to as grids or blocks. A fraction of the pixels are sampled from each grid to capture the grayness information. A number or a digest corresponding to the distribution of grayness values is generated for each grid. The digest for the entire image is a composition of the digests computed for each grid. The digest generated for visually related images generate digests that are close to one another numerically. These digests can be used as keys to store and retrieve their corresponding images.

Index Terms—Image Digest, Image Segmentation, Image Annotation, Automatic Image Annotation, Image Retrieval, Image Analysis, Image Processing

I. INTRODUCTION

Historically images have been annotated and tagged manually which is a time consuming and tedious task. In contrast, Unsupervised or Automatic image annotation relieves the bur-den of the user. Automatic image annotation and retrieval is an interesting research topic. This field has its roots in computer vision and has dependence on one more of the following related areas: Image segmentation, Pattern Recognition and Content Based Image Retrieval.

Digest Generation for a given block of message is an activity that has a history of its own. It has found several interesting applications in the field of Computer Cryptography. Specifically it has been used to generate Digital Signatures, Message Authentication Codes and so on, mainly to verify the integrity of the message.

This work borrows the idea of digest generation from the Cryptography field and solves the annotation problem of digital images. Our algorithm treats the pixel intensity values as a distribution and generates a digest for the given image. For a given image of a particular distribution of intensities, the digest generated is unique. This property can be used to match a digest with a given image.

We call this algorithm Image Digest II since this is a follow up to an earlier algorithm called Image Digest. The predecessor of this algorithm treated the entire image space as one locality or region in calculating the digest. The advantage with this technique is the speed and the space efficiency that results. However there is an error introduced with that approach. Two different, unrelated images having the same distribution of intensities can generate digests that are close numerically.

Image Digest II algorithm, attempts to improve on this by considering the region or locality information in calculating the digest. This improves the situation by mapping a unique digest to an image with only a marginal overhead in computation but no extra requirements in terms of storage.

The generated digest has several interesting applications in the image annotation and retrieval field. We can use the digest as a key to store and retrieve images from image catalogs or databases. We can also use the generated digests from two images to compare and classify them.

Since segmentation is an important step of the solution in any image processing problem we have devoted the next section to discuss the state of the art in the image segmentation domain.

II. IMAGE SEGMENTATION

Image segmentation can be defined as a process of discovering homogeneous areas of an image. Homogeneity is determined with respect to certain properties of the image such as grayness, texture and color. It is an essential first step in image analysis and pattern recognition and also happens to be one of the most difficult aspects of it. Image segmentation has been an active field of research and as of date numerous segmentation methodologies or techniques have been proposed. In an effort to understand the various techniques, researchers have tried to classify the various segmentation methods available.

Segmentation methodologies can be classified at the highest level as similarity and discontinuity methods. The properties of similarity and discontinuity are based on the pixel value differentials with their local neighborhood pixels. Methods based on the similarity property are called region-based methods and methods based on the discontinuity property are called boundary-based methods.

Traditionally, image segmentation techniques or algorithms have been classified into the following three classes: Thresholding and Clustering, Edge Detection and Region Extraction. Thresholding techniques are based on the gray level histogram and structural characteristics of the image. Clustering techniques are an extension of Thresholding techniques. They use multiple features of the image such as the mean grayness level and texture together. Edge detection algorithms are based on detecting edges or boundaries in an image. An edge could be defined as a sudden or abrupt change in the grayness value. Region based segmentation methods group pixels of similar characteristics into a region.

The more recent texts have classified the segmentation techniques into the following three classes: Edge based, Region based and Cooperative methods. Edge based methods are based on finding edges in the image. Edge methods are further classified as Parallel edge detection methods and Sequential edge detection methods. In Parallel edge detection, the decision of including a set of points in an edge is independent of the other points that lie on the edge or not. On the other hand in Sequential edge detection, the decision of including a point depends on the previously examined points.

Region based techniques group a set of pixels with similar characteristics into a region. Region based methods are further classified into Region Growing, Region Splitting and Split and Merge techniques. The Region Growing methods start with a region, which includes a single pixel called the seed and expands it by adding pixels in the neighborhood with similar characteristics. The Region Splitting methods operate in the other direction. They start with a region which includes the whole image and apply similarity and discontinuity criteria to recursively split the region into several homogeneous regions. The Split and Merge technique applies both the region growing and region splitting operations based on homogeneity criteria. That is the technique works in several recursive steps, and in each step, regions are split or merged based on the prevailing region characteristics.

Unfortunately Edge based and Region based methods alone do not produce satisfactory results on their own. Cooperative Methods combine the effectiveness of both Edge and Region methods to achieve good results. Embedded Integration is a popular cooperative method that starts with an edge as a seed and grows regions from it. Another popular cooperative method is Post-processing Integration where an image is analyzed independently using Edge and Region methods and the results are integrated in a post-processing step.

Our Image Segmentation technique can be classified as a Thresholding technique that uses pixel grayness values in capturing the image characteristics. Details of the methodology are given in section IV.

III. RELATED WORK

References [1][2][3][4][5][6] is a good collection of papers on the topic of image segmentation. Reference [1] presents a segmentation methodology which starts with a region of the image and recursively splits the region using histogram of the feature values in the region. Reference [2] presents a split-and-merge segmentation technique. Reference [3] presents Picture Segmentation by a Tree Traversal Algorithm. Reference [4] presents a segmentation algorithm that is based on the concept of the lattice of connective criteria. Reference [5] presents a segmentation algorithm which builds segments through random walks from unlabelled pixels to pixels with predefined labels. Reference [6] presents a segmentation algorithm based on graph partitioning. Reference [7] is a good treatise on the subject of image composition. References [8][9][10][11][12] are excellent write-ups on the topic of image captioning and annotation. Reference [8] presents a real-time image annotation method based on the generative modeling approach. Reference [9] presents a image annotation approach based on a statistical modeling approach. Reference [10] discusses a method to discover correlations between image features and keywords. Reference [11] Presents a two layer learning system that is image content sensitive. Reference [12] presents a probabilistic model of image annotation that supports image classification and annotation together. Reference [13] provides a good introduction to the topic of variable and feature selection. Reference [14] discusses significant challenges involved in the adaptation of existing image retrieval techniques. Reference [15] presents a survey that reviews recent articles on content-based multimedia information retrieval. Reference [16] is an earlier work by the authors of this paper on the topic of image digest. References [17][18][19][20] provide good coverage of the message digest algorithms md2, md4, md5 and md6. Reference [17] presents a technique to generate a 128 bit fingerprint for a message of arbitrary length targeted for digital signature applications. Reference [18] presents a message digest algorithm that is a follow up to md2 that is expected to perform better on 32 bit machines. Reference [19] presents the MD5 algorithm which does not require any large substitution tables and the algorithm can be coded quite compactly. Reference [20] presents a message digest algorithm for 64 bit machines that is simple and flexible and generates digest of size up to 512 bytes.

IV. IMAGE DIGEST II

In this section, we elaborate on our image digest generation technique. The given monochrome image of VGA resolution is first partitioned in to grids, blocks or segments of equal size. For instance we could choose to have 32 grids. For efficiency purposes we could sample only 50% of the total pixels in the image. Considering the dimension of 640X480 for VGA we have total of 307200 pixels. Sampling only 50% of the pixels we have 153600 pixels which works out to 4800 pixels for each grid. These pixels form the distribution from which the grayness component for the grid is computed.

We need 8 bits to capture the grayness component (256 levels) for a grid. This component is the sum of individual weighted grayness value for each grayness level. The grayness level for each pixel ranges from 0 to 255. Suppose block 1 has the following distribution of grayness levels: 20% of the pixels have a value of 66, 10% of the pixels have a value of 99 and 70% of the pixels have a value 127. The grayness value or the digest component for the block is computed as follows,

$$GV = 0.2 * 66 + 0.1 * 99 + 0.7 * 127 = 112$$

This value takes the field for block 1 in the digest. Since we have 32 blocks and each block takes up 8 bits we can represent the digest for the image in a 128 bit number. This is available on most modern machine architectures and programming languages as a data type of double.

32	24	14	15	16	5	17	25
31	23	13	4	1	6	18	26
30	22	12	3	2	7	19	27
29	21	11	10	9	8	20	28

Figure-1: An Image divided in to 32 equal sized grids

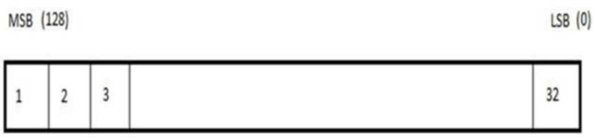

Figure-2: Structure of the Image Digest

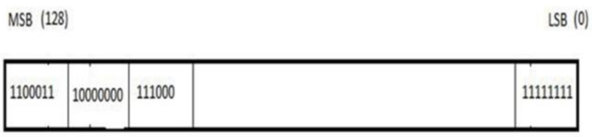

Figure-3: Example of an Image Digest

Figure-1 shows a digital image divided in to 32 grids, blocks or segments. Each grid contributes pixel samples for the calculation of the image digest component for the grid. The digest for the entire image is a composition of the digests for all the individual grids. The central portions of the image are given higher significance when compared to the peripheral portions of the image. This makes sense for most images since the important features of the image are clustered towards the center when compared to the edges. So blocks are numbered starting from the center and radiates in spiral fashion towards the periphery of the image. This arrangement ensures that central blocks contribute to the MSB of the digest and the edges contribute to the LSB. Thus block numbers determine which fields of the image digest a particular block contributes its digest component. This scheme produces closely related digests (numerically) for images that are related visually especially in the central parts of the image.

Figure-2 illustrates the structure of the image digest for a particular grid. It shows bit numbers from 0 to 127. Each field is 8 bits wide and holds the digest for a block. The digest for block 1 occupies the field in the most significant position (MSB) and the digest for block 32 occupies the least significant position (LSB).

Figure-3 illustrates a case where the individual digests for block 1, block 2, block 3 and block 32 are respectively, 99, 128, 56 and 255. The corresponding numbers in binary are: 1100011, 10000000, 111000 and 11111111. The bit values for other fields of the digest are calculated in a similar fashion.

Since the digest generated for the image is a composition of individual digests for grids or blocks of the image, region or locality information is captured. On the other hand, a technique that operates on the image ignoring the locality information would have flagged two different images with the same distribution of grayness values as same. Image Digest II is superior in that sense.

It is to be noted that the choice of 32 segments is driven by the need to keep the digest under 128 bits. If higher accuracies are desired we can choose to divide the image into more segments. The digest generated would proportionally require more bits of storage. For instance if the number of segments chosen is 64, the resulting digest would require 256 bits and so on.

V. CONCLUSION

In this paper we presented a methodology to generate a digest for a given digital image using region or locality information. The method operates on monochrome VGA images. So images of higher resolution are first converted to VGA. Color images are converted to their monochrome equivalents. The given image is divided into equal sized grids, blocks or segments. A fraction of the pixels in each grid is sampled for digest generation. The distribution of grayness values in the grid is used to generate a number or a digest for the grid. The digest for the entire image is composed of individual digests of the grids forming the image.

Segments are numbered starting from the center and they radiate towards the edges of the image. This arrangement gives importance to the pixels in the center of image over the ones that are in the periphery. Digests for similar images are numerically related to one another. This property can be used to solve problems in image annotation and retrieval areas. For instance, the digest can serve as a key to store and search related images in a database of digital images. The digest can also be used for image classification purposes.

This work considered the grayness distribution of an image for calculating the digest. We have consciously avoided considering other image features such as texture and color to keep the algorithm simple. However future follow up work in the area, needs to consider these other image features for calculating the image digest.

ACKNOWLEDGMENT

The authors would like to thank Dr. G. S. Nagaraja and Dr. N. K. Srinath for their support and encouragement.

REFERENCES

[1] R. Ohlander, K. Price, and D. R. Reddy, “Picture segmentation using a recursive region splitting method,” Computer Graphics and Image Processing, vol. 8, no. 3, pp. 313 – 333, 1978. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0146664X78900606

[2] S. L. Horowitz and T. Pavlidis, “Picture Segmentation by a directed split-and-merge procedure,” Proceedings of the 2nd International Joint Conference on Pattern Recognition, Copenhagen, Denmark, pp. 424– 433, 1974.

[3] S. L. Horowitz and P. Theodosios, “Picture segmentation by a tree traversal algorithm,” J. ACM, vol. 23, no. 2, pp. 368–388, Apr. 1976. [Online]. Available: http://doi.acm.org/10.1145/321941.321956

[4] [1]J. Serra, “A lattice approach to image segmentation,” J. Math. Imaging Vis., vol. 24, no. 1, pp. 83–130, Jan. 2006. [Online]. Available: http://dx.doi.org/10.1007/s10851-005-3616-0

[5] [1]L. Grady, “Random walks for image segmentation,” Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 28, no. 11, pp. 1768–1783, 2006.

[6] L. Grady and E. Schwartz, “Isoperimetric graph partitioning for im-age segmentation,” Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 28, no. 3, pp. 469–475, 2006.

[7] J. Koenderink, “The structure of images,” Biological Cybernetics, vol. 50, no. 5, pp. 363–370, 1984. [Online]. Available: http://dx.doi.org/10.1007/BF00336961

[8] [6]J. Li and J. Wang, “Real-time computerized annotation of pictures,” Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 30, no. 6, pp. 985–1002, 2008.

[9] L. Jia and J. Wang, “Automatic linguistic indexing of pictures by a statistical modeling approach,” Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 25, no. 9, pp. 1075–1088, 2003.

[10] J.-Y. Pan, H.-J. Yang, P. Duygulu, and C. Faloutsos, “Automatic image captioning,” in Multimedia and Expo, 2004. ICME ’04. 2004 IEEE International Conference on, vol. 3, 2004, pp. 1987–1990 Vol.3.

[11] E. Akbas and F. Yarman Vural, “Automatic image annotation by ensemble of visual descriptors,” in Computer Vision and Pattern Recognition, 2007. CVPR ’07. IEEE Conference on, 2007, pp. 1–8.

[12] C. Wang, D. Blei, and F. fei Li, “Simultaneous image classification and annotation,” in Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, 2009, pp. 1903–1910.

[13] [6]I. Guyon and A. Elisseeff, “An introduction to variable and feature selection,” J. Mach. Learn. Res., vol. 3, pp. 1157–1182, Mar. 2003. [Online]. Available: http://dl.acm.org/citation.cfm?id=944919.944968

[14] R. Datta, D. Joshi, J. Li, and J. Z. Wang, “Image retrieval the new age,” ACM Ideas, influences, and trends of Comput Surv., vol. 40, no. 2, pp. 5:1–5:60, May 2008. [Online]. Available: http://doi.acm.org/10.1145/1348246.1348248.

[15] M. S. Lew, N. Sebe, C. Djeraba, and R. Jain, “Content-based multimedia information retrieval: State of the art and challenges,” ACM Trans. Multimedia Comput. Commun. Appl., vol. 2, no. 1, pp. 1–19, Feb. 2006. [Online]. Available: http://doi.acm.org/10.1145/1126004.1126005

[16] S. Kalyur and G. Nagaraja, “Image digest: A numbering scheme for classifying digital images,”

in International Conference on Emerging Computation and Information Technologies, 2013.

[17] B. Kaliski, “The md2 message-digest algorithm,” Internet RFC 1319, ISSN 2070-1721, 1992.

[18] R. Rivest, “The md4 message-digest algorithm,” Internet RFC 1320, ISSN 2070-1721, 1992.

[19] R. L. Rivest, “The md5 message-digest algorithm,” Internet RFC 1321, ISSN 2070-1721, 1992.

[20] R. L. Rivest, B. Agre, D. V. Bailey, C. Crutchfield, Y. Dodis, K. Elliott, F. A. Khan, J. Krishnamurthy, Y. Lin, L. Reyzin, E. Shen, J. Sukha, D. Sutherland, E. Tromer, and Y. L. Yin, “The md6 hash function a proposal to nist for sha-3,” 2008.

❖❖❖