Lossless Location Coding for Image Feature Points

Gyeongmin Choi, Hyunil Jung, and Haekwang Kim

Abstract—**Image retrieval research activity has moved its focus from global descriptors to local descriptors of feature point such as SIFT. Currently MPEG is working on standardization of effective coding of location and local descriptors of feature point in the context mobile based image search driven application in the name of MPEG-7 CDVS (Compact Descriptor for Visual Search). While CDVS is dealing with lossy compression of location information, this paper presents various lossless compression methods of locations and provides comparative experimental results for applications requiring fine matching precision. Among 4 methods presented in this paper, Experimental results show that Block-based lossless location coding using circular scan order method shows the best compression results with compression ratio achieving 2.5 to 1 with reference to Fixed-Length lossless location coding method.**

Index Terms—**Local Feature, location coding, compression, lossless**

I. INTRODUCTION

Image retrieval technology is a critical tool for effectively using enormous amount of images as text-based search such as google and yahoo brings information from the huge web pages on the internet to your demand. Research on Content-Based Image Retrieval (CBIR) [1] started at the end of 1980s and came to its climax with the MPEG-7 [2] international stand efforts in the early of 2000s. A typical usage of CBIR is that a user provides an image as query and similar images to the query are retrieved and sent to the user. Most of CBIR research focused on global descriptors [3] of color, texture and shape features extracted from the whole given image for the similarity measure. Global descriptor has inherent problem that it is incapable of distinguishing two different images with same descriptors. For example two to-tally different images may produce the same color histogram descriptor.

Recently research on local descriptors such as SIFT [4] and SURF [5] has gained heavy attention with its meaningful performance results and various useful applications. Local descriptors are extracted from feature points detected in the image. Local descriptors can be used for computer vision applications such as 3D reconstruction as well as image retrieval applications as it basically provides corresponding feature point matching functionality between two given images.

MPEG inaugurated another MPEG-7 international standardization work called CDVS (Compact Descriptor for Visual Search) [6] on the basis of local descriptors. The main scope of standard is effective and efficient coding of location and local descriptors for feature points in a given image. The target scenario of this standard is that you take a picture of an object of interest with your mobile phone camera and send the extracted feature information through mobile network to the server. The server identifies the object by comparing received CDVS features with its CDVS feature database already extracted from images of objects. The server sends identified object elated information back to the user.

Efficient coding is key issue in this CDVS scenario because an economic usage of network capacity is required due to mobile network's constraint. MPEG-7 standard currently addresses lossy coding issue for feature location for the trade of matching precision and bandwidth. In this paper, a method of lossless coding is presented with its results as some applications require matching precision as fine as possible.

In this paper, various lossless coding schemes for feature locations are presented in Section 2 and their experimental results are shown in Section 3 with analysis. Section 4 concludes the paper.

II. METHODS OF LOSSLESS FEATURE LOCATION CODING

A. Feature Location

Fig. 1 shows an example of feature points extracted from an image using SIFT local descriptor methods. The location of each feature point is expressed in (x,y) coordinate. All lossy location coding algorithms exist one feature point as representing coordinate(x,y) in each block for compression. However, the goal of this re-search is compress the location information as much as possible without the feature point representative. Four different compression methods are presented and com-pared in this paper.
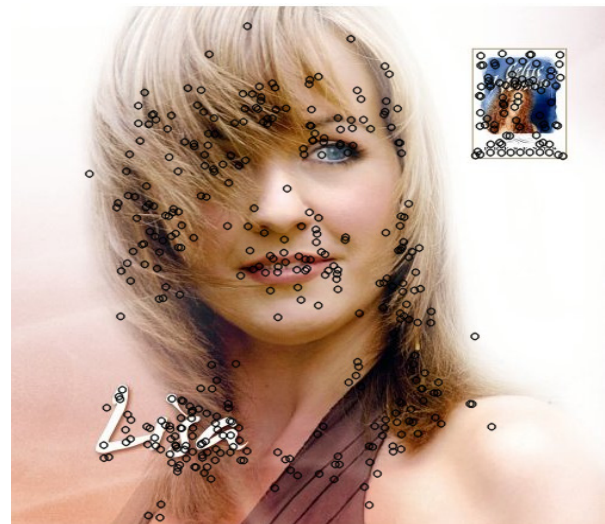

Fig. 1. Feature points in CD case database extracted by using SIFT

B. Fixed-Length Lossless Location Coding

Fixed-Length lossless location coding method is the most basic idea for compression of location(x,y). The absolute values of x and y coordinates for each extracted feature point is needed same bit-quantity for encoding. For example, if one image has 500 feature points and each feature point needs 20 bits, the encoded output bins are exactly 10000 bits. For an image of width w and height h image, bit length for x, bl_x and bit length for y, bl_y are calculated as follows:

$$bl\_x = \text{Ceil (log2w)} \tag{1}$$

$$bl\_y = \text{Ceil (log2h)} \tag{2}$$

This method is the most straight forward, inefficient method.

C. Pixel-Based Lossless Location Coding

Pixel-based lossless location coding method is simple way to compress location information using arithmetic coder. From left-top pixel to right-bottom pixel, All pixels are encoded sequentially pixel's location information as 1 or zero using raster scan order. Commonly, bin 1 means non-empty block and zero means empty block in image. Previously, Stanford [7] investigated efficient context model 'sum-based context model' [8] for location coding. That is why, we use CABAC [9] including sum-based context model.

When one pixel is encoded, look surrounding context and encoded as 1 or zero using CABAC. This is easy method, but still have problem that many bins go to waste for compression at sparse area.

D. Block-Based Lossless Location Coding

There are several ways for improved lossless location coding. One of methods is the Block-based lossless location coding which is explained that all feature points are divided into block histogram and encoding.

When we use block histogram for compression, there is a particular advantage that we can reduce local redundancy about empty blocks in images. If encoding all feature point's location information as pixel-based with-out block histogram, we have to encode every single area including unnecessary region for compression. Which means bit rate is increasing because of useless empty areas.

We can explain this algorithm in three steps. In the first, when having an input image that already included features extracted before, it is divided into histogram block as specific size such as 16×16, 12×12. Second, encode block histogram information called 'header in-formation' as 1 or zero using raster scan order by using CABAC including sum-based context model. Third, for reference of 'header information' already encoded, make bins when every 1 block in header by using CABAC including sum-based context model in shown Fig. 2.

E. Block-Based Lossless Location Coding Using Circular Scan Order

Block-based lossless location coding using circular scan order method [10] is on the basis of block-based lossless location coding mentioned above. Core algorithm for encoding is almost same, but it can more care of sparse area by using circular scan order in image. The reason we can obtain high efficiency bit rate using circular scan order

instead of raster scan order is inspired by distribution of feature points extracted from SIFT in TM3.0 as standardization work software[11]. More concretely speaking, if there are enough feature points extracted in one image for matching, feature point selection algorithm in TM3.0 is selecting certain feature points as close as possible to center in image. Moreover most main objects are placed in the center generally. So that if starting binning from center part using circular scan order, it can save more bits as compared with the raster scan order method in shown Fig. 3.
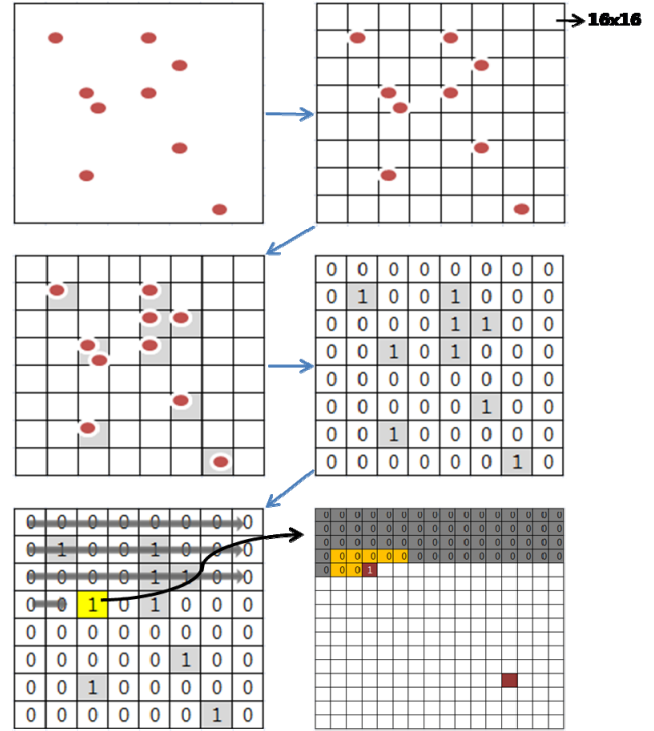

Fig. 2. Input image including features(left-top) and divided image into 16x16 block histogram(right-top) and making histogram header information and encoding as 1 or zero (middle row) , encoding using raster scan order and CABAC including sum-based context model by reference header information (bottom row).
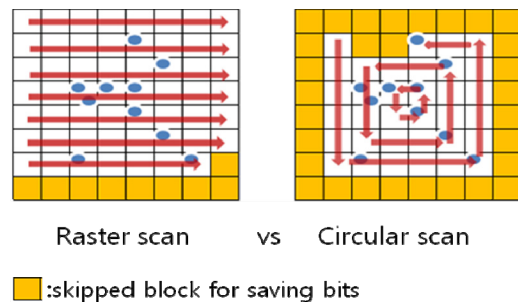

Fig. 3. Compare raster scan order and circular scan order

III. EXPERIMENTAL RESULTS

We apply the proposed four schemes to compress location coordinate(x,y) and evaluate results. In our experiments, we consider two criteria that how compression of location information and correspondence error and all experiment work on TM3.0 software.

Compression rates of each algorithm are evaluated as 'bits per feature' in one category database and correspondence error is checked as true positive ratio and localization accuracy ratio for different operating point.

We use Graphic category in database being used for CDVS standardization including 2505 images in. And for comparison we test on 5 operating points which have different descriptor lengths in shown Table I.

TABLE I: OPERATING POINTS FOR EXPERIMENT

	8mode	9mode	10mode	11mode	12mode
descriptor lengh	1k	2k	4k	8k	16k
maximum point selected	225	255	511	1023	1023

As a result, we examine the coding performance of the different coding methods described above by Fig. 4. We found the best efficient block size is 12 ×12 empirically and 12 ×12 Block-based lossless coding using circular scan order method shows the highest performance and can reduce more 10 bits as compared to fixed-length lossless coding.

Of course, all lossless coding methods have same true positive ratio and localization accuracy ratio. So we compared with lossy location coding result from TM3.0 as shown Fig. 5 and Fig. 6.And the presented loss-less coding method is 0.5% more accurate than lossy coding method.
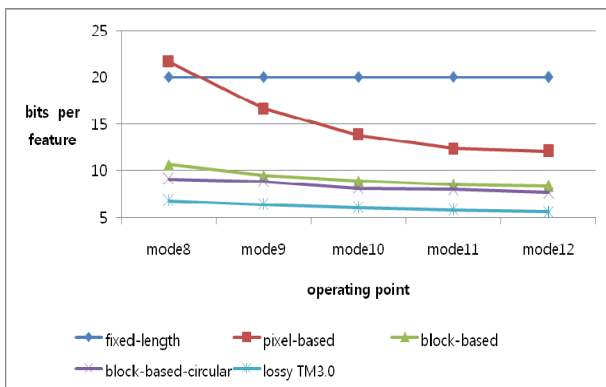

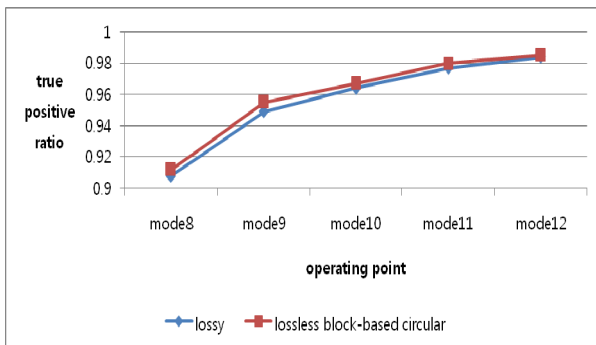Fig. 4. Comparison of Compression rates for graphic category database

Fig. 5. Comparison of true positive ratio between lossy coding and block-based circular scan order lossless coding.
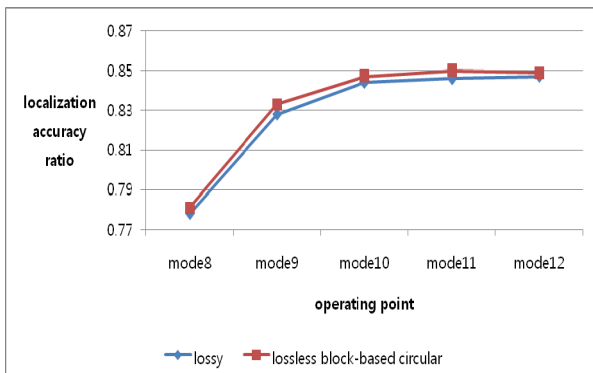

Fig. 6. Comparison of localization accuracy ratio between lossy coding and block-based circular scan order lossless coding

IV. CONCLUSIONS

We have proposed four methods of compressing location information lossless. In the future, lossless location coding algorithms are used frequently because network's constraint is moderated and mobile devices performance in better.

From our experimental results, it is shown that 0.5% more accurate than lossy location coding scheme and Block-based lossless location coding using circular scan order method is the best novel method among four other methods. It can reduce more 10 bits per feature than fixed-length lossless location coding and 5~10 bits per feature than pixel-based lossless location coding.

REFERENCES

[1] J. Eakins and M. Graham, "Content-based image retrieval," Technical Report, JISC Technology.
[2] *ISO/MPEG N4320, MPEG-7 Requirements Document*, vol. 15, F. Pereira, ed., MPEG Requirements Group,Sydney, July 2001.
[3] E. Ch. Avez, G. Navarro, R. Baeza-Yates, and J. L. Marroqu, "Searching in metricspaces," *ACM Computing Surveys*, vol. 33, no. 3, 273-321, 2001. Visual Image Search: Feature Signatures or/and Global Descriptors.
[4] D. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91-110, Nov. 2004.
[5] H. Bay, T. Tuytelaars, and L. V. Gool, "SURF: speeded up robust features," in *Proc. European Conferenceon Computer Vision*, Graz, Austria, May 2006, pp. 404-417.
[6] Y. Reznik, *Compact Descriptors for Visual Search: Applications and Use Scenarios Requirements Subgroup*, MPEG N11529. 2010.
[7] S. Tsai, D. Chen, G. Takacs, V. Chandrasekhar, J. P. Singh, and B. Girod, "Location Coding for Mobile Image Retrieval," in *Proc. International Mobile Multimedia Communications Conference*, Sept. 2009.
[8] ISO/IEC JTC1/SC29/WG11/M24758, "Improvements to location coding : reponse to CE3 - feature point location coding," Geneva, CH, April 2012
[9] D. Marpe, H. Schwarz, and T. Wiegand, "Context-adaptive binary arithmetic coding for H.264/AVC," *IEEE Trans. Circuits Syst. VideoTechnol.*, vol. 13, pp. 620–636, July 2003.
[10] M. Azimipour, M. Eshghi, and A. Khademzadeh, "A Modification to Circular-Scan Architecture to improve test data compression," in *IEEE CS Proc. of 15th Int. Conf. On advance computing and communication*, 2007
[11] ISO/IEC JTC1/SC29/WG11/W12929, "Test Model 3: Compact Descriptors for Visual Search," Stockholm, SE, July 2012.

Gyeong Min Choi was born in Seoul, Korea in 1986 and received the B.E. degree in computer engineering at Sejong university in Seoul, Korea in 2011. Presently he is a master student at the Department of Computer Engineering, Sejong University. His current research interests include mobile image retrieval, and video coding.
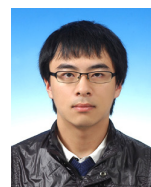
Hyunil Jung was born in Pucheon, Korea in 1986. Presently he is a bachelor student at the Department of Computer Engineering, Sejong University He is enrolled in master degree at the Department of Computer Engineering, Sejong University. His current research interests include mobile image retrieval, and video coding.

Hae Kwang Kim was born in Pohang, Korea in 1963 and received the B.S. degree in electronics engineering at Hanyang university in Seoul, Korea in 1986 and the D.E.A and Ph.D degrees in informatics at Paul Sabatier University in Toulouse, France in 1994 and 1997, respectively. He was a Research Engineer with the Central R&D center, Samsung Electronics from 1986 to 1992 developing optical disk systems and multimedia applications. He was a Principal Research Engineer at Information and Telecommunications R&D center, Hyundai Electronics performing research work on multimedia coding and metadata representation. Since 2000, he has been with Sejong University, Korea. Presently he is a Professor at the Department of Computer Engineering, Sejong University. His current research interests include mobile image retrieval and video coding.