

BoVW model in Content Based Image Retrieving

Linta John¹ Smita C Thomas²

^{1,2}Mount Zion College of Engineering Kadamannitta, Pathanamthitta, India

Abstract— A scalable approach for sharing multimedia data across the network is the Peer-to-peer networking. For a large amount of data which is distributed among different nodes, to perform content-based retrieval in peer-to-peer networks is an important but challenging problem. But many of the existing methods are focus on the indexing high dimensional visual features. This type of methods has scalability limitations. By using the bag-of-visual words model found a scalable approach for content-based image retrieval in peer-to-peer networks. By making a comparison with centralized ones, the key problem is to efficiently and precisely obtain a codebook which is global. For a peer-to-peer network, which is often evolves dynamically, and that makes a static codebook which is less effective for the retrieval tasks. Therefore, have to propose a dynamic codebook updating method. This is by optimization of the mutual information between the codebook which is produced and important information. To manage the different code words the workload balance among nodes is necessary. Indexing pruning techniques are used for the improvement in the retrieval process and for reducing the network cost.

Key words: BoVW, DHT, CBIR, Peer to Peer network

I. INTRODUCTION

Peer to peer (P2P) networks is a network which consists of nodes, which are self-organized and have equal privilege. This is the most favourable architectures for data sharing. Networks such as eDonkey1 is the most popular file sharing network which counts millions of users and hundreds of millions of files. Web pages which mostly consist of textual documents like news, articles or forums. The multimedia files have a vital role in most P2P networks which contain multimedia data. For large scale multimedia retrieval applications such as content-based image sharing, and copyright infringement detection, there is a need of P2P because of the computational power of P2P networks. In file sharing P2P networks are well known for their efficiency, scalability and robustness. Also P2P provides search functionality such as content-based image retrieval (CBIR). In comparison with the centralized ones, the data here is distributed among different nodes, so for a CBIR algorithm needs to search for images in a distributed manner. But the nodes in P2P networks have only limited network bandwidth and computation power; because of this thus the algorithm needs to keep the network cost low and the workload. In P2P networks the nodes may join/leave at any time, so need to be update dynamically in order to adapt such changes. Distributed Hash Tables (DHTs) are often implemented on top of a physical network for support content indexing and avoid message flooding. For the efficient routing of messages, the nodes are organized in a structured way. For the functionality of CBIR, many of the existing systems adopt a global approach that is, an image is represented by its high dimensional feature vector like colour histogram, and also the similarity between files and is

measured by using the distance between two feature vectors. Mainly, distributed high-dimensional index or Locality Sensitive Hashing (LSH) is used to represent feature vectors and which is on over the DHT overlay. The bag-of-visual-words (BoVW) model is successfully used for large scale image retrieval. In this BoVW model, every image is represented with a collection of local features. Generally, for the employment of the BoVW model, three steps are required, ie, a number of local regions or key points will be identified from an image. With a high dimensional descriptor every region or keypoint is represented. Scale-Invariant Feature Transform (SIFT) descriptor is the widely used one. Secondly, a codebook is generated, which is for the quantizing of the feature vectors into discrete codewords,. That is an image can be interpreted as a set of feature codewords. Nearest-neighbor quantization is the widely used one. In this quantization each feature vector is represented by its nearest codeword centroid. And at last the codebook forms a Voronoi partitioning of the feature space. Last step is similar to the BoW model, statistical distributions of the codewords in a given image is utilized to represent the image.

II. LITERATURE SURVEY

In this paper [1], each image in bag of visual model represents with a quantized code words which is derived from the images local features. The similarity between images with the BoVW is measured. By an inverted index the retrieval process is supported. There are two categories to distribute index tuples which is document partition and term partition. For document partition, every node maintains an index. By combining the lists of candidate documents returned from each node, this is achieved by, a query will be sent to all index nodes. And with the term partition, every node maintains an index for a subset of terms. Here the query will only send to the nodes which manage corresponding terms. By combining the inverted list returned from them. So document partition has higher network cost while comparing with term partition. Because of this, term partition is a more popular choice in the case of P2P networks. In P2P networks, a BoVW based CBIR system in distributed servers is most relevant to the work. CBIR systems build an inverted index among distributed servers with term partition. To reduce network cost, learning processes are used to first filter the terms and then distribute the terms into different servers to improve workload balance.

For the BoVW retrieval process, the system builds inverted indices over the DHT's hash interface. DHT is a class of structured Peer to Peer overlay networks that provides GET (k) and PUT (k, v) operations. These operations is similar to a hash table, where k, v. And these are the key and value of a table entry. Additionally, DHT handles most issues in node management, including redundancy and failure recovery mechanisms in cases of nodes joining/leaving/ failing, and caching. Therefore, DHT

forms an infrastructure that can be used to build more complex applications.

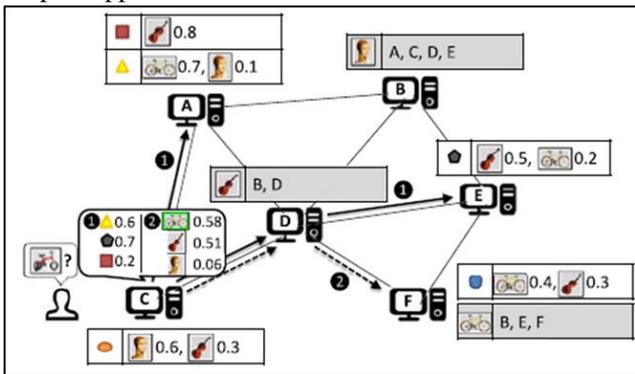


Fig. 1: CBIR Process of a Query over the DHT Overlay Network

In this paper [2], the key points which is extracted from the images. Every image can be represented as a “bag of visual words”. And this representation has been used in scene classification. The representation choices include the dimension, selection, and weighting of visual words. Similar to terms in a text document, an image has local interest points or key points defined as salient image patches or small regions that contain the local information of the image. key points are usually seen around the corners and edges of image objects. That means the edges of the map and around people’s faces. Here use the Difference of Gaussian (DoG) detector to automatically detect key points from images. The detected keypoints are depicted using PCA-SIFT descriptor. This descriptor is a 36-dimensional real-valued feature vector. An image can be represented by a set of key point descriptors, but this set varies in cardinality and lacks meaningful ordering. The bag-of-visual-words representation can be converted into a visual-word vector similar to the term vector of a document. The visual-word vector may contain the relevant information of each visual word in the image.

In this paper[3], the Bag-of-Features (BoF) model has become a popular solution to scalable content-based image retrieval (CBIR) the widely used Scale Invariant Feature Transform (SIFT) descriptor is employed. Creating a codebook: Since the local features are of continuous values, clustering (e.g. k-means) will be applied to them for creating a feature codebook where each codeword corresponds to a cluster. Therefore, an image can be understood as a set of feature codewords. Representing an image: Similar to the BoW model, statistical distributions of the codewords (by assigning each local feature to the closest center of the clusters) in a given image will be utilized as the representation. In this paper we utilize the well-studied *tf-idf* weighting scheme and Cosine distance as the similarity measurement.

III. CONCLUSION

In this paper, for content based image retrieval in peerto-Peer networks, present a bag-of-visual-words model. To overcome the difficulty for generating and maintaining a codebook which is global, BoVW model is used in P2P networks. This method is scalable for sharing images in a

P2P network. For improving the retrieval performance of the proposed method and reduce network cost.

ACKNOWLEDGMENT

We would like to thank, first and foremost, Almighty God, without his support this work would not have been possible. We would also like to thank all the faculty members of Mount Zion College of engineering, for their immense support.

REFERENCES

- [1] L. Zhang, Z. Wang, and D. Feng, “Content-based image retrieval in P2P networks with bag-of-features,” in Proc. IEEE Int. Conf. Multimedia Expo Workshops, Jul. 2012, pp. 133–138.
- [2] J. Yang, Y.-G. Jiang, A. G. Hauptmann, and C.-W. Ngo, “Evaluating bag-of-visual-words representations in scene classification,” in Proc. ACM Int. Workshop Multimedia Inf. Retrieval, 2007, pp. 197–206.
- [3] L. Zhang, Z. Wang, and D. Feng, “Content-based image retrieval in P2P networks with bag-of-features,” in Proc. IEEE Int. Conf. Multimedia Expo Workshops, Jul. 2012, pp. 133–138.
- [4] J. Yang, Y.-G. Jiang, A. G. Hauptmann, and C.-W. Ngo, “Evaluating bag-of-visual-words representations in scene classification,” in Proc. ACM Int. Workshop Multimedia Inf. Retrieval 2007, pp. 197–206.
- [5] C. Tang, Z. Xu, and S. Dwarkadas, “Peer-to-peer information retrieval using self-organizing semantic overlay networks,” in Proc. ACM Conf. Appl., Technol., Archit., Protocols Comput. Commun., 2003, pp. 175–186.
- [6] P. Reynolds and A. Vahdat, “Efficient peer-to-peer keyword Searching,” in Proc. ACM/IFIP/USENIX Int. Conf. Middleware, 2003, pp. 21–40.