Contents lists available at ScienceDirect

Future Generation Computer Systems

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

A distributed image-retrieval method in multi-camera system of smart city based on cloud computing

Jiachen Yang [a], Bin Jiang [a,*], Houbing Song [b,*]

[a] School of Electrical and Information Engineering, Tianjin University, Tianjin, China
[b] Department of Electrical, Computer, Software, and Systems Engineering, Embry-Riddle Aeronautical University, Daytona Beach, FL, United States

HIGHLIGHTS

- The introduction of DFP can improve the processing rate based on cloud computing.
- We put information processing on the platform, and achieve fast image retrieval.
- This paper mainly adopts the method based on fault tolerance mechanism.

ARTICLE INFO

Article history:
Received 21 September 2017
Received in revised form 25 October 2017
Accepted 7 November 2017
Available online xxxx

Keywords:
Smart city
Cloud computing
Distributed image-retrieval method
Multi-camera system

ABSTRACT

In smart city, accurate video sensing information is very critical for the whole city system. However, huge amount of information is the main characteristic of multi-camera system, and we have truly been in big data era for image/video processing. In the development of modern smart city, how to make accurate image retrieval in multi-camera system should be considered seriously. In previous researches, many approaches have been put forward. To the best of our knowledge, there is few research focusing on the method based on distributed fault-tolerant processing (DFP) method. The introduction of DFP will greatly improve the processing rate based on cloud computing, so it will be beneficial to the improvement of this issue. In this paper, we propose a distributed image-retrieval method designed for cloud-computing based multi-camera system in smart city. Through the combination of the cloud storage technology, data encryption and data retrieval technology, we achieve efficient integration and management of multi-camera resources. In this way, the cloud computing network data will be released more quickly, which can provide convenient storage service for users. What is more, experimental results show the scalability and effectiveness of the proposed method, compared with previous processing methods.

1. Introduction

In smart city, urban public monitoring system (UPMS) occupies a very important position [1–5]. For image/video processing issue, researchers made full use of the Internet of things, cloud computing, artificial intelligence, automatic control and mobile Internet technology to finish the large amount of information collection [6–12]. In management support for urban public monitoring system, it requires the ability to perceive, interconnect, analyze, predict, and control image/video processing systems in regions, cities, even greater spatial and temporal boundaries [13,14]. In this way, it can fully guarantee city safety, give full play to the efficiency of all the city infrastructures, improve the operation efficiency of the transportation system, and serve the unobstructed public activity with sustainable economic development [15].

As a very important research topic, image retrieval has been paid more and more attention. Specially, content based image retrieval(CBIR) system has a very broad application prospects. At this stage, the field of content-based image retrieval technology has been widely applied in medical image retrieval, digital library and face recognition. In the research of multi-camera sensing image retrieval, it often combines remote sensing image processing, network transmission technology, image data database technology, pattern recognition, computer vision and other cutting-edge research. As an important source of information, distributed image-retrieval method has become the new hot research field [16].

Image retrieval is a branch of information retrieval technology, which studies how to query relevant images quickly and accurately from a large image dataset [17–21]. The image retrieval technology is mainly divided into text based image retrieval (TBIR), content based image retrieval (CBIR) and semantic based image retrieval

* Corresponding authors.
E-mail addresses: jiangbin@tju.edu.cn (B. Jiang), Houbing.Song@erau.edu (H. Song).

https://doi.org/10.1016/j.future.2017.11.015
0167-739X/

Please cite this article in press as: J. Yang, et al., A distributed image-retrieval method in multi-camera system of smart city based on cloud computing, Future Generation Computer Systems (2017), https://doi.org/10.1016/j.future.2017.11.015.

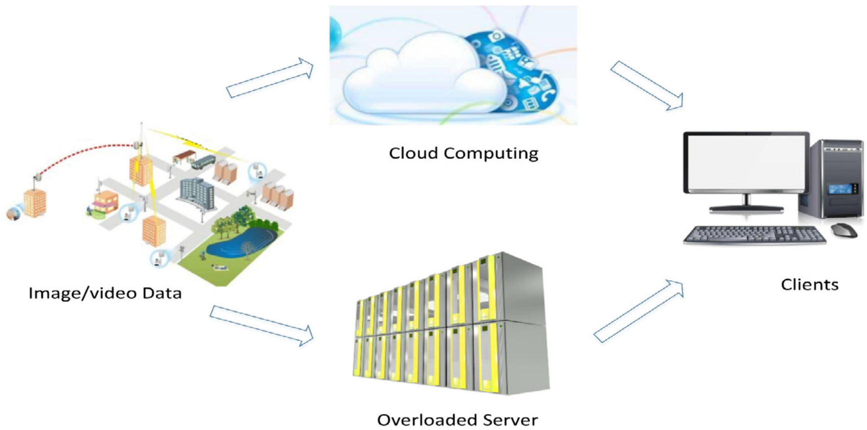

Fig. 1. Multi-camera system of smart city.

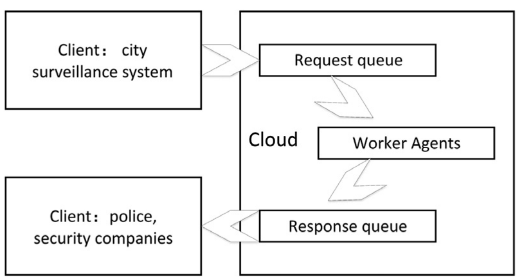

Fig. 2. System diagram.

(SBIR) [22]. In this paper, we mainly consider the CBIR. Different from the traditional data retrieval, image-retrieval in multi-camera system has it own characters. In order to better manage the multi-camera system of smart city, cloud computing will be the best choice, which is shown in Fig. 1. With the increasing size of the cloud computing system, most of the storage nodes are cheap with poor reliability in order to save cost. How to effectively protect the reliability of data has also become the primary concern for the current distributed storage system. In order to provide a reliable storage service, the distributed storage system can improve the fault tolerance of the system by introducing redundant information. This redundancy storage mode can make the system to tolerate a certain number of node failure [23,24]. And the system also needs a good node repair mechanism, which can repair the failure data quickly in the event of failure and keep the maintenance of system redundancy. Large scale distributed storage system has been widely used because of its high storage capacity, high throughput, high availability and low cost [25].

For a typical cloud computing system, it can be divided into three parts: (1) Clients send requests for the special purpose to the cloud computation resource. (2) Cloud computation takes the task and make decisions or retrieval. (3) The final calculation information and analysis results are returned to the clients. The whole process is shown in Fig. 2.

In this paper, image-retrieval method designed for multi-camera system based on distributed fault-tolerant processing (DFP) is proposed, which is processed based on the cloud computing platform. There are three main contributions can be summarized as follows.

(1) To the best of our knowledge, there is few research focusing on the method based on distributed fault-tolerant processing (DFP) method. The introduction of DFP will greatly improve the processing rate based on cloud computing, so it will be beneficial to the improvement of this issue.

(2) The problem of image management in smart cities has been puzzling researchers. In this paper, we put the information processing on the cloud platform, and solve the problem of fast image retrieval by using cloud computing.

(3) In previous studies, the extraction of image features and the formation of feature vectors are important steps. However, the processing speed of this method is not satisfactory. In the specific retrieval process, this paper mainly adopts the method based on fault tolerance mechanism.

The rest in this paper is organized as follows. Section 2 will illustrate the background and motivation based on the related work. In Section 3, the special algorithm framework will be given. In Section 4, the experimental design and results are shown. At last, conclusion will be given in Section 5.

2. Background and motivation

Multi-camera system integrates the public security monitoring, traffic monitoring, road alarm and electronic map system. It is highly integrated and needs to be distributed by way of centralized control with decentralized management. Because of the different infrastructures in different cities, the city monitoring system's size requirements are not the same. In this way, it requires that the system module should have a high degree of independence [26].

In addition, the whole multi-camera system brings a big challenge on the computational capacity for the management. Under normal conditions, the frame rate of the camera is about 20 frames. In this model, the image data of the whole city will be very large, which is more than the capacity of traditional image processing equipment. Cloud computing can solve the problem by offering infinite computing resources apparently [27,28].

In order to improve the whole scalability for the basic framework, many previous works have been done. The proposed image-retrieval method designed for smart city will be based on the special cloud computing workflow. Generally, four main method can be referred for the distributed fault-tolerant system in the multi-camera workflow for cloud computing. So the special content will be discussed as the preparation for the design.

2.1. Multi-camera system based on erasure code

Erasure code is proposed in the field of communication transmission, which can be used in the design of multi-camera system [29,30]. Due to its mathematical characteristics, it has been gradually applied to large scale storage systems, especially in distributed storage environment for data redundancy protection. Compared with the replication strategy, erasure codes can be used

Please cite this article in press as: J. Yang, et al., A distributed image-retrieval method in multi-camera system of smart city based on cloud computing, Future Generation Computer Systems (2017), https://doi.org/10.1016/j.future.2017.11.015.

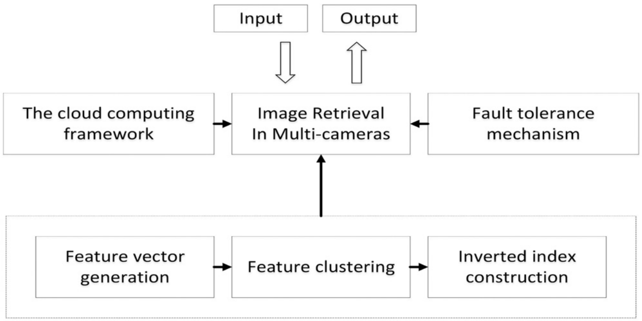

Fig. 3. Proposed workflow.

to minimize redundant storage under the same reliability, and the erasure codes have been widely used in distributed file systems. The basic principle of erasure code is as followed. At first, the original file storage is O, and it is divided into k data blocks at first, which can be denoted as $O_1, O_2, \ldots, O_k$, and then they are encoded for the generation of n coding blocks, which can be denoted as $B_1, B_2, \ldots, B_k$, $n > k$. Finally the n coding blocks are stored in different nodes in accordance with a certain set of rules. When the system has a storage node failure, it can recover the lost data and keep the system redundancy as long as the left encoding block is enough to utilize the remaining encoding block. If any k blocks in the n encoding blocks can reconstruct the original file, the erasure codes can satisfy the maximum distance separable (MDS) characteristics. In addition, the most commonly used encoding method is RS code [31].

2.2. Multi-camera system based on regenerative code

In multi-camera system, the use of regenerative code can improve the robustness of cloud computing storage system. In the last subsection, the erasure codes are based on degree limiting methods, but the regenerative code do not limit the degree of data blocks and redundant blocks. In addition, it can construct the generated matrices by selecting special coding coefficients.

In multi-camera system, it can be used to improve the operation quality of the whole system. Regenerative codes are used to improve the traditional erasure codes so as to reduce the bandwidth consumption of the restoration failure nodes. In principle, the idea of network coding is introduced into regenerating codes. When the failure node is repaired, the nodes involved in the repair process will first make the data in the node linear combination and then upload it. In this mode, the final restoration bandwidth consumption will be controlled at minimum. There are two extreme points in the optimal curve, representing the optimal storage effect and the minimum bandwidth repair effect, which can achieve minimum storage regeneration code and minimum bandwidth regeneration [32].

2.3. Multi-camera system based on replication strategy

Replication strategy is the most simple method to introduce redundancy to multi-camera system. The basic idea for copy strategy is to establish a number of identical copies for each data object in the system and put the copy stored in different nodes. When the data is corrupted or invalid which cannot be used normally, it can be accessed through the nearest storage node to obtain the data backup which is consistent with the original [33,34].

If there is a survival copy for the data object, the distributed storage system can have normal operation. Repair process is also very simple and efficient, as long as requesting the nearest node in the storage [35]. The replication strategy in multi-camera system is simple, which is easy to implement, easy to repair and easy to expand. In addition, multiple copies of the store can also share the load by reading the file, such as the more copies are equipped with hot file configuration to support concurrent read operations. But in complex storage structure with a large number of nodes and large-scale distributed system, we must solve three problems in order to achieve fast and efficient fault tolerant technology: setting replica number, the placement of replica and the replica repair strategy [36].

2.4. Multi-camera system based on locally repairable code

Locally repairable code(LRC) is proposed to improve the storage system. Generally, there are only two types of EC chunks: data chunk and code chunk, each of which is a linear combination of all other data chunk.

In this part, we will give an implementation based on LRC technology [37–39]. It can be seen that the LRC technology reduces the repair bandwidth of RS codes at the expense of an additional 14 percents for storage overhead. But its encoding is still RS code, so the coding efficiency is not improved. In addition, LRC coding does not meet the MDS characteristics, and the system also need to add additional information to mark the two coding data. When the repair of a node is in failure, LRC has very good repair ability. For the repair of two or more than two nodes, they need to connect the k nodes. As the storage system becomes larger and larger, the probability of node failures increases. In addition, in order to solve the problem of fault tolerance in large data storage system, we have improved the structure of memory encoding and obtained lower redundancy cost and more efficient repair [40,41].

3. Distributed image-retrieval method in multi-camera system

For cloud storage system, the distributed cloud storage technology, data encryption technology and data retrieval technology are all very important. Combination of the above key technology not only can achieve the efficient integration and management of network storage resources, but also can provide a friendly external connection window, so that the cloud computing network data can be released more quickly to provide convenient storage service for users. In this paper, we propose a new method for the cloud storage in multi-camera retrieval system, and the following is a detailed description.

In this paper, a distributed image-retrieval method in multi-camera system of smart city based on cloud computing is proposed, which is shown in Fig. 3. And we explain it in three aspects. First of all, the cloud computing framework for multi-camera system is set up. And the designed image retrieval method is given. At last, it will analyze the fault tolerance mechanism in the whole system.

3.1. The cloud computing framework for multi-camera system

For multi-camera system, the cloud computing framework is very important, and it should be set as the special construction. In traditional image processing system, the storage and management of big data is regarded as the core task, which can provide relevant data storage and business access functions for the outside world [42,43].

The distributed storage technology designed for big data cloud storage system makes use of the transmission and storage of data, while the distributed storage resources are formed as a virtual storage device [44,45]. In this section, the distributed cloud storage technology is analyzed based on the directed random rule. According to the rule of directed random walk, the cloud storage system's source data packet arrives at the network node v, and the node u is randomly selected from all the neighbor nodes of v. In detail, the special characters for the four different framework layers will be listed as following.

3.1.1. Display layer

In most data center cloud computing architecture, the display layer is mainly used to show the user's content and service experience in a friendly manner, and it will take advantage of the services provided by the middleware layer below. In the framework of this design, such arrangements are also necessary. In this section, any front-end language can meet the needs of the system, so there is not much to describe here.

3.1.2. Intermediate layer

In the design of the intermediate layer, it provides a variety of services, based on the resources provided by the infrastructure layer, such as caching services and REST services. In addition, these services can be used to support the display layer, or directly to the user call. In this part of the design, we use a multi tenant and parallel processing model to design. Among them, multi tenant is to make a single camera can be more than one service demand, but also to maintain isolation and safety, and by this technology, can effectively reduce the purchase cost and maintenance cost. In addition, parallel processing can process huge amounts of data.

3.1.3. Infrastructure layer

The infrastructure layer provides the computing and storage resources needed for the intermediate layer or the user. In this paper, the cloud computing architecture, including the following special design, in order to better adapt to the multi-camera system. First of all, virtualization can generate multiple virtual machines on a physical server, and can achieve full isolation between these virtual machines, it can not only reduce the purchase cost of the server, but also reduce the maintenance costs of the server. In order to carry large amounts of data and ensure the manageability of these data, the design of the whole architecture needs a whole set of distributed storage system.

3.1.4. Management layer

The management design provides a variety of management and maintenance techniques for the above three layers. Specifically, it can facilitate user login in the safe condition through the good account management technology, and convenient administrator to account management. In addition, it can monitor the performance of virtual machines, services, and applications running at all levels so that they can run in the case of SLA Agreement Level (Service). Other types of management also need to be considered, which are no longer listed here.

3.2. Image retrieval method in multi-camera system

Traditional image retrieval methods are confronted with a lot of problems in the presence of large amounts of image data. In this paper, we propose a large scale image retrieval method based on MapReduce computing model.

In this method, our image retrieval is based on feature vector generation, feature clustering, and image vector inverted index construction in three steps.

3.2.1. Feature vector generation

Since the feature detection and description between images are independent with each other, it is necessary to simply encapsulate the above calculation process into the Map function, and this phase only needs Map part to complete [46]. Here's the MapReduce design for this phase: (1) Map. Enter images that are shaped with image-ID and image-data. The Map function performs SURF algorithm on each image input, extracts the feature vector, and calculates the feature maps. This characteristic number is used to normalize the posterior term frequency. Its output form the construction with image-id. (2) Reduce. The Reduce function acts like an identity, passing only each key pair to the output section. At the end of this phase, a description file containing the feature vectors is obtained [47].

3.2.2. Feature clustering

K-means algorithm is a hard clustering algorithm, whose objective function is a typical clustering method based on the prototype's representative. The prototype data is processed by the optimization objective function, and the extremum method is based on iterative adjustment rules [48]. In this paper, the K-means algorithm takes Euclidean distance as similarity measure. It is the optimal classification of the initial clustering center vectors, which makes the evaluation index J minimum. In addition The algorithm uses the square error sum criterion function as the clustering rule [49,50].

$$V = \sum_{i=1}^{k} \sum_{x_j \in S_i} (x_j - \mu_i)^2 \tag{1}$$

In the retrieval operation of the city image database, the calculation process is as followed: first, it randomly selects samples as the initial cluster center. According to the distance from the center, cluster can be made assigned to each sample, recalculating the K new cluster centers. Then, each sample according to the distance will be assigned to each new cluster. So it will iterate until the target function converges or iterates to a fixed number of steps. This iterative process can be accomplished by repeating the call to the MapReduce task, each starting a MapReduce calculation corresponding to an iteration.

3.2.3. Inverted index construction

In the image retrieval process, each image will be represented as a vector, each of which corresponds to the visual word in the visual vocabulary.

$$TF - IDF_{t,d} = TF_{t,d} * IDF_{t,d} = \frac{n_{t,d}}{n_d} * log\frac{N}{DF_t} \tag{2}$$

When the images are represented as vectors, the similarity between them can be computed by cosine similarity:

$$sim(d_1, d_2) = \frac{V(d_1) * V(d_2)}{|V(d_1)|\,|V(d_2)|} \tag{3}$$

3.3. Fault tolerance mechanism and encryption algorithm in the whole system

In the proposed distributed image-retrieval method in multi-camera system, the special fault tolerance mechanism and encryption algorithm in the whole system is very important.

Data retrieval technology is one of the key technologies of cloud storage system. The performance of the retrieval system not only determines the reliability of the system, but also determines the efficiency of data access [51].

When the user retrieves the cloud storage data, it can restore the original data by accessing any k in the n cloud servers. In this way, the user can retrieve all the selected k cloud servers by coding vector and the coding mark for verification.

If an encoding tag fails, the user will send a report to the third party server while accessing the alternate cloud storage server.

Please cite this article in press as: J. Yang, et al., A distributed image-retrieval method in multi-camera system of smart city based on cloud computing, Future Generation Computer Systems (2017), https://doi.org/10.1016/j.future.2017.11.015.

Table 1
Definition of precision and recall.

	Relevant	Irrelevant	Total
Retrieved	True Positive	False Positive	Predicted Positive
Not retrieved	False Negative	True Negative	Predicted Negative
Total	Actual Positive	Actual Negative	$TP+FP+FN+TN$

If the k storage server code tags are all verified, the user only carries out the belief vector propagation decoding algorithm. At the same time, the data user retrieves the corresponding coding group and the retrieval mark, and it completes the verification of the coding packet integrity. All the original packets of the file M can be recovered by the same calculation process as the coding group. Finally, the M decryption operation is performed on the file. At the same time, the plaintext data is obtained to complete the retrieval of the cloud storage system.

As mentioned before, the data will be stored in a number of data centers, so the data security has become an urgent problem to be solved [52,53]. In this paper, the RAS encryption algorithm will be used. RSA encryption algorithm is a typical public key cryptography algorithm [54].

4. Experimental results and analysis

4.1. Experimental protocol

In the traditional image retrieval system, how to make accurate evaluation on the retrieval is very important. In the previous research, some methods were proposed to assess the experimental results based on different algorithms.

In this paper, two important indexes are introduced for the experiments, which are called as precision and recall. For precision, it can be computed as the ratio of TP and TP +FP. In this concept, TP represents the final true retrieval images number. But FP represents the final false images number. In other word, precision is used to measure the success probability in a large image retrieval system. For recall, it can be computed as the ratio of TP and $TP+FN$. In this concept, TP represents the final true retrieval images number. But FN represents all the other true images in the database. In other word, recall is used to measure the retrieval percentage in a large image retrieval system.

In order to express the two indexed in more detail, we list them in Table 1. In addition, it can be expressed as mathematical equations.

$$precision = \frac{TP}{TP+FP} \tag{4}$$

$$recall = \frac{TP}{TP+FN} \tag{5}$$

Comprehensive evaluation index , $F - measure$, is one of the common evaluation criteria in the field of information retrieval. When we consider the effects of both the precision and the recall, the $F - measure$ is obtained, $F - measure$ is calculated as follows.

$$F - measure = \frac{(\varphi^2+1)\times precision \times recall}{\varphi^2 \times (precision + recall)} \tag{6}$$

In the above formula, φ It is used to adjust the weighted proportion of both the precision and the recall. If $\varphi = 1$, F-measure is $F1$,

$$F - measure = \frac{2 \times precision \times recall}{precision + recall} \tag{7}$$

Obviously, $F1$ is the perfect combination of both the precision and the recall . If $F1$ is greater, the ability of the retrieval system is stronger.

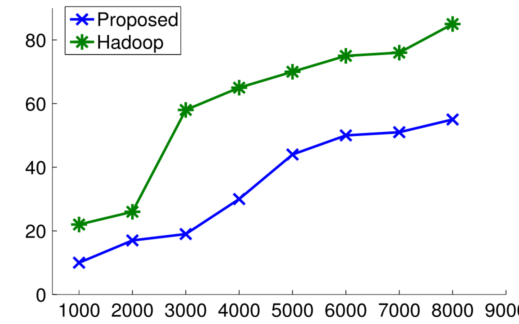

Fig. 4. When the amount of data is small, comparison results for the consumption time.

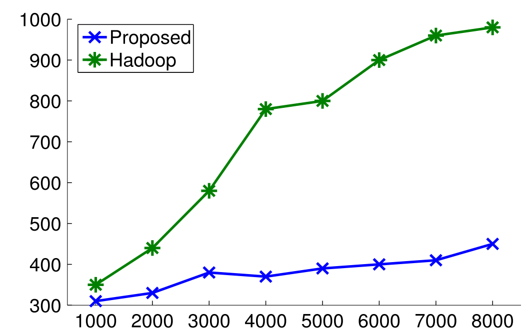

Fig. 5. When the amount of data is large, comparison results for the consumption time.

4.2. Overall performance and comparison results

In multi-camera system, the data scale is very large and it requires the system to be capable in fast image retrieval. In this paper, we choose ImageNet as the simulation database. ImageNet has more than one Nodes, which can be seen as a network. Specially, each node can be regarded as equivalent to a item or subcategory. Currently, a node contains at least 500 corresponding objects for training images. It is actually a large image database for image training. The overall performances are shown in Table 2, and it can be seen that the proposed method can get more previous retrieval results.

4.3. Run time analysis

In the case of a small amount of data, with the gradual increase of the data storage, the method proposed in this paper will be compared with Hadoop system for time consumption. In this part, the time consumed mainly includes reading time, storage time and writing time.

As shown in Figs. 4 and 5, when the data is small, there is little difference between the proposed system and the Hadoop system. But when processing a large amount of data, the time consumed by the system will be significantly lower than that of Hadoop system, which is because the system designed specifically for big data, and the time is very obvious advantages. In addition, the detailed statistic is shown in Table 3.

4.4. System availability analysis

When the number of data nodes is 100, we made statistics on the data distribution of the system and Hadoop system , and the results are shown in Figs. 6 and 7.

Please cite this article in press as: J. Yang, et al., A distributed image-retrieval method in multi-camera system of smart city based on cloud computing, Future Generation Computer Systems (2017), https://doi.org/10.1016/j.future.2017.11.015.

Table 2
The overall performance based on precision.

Data	RSA system	NSA system	BSA system	Hadoop system	Proposed
1000	0.733	0.746	0.762	0.784	**0.793**
1500	0.745	0.762	0.772	0.791	**0.797**
2000	0.752	0.782	0.783	0.803	**0.806**
2500	0.767	0.792	0.791	0.814	**0.817**
3000	0.783	0.810	0.801	0.821	**0.827**
3500	0.791	0.821	0.814	0.832	**0.833**
4000	0.828	0.843	0.824	0.842	**0.845**
4500	0.832	0.851	0.831	0.853	**0.856**
5000	0.841	0.863	0.845	0.871	**0.881**
5500	0.852	0.871	0.864	0.882	**0.892**

Table 3
The result of run time analysis.

Data	RSA system	NSA system	BSA system	Hadoop system	Proposed
1000	19.43	18.32	16.44	13.88	10.91
1500	23.13	21.86	20.38	17.22	15.51
2000	25.37	23.87	21.05	20.14	19.81
2500	26.93	25.32	24.31	21.44	19.62
3000	37.43	34.34	31.44	28.32	21.71
3500	49.43	34.53	32.42	30.12	23.12
4000	51.44	40.23	38.23	31.34	24.68
4500	55.23	45.39	41.33	35.32	27.62
5000	58.53	49.54	45.32	40.23	37.87
5500	59.43	50.56	49.43	43.23	38.54
6000	62.45	58.23	53.23	49.34	42.83
6500	63.56	60.34	57.34	51.32	43.53
7000	66.33	62.34	60.15	53.33	49.22
7500	67.35	65.34	63.83	58.43	53.21
8000	69.34	68.54	61.42	59.23	56.24

Table 4
The result of system availability analysis.

Data	RSA system	NSA system	BSA system	Hadoop system	Proposed
1	61	45	84	59	85
5	41	57	83	94	46
9	45	82	94	46	74
13	50	47	91	60	94
17	54	157	83	94	46
21	43	35	45	47	87
25	53	36	83	47	68
29	95	37	46	84	75
33	74	84	57	35	45
37	56	56	34	58	57
41	45	84	93	84	84
45	73	46	34	35	57
49	84	84	67	36	84
53	46	46	34	93	56
57	93	77	94	46	47
61	35	59	45	36	67
65	83	83	46	98	46
69	46	94	24	57	47
73	94	57	45	84	85
77	85	84	48	45	46
81	46	94	95	92	47
85	94	37	45	36	85
89	67	87	39	74	73
93	46	47	73	54	46
97	46	73	35	44	67

It can be seen from Figs. 6 and 7 that the proposed method can get better result. Compared with the Hadoop system, the data distribution of the system in this paper is more uniform. In other words, the system has a high availability. The distribution of the data greatly affected on the performance of the entire system, if the data distribution is not uniform, unbalanced load will come, which will greatly reduce the usability of the system. In addition, the detailed statistic is shown in Table 4.

4.5. Overall energy consumption of the network

The overall energy consumption of the system is compared with other system such as the Hadoop system , and the results are shown in Fig. 8.

As can be seen in Fig. 8, when the number of stored data is not the same, the overall energy consumption of the two systems varies. The total consumption of this system was significantly lower than Hadoop system. It is mainly because the Hadoop system

Please cite this article in press as: J. Yang, et al., A distributed image-retrieval method in multi-camera system of smart city based on cloud computing, Future Generation Computer Systems (2017), https://doi.org/10.1016/j.future.2017.11.015.

Table 5
The result of overall energy consumption.

Data	RSA system	NSA system	BSA system	Hadoop system	Proposed
1000	29.34	28.75	26.84	23.64	20.83
1500	33.67	31.57	30.34	27.32	25.38
2000	35.44	33.83	31.56	30.56	29.81
2500	36.56	35.57	34.93	31.34	29.56
3000	47.45	44.35	41.73	38.75	31.93
3500	59.56	44.57	42.35	40.25	33.36
4000	61.65	50.93	48.56	41.36	34.46
4500	65.67	55.35	51.93	55.43	37.35
5000	68.79	59.85	55.36	50.57	47.83
5500	69.44	60.46	69.93	53.93	48.34
6000	72.34	68.36	63.36	59.48	52.83
6500	83.78	70.64	67.94	61.73	63.35
7000	86.57	72.26	70.98	63.39	49.47
7500	87.47	85.92	83.46	68.48	63.46
8000	89.26	88.24	81.83	79.39	76.23

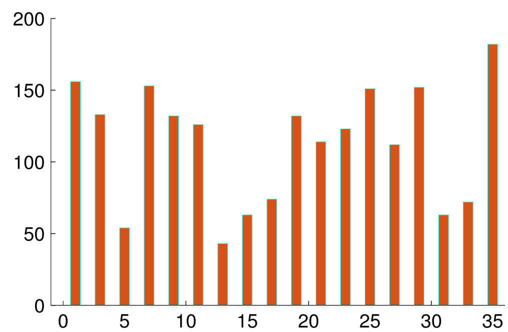

Fig. 6. The data distribution of the method in this paper.

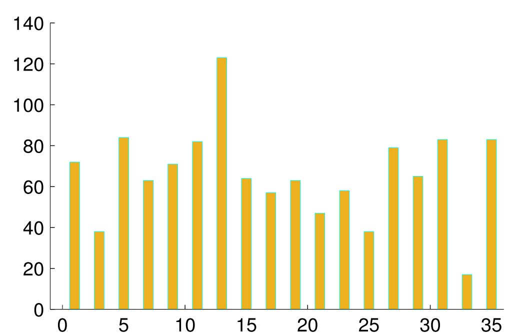

Fig. 7. The data distribution of the method in Hadoop system.

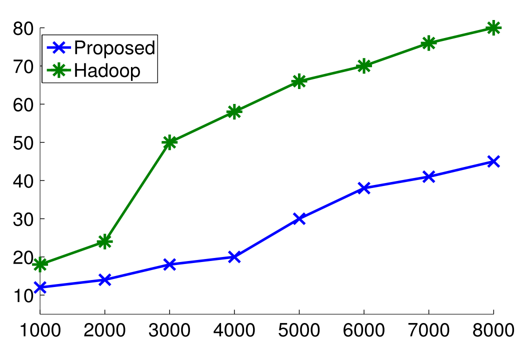

Fig. 8. Comparison of energy consumption for the two systems.

needs to find storage node in the cloud system, which greatly increasing the storage time required to improve the overall energy consumption of the network. In addition, the detailed statistic is shown in Table 5.

5. Conclusion and discussions

In the development of modern smart city, how to make accurate image retrieval in multi-camera system should be considered seriously. In this paper, image-retrieval method designed for multi-camera system based on distributed fault-tolerant processing (DFP) is proposed, which is processed based on the cloud computing platform. There are three main contributions can be summarized as follows. To the best of our knowledge, there is few research focusing on the method based on distributed fault-tolerant processing (DFP) method. The introduction of DFP will greatly improve the processing rate based on cloud computing, so it will be beneficial to the improvement of this issue. The problem of image management in smart cities has been puzzling researchers. In this paper, we put the information processing on the cloud platform, and solve the problem of fast image retrieval by using cloud computing. In previous studies, the extraction of image features and the formation of feature vectors are important steps. However, the processing speed of this method is not satisfactory. In the specific retrieval process, this paper mainly adopts the method based on fault-tolerant processing. Experimental results show the scalability and effectiveness of the proposed method, compared with previous processing methods.

Acknowledgments

This research is partially supported by the Natural Science Foundation of China (No. 61471260), and Natural Science Foundation of Tianjin (No. 16JCYBJC16000).

References

[1] Z. Liu, S. Jiang, P. Zhou, M. Li, A participatory urban traffic monitoring system: The power of bus riders, IEEE Trans. Intell. Transp. Syst. PP (99) (2017) 1–14.
[2] H. Song, R. Srinivasan, T. Sookoor, S. Jeschke, Smart Cities: Foundations, Principles and Applications, Wiley, Hoboken, NJ, 2017, pp. 1–906.
[3] X. Li, J. Niu, S. Kumari, F. Wu, K.K.R. Choo, A robust biometrics based three-factor authentication scheme for global mobility networks in smart city, Future Gener. Comput. Syst. (2017).
[4] L. Wu, Y. Zhang, K.K.R. Choo, D. He, L. Wu, Y. Zhang, K.K.R. Choo, D. He, Efficient identity-based encryption scheme with equality test for smart city, pp(99) (2017) 1–1.
[5] J. Liu, X. Yu, Z. Xu, K.K.R. Choo, L. Hong, X. Cui, A cloud-based taxi trace mining framework for smart city, Softw. Pract. Exp. 47 (8) (2017) 1081–1094.
[6] Z. Xiao, Q. Chen, H. Luo, Automatic scaling of internet applications for cloud computing services, IEEE Trans. Comput. 63 (5) (2014) 1111–1123.
[7] P. Samimi, Y. Teimouri, M. Mukhtar, A combinatorial double auction resource allocation model in cloud computing, Inform. Sci. 357 (2014) 201–216.
[8] S.A. Miraftabzadeh, P. Rad, K.-K.R Choo, M. Jamshidi, A privacy-aware architecture at the edge for autonomous real-time identity re-identification in crowds, IEEE Internet Things J. (2017).
[9] M. Roopaei, P. Rad, K.K.R. Choo, Cloud of things in smart agriculture: Intelligent irrigation monitoring by thermal imaging, IEEE Cloud Comput. 4 (1) (2017) 10–15.

Please cite this article in press as: J. Yang, et al., A distributed image-retrieval method in multi-camera system of smart city based on cloud computing, Future Generation Computer Systems (2017), https://doi.org/10.1016/j.future.2017.11.015.

[10] L. Wang, P. Liu, W. Song, K.K.R. Choo, DUK-SVD: dynamic dictionary updating for sparse representation of a long-time remote sensing image sequence, Soft Comput. (11) (2017) 1–12.
[11] P. Liu, K.K.R. Choo, L. Wang, F. Huang, SVM or deep learning? A comparative study on remote sensing image classification, Soft Comput. (2016) 1–13.
[12] Y.C. Hu, K.K.R. Choo, W.L. Chen, Tamper detection and image recovery for BTC-compressed images, Multimedia Tools Appl. (2016) 1–29.
[13] Y. Li, W. Dai, Z. Ming, M. Qiu, Privacy protection for preventing data over-collection in smart city, IEEE Trans. Comput. 65 (5) (2016) 1339–1350.
[14] Y. Yamagata, H. Seya, Simulating a future smart city: An integrated land use-energy model, Appl. Energy 112 (4) (2013) 1466–1474.
[15] R. Ranjan, L. Zhao, Peer-to-peer service provisioning in cloud computing environments, J. Supercomput. 65 (1) (2013) 154–184.
[16] M.A. Sharkh, M. Jammal, A. Shami, A. Ouda, Resource allocation in a network-based cloud computing environment: design challenges, IEEE Commun. Mag. 51 (11) (2013) 46–52.
[17] P. Srivastava, N.T. Binh, A. Khare, Content-based image retrieval using moments of local ternary pattern, Mob. Netw. Appl. 19 (5) (2014) 618–625.
[18] D. Quick, K.K.R. Choo, Digital forensic intelligence: Data subsets and open source intelligence (DFINT + + mathcontainer loading mathjax OSINT): A timely and cohesive mix, Future Gener. Comput. Syst. (2016).
[19] D. Quick, K.K.R. Choo, Pervasive social networking forensics: Intelligence and evidence from mobile device extracts, J. Netw. Comput. Appl. (2016).
[20] D. Quick, K.K.R. Choo, Big forensic data management in heterogeneous distributed systems: quick analysis of multimedia forensic data, Softw. Pract. Exp. 47 (2016).
[21] D. Quick, K.K.R. Choo, Data reduction and data mining framework for digital forensic evidence: Storage, intelligence, review and archive, Trends Issues Crime Criminal Justice AIC (480) (2014) 1–11.
[22] A. Farruggia, R. Magro, S. Vitabile, A text based indexing system for mammographic image retrieval and classification, Future Gener. Comput. Syst. 37 (7) (2014) 243–251.
[23] Y. Huang, H. Huang, J. Zhang, A noisy-smoothing relevance feedback method for content-based medical image retrieval, Multimedia Tools Appl. 73 (3) (2014) 1963–1981.
[24] Y. Zhuang, N. Jiang, Z. Wu, Q. Li, D.K.W. Chiu, H. Hu, Efficient and robust large medical image retrieval in mobile cloud computing environment, Inform. Sci. 263 (3) (2014) 60–86.
[25] Z. Xia, N.N. Xiong, A.V. Vasilakos, X. Sun, EPCBIR: An efficient and privacy-preserving content-based image retrieval scheme in cloud computing, Inform. Sci. 387 (2017) 195–204.
[26] W. Zhu, C. Luo, J. Wang, S. Li, Multimedia cloud computing, IEEE Signal Process. Mag. 28 (3) (2011) 59–69.
[27] D. Mazza, D. Tarchi, G.E. Corazza, A unified urban mobile cloud computing offloading mechanism for smart cities, IEEE Commun. Mag. 55 (3) (2017) 30–37.
[28] S. Hosseinimotlagh, F. Khunjush, R. Samadzadeh, SEATS: smart energy-aware task scheduling in real-time cloud computing, J. Supercomput. 71 (1) (2015) 45–66.
[29] H.Y. Lin, W.G. Tzeng, A secure erasure code-based cloud storage system with secure data forwarding, IEEE Trans. Parallel Distrib. Syst. 23 (6) (2012) 995–1003.
[30] H.Y. Lin, W.G. Tzeng, A secure decentralized erasure code for distributed networked storage, IEEE Trans. Parallel Distrib. Syst. 21 (11) (2010) 1586–1594.
[31] E. Paolini, G. Liva, B. Matuz, M. Chiani, Maximum likelihood erasure decoding of LDPC codes: Pivoting algorithms and code design, IEEE Trans. Commun. 60 (11) (2012) 3209–3220.
[32] M. Naraghi-Pour, M. Hegde, N. Arora, DPCM encoding of regenerative composite processes, IEEE Trans. Inform. Theory 40 (1) (2002) 153–160.
[33] J. Lygeros, D.N. Godbole, M. Broucke, A fault tolerant control architecture for automated highway systems, IEEE Trans. Control Syst. Technol. 8 (2) (2000) 205–219.
[34] M. Bsoul, A. Al-Khasawneh, Y. Kilani, I. Obeidat, A threshold-based dynamic data replication strategy, J. Supercomput. 60 (3) (2012) 301–310.
[35] C. George, S. Vadhiyar, Fault tolerance on large scale systems using adaptive process replication, IEEE Trans. Comput. 64 (8) (2015) 2213–2225.
[36] B. Nazir, K. Qureshi, P. Manuel, Replication based fault tolerant job scheduling strategy for economy driven grid, J. Supercomput. 62 (2) (2012) 855–873.
[37] D.S. Papailiopoulos, A.G. Dimakis, Locally repairable codes, IEEE Trans. Inform. Theory 60 (10) (2012) 5843–5855.
[38] I. Tamo, D.S. Papailiopoulos, A.G. Dimakis, Optimal locally repairable codes and connections to matroid theory, IEEE Trans. Inform. Theory 62 (12) (2016) 6661–6671.
[39] A.S. Rawat, O.O. Koyluoglu, N. Silberstein, S. Vishwanath, Optimal locally repairable and secure codes for distributed storage systems, IEEE Trans. Inform. Theory 60 (1) (2014) 212–236.
[40] N. Silberstein, A.S. Rawat, S. Vishwanath, Error-correcting regenerating and locally repairable codes via rank-metric codes, IEEE Trans. Inform. Theory 61 (11) (2015) 5765–5778.
[41] M. Shahabinejad, M. Khabbazian, M. Ardakani, An efficient binary locally repairable code for hadoop distributed file system, IEEE Commun. Lett. 18 (8) (2014) 1287–1290.
[42] M. Ali, S.U. Khan, A.V. Vasilakos, Security in cloud computing: Opportunities and challenges, Inform. Sci. 305 (2015) 357–383.
[43] S.K. Garg, S. Versteeg, R. Buyya, A framework for ranking of cloud computing services, Future Gener. Comput. Syst. 29 (4) (2013) 1012–1023.
[44] A. Fox, Computer science. Cloud computing–what's in it for me as a scientist? Science 331 (6016) (2011) 406–407.
[45] H. Zhang, H. Jiang, B. Li, F. Liu, A. Vasilakos, J. Liu, A framework for truthful online auctions in cloud computing with heterogeneous user demands, IEEE Trans. Comput. 65 (3) (2016) 805–818.
[46] H. Yamasaki, T. Shibata, A real-time image-feature-extraction and vector-generation vlsi employing arrayed-shift-register architecture, IEEE J. Solid-State Circuits 42 (9) (2007) 2046–2053.
[47] W.C. Lin, C.F. Tsai, Z.Y. Chen, S.W. Ke, Keypoint selection for efficient bag-of-words feature generation and effective image classification, Inform. Sci. 329 (2016) 33–51.
[48] J.-Y. Jiang, R.-J. Liou, S.-J. Lee, A fuzzy self-constructing feature clustering algorithm for text classification, IEEE Trans. Knowl. Data Eng. 23 (3) (2011) 335–349.
[49] X. Huang, L. Zhang, B. Wang, F. Li, Z. Zhang, Feature clustering based support vector machine recursive feature elimination for gene selection, Appl. Intell. (10) (2017) 1–14.
[50] R.F. Xu, S.J. Lee, Dimensionality reduction by feature clustering for regression problems, Inform. Sci. 299 (C) (2015) 42–57.
[51] R. Porter, A. Ronen, Y. Shoham, M. Tennenholtz, Fault tolerant mechanism design, Artif. Intell. 172 (15) (2008) 1783–1799.
[52] I.P. Egwutuoha, D. Levy, B. Selic, S. Chen, A survey of fault tolerance mechanisms and checkpoint/restart implementations for high performance computing systems, J. Supercomput. 65 (3) (2013) 1302–1326.
[53] M.E. Gomez, N.A. Nordbotten, J. Flich, P. Lopez, A. Robles, J. Duato, T. Skeie, O. Lysne, A routing methodology for achieving fault tolerance in direct networks, IEEE Trans. Comput. 55 (4) (2006) 400–415.
[54] A.S. Prasad, S. Rao, A mechanism design approach to resource procurement in cloud computing, IEEE Trans. Comput. 63 (1) (2014) 17–30.

Jiachen Yang received the M.S. and Ph.D. degrees in communication and information engineering from Tianjin University, Tianjin, China, in 2005 and 2009, respectively. He is currently a professor at Tianjin University. He was a Visiting Scholar with the Department of Computer Science, School of Science, Loughborough University, U.K. His research interests include image processing, multimedia quality assessment and cloud computing.

Bin Jiang received the B.S. and M.S. degree in communication and information engineering from Tianjin University, Tianjin, China, in 2013 and 2016. He is currently pursuing the Ph.D. degree at the School of Electrical and Information Engineering, Tianjin University, Tianjin, China. His research interests include image processing, multimedia quality assessment and cloud computing.

Houbing Song received the Ph.D. degree in electrical engineering from the University of Virginia, Charlottesville, VA, in August 2012. In August 2017, he joined the Department of Electrical, Computer, Software, and Systems Engineering, Embry-Riddle Aeronautical University, Daytona Beach, FL, where he is currently an Assistant Professor and the Director of the Security and Optimization for Networked Globe Laboratory (SONG Lab, www.SONGLab.us). He was a faculty member of West Virginia University from August 2012 to August 2017. He has served as an Associate Technical Editor for IEEE Communications Magazine since 2017. He is the editor of 4 books and the author of more than 100 articles. His research interests include cyber-physical systems, internet of things, cloud computing, big data analytics, connected vehicle, wireless communications and networking, and optical communications and networking. Dr. Song is a senior member of the ACM.

Please cite this article in press as: J. Yang, et al., A distributed image-retrieval method in multi-camera system of smart city based on cloud computing, Future Generation Computer Systems (2017), https://doi.org/10.1016/j.future.2017.11.015.