International Conference on Information Engineering, Management and Security 2016 [ICIEMS]

ISBN	978-81-929866-4-7		**VOL**	01
Website	iciems.in		**eMail**	iciems@asdf.res.in
Received	02 – February – 2016		**Accepted**	15 - February – 2016
Article ID	ICIEMS002		**eAID**	ICIEMS.2016.002

Beyond Bag of Features: Adaptive Hilbert Scan Based Tree for Image Retrieval

Mo Xu[1], Fuyan Liu[1], Qieshi Zhang[2], Sei-ichiro Kamata[2]
[1]School of Computer Engineer and Science, Shanghai University
[2]Graduate School of Information, Production and Systems, Waseda University

Abstract- *One fundamental problem in large scale image retrieval with the bag-of-features is its lack of spatial information, which affects accuracy of image retrieval. Depending on distribution of local features in an image, we propose a novel adaptive Hilbert-scan strategy which computes weight of each path at increasingly fine resolutions. Owing to merits of this strategy, spatial information of object will be preserved more precisely at Hilbert order. Extensive experiments on Caltech-256 show that our method obtains higher accuracy.*

Keywords : *Hilbert-Scan; image retrieval; bag-of-features; feature representation*

I. INTRODUCTION

In last decade, the bag-of-features (BOF) [1] model has become very popular in image retrieval and object classification because of its simplicity and good performance. However, the BOF model loses sight of spatial order of local descriptors, which severely limits the descriptive ability of image representation. Hence, it is incapable of capturing shapes or locating an object in image. To overcome this drawback, many extensions of BOF model were proposed such as SPM [2], Spatial BOF [3], Spatial Weighting BOF [4] and HS-BOF [5].

In our research, we focus on the Hilbert-Scan based tree (HSBT) [5] approach which can do retrieval quickly yet still loss some spatial information of interest points. We aim to construct a mechanism to select scanning path for each image automatically. Generally, two factors are considered in our proposed method. One is the total number of interest points in two adjacent blocks in an image. Another one focuses on comparing the amount of interest points between these two blocks. In order to combine these two factors effectively, a weighing coefficient is proposed to control the relative significance of them. Furthermore, inspired by the generative method of Hilbert-Scan, a hierarchical strategy is performed from global geometric distribution of interest points from the local geometric distribution of them. Since the mass of interest points are closer in linear sequence after mapping, the appearance of key objects can be captured more quickly. The merging error and the number of layers in HSBT will be reduced.

II. Related Works

A. Hilbert-Scan

A Hilbert curve is a continuous fractal space-filling curve first described by the German mathematician David Hilbert in 1891 [6]. Hilbert space filling curve has the property to preserve the locality between objects of multidimensional space in the linear space. If the distance between two points in the 2-D image is small, the distance between the same pair of points in the 1-D sequence is also small in

Cite this article as: Mo Xu, Fuyan Liu, Qieshi Zhang, Sei-ichiro Kamata. "Beyond Bag of Features: Adaptive Hilbert Scan Based Tree for Image Retrieval". *International Conference on Information Engineering, Management and Security 2016*: 08-12. Print.

most cases. In the application of data analysis, it is used for scanning data in two dimensional spaces. This scanning way is called Hilbert-scan. Original Hilbert-Scan requires square-sized image. To solve this problem, Reference [7] proposed Pseudo Hilbert-Scan which can be applied for arbitrarily-sized image. Fig. 1 and Fig. 2 show the Hilbert-Scan and Pseudo Hilbert-Scan.

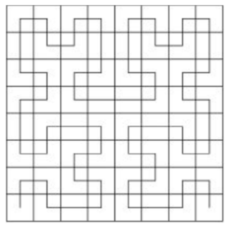


Figure 1. 8 × 8 Hilbert curve in 2-D space

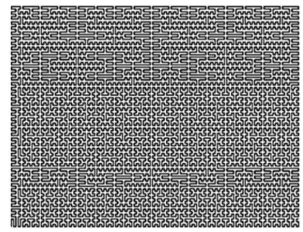


Figure 2. Pseudo Hilbert-Scan for arbitrarily-sized rectangle

B. Hilbert-Scan Based Tree (HSBT)

We set as the resolution of an image. After detecting the interest points, we use Pseudo Hilbert-Scan [7] to map all this interest points from 2-D space to 1-D space. Then this linear sequence is divided into many segments averagely which called sub regions. The *j*-th region in *i*-th grouping is labeled by and it is made of four data: the number of local features in this region ( ), the region's gravity center ( ), the set of descriptors ( ), the clustering center of this region ( ). Regions in *i*-th grouping are denoted as . There exist three steps in grouping stage: initialization, region selection and region merging.

1) Initialization: Linear sequence *S* is firstly divided into segments by a factor .

2) Region selection: Firstly, we sort regions depended on the number of interest point. After sorting, can be changed as . For sorted set, . Finally, ; .

3) Merging step: For example, there exist three adjacent regions in the *i*-th grouping: , are main regions and is the rest region, . There comes a question that which main region should be merged into. The merging rule [5] is :

(1)

III. Adaptive Hilbert-Scan

(a) Error path

(b) Correct path

Figure 3. Drawback of Hilbert-scan based tree structure

HSBT can add the spatial information of interest points into nodes without any labeling and manual handing. However, four different kinds of paths can be utilized for image scan. When choosing different scanning path, the order of each block in an image is different in 1-D sequence. For example, see Fig. 3, the region containing object will be separated into two parts in liner sequence by using path1. Thus, many uncorrelated interest points (blue and purple dots in Fig. 3) will be mistakenly merged into this region when building HSBT. So our target is to make sure that those interest points extracted from local appearance of object are as close as possible after

Cite this article as: Mo Xu, Fuyan Liu, Qieshi Zhang, Sei-ichiro Kamata. "Beyond Bag of Features: Adaptive Hilbert Scan Based Tree for Image Retrieval". *International Conference on Information Engineering, Management and Security 2016*: 08-12. Print.

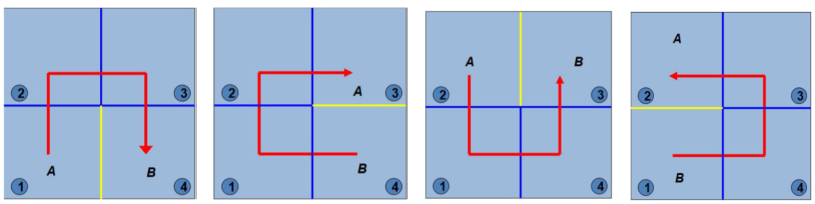
mapping them from 2-D space to 1-D space. Hence, a novel hierarchical path selection strategy is proposed to choose correct path for each image.

C. Path selection

As we mentioned in the last part, the region which contains the majority of interest points should be treated as the main region, for example, sub-block 2 and sub-block 3 in Fig. 4. So the first factor focuses on the number of interest points in sub-blocks on the both sides of split edge (see Fig. 4 the yellow lines represent the split edge). The formula is given as follow,

(2)

In this formula, and denote the number of interest points in sub-blocks on the both sides of split edge respectively, denotes the total number of interest points in this image, *s* denotes the *s*-th scanning path. If , sub-block 1 and sub-block 2 shouldn't be separated because they contain more interest points.



(a) (b) (c) (d)

Figure 4. Illustration of four kinds of scanning path in an image

At this time, there comes a problem that if most of interest points in split region distribute in one sub block (e.g. sub-block 1 or sub-block 2 in Fig 4 (a)), then these interest points are still close after mapping them to 1-D space. To solve this problem, another factor is proposed as follow,

(3)

To combine these two factors effectively, we set a weighting coefficient λ to control the relative importance of them. And the final formula is as follow:

(4)

Inspiring by the generative method of Hilbert Curve [7], we proposed to perform this strategy at increasingly fine resolutions. Thus, this hierarchical strategy can be represented as follow, where *i* denotes the *i*-th division, *k* is the total number of division, *j* denotes the *j*-th sub-block at *i*-th division, *c* denotes the total number of interest points in image, denotes the number of interest points in *j*-th sub block at *i*-th division, represents the weight of *j*-th block at *i*-th division relative to the whole image. Because this novel strategy can help select correct scanning path for each image automatically, we call this method Adaptive Hilbert-Scan (AHS).

— (5)

After choosing path and constructing HSBT for each image, the BOF model will be combined with adaptive Hilbert-Scan based tree (AHSBT) to form the final model Adaptive Hilbert-Scan based Bag-of-Features (AHS-BOF). Finally, we are able to obtain a descriptive histogram representation for each image by AHS-BOF.

IV. Experiments

In our experiments, we evaluate this approach on a challenging object dataset -- Caltech-256. Mean average precision (mAP) is used to evaluate our proposed method. We select SIFT [8] to extract local features. To train the vocabulary, we randomly choose 50 images from each category (totally 12800 images) as the training set. Then, K-means is used to generate the vocabulary. For testing, 5 images per category are randomly selected from the rest images in each category. We choose the same vocabulary size as the previous work [5]. We set =500 and Th = 0.8. In Fig. 5, it can be clearly seen that our method outperforms than other methods. Table I compares AHS-BOF with previous work HS-BOF under different vocabulary sizes (10k, 20k, 50k, and 100k) in terms of mAP. When vocabulary size is 100k, the number of visual words is almost equal to the numbers of local descriptors. Hence, the histogram representations become less discriminative, which affect the retrieval precision. So no matter which kind of scanning path we select,

Cite this article as: Mo Xu, Fuyan Liu, Qieshi Zhang, Sei-ichiro Kamata. "Beyond Bag of Features: Adaptive Hilbert Scan Based Tree for Image Retrieval". *International Conference on Information Engineering, Management and Security 2016*: 08-12. Print.

the retrieval result is nearly the same. The result of level 3 in Table II has a little degeneration which indicates paying more attention on local details of interest points will lose the appearance of objects in building AHSBT

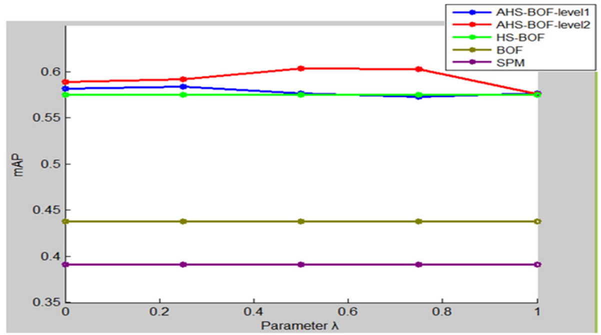


Figure 5. Illustration of performance under different λ

Table I Comparison of mAP on Caltech-256 with different vocabulary size

Word size	BOF[11]	SPM[11]	HS-BOF	AHS-BOF
10k	0.438	0.391	0.576	0.604
20k	0.541	0.437	0.625	0.653
50k	0.573	0.472	0.635	0.657
100k	0.604	0.499	0.595	0.596

Table II Results under different level

L (level of split)	AHS-BOF (λ=0.5)	HS-BOF
1	0.576	0.575
2	0.604	
3	0.592	

V. Conclusions

In our study, we propose a novel AHS method depended on the distribution of interest points in image. The AHS can choose correct path for each image automatically. There are two main contributions in our method: a) it can reduce the merging error and recover the shape of objects more quickly; b) The computing time of constructing tree is reduced because of fewer layers. Evaluations on public database Caltech-256 have demonstrated the effectiveness of this method.

References

1. J. Sivic and A. Zisserman, "Video Google: A text retrieval approach to object matching in videos," *Proceedings of Ninth IEEE International Conference on Computer Vision* (ICCV), pp. 1470-1477, 2003.
2. S. Lazebnik, C. Schmid, and J. Ponce, "Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories," *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition* (CVPR), vol. 2, pp. 2169-2178, 2006.
3. Y. Cao, C. Wang, Z. Li, *et al.*, "Spatial-bag-of-features," *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition* (CVPR), pp. 3352-3359, 2010.
4. M. Marszaek and C. Schmid, "Spatial weighting for bag-of-features," *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition* (CVPR), vol. 2, pp. 2118-2125, 2006.
5. P. Hao and S. Kamata, "Hilbert scan based bag-of-features for image retrieval," *IEICE Transactions on Information and Systems*, vol. 94, no. 6, pp. 1260-1268, 2011.
6. D. Hilbert, "Über die stetige Abbildung einer Linie auf ein Flächenstück," in *Dritter Band: Analysis·Grundlagen der Mathematik·Physik Verschiedenes*. Springer Berlin Heidelberg, pp. 1-2, 1935.

Cite this article as: Mo Xu, Fuyan Liu, Qieshi Zhang, Sei-ichiro Kamata. "Beyond Bag of Features: Adaptive Hilbert Scan Based Tree for Image Retrieval". *International Conference on Information Engineering, Management and Security 2016*: 08-12. Print.

7. J. Zhang, S. Kamata, and Y. Ueshige, "A pseudo-Hilbert scan for arbitrarily-sized arrays," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. 90, no. 3, pp. 682-690, 2007.
8. D. Lowe, "Object recognition from local scale-invariant features," *Proceedings of Seventh IEEE International Conference on Computer Vision* (ICCV), vol. 2, pp. 1150-1157, 1999.

Cite this article as: Mo Xu, Fuyan Liu, Qieshi Zhang, Sei-ichiro Kamata. "Beyond Bag of Features: Adaptive Hilbert Scan Based Tree for Image Retrieval". *International Conference on Information Engineering, Management and Security 2016*: 08-12. Print.