Signal Processing 112 (2015) 17–26

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Inertial sensors supported visual descriptors encoding and geometric verification for mobile visual location recognition applications

Yan Zhang [b], Tao Guan [a,*], Liya Duan [b], Benchang Wei [a], Juan Gao [a], Tan Mao [a]

[a] School of Computer Science and Technology, Huazhong University of Science & Technology, China
[b] School of Hydropower & Information Engineering, Huazhong University of Science & Technology, China

ARTICLE INFO

Article history:
Received 21 January 2014
Received in revised form
8 August 2014
Accepted 16 August 2014
Available online 6 September 2014

Keywords:
Mobile visual location recognition
On-device
Location aware encoding
Transform coding
Geometric verification

ABSTRACT

In this paper, two contributions are introduced to improve the accuracy of city scale on-device Mobile Visual Location Recognition (MVLR) systems. Firstly, to compress image descriptors, we design an improved Transform Coding (TC) algorithm for the implementation of a location aware encoding strategy. Compared with traditional encoding algorithms, our algorithm can provide reasonable searching accuracy with very low memory consumption, which makes the implementation of the location aware image descriptors encoding directly on a mobile device come true. Secondly, to perform Geometric Verification (GV) directly on a mobile device, we design a gravity aligned geometric encoding algorithm. The algorithm is not only memory and computation efficient, but also can improve the location recognition accuracy obviously. Experiments on city scale datasets demonstrate the effectiveness of the proposed algorithms.

1. Introduction

With the development of mobile computing techniques, mobile visual location recognition has attracted more and more attention. Due to the limited memory size and processing power of mobile devices, most of the current systems perform MVLR by searching a large database on a remote server. However, some negative factors such as network delay will easily affect the users' experience. Recently, some researchers [1–3] propose to perform MVLR directly on a mobile device, which can provide numerous benefits such as network independent and low-latency image retrieval. For example, in [2], the authors have designed an on-device MVLR system by encoding each image descriptor into a short code. While promising, there are still some problems to be solved to further improve the location recognition accuracy.

Firstly, a global codebook is used to encode database image descriptors. Although the method is compact enough for memory limited mobile devices, it will decrease the searching accuracy obviously.

Secondly, only image descriptors generated from the local features are used in retrieving process. Global information such as geometric information which is proved to be useful for improving retrieval accuracy is neglected absolutely.

In view of the above problems, we make two contributions to improve the accuracy of the MVLR method proposed in [2].

Firstly, we design a modified Transform Coding (TC) algorithm based on which a location aware image descriptors encoding strategy is implemented to improve searching accuracy. Compared with traditional algorithms, our algorithm is more efficient in memory consumption, which

* Corresponding author. Tel.: +86 27 87556154.
E-mail address: qd_gt@126.com (T. Guan).

http://dx.doi.org/10.1016/j.sigpro.2014.08.029
0165-1684/

makes the implementation of location aware encoding strategy directly on a mobile device come true.

Secondly, we design a gravity aligned geometric encoding algorithm for the use of performing geometric verification directly on a mobile device. The algorithm is not only fast and compact enough for on-device implementation, but it also can improve the location recognition accuracy obviously.

The rest of the paper is organized as follows: Section 2 presents a survey of related works. Section 3 presents the location aware encoding method. Section 4 describes the gravity aligned geometric encoding algorithm. Section 5 describes the location recognition algorithm. Section 6 presents experiments and results, and Section 7 draws some conclusions.

2. Previous work

Due to the memory limitation of mobile devices, current MVLR systems [4–8,19–21] commonly rely on client-server architecture in which the city scale image database is stored and searched on a remote server. For example, Chen et al. [4] develop several methods such as histogram equalizing and computing upright descriptors to improve the accuracy of MVLR systems. Ji et al. [5] design a multiple-channel coding scheme to extract compact visual descriptors for low bit rate mobile visual search. Yap et al. [6] propose an efficient mobile landmark recognition framework by incorporate saliency information in the processes of local descriptor generation, vocabulary tree construction and image representation. Chen et al. [7] fuses visual and inertial sensors to improve the location recognition accuracy. Xia et al. [8] design a geometric context preserving progressive transmission method to reduce retrieval latency.

Recently, researchers have proposed to perform image searching directly on a mobile device to obtain low-latency mobile visual location recognition. Schroth et al. [3] propose to perform location recognition directly on a mobile device by downloading the quantizers and inverted files from a remote server in an incremental way. Chen et al. [1] design a computationally efficient image signature called Residual Enhanced Visual Vector (REVV) to enable fast and compact on-device mobile visual search. Guan et al. [2] design a memory efficient image search engine by compress each database image into several bytes at the cost of sacrificing retrieval accuracy to some extent. In this research, we design a modified transform coding algorithm and implement a location aware encoding strategy to improve the searching accuracy of the method proposed in [2].

Geometric Verification (GV) [9] is widely used in Content-Based Image Retrieval (CBIR) systems to improve retrieval accuracy. In GV, features of the query image are firstly matched with features of the database images, and then a RANSAC process is carried out to estimate a geometric transformation between the query image and each candidate image for re-ranking use. However, the RANSAC based GV methods are computation and memory inefficient, which makes them cannot be used for on-device MVLR applications. Recently, some researchers have investigated different ways to facilitate the GV process. For example, Jegou et al. [10] use WGC (weak geometric consistency constraints) to filter local descriptors that are not consistent in terms of angle and scale. Xie et al. [11] design a P-WGC (pairwise weak geometric consistency constraint) method to handle view point changes and non-rigid deformations. Zhou et al. [12] propose a geometric coding algorithm to encode the spatial relationship among local features for large scale image retrieval. Tsai et al. [13] propose a location geometric similarity scoring method that is invariant to rotation, scale, and translation, and can be easily incorporated in mobile visual search systems. While promising, the above methods need to store the information of detected local features, which is impossible for memory limited mobile devices since hundreds of millions of, even billions of local features will be obtained from the city scale workspace. In view of that, we design a memory and computation efficient geometric re-ranking algorithm to improve the accuracy of on-device MVLR systems.

3. Location aware image descriptors encoding based on improved transform coding

3.1. Preliminary discussion

In [2], authors propose an efficient vector encoding algorithm called Transform Coding–Residual Vector Quantization (TC–RVQ) which can provide more accurate searching results than the traditional Transform Coding [14] (TC), Product Quantization [15] (PQ) and Residual Vector Quantization [16] (RVQ) algorithms. However, the method does not fully consider the diversity of city scale database and use a single codebook to fulfill the encoding task. In MVLR systems, the city scale database is usually divided (by considering GPS) into different regions to accelerate the searching process. In view of that, we design a location aware encoding strategy in which the vector quantization and encoding is performed in different regions separately. We think that our strategy will provide more accurate quantization results because the visual variances of a small and geographically adjacent image set will be much smaller than the variances of a large and geographically dispersed database. In fact, the location aware encoding strategy has previously been used in [17] to obtain a compression function within each region for low bit rate transmission purpose. However, the method is not compact enough for on-device MVLR because only the query vector will be encoded to reduce the transmission delay and the index structure on the server will still consume a large amount of memory in case of dealing with city scale database. Our approach encodes both the query and database images, thus can reduce the memory usage of index structures significantly.

To implement location aware encoding directly on a mobile device, we need a memory and computation efficient quantization algorithm to fulfill the vector encoding task within each region. We carry out an experiment to test the computation time and memory consumption of several different algorithms by dividing the Wuhan [2] dataset into 64 regions. In each algorithm, we use 10 bytes to encode each 128 dimensional PCA compressed VLAD [18] descriptor.