An Efficient Visual Loop Closure Detection Method in a Map of 20 Million Key Locations

Junjun Wu[1], Hong Zhang[2] and Yisheng Guan[3]

Abstract— **An important problem in robot simultaneous localization and mapping (SLAM) is loop closure detection. Recent studies of the problem have led to successful development of methods that are based on images captured by the robot. These methods tackle the issue of efficiency through data structures such as indexing and hierarchical (tree) organization of the image data that represent the robot map. In this paper, we offer an alternative approach and present a novel method for visual loop-closure detection. Our approach uses an extremely simple image representation, namely, a down-sampled binarized version of the original image, combined with a highly efficient image similarity measure - mutual information. As a result, our method is able to perform loop closure detection in a map with 20 million key locations in about 2.38 seconds on a commodity computer. The excellent performance of our method in terms of its low complexity and accuracy in experiments establishes it as a promising solution to loop closure detection in large-scale robot maps.**

I. INTRODUCTION

Visual loop-closure detection is an important problem since a camera has become a popular choice of perception in robot simultaneous localization and mapping (SLAM). SLAM is an incremental process in which loop closure detection determines if the robot has returned to a previously visited place, and this information is critical for creating a topologically correct map and for improving the metric information about the map. In recent years, robotics researchers have invested a considerable amount of effort to address the problem of visual loop closure detection due to its importance.

Successful approaches to visual SLAM exist, and they all benefit from two key properties of visual sensing: low-cost of the sensor hardware, and rich textural information of a visual image. In particular, compared with range sensors, vision is more convenient to utilize for making the decision on loop closure detection due to the rich features in an image such as color, texture and shape of the objects in the environment. In addition, extensive literature exists in contents-based image retrieval (CBIR). By formulating visual loop closure detection as one of image matching between the view of the robot at the current location and images observed by robot at previously visited places, known as key locations, many algorithms in CBIR can be exploited in detecting loop closures. The efficiency and accuracy of computing similarity between an image pair is crucial for a robust solution to visual loop closure detection due to the real-time nature of the SLAM problem.

In this paper, we present a novel, highly efficient method that addresses the issue of detecting loop closure in large-scale maps. Computationally, our method is capable of dealing with maps on the order of 20 million keyframes or key locations in 2.38 seconds for their real-time visual loop-closure detection. Distinguished from the exiting approaches, our method explores image representation as a way of achieving an efficient loop closure detection algorithm, as opposed to efficient data structures for searching image descriptors. Specifically, the success of our method is derived from two key steps: representation of an image in the form of a short binary code (e.g. 300 bits) and the computation of similarity between images using mutual information. As a result, we are able to compute an image similarity value with just a few instructions in SSE4 (SIMD Extensions 4 used in the Intel Core microarchitecture), giving our method an extremely high performance in terms of its scalability to a large-scale environment map. This highly efficient and accurate similarity measure produces a short list of highly-ranked keyframes as our loop-closure candidates, and they are subsequently verified through a rigorous method such as feature matching and multi-view geometry, a common step of constant complexity in all loop-closure detection methods. It is important to point out that although the complexity of our method grows linearly with the number of images that define the robot map, the similarity computation between images is performed in a fraction of a microsecond, so that it could detect loop closure in real-time even along a route thousands of kilometers in length.

The rest of this paper is organized as follows. In Section 2, we review the exiting research in visual loop-closure detection. We then describe the details of the similarity measure for image comparison in our study, selection of loop-closure candidates and their verification in Section 3. Section 4 illustrates and evaluates the proposed method using a popular experimental dataset and a large-scale outdoor dataset. We then conclude with remarks on our method and outline future work in Section 5.

This work was supported in part by NCRFN and NSFC.

[1]Junjun Wu is with Department of Software, Guangdong Food and Drug Vocational College, Guangzhou, China, wujj@gdyzy.edu.cn.

[2]Hong Zhang is with Department of Computing Science, University of Alberta, Edmonton, Canada, hzhang@ualberta.ca.

[3]Yisheng Guan is with School of Electro-mechanical Engineering, Guangdong University of Technology, Guangzhou, China, ysguan@gdut.edu.cn.

II. RELATED WORK

Visual loop-closure detection is crucial for SLAM algorithms, and it is similar to the global localization problem in a known map. An early method [1] used color histogram as an image descriptor and performed image matching by a voting scheme. The idea of using visual BoW to detect loop closure in recent years is considered as the leading approach in visual SLAM. Angeli et al. [2] [3] utilized BoW to describe an image and calculated the loop-closure probability of the current view with respect to previous observed images in a Bayesian framework. Cummins and Newman [4] also used BoW to develop the FAB-MAP framework that is considered as a representative achievement. In their follow-up work, FAB-MAP 2.0 [5] used a randomized forest of kd-trees to speed up visual vocabulary generation and the largest map has been studied thus far covers a route of 1,000 km with 100K images using a bag-of-words based approach. More recently, computer vision researchers proposed binary feature descriptors [6] [7], which have shown competitive performance in terms of real-time efficiency. Galvez-Lopez and Tardos [8] presented a bag of binary words with BRIEF descriptors, as a derivative of BoW.

The BoW-based representation of an image has shown its effectiveness for selecting loop-closure candidates, but it suffers from perceptual aliasing [9] due to vector quantization and its accuracy can be dependent on the training images. As an alternative to BoW, Liu and Zhang [10] used Gabor-Gist as a compact image descriptor to detect loop closures without the need to detect keypoints and vector quantization to measure similarities among images.

To improve the efficiency of loop-closure detection, efficient data structures (e.g. hierarchical k-means, kd-tree [5] and locality sensitive hashing [11]) are also employed in order to manage the complexity of handling a large-scale map. The former is essentially a recursive partition of the feature space along hyper-planes orthogonal to the coordinate axes to provide an efficient means to match features approximately [12]. The latter offers sub-linear time by hashing highly similar data points into the same bucket of a hash table with high probability. Shahbazi et al. [11] presented an application of locality sensitive hashing to real-time loop closure detection, which used SIFT and E2LSH as a way to selecting the loop-closure candidates.

Recently, CBIR researchers have turned towards techniques using a binary code for images [13]. Although its complexity grows with the number of images, it can still be faster than tree-based or hashing-based techniques for searching the candidates on a large image database with millions of images due to compact binary code that represents an image in a few hundred bits. Although the exiting approaches for generating a small binary code may not be suitable to loop-closure detection, the idea motivates us to adopt a similar approach for measuring image similarity, in order to make it possible to detect loop-closure efficiently without using a complex data structure.

Our decision to work with the whole image rather than its local features is also encouraged by the success of using raw pixels at a reduced dimension to represent an image [14] where images are down-sampled to 64x32 and pixel-wise differences are used as a basis for similarity measure. However, our work takes the above approach two steps further. First we use a binarized version of the down-sampled image and secondly, we use mutual information (MI) as a similarity measure, which is more accurate and more efficient to calculate than raw-pixel differences.

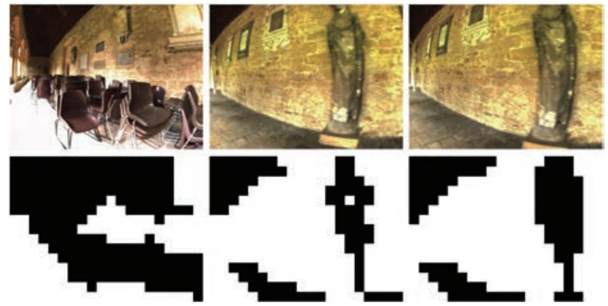


Fig. 1. The first row shows original keyframes observed by robot; the second row shows corresponding binarized appearance at level=6 in scale space.

The MI between two random variables is a concept with roots in information theory and it essentially measures the amount of information that one variable contains about the other. MI was introduced as a similarity measure of image pairs in [15], and has been widely used in image registration applications [16]. In robot visual homing, Dame and Marchand [17] have achieved successfully a navigation task for a non-holonomic vehicle by building a control law directly from the maximization of the shared information between the current image and the next keyframe in the visual path. However, MI has not been applied to visual loop-closure detection.

III. IMAGE DESCRIPTION, SIMILARITY MEASURE AND LOOP-CLOSURE DETECTION

In this section, we will detail our efficient method for visual loop-closure detection in robot SLAM. As mentioned, our approach achieves its efficiency through two important steps: description of an image with a short binary code and similarity computation between images with mutual information. Our method does not involve the expensive step of feature or keypoint detection and description. By directly selecting loop-closure candidates with the top scores of the proposed similarity measure, we have been able to achieve nearly 100% recall in two visual SLAM datasets. Our method can be easily implemented in real time, with the ability to perform loop-closure candidate selection in a large-scale map. The selected candidates should be verified in a subsequence step of loop closure detection, with a complexity that is independent of the map size. The high efficiency and accuracy of our proposed visual loop-closure method will be demonstrated in the experimental result section.

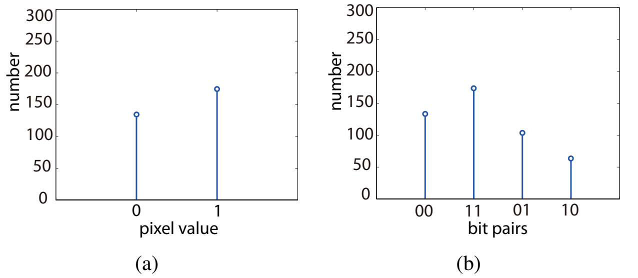


Fig. 2. (a) shows counting bits in a binary patch for calculating entropy; (b) shows counting bit-pairs of two binary patches for calculating joint entropy.

A. *Image Description*

The purpose of an image descriptor is to capture sufficient details of an image in a compact fashion in order to be able to compare images efficiently and accurately. In the proposed method, an image is first smoothed with a Gaussian filter and then down-sampled to reduce the image size. The reduced image is then binaried with a thresholding algorithm such as Otsu's method [18], to produce a binary code on the order of a few hundred bits. Note that other binarization algorithms such as local adaptive thresholding techniques or binary image descriptors such as BRIEF or BRISK are also possible, although we do not consider them in this paper. Typically results of this smoothing, down-sampling and thresholding process are shown in Fig. 1, where the left image is from one location, and the middle and right images are two different visits to the same location. The similarity between the middle and the right images is apparent, both in the original color space and in the binary space, as is the dissimilarity between the left image and the middle (or right) image.

B. *Similarity Measure and Loop Closure Detection*

The mutual information of an image pair considers both the joint information entropy $h(x, y)$ and the individual entropies $h(x)$ where:

$$h(x) = -\sum_{x} p(x) \log[p(x)] \tag{1}$$

$$h(x, y) = -\sum_{x} \sum_{y} p_{xy} \log[p_{xy}(x, y)] \tag{2}$$

where $p(x)$ is the histogram of image x, $p(x, y)$ is the joint histogram of image x and image y. MI is defined in terms of joint and individual entropies as:

$$\begin{aligned} MI(x, y) &= h(x) + h(y) - h(x, y) \\ &= \sum_{x} \sum_{y} p_{xy} \log \frac{[p_{xy}(x, y)]}{p_x(x) * p_y(y)} \end{aligned} \tag{3}$$

For binary images, the computation of mutual information has an extremely simple form to make it highly efficient to compute, as will be shown below. Specifically when each image is encoded with a binary vector, its intensity histogram has only two bins, for the number of 0's and 1's in the vector. Similarly, the joint histogram of two such vectors has four bins, for the four combinations of (0, 0), (0, 1), (1, 0), and (1, 1) pairs that occur in the corresponding positions of the two vectors. As a result of this simple representation of the marginal and joint histograms, we can compute the three relevant histograms, $p(x)$, $p(y)$ and $p(x, y)$, of the two binarized images in a few logic instructions and the POPCNT instruction in SSE4 for counting the number of 1's in a binary vector. Specifically, to calculate the histogram of a binarized image, we need one POPCNT instruction. For the joint entropy of an image-pair, our method utilizes three instructions, shown in Fig. 2, where we need two negation instructions to compute the complements of the image vectors and three logical AND instructions, followed by three POPCNT instructions. Once the histograms and the joint histograms are converted to their corresponding probability mass functions, the MI of the two binarized images can be readily computed with a total of 12 muls/divs and four logarithmic instructions, as detailed in Algorithm 1. Most importantly, the computation of similarity between two images involves only scalar arithmetic instructions on a SIMD machine, and this translates into a computation time of a fraction of a microsecond.

Algorithm 1 The similarity measure of an image-pair

Input: the candidate keyframe, x; a keyframe, y;
Output: the similarity, $s(x, y)$;
% the information entropy of a binary image, $h(x)$
  $[m,n]$=size(x);
  q_1=$popcnt(x)$;
  q_0=mn-q_1;
  q=$[q_0,\ q_1]$;
  $\overline{q}$=$q./(mn)$;
  $h(x) = -\sum_{i=1}^{2} \overline{q}(i) \log \overline{q}(i)$;
% the joint information entropy of image-pair, $h(x, y)$
  $p_{00} = popcnt(\sim x \& \sim y)$;
  $p_{11} = popcnt(x \& y)$;
  $p_{01} = popcnt(x \& \sim y)$;
  $p_{10} = mn - p_{00} - p_{11} - p_{01}$;
  $\overline{p} = [p_{00}, p_{11}, p_{01}, p_{10}]\ ./(mn)$;
  $h(x, y) = -\sum_{j=1}^{4} \overline{p}(j) \log \overline{p}(j)$;
% the mutual information of image-pair, $mi(x, y)$
  $mi\,(x, y) = h\,(x) + h\,(y) - h\,(x, y)$;
% the similarity of image-pair, $s(x, y)$
  $s(x, y) = mi(x, y)$;
return s;

For loop closure detection, the keyframes in an appearance map with the highly-ranked MI values to the current view of the robot can be considered as loop-closure candidates. From the candidates, we need to verify the occurrence of a true loop-closure event by using a stringent verification step. For this purpose, the conventional approach is to use multi-view geometry to determine if the current view is geometrically consistent with the map view of the loop-closing key location. For our study in this paper, we simply

(a) Experimental Scene (b) Ground Truth

Fig. 3. (a) an image in New College dataset shows the experimental scene; (b) the ground truth.

use the BoRF method [21] with a small set of key points with a sufficiently high threshold to filter the false positives. Please note that although simple feature matching is not the most robust way to verify a loop closure, we use it in this paper only for its simplicity, and that how the candidates are verified is a common issue to all loop closure detection methods, but beyond the scope of this paper.

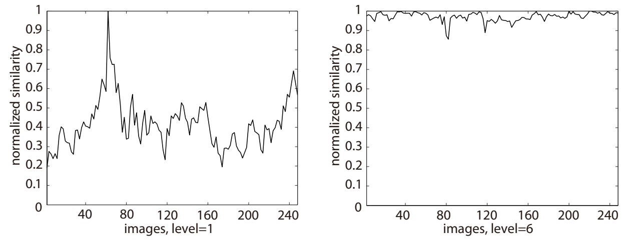
C. Complexity Analysis

In general, the complexity of a loop-closure detection method is due to feature extraction, loop-closure candidate selection, and candidate verification. The feature extraction step, which is expensive on the order of a second, is employed in most existing approaches to visual loop-closure detection, and its cost depends on the features used and is the same for all algorithms. In our method the cost of down-sampling and binarization for current new observation does not grow with the map size, and the loop-closure candidates are selected directly with the binary version of MI computation. Selected features at coarse scale are used only in the verification step of our method, and we avoid the processing of features such as vector quantization and indexing as is necessary in BoW-based approaches.

Although the selection of the loop closure candidates in our method is through linear search, the search is performed in a one-dimensional space of similarities, each of which is computed with only scalar instructions on a SIMD machine. Due to the high accuracy of our similarity measure, as will be shown in the next section, we are able to propagate just a few top candidates to the verification step, and still achieve a high recall. As a result, the complexity of this verification step is constant and independent of the map size.

IV. EXPERIMENTAL VALIDATION

In this section, we will describe the experimental validation of the proposed method for visual loop-closure detection. We will first describe the two datasets used in our experiments, and select the scale at which to binarize an image to produce its descriptor. Subsequently, we will present the results of the performance of loop-closure candidate selection in terms of recall, precision, F-measure and execution time. Finally, the accuracy of the overall loop closure detection algorithm will be presented, using feature matching in the verification step.



(a) (b)

Fig. 4. This figure shows the similarity based the MI in the original color space between the 249th image and the previous observed images at two different scales for the New College dataset. Level 1 represents the original image size (640x480), and level 6 is a down-sampled version by a factor of 32 in each image dimension (20x15).

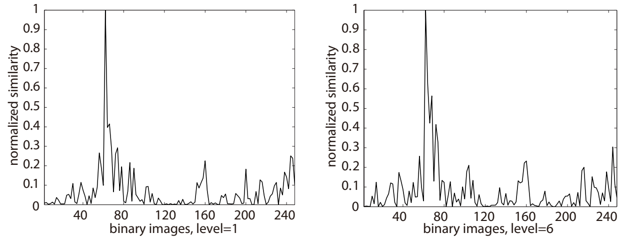
A. Experimental Datasets

We ran our experiments on two datasets. One is the popular New College dataset introduced in FAB-MAP [19], which contains 1237 pairs of images. At a resolution of 640x480, each pair of the images were taken by the left and right cameras on the robot in Oxford City. Fig. 3 (a) and (b) show a representative image and the ground truth matrix with visible off-diagonal line indicating loop closures. The other dataset is a Google Street View dataset, with 49224 images in an image sequence collected by a moving vehicle. There are four images for four distinct cameras orientations at each location, and we use a sample from the dataset with 2x12556 images captured by left and right cameras and ground truth for this set is constructed manually in our experiment.

B. Image Scale Selection

Our method calculates the mutual information of binary image pairs at a particular scale in the scale space of the images. In our case, we generate an image at the scale of interest by simply down-sampling the original image after an initial Gaussian smoothing. To identify this proper scale, we have analysed the effectiveness of MI similarity measure at six scales for both the original images and the binarized images, and concluded that, as a similarity measure, MI is not affected significantly by down-sampling if the images are binarized, unlike color or grayscale images. To illustrate this we take image 249 in New College dataset as an example. According to the ground truth, image 61 is the loop-closure location of image 249.

As shown in Fig. 4, with color images, the larger the scale at which to calculate MI between image 249 and the images in the map, the less discriminating MI becomes as a similarity measure. In contrast, the MI of our binarized codes still is relatively insensitive to scale in terms of its discriminative power. In fact, as the scale becomes larger the likelihood according to MI remains consistent with the ground truth, as shown in Fig. 5. Hence, in our experiments, we are able to down-sample an image from a resolution of 640x480 by a factor of 32 in each dimension to a final size of 20x15, for a binary vector with 300 bits or a reduction of the storage requirement by almost four orders of magnitude (8,192).



(a) (b)

Fig. 5. This figure shows the similarity measuring based on MI between the 249th binary code and the previous binary codes at two different scales for the New College dataset. Level 1 represents the original image size (640x480), and level 6 is a down-sampled version by a factor of 32 in each image dimension (20x15).

TABLE I

THE COMPARISON OF PERFORMANCE BETWEEN BOW AND THE PROPOSED METHOD IN TERMS OF F-MEASURE (F), PRECISION (P), AND RECALL (R)

Method	F	P	R	Words	Threshold
BoW method	0.80	76.1%	86.3%	700	–
	0.89	82.1%	97.2%	4100	–
	0.92	95.1%	93.0%	7500	–
proposed method	0.93	86.9%	100%	–	0.017
	0.97	94.3%	100%	–	0.02
	0.91	96.7%	85.3%	–	0.035

C. Loop-closure Candidate Selection and Loop-closure Detection

For the loop closure detection in an appearance map in which each key location is characterized by an image, it is critical to produce a high recall of the loop-closure candidates and do so in real time. In this sub-section, we analyze the recall of our proposed method for selecting the loop-closure candidates on the two experimental datasets. We adopt a common practice in image retrieval and examine how often the correct loop-closure candidates are contained in the top k ranked retrieval results, i.e., those with the top-k similarity values to the current robot view. We also compare the effectiveness of the proposed method with the visual BoW method in terms of precision and recall. Finally, we measure the computational efficiency of our method using the second large scale dataset.

Fig. 6 summarizes the results when the binary images are represented at six deferent scales, all with recall values at 100% when the k is greater than 7 for the New College dataset. In addition, Fig. 7 shows that recall on the Google Street View dataset is nearly 100% for the images in RGB space at the finest scale and the binary images represented at the finest scale and the 5th scale, when the k is greater than 12. When the images in the RGB space are represented at the 5th scale, recall becomes weak as corroborated by Fig. 4. Hence, in the hamming space we are able to *increase* the performance of the loop detection algorithm by *decreasing* the computational and memory cost of evaluating image similarity.

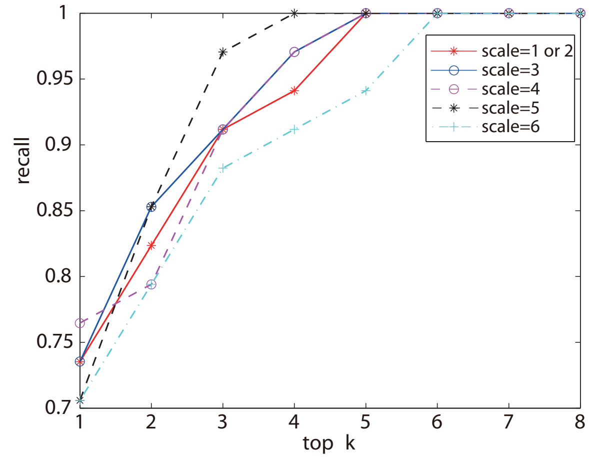


Fig. 6. The figure shows the relationship between recall and top k when the binary images are at different six scales for New College dataset.

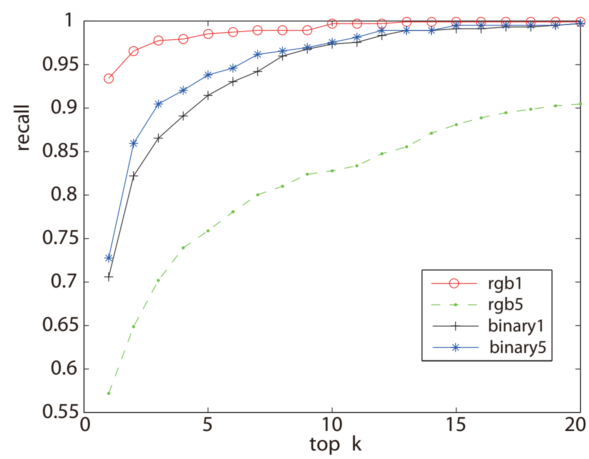


Fig. 7. The figure shows the relationship between recall and top k for the sample from the Google Street View dataset. "rgb1" and "rgb5" represent the original images and images at the 5th scale respectively, "binary1" and "binary5" are corresponding binarized images .

To compare the performance of visual BoW with our MI-based method in this paper, firstly, we set the different numbers of visual words to test the BoW method on the same experimental dataset with the help of MVT3.3 package [20]; secondly, in the proposed method we will select k=12 keyframes and then the further verify loop-closure candidates by matching scale invariant features directly at coarse scales [21]. As shown in Table I, the accuracy of the BoW method is dependent of the number of words in the visual vocabulary. In contrast, the proposed method with a low computational cost shows 100% recall with high precision, where thresholds are used in verification step. In fact, if one is to employ verification using multi-view geometry, a precision approaching 100% for both methods can be expected.

Regarding the computational efficiency of our proposed algorithm, the average computation time of loop-closure detection for different map sizes is measured on a commodity PC (C++ in Visual Studio 2010, with Intel(R) Core(TM) i7-3770 CPU at 3.40GHz and 8GB memory). Loop closure

TABLE II

TIME COST ON DIFFERENT SIZE MAP

Map Size	2.4k (20%)	4.8k (40%)	7.2k (60%)	9.6k (80%)	12k (100%)
Time(*ms*)	0.28	0.58	0.90	1.18	1.47

detection in general begins with image capture and, in our case, it is followed by image down-sampling, binarization, MI computation and finally verification of the top-k candidates. Since we are most interested in the scalability of our algorithm, we focus on the time complexity of MI computation step between the current view and map images, since it is the only step in our algorithm that is dependent on the map size and all the other steps are constant in time complexity. We run the proposed method on maps of different sizes using the Google Street View dataset, and Table II summarizes the average execution time, which includes some of the overhead steps independent of the map size. From these figures, it is easy to infer that the MI computation on 20% of the entire map or 2.4K images takes roughly 0.3 ms and that on 80% of the map or 9.6K images takes roughly 1.19 ms, etc. On average, our algorithm is therefore able to handle approximately 8.4 million images per second or a map of 20 million images in 2.38 seconds, well within real-time constraint of applications that require loop closure detection. To put these numbers in perspective, if the distance traveled by a robot between two consecutive keyframes is one meter, the proposed approach is able to handle an appearance map that covers a distance of 20,000 kilometers in real time.

V. CONCLUSION AND FUTURE WORK

This paper presents a novel and simple method for loop-closure detection with the MI of binarized images at a low resolution, and the method has been shown to be suitable to handle an appearance map with as many as 20 million images in slightly over two seconds. The proposed method does not require offline visual vocabulary construction, as do the popular approaches in visual loop-closure detection based on visual BoW, and the proposed similarity measure is capable of achieving a recall of near 100% with a small top k value with respect to our experimental datasets. This makes our method a competitive choice to detect loop closure on maps with key locations on the order of millions.

Our results and observations are limited by the datasets used in this study. In the future, we will further evaluate our method in other types of environments and employ verification based on multi-view geometry or integrate our method within a Bayesian framework of a complete and practical SLAM system.

REFERENCES

[1] I. Ulrich and I. Nourbakhsh. Appearance-based place recognition for topological localization. IEEE International Conference on Robotics and Automation, San Francisco, USA, pp. 1023-1029, 2000.

[2] A. Angeli, D. Filliat, et al. Real-time visual loop-closure detection. IEEE International Conference on Robotics and Automation, Pasadena, USA, pp. 1842-1847, 2008.

[3] A. Angeli, D. Filliat, et al. Fast and incremental method for loop-closure detection using bags of visual words. IEEE Transactions on Robotics, 24(5), pp. 1027-1037, 2008.

[4] M. Cummins and P. Newman. FAB-MAP: Probabilistic localization and mapping in the space of appearance. The International Journal of Robotics Research, 27(6), pp. 647-665, June 2008.

[5] M. Cummins and P. Newman. Appearance-only SLAM at large scale with FAB-MAP 2.0. The International Journal of Robotics Research, 30(9), pp. 1100-1123, 2010.

[6] M. Calonder, V. Lepetit, et al. BRIEF: Binary robust independent elementary features. The European Conference on Computer Vision, pp. 778-792, 2010.

[7] S. Leutenegger, M. Chli, et al. BRISK: binary robust invariant scalable key points. IEEE International Conference on Computer Vision, Barcelona, pp. 2548-2555, 2011.

[8] D. Galvez-Lopez and J. D. Tardos. Real-time loop detection with bags of binary words. IEEE International Conference on Intelligent Robots and Systems, San Francisco, USA, pp. 51-58, 2011.

[9] Paul Newman, D. Cole, et al. Outdoor slam using visual appearance and laser ranging. IEEE International Conference on Robotics and Automation, Orlando, FL, pp. 1180-1187, 2006.

[10] Liu Y, Zhang H. Visual loop closure detection with a compact image descriptor. IEEE International Conference on Intelligent Robots and Systems, Piscataway, USA, pp. 1051-1056, 2012.

[11] H. Shahbazi and H. Zhang. Application of locality sensitive hashing to real-time loop closure detection. IEEE International Conference on Intelligent Robots and Systems, San Francisco, USA, pp. 1228-1233, 2011.

[12] Y. Liu and H. Zhang. Indexing visual features: real-time loop closure detection using a tree structure. IEEE International Conference on Robotics and Automation, St. Paul, USA, pp. 3613-3618, 2012.

[13] A. Torralba, R. Fergus, et al. Small codes and large databases for recognition. IEEE Conference on Computer Vision and Pattern Recognition, Anchorage, AK, pages 1-8, 2008.

[14] M. Milford, G. Wyeth. SeqSLAM: Visual route-based navigation for sunny summer days and stormy winter nights. IEEE International Conferece on Robotics and Automation, River Centre, Saint Paul, Minnesota, pp. 1643-1649, 2012.

[15] H. Hirschmu. Accurate and efficient stereo processing by semi-global matching and mutual information. IEEE International Computer Vision and Pattern Recognition, pp. 807-814, 2005.

[16] F. maes, A. Collignon, et al. Multimodality image registration by maximization of mutual information. IEEE Transactions on Medical Imaging, 16(2), pp.187-198, 2002.

[17] A. Dame and E. Marchand. A new information theoretic approach for appearance-based navigation of non-holonomic vehicle. IEEE International Conference on Robotics and Automation, Saint Paul, MN, pp. 2459-2464, 2011.

[18] Otsu, Nobuyuki. A threshold selection method from gray-level histograms. Automatica. pp.285-296, 1975.

[19] New College Dataset, http://www.robots.ox.ac.uk/ mobile/IJRR_2008 _Dataset/dataset.html.

[20] MVT3.3 package, http://www.petercorke.com/Machine_Vision_Toolbox .html.

[21] H. Zhang. BoRF: Loop-closure detection with scale invariant visual features. IEEE International Conference on Robotics and Automation, Shanghai, China, pp. 3125-3130, 2011.