Neurocomputing 181 (2016) 122–131

Contents lists available at ScienceDirect

Neurocomputing

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

TMR: Towards an efficient semantic-based heterogeneous transportation media big data retrieval

Kehua Guo [a], Ruifang Zhang [a], Li Kuang [b,*]

[a] School of Information Science and Engineering, Central South University, Changsha, China
[b] School of Software, Central South University, Changsha 410075, China

ARTICLE INFO

Article history:
Received 11 February 2015
Received in revised form
8 May 2015
Accepted 1 June 2015
Available online 5 December 2015

Keywords:
Heterogeneous media retrieval
Ontology
Fuzzy matching
Information retrieval
Big data
Intelligent transportation

ABSTRACT

In media retrieval system for intelligent transportation, media data variety and heterogeneity have been one of the most critical features. Documents with different formats may express similar semantic information, thus, searching documents reflecting users' intention has been a crucial and important task. For solving this problem, this paper proposes a novel semantic-based heterogeneous transportation media retrieval (TMR) approach to improve the performance. TMR supports the function of retrieving various media types such as image, video, audio and text by using a single media type. Firstly, semantic fields are extracted from the user annotating and automatic learning to express the users' intention. Secondly, ontology is used to represent the semantic fields of a media, and the ontology represented semantic information is saved together with the media document data. Thirdly, the semantic field adjustment process is described. Finally, fuzzy matching is employed to measure the similarity between the users' intention and media documents. For the returned results, we carry out the performance evaluation models in comparison with the existing approaches. Experimental result indicates the superiority of TMR in term of precision rate, computing speed, storage cost and user experience.

1. Introduction

In intelligent transportation system, due to the increase in the number of various media documents such as images, videos, audios and text files, media storage and retrieval has become a crucial function. With the emerging development of information technology, the transportation media contents are used growing frequently in every day computing. Retrieving the information reflecting users' intention from huge amounts of media data has been a crucial issue.

In transportation field, the media documents are various and heterogeneous [29]. Based on this particularity, the current media information retrieving techniques are facing two major challenges. Firstly, to meet the users' query intention, heterogeneous media based retrieval has become a crucial requirement. Although traditional approaches can return the result by low-level feature based extraction and matching, low-level features are facing great challenges because of the type diversity [40]. The type of media documents could be different due to the devices, methods or even applications which the users choose. Even in the same media type, the data format may be various [26,25]. Secondly, the expression of users' query intention has been a difficulty [15]. For example, in some widely used media retrieval systems, the query intention is represented by text. However, the text-based query intention expression is extremely limited in most of cases. The users seldom type long keywords to express their intention, so the insufficient keywords may cause ambiguity. In other cases, the query intention is described by using content (e.g. searching traffic scenes by uploading similar images). Nevertheless, such content-based retrieval does not express the personal understanding, since semantic information is not described by the physical features such as color, texture or shape. Thus, how to search heterogeneous media documents reflecting users' query intention is a crucial question in media retrieval research.

Generally speaking, users' query intention is a fuzzy definition and not able to be accurately expressed, even some information is incomplete [37,14]. In transportation media retrieval system, a large number of documents which reflect the query intention will be returned. Since these documents will be displayed to users, how to measure the similarity between users' intention and media documents has been a crucial issue. Some approaches such as Euclidean distance, probability and Support Vector Machine (SVM) have indicated positive contributions [19,38]. However, the approach requires a host of parameters to be adjusted heuristically for the optimal result and how to deal with the incomplete knowledge [30,39] is a difficulty.

Therefore,in transportation media retrieval, it becomes a significant issue to solve type heterogeneity for achieving perfect retrieval performance. At present, semantic-based retrieval and fuzzy approaches have been widely used into the information retrieval to deal with ill-defined features [36]. These approaches

* Corresponding author.
E-mail address: kuangli@csu.edu.cn (Li Kuang).

http://dx.doi.org/10.1016/j.neucom.2015.06.101
0925-2312/

have been proved as a powerful tool for representing and dealing with imprecision, vagueness and ambiguity arising from measurements and judgments. Several contributions have been achieved using these approaches to weight or rank inaccurate values [23,2]. In consideration of the inaccuracy of users' query intention, it is believed that semantic-based representation and fuzzy matching are suitable for measuring the similarity of inaccurate semantic information to obtain more expected results.

In this paper, we report a transportation media retrieval approach called TMR, which extract semantic information and query media files by using fuzzy ontology matching. The proposed TMR differs from other existing methods in several ways. The main innovations of this approach are summarized as follows: (1) TMR supports heterogeneous retrieval for any uploaded images, videos, audios or text documents; (2) TMR does not require low-level features extraction in retrieval process, which may increase processing speed; and (3) TMR remains semantic information even if the media documents are copied, cut or leave the media database. The experiment result shows that TMR can search heterogeneous media documents reflecting users' query intention more effectively compared with some traditional approaches.

The remaining sections of the paper will discuss about the details of TMR. Section 2 details about related work of current media retrieval approaches. Section 3 details down the description of TMR and the procedure of media retrieval step by step. Section 4 discusses about the experiment result and performance evaluation. Finally, Section 5 concludes this paper and proposes the future work.

2. Related work

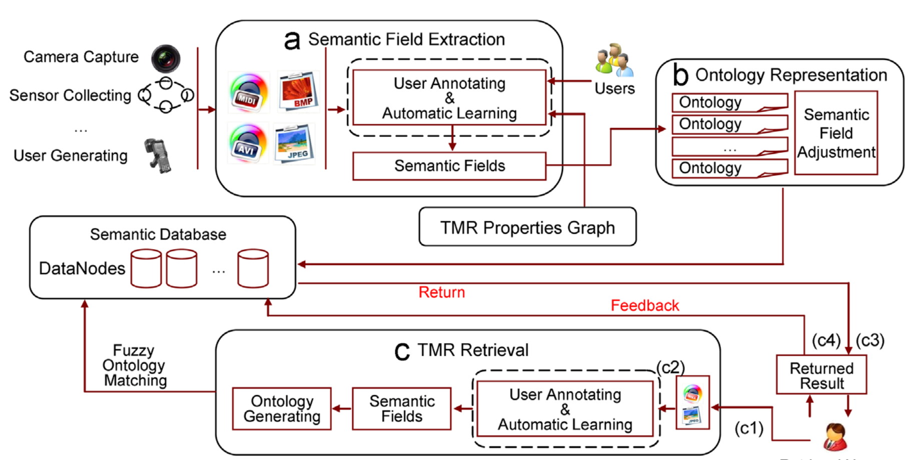
In the past decades, heterogeneous media retrieval was mainly using text-based approaches. But in fact, this retrieval is only based on the text related to media documents in host files. Although we can type keywords to obtain various types of media documents, these methods are intrinsically not heterogeneity supported. Because the retrieval could not be executed if the text is unavailable, the current research on text-based approaches is focused on combining the text-based approach with other methods [32,24]. Nevertheless, certain disadvantages exist in this approach [20,21]. Firstly, in many cases, because only short or ambiguous keywords are used, the users' query intention is not able to be correctly described. For example, if a user inputs the keyword 'John F. Kennedy', the retrieval system may return results with airports, roads or individuals. Secondly, this method cannot capture the documents in databases that do not contain explanatory texts to the media contents.

Meanwhile, some other retrieval systems developed content-based retrieval. In this approach, user uploads a file (e.g. image), the retrieval system uses content-based approaches to search the documents which are similar to it. This approach will suffer from three disadvantages. Firstly, content-based approach cannot support heterogeneous search (e.g. upload an image to search a video). Secondly, the retrieval system will ignore the users' query intention and cannot get any expected similar results. Thirdly, feature extraction will cost much computation resource. Hence, heterogeneous retrieval is arduous to be realized based on feature extraction approaches [31,41], and full heterogeneous media retrieval has not been achieved.

Actually, feature recognition is based on not only the low-level visual features such as color, texture or shape, but also the semantic information [28,9]. Semantic annotation was used in some researches to improve the media retrieval performance. Datta et al. [6] presented comprehensive surveys for the user feedback in image retrieval. Representative work includs active learning approach [13], local geometrical graph approach [4] and Biased Discriminant Analysis (BDA) analysis [35], etc. Some papers have obtained promising results and developed prototype systems [16,1].

In recent years, some researches using semantic-based approaches to support heterogeneous media retrieval have been proposed. References [20,21] proposed a scheme that is able to enrich textual answers with appropriate media data which can automatically determine which type of media information should be added for a textual answer. To achieve various types supported retrieval, Lu et al. [18] proposed IBCR (Indexing-based Cross-Media Retrieval) approach. For text-image retrieval, Costa et al. [5] learned the correlations between text and image with canonical correlation analysis. However, these approaches only support the heterogeneous retrieval between image and text documents. Some other approaches dealing with heterogeneous media retrieval are proposed [33,34], but they may suffer from the drawback of computation resource cost such as visual features extraction and sample training.

At present, a multitude of contributions are proposed for the similarity measurement. In these approaches, Euclidean distance [17] is the most widely used one in various fields. In addition, Mahalanobis distance [22] and multiple neighborhood similarities [27] have also been applied into the similarity measurement. Although these approaches have achieved promising results, to ontology similarity measurement, sometimes their performance is not satisfactory to reflect users' intention due to the incomplete information. In the retrieval system, users prefer to provide descriptions in inaccurate expressions rather than in numeric values. To solve this problem, a feasible solution is to use fuzzy matching. In fact, fuzzy approaches have been applied in many fields of information retrieval such as



The architecture contains step (a),(b) and (c) , step (c) includes 4 subsections.

Fig. 1. TMR overview.

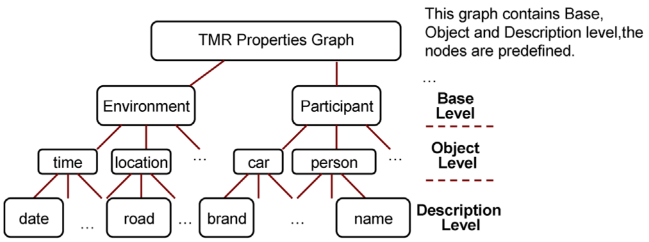


Fig. 2. TMR properties graph overview.

interactive image retrieval [8], content-based image retrieval [3], decision support system [2] and Web resources retrieval [7].

In current approaches, semantic features are stored in the server's database and separated from the media documents. In this case, if the media documents leave the knowledge base, rebuilding the semantic information would be arduous. Therefore, designing a semantic-based media retrieval approach which does not depend on the knowledge base has been a key problem. The proposed architecture will save semantic fields together with the media document data and use fuzzy matching approach to support the heterogeneous media retrieval in transportation information system.

3. Methodology

TMR adopts a three-step architecture as shown in Fig. 1. The architecture consists of following steps: semantic field extraction (Fig. 1(a)), ontology representation (Fig. 1(b)) and TMR retrieval (Fig. 1(c)).

In the first step, the media files will be obtained from various generating source, such as camera capture, sensor collecting and user generating, etc. The media types may include images, videos, audios and text with various formats. The semantic information will be initialized by two ways: (1) user annotating, which means extracting semantic fields from annotations provided by system users; and (2) automatic learning, which denotes analyzing semantic information from media features. After the semantic extraction, the semantic fields together with the media location will be represented by ontology in the second step. A semantic field adjustment is employed to adjust the weights of semantic fields.

In the third step, users may upload annotated media documents with arbitrary format to execute the heterogeneous media retrieval (Fig. 1 (c1)). The semantic information of the uploaded file(s) can also be extracted by user annotating and automatic learning (Fig. 1 (c2)). Then TMR will execute the ontology generating and fuzzy ontology matching to perform retrieval. After the retrieval, the engine will return the results as the thumbnails with the file locations (Fig. 1 (c3)). Finally, system users will be asked to give additional feedback to the media documents they selected (Fig. 1 (c4)) to make the annotations more abundant and accurate.

3.1. Semantic field extraction

In the system addressed above, we assume that there are N media documents represented as $C = M_1, M_2, \ldots, M_N$. An original media document can be defined as follows:

Definition 3.1. $\forall M_i \in C$, M_i is represented as $\{M_i(d), M_i(t), M_i(p)\}$, where $M_i(d)$ is the media data of M_i, $M_i(t)$ is the text description of M_i and $M_i(p)$ is the storage path of M_i.

Given $M_i \in C$, semantic field extraction process gets the semantic fields through user annotating and automatic learning. During the using of TMR, users can manually annotate media documents using the software interfaces proposed in our previous research [10,11]. In order to restrict the scope of annotation, we use a predefined properties information (TMR properties graph) to identify the semantic fields. This graph contains a multitude of property names related to transportation. Fig. 2 shows an overview of the TMR properties graph.

It is obvious that not all the semantic fields can accurately describe the semantic information of the media document. In order to dynamically update the semantic fields, a weight wu is defined for each semantic field. After the user annotating, we can construct a semantic collection as

$$S_u = (su_i, wu_i) | 1 \le i \le m_1 \tag{1}$$

where m_1 is the number of the semantic fields, su_i is a semantic field. $\forall wu_i \in S_u$ the initial value of wu_i satisfies:

$$wu_i = \frac{1}{m_1} \tag{2}$$

After removing stop words (e.g. 'and', 'for', etc.) from all the nodes in TMR properties graph, $M_i(t)$, which is supplied by the content provider and embedded in the host document, will be processed by an algorithm to automatically annotate the media file. The algorithm is presented as Algorithm 1.

Algorithm 1. Automatic Semantic Extraction.

input : $M_i(\{M_i(d), M_i(t), M_i(p)\}) \in C$;
  G; //TMR Properties Graph
output: $S_l = \{(sl_i, wl_i) | i \ge 1\}$
$q = null$; //a queue
Search G Using Breadth First Algorithm;
for *each node t in G* **do**
  $temp = tf(\frac{t}{M_i(t)})$;
  if $temp > 0$ **then**
    $q.add(t)$;
  end
end
descended sort q according to $tf(\frac{t}{M_i(t)})$;
$m_2 = sizeof(q)$;
for $i = 1$ *to* m_2 **do**
  $sl_i = q.pop()$;
end
return sl;

In Algorithm1, $tf(\frac{t}{M_i(t)})$ represents the term frequency of t in $M_i(t)$.

$\forall sl_i \in S_l$, wl_i is assigned as follows:

$$wl_i = \frac{tf\left(\frac{sl_i}{M_i(t)}\right)}{\sum_{j=1}^{m_2} tf\left(\frac{sl_i}{M_i(t)}\right)} \tag{3}$$

After the extraction, $\forall M_i \in C$, we can combine S_u and S_l to construct a semantic matrix S for a media document as follows:

Definition 3.2. Semantic Matrix(SM): Given $M_i \in C$, an SM is a matrix satisfying:

$$S = \begin{pmatrix} s_1 & s_2 & \dots & s_n \\ w_1 & w_2 & \dots & w_n \end{pmatrix} \tag{4}$$

where $n = m_1 + m_2$.

3.2. Ontology representation

TMR uses ontology technology to describe the media semantic information. In ontology, global features of an object are expressed as a single ontology, then local features are described as sub-ontology objects, the definition of ontology meets the recursive and hierarchical relation [12].

Ontology representation aims at mapping semantic matrix S to an ontology object. Ontology nodes at the previous level are used to represent the most obvious features. The next and other levels of semantic fields will be provided based on the previous levels. Fig. 3 shows the ontology generating process of an image.

We can see from Fig. 3 that the original media document is annotated by users annotating and automatic learning, it satisfies tree structure. Every node in the tree has been assigned a weight according to the previous sections. Then the semantic information can be represented as ontology nodes. Finally, to represent the recursive and hierarchical structure, TMR adopts the tree structure and the composite pattern as the data representation. In this pattern, client can use the same method to deal with complex elements as with simple elements, so that the internal structure of the complex elements will be independent with the client program [10].

The definition of semantic tree is proposed as follows.

Definition 3.3. Semantic Tree(ST): Given $M_i \in C$, an ST is a tree satisfying:

1. $\forall node \in ST$, $node = (s, w)$, where w is the weight of semantic field s.
2. $\forall node_1, node_2 \in ST$, the relationship between $node_1$ and $node_2$ satisfies TMR Properties Graph.

The process to construct the ontology is presented as Algorithm 2.

Algorithm 2. Construct the Ontology.

input : $S = \{(s_i, w_i) | 1 \le i \le n\}$;
G //TMR Properties Graph
output: ST //Semantic Tree
$ST = null$;
Search G Using Breadth First Algorithm;
for *each node t in G* **do**
  if $t \in S$ **then**
    $node = (t, w)$;
    $ST.add(node)$;
    $node.children = t.children$;
  end
end
return ST;

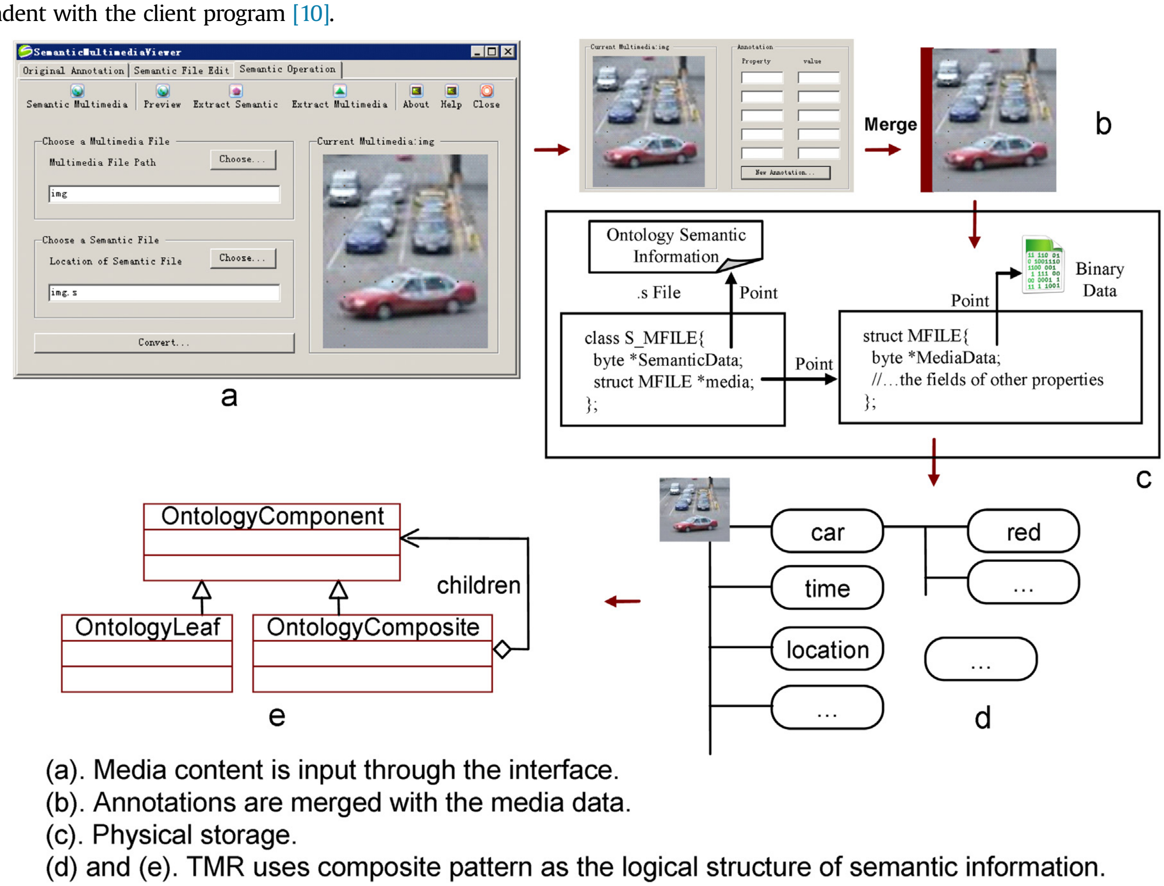


Fig. 3. An Example of Ontology generating process

3.3. Semantic field adjustment

In TMR, ontology objects are stored together with the corresponding media document. For saving storage and guaranteeing the processing speed, we directly save the binary data of ontology object at the head of the media document. In the media retrieval process, we extract the ontology data for matching. Hence, through the annotation and retrieval interface, users can visit the database and upload an annotated media document with arbitrary format to execute the heterogeneous media retrieval. The path of real media files will be returned from the media server.

It is obvious that more frequently used fields can express semantic information more accurately, and they should have a relatively higher weight. Therefore, an adjustment strategy is designed to dynamically change the weight w_i of the annotation after a retrieval R. The strategy is described as Algorithm 3.

Algorithm 3. Semantic Field Adjustment.

input: $S = \{(s_i, w_i) | 1 \le i \le n\}$;
ST; //Semantic Tree
Adjustment:
update w_i:

$$w_i = w_i + \frac{y_i}{n} \tag{5}$$

Feedback:
for all newly added s:

$$w = \frac{1}{n} \tag{6}$$

Refinement:
$\forall s_i \in S$, eliminate s_i when w_i satisfies:

$$w_i < \frac{\theta}{n}\sum_{j=1}^{n} w_j \tag{7}$$

In Algorithm 3, θ is a parameter and $0 \le \theta \le 1$. y_i satisfies:

$$y_i = \begin{cases} 1 & M \text{ is retrieved by } s_i \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

After the semantic field adjustment,the corresponding weight will be updated in ST. We can see from the algorithm that a refinement is conducted to eliminate the annotations which are wrong or less frequently used, and retain the annotations which are used at high frequency. The refinement step in the algorithms will cost much computation resource, we will execute it in a long time interval (e.g. every 24 hours) as background threads.

3.4. Fuzzy ontology matching

Similarity measurement is the most crucial issue in retrieval process. The matching between two ontology objects is arduous because semantic fields are inaccurate to express users' intention and different fields have different weight values.

In TMR, fuzzy matching is employed to measure similarity between two ontology objects. Every semantic field cannot accurately express all the features of the media document, so we map the ontology as a fuzzy set as follows:

Definition 3.4. Ontology Fuzzy Set(OFS): Given S and the corresponding ontology object O, an OFS $\tilde{O}$ is a fuzzy set satisfying:

$$\tilde{O} = \{(s, \mu_{\tilde{O}}(s)) | s \in S\} \tag{9}$$

where $\mu_{\tilde{O}}(s)$ is the membership function of s:

$$\mu_{\tilde{O}}(s) = w(s) \tag{10}$$

The definition shows that traditional definition of fuzzy set has been slightly modified in TMR. We can see from the formula that the value range of the membership function $\mu_{\tilde{O}}(s)$ is (0,1), not [0, 1]. However, $\mu_{\tilde{O}}(s)$ can guarantee that the larger value is able to possess the more accurate capability of expressing semantic information.

A multitude of weight values are assigned to apply the similarity measurement. In fact, they have a possibility distribution. We propose that the measurement can be represented by fuzzy similarity. Given two ontology objects O_1 and O_2, we define the intersection set of the two fuzzy sets as follows:

$$IS(\tilde{O_1}, \tilde{O_2}) = \{s | s \in \tilde{O_1}, s \in \tilde{O_2}\} \tag{11}$$

The fuzzy similarity function is defined as follows:

$$sim(\tilde{O_1}, \tilde{O_2}) = \sum(\mu_{\tilde{O_1}}(s) + \mu_{\tilde{O_2}}(s)) \tag{12}$$

where $s \in IS(\tilde{O_1}, \tilde{O_2})$.

We perform the fuzzy ontology matching as Algorithm 4.

Algorithm 4. Fuzzy Set Matching.

input : $C = \{\tilde{O}_1, \tilde{O}_2, ..., \tilde{O}_N\}$;
$\tilde{O}$; // a sample
output: q; //a queue
Match:
$q = null$;
for i=1 to N **do**
  $sim = sim(\tilde{O}_i, \tilde{O})$;
  $q.add(\tilde{O}_i, sim)$;
end
Rank:
descended sort q according to sim;
return q;

From the algorithm, we can see that given a sample ontology object O, we want to retrieve the result from the database, the similarity value between O and every ontology will be stored in a queue. Obviously, greater value represents greater similarity. We will sort the queue in descending order. Because the ontology is saved together with the media document, the result will be displayed in the same order as the sorted queue.

4. Experiment result

4.1. Datasets and experiment tools

A wide variety of media types are required to conduct the heterogeneous transportation media retrieval applications. We use various quintessential transportation media files together with the related annotations in TMR. The experiment uses the media database which includes 40,000 media documents, including 10,000 images, 10,000 videos, 10,000 audios and 10,000 text documents. The documents are gathered from the transportation media database of a city, all the documents are from real environment. The annotations are obtained through users annotating and automatically analyzing the text from the host document.On the one hand, users can provide text tags on the transportation media files in the proposed interfaces. On the other hand, the approach proposed in Algorithm 1 is employed to analyze the text descriptions to extract the semantic fields. For each media file, the document data, file location and semantic information are used to establish the dataset.

In order to verify the effectiveness of TMR method, this paper develops some software modules. In TMR, the software tools and running environment are as follows. (1) Annotation interface. The PC (Personal Computer) version is based on Microsoft Foundation Classes (2.0 GHZ CPU, 1 GB RAM (PC)) and the MD (Mobile Device) version is based on Android4.0(1.2 GHZ CPU, 1 GB RAM). (2) Retrieval interface. The retrieval interface is based on HTML5 (Browser) and Android 4.0 (MD). (3)Server configuration. Media database is based on Oracle 11 g(Xeon E7-4820 2 GHZ,16 GBRAM), Web Server is Tomcat 6.0. In real environment, there is a time period with less traffic every day, to simulate this case,in the server, background processes will be executed every 24 hours. Considering the statistical proportion of annotation noise, θ in Algorithm 3 is selected as 0.8.

4.2. Heterogeneous media retrieval result

In the retrieval, an annotated media document (e.g. the sample document shown in Fig. 3) is firstly uploaded to the retrieval system. Then all the semantic media documents with similar semantic information to the sample document are searched by the engine.

We use precision rate to evaluate the effectiveness of TMR. For each retrieval process, the retrieved result set is defined as set D_t and all the relevant media documents are defined as set D_l. The precision rate is computed by the proportion of retrieved relevant media documents in total retrieved documents. Therefore, the precision ratio p can be defined as follows:

$$p = \frac{|D_l \cap D_t|}{|D_t|} \tag{13}$$

In order to demonstrate the precision rate of TMR, in the first experiment, we perform the retrieval process by submitting one media type to search the four types (e.g. use video to search images, videos, audios and text documents). For every media type, we perform 10 different retrievals and compute the average precision rates. The 10 sample files are randomly chosen from the database. The average precision rates are demonstrated in Fig. 4.

Fig. 4 illustrates that even in the retrieval process between different media types, the precision rates are not reduced (precision rates are stably from 75% to 85%). This is because TMR completely abandons the physical feature extraction, and executes the retrieval process based only on semantic fields.

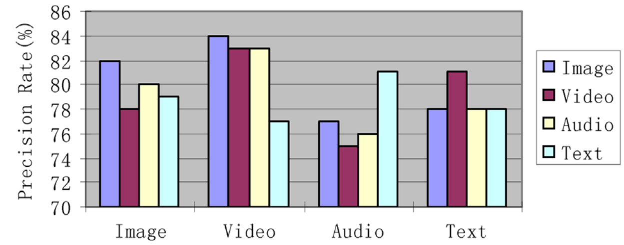


Fig. 4. Averageprecision rates of heterogeneous retrieval

4.3. Performance comparison with existing approaches

The second experiment will illustrate the precision comparisons between our fuzzy matching and two other quintessential measurements: Euclidean distance (ED) and Mahalanobis distance (MD). We use four media types (image, video, audio and text) in the experiment, for every type, 10 heterogeneous retrievals are executed. For different scale of returned result, we compute the average precision rates. Feedback is not considered in all retrievals. The average precision rates are illustrated in Fig. 5.

It can be seen from Fig. 5 that for every sample type, TMR achieves better retrieval precision rates compared with some other similarity measurements. In particular, with the increase of $|D_t|$, the advantages of TMR will become more apparent.

4.4. Time cost evaluation

Time cost includes two factors. The first factor is the time cost of data process. In TMR, several background processes are time consuming. The background process time is defined as follows:

$$t_b = t_{pre} + t_{adj} \tag{14}$$

where t_{pre} is the preprocess time (convert the semantic fields to ontology) and satisfies

$$t_{pre} = \sum_{i=1}^{N} t_{pre}^{i} \tag{15}$$

t_{adj} represents the semantic field refinement time (eliminate the redundant or error semantic information and add the new semantic information from the feedback). As discussed above, in order to carry out the retrieval process, TMR has to perform several background processes whose time cost is t_b, which includes t_{pre} and t_{adj}. Fig. 6 shows the time cost of the background processes.

It can be seen from the figure that t_{pre} and t_{adj} cost a host of seconds (t_{pre} maximum close to 11 seconds and t_{adj} maximum close to 4 s). However, the background processes are not invariably executed. Preprocess is executed once for initialization, and semantic field refinement is executed every 24 h in background threads.

The second factor is retrieval time. Define t_r as the time cost for retrieval. In fact, t_r includes extraction time (extract semantic information from the database) and the matching time (match the semantic similarity between the sample document and the stored files). Next, we measure the retrieval time t_r. We specially record the time cost of 40 retrieval processes. For every document type (image, video, audio and text), we perform 10 different retrievals (the samples are numbered from 01 to 10). In every retrieval, t_r is recorded respectively. Detailed time cost of 40 retrievals is listed in Fig. 7.

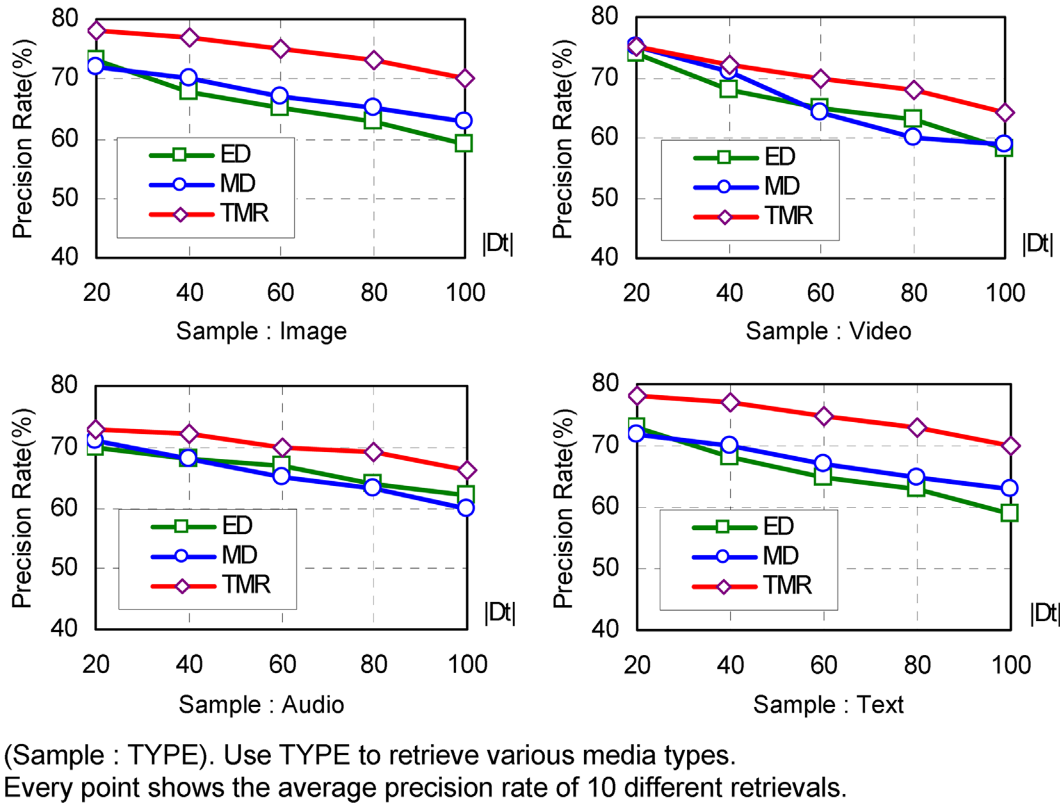


Fig. 5. Precision rate comparison (no feedback)

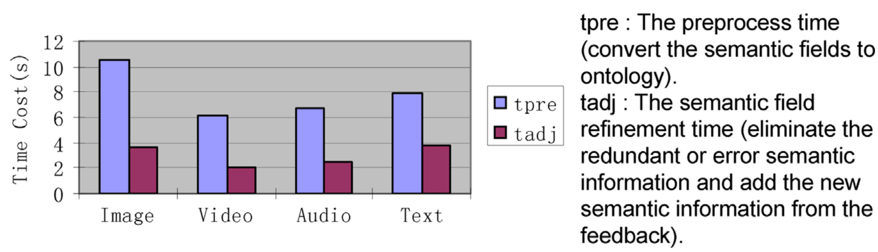


Fig. 6. Time cost of background processes

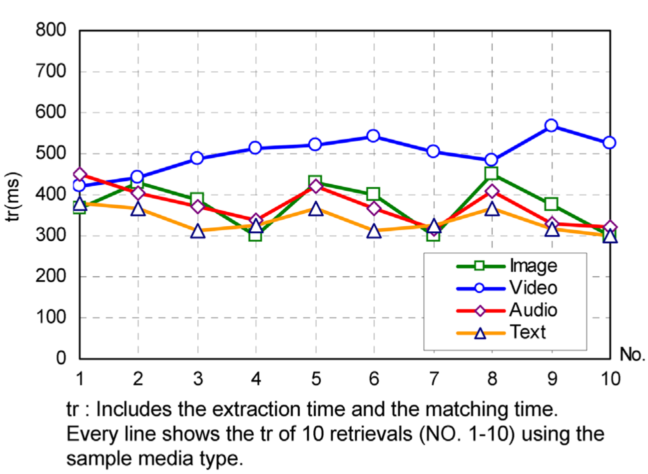


Fig. 7. Time cost of 40 retrievals

Fig. 7 shows that the semantic information extraction costs only an extremely short period of time (the vast majority of the t_r are less than 500 ms), this is because we only need to directly extract the semantic segment from the sample document.

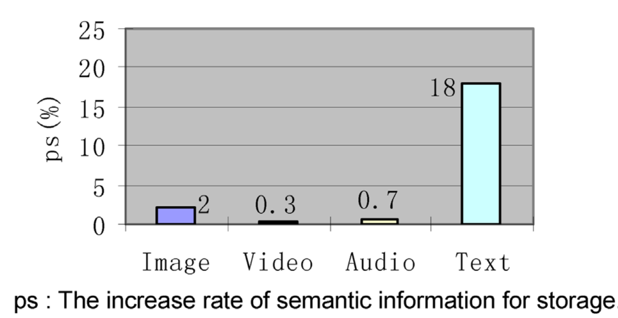


Fig. 8. Storage space cost of TMR.

4.5. Storage cost evaluation

Because the database stores the ontology information, the storage cost has to be taken into consideration. The increase rate of storage p_s is defined as follows:

$$p_s = \frac{S_{ont}}{S_{org}} \tag{16}$$

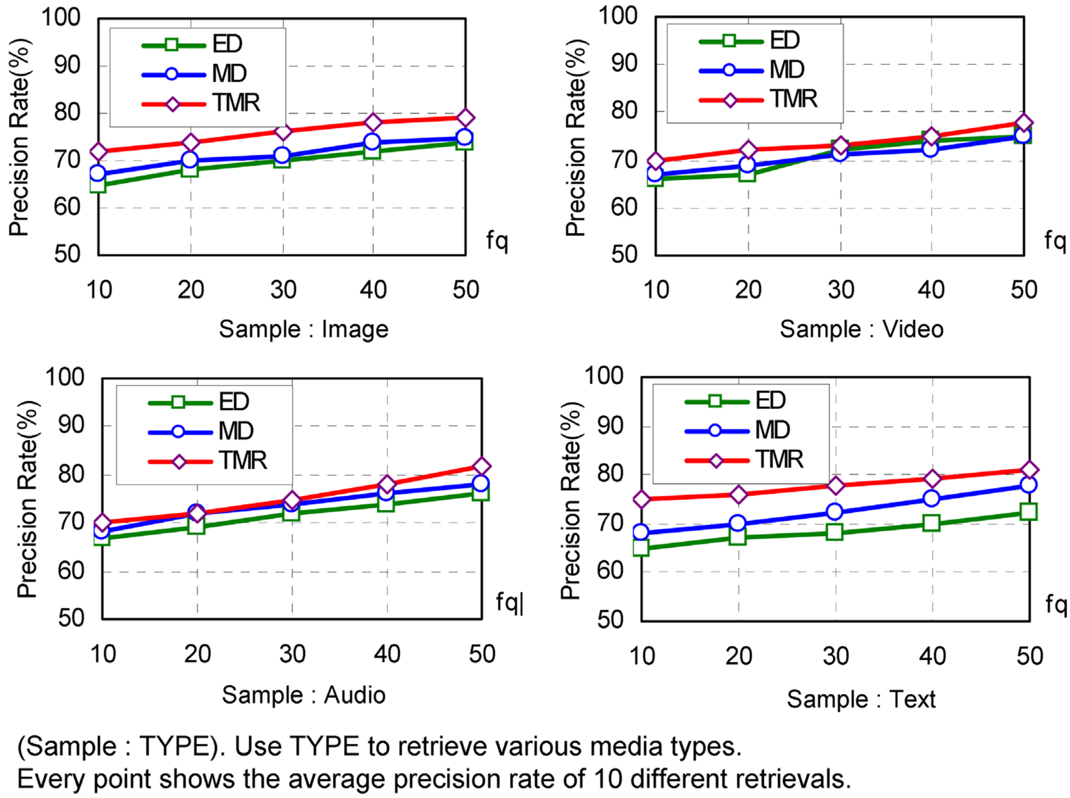


Fig. 9. Precision rate comparison (after feedback).

where s_{ont} and s_{org} are respectively the total size of ontology files and media documents:

$$\begin{cases} s_{ont} = \sum_{i=1}^{N} sizeof(ST_i) \\ s_{org} = \sum_{i=1}^{N} M_i(d) \end{cases} \tag{17}$$

Fig. 8 shows the storage space cost in TMR.

We can see from Fig. 8 that the semantic file size has almost not increased for image, video and audio types (p_s respectively is about 2% for image, 0.3% for video and 0.7% for audio). The semantic information file size occupies 18% for text type, this is because the semantic information in text files is abundant.

4.6. Performance comparison after feedback

In order to demonstrate the effectiveness of user feedback, in the next experiment, we specially record the precision rates considering asking the user to give feedback annotations to the returned documents. We define feedback quantity to represent the quantity of feedback annotations. For every feedback quantity fq= (10, 20, 30, 40 and 50), we perform 10 different retrievals for every fq and compute the average precision rates. The precision rates after feedback are illustrated in Fig. 9.

Fig. 9 demonstrates that after the user feedback, the precision can be increased. With the increase of the feedback quantity, the gap will be growing greatly. From the comparison, we can see that the effectiveness of TMR outperforms some other approaches in the case of user feedback.

4.7. Discussion

It can not be ignored that some traditional technologies which support content-based retrieval have promising performances too. These approaches actually can get perfect results reflecting users' intention in content-based retrieval. However, they have the following disadvantages in comparison with TMR: (1) Traditional search pattern can not support heterogeneous retrieval because of the physical feature extraction; and (2) Compared with physical features, annotations can better represent the users' query intention.

Therefore, TMR can get a higher speed and more accurate results in the case of the semantic media retrieval. If the documents contain more abundant annotations, the retrieval performance would be better. In addition, TMR has the advantage of promising speed because of skipping the physical feature extraction.

5. Conclusions

In this paper, a new approach named TMR supporting heterogeneous media retrieval and reflecting users' retrieval intention has been proposed. We designed the framework of TMR and described semantic information extraction, ontology-based representation, semantic field adjustment and fuzzy ontology matching. The framework solves two critical problems. Firstly, we use ontology represented semantic fields to express the users' intention. Secondly, fuzzy matching is employed to measure the similarity of ontologies.

We have applied several experiments on TMR. Experiment result shows that TMR can achieve a promising performance for the heterogeneous media retrieval reflecting the users' intention. In comparison with some traditional similarity measurements, TMR has its superiority.

However, in this paper, we just compare the precision rate of our approach with traditional measurements. In the future work, we will plan to explore several improvements of TMR, including increasing the retrieval speed and performing the experiments in real transportation environment.

Acknowledgement

This work is supported by the Major Science and Technology Research Program for Strategic Emerging Industry of Hunan (2012GK4106), International Science and Technology Cooperation Special Projects of China (2013DFB10070), Hunan Science and Technology Plan (2012RS4054), National Natural Science Foundation of China (61202341,61202095) , Key Laboratory of Intelligent Perception and Systems for High-Dimensional Information of Ministry of Education Innovation Fund (JYB201502), and Project of Innovation-driven Plan in Central South University (2015CXS010).

References

[1] Marco Attene, Francesco Robbiano, Michela Spagnuolo, Bianca Falcidieno, Semantic annotation of 3d surface meshes based on feature characterization. In semantic multimedia, Springer (2007), p. 126–139.
[2] Shahina Begum, Mobyen Uddin Ahmed, Peter Funk, Ning Xiong, Bo Von Schéele, A case-based decision support system for individual stress diagnosis using fuzzy similarity matching, Comput. Intell. 25 (3) (2009) 180–195.
[3] Tamalika Chaira, AK Ray, Fuzzy measures for color image retrieval, Fuzzy Sets Syst. 150 (3) (2005) 545–560.
[4] R Chen, YF Cao, Hongbin Sun, Active sample-selecting and manifold learning-based relevance feedback method for synthetic aperture radar image retrieval, IET Radar Sonar Navig. 5 (2) (2011) 118–127.
[5] Jose Costa Pereira, Emanuele Coviello, Gabriel Doyle, Nikhil Rasiwasia, Gert RG Lanckriet, Roger Levy, Nuno Vasconcelos, On the role of correlation and abstraction in cross-modal multimedia retrieval, IEEE Trans. Pattern Anal. Mach. Intell. on 36 (3) (2014) 521–535.
[6] Ritendra Datta, Dhiraj Joshi, Jia Li, James Z Wang, Image retrieval: Ideas, influences, and trends of the new age, ACM Comput. Surv. 40 (2) (2008) 5.
[7] Carmen De Maio, Giuseppe Fenza, Vincenzo Loia, Sabrina Senatore, Hierarchical web resources retrieval by exploiting fuzzy formal concept analysis, Inf. Process. Manag. 48 (3) (2012) 399–418.
[8] Hichem Frigui, Interactive image retrieval using fuzzy sets, Pattern Recognit. Lett. 22 (9) (2001) 1021–1031.
[9] Arjan Gijsenij, Theo Gevers, Color constancy using natural image statistics and scene semantics, IEEE Trans. Pattern Anal. Mach. Intell. on 33 (4) (2011) 687–698.
[10] Kehua Guo, Jianhua Ma, Guihua Duan, Dhsr: a novel semantic retrieval approach for ubiquitous multimedia, Wirel. Pers. Commun. 76 (4) (2014) 779–793.
[11] Kehua Guo, Wei Pan, Mingming Lu, Xiaoke Zhou, Jianhua Ma, An effective and economical architecture for semantic-based heterogeneous multimedia big data retrieval, J. Syst. Softw. 102 (2015) 207–216.
[12] Kehua Guo, Shigeng Zhang, A semantic medical multimedia retrieval approach using ontology information hiding, Comput. Math. Methods Med. (2013).
[13] Xiaofei He, Laplacian regularized d-optimal design for active learning and its application to image retrieval, IEEE Trans. Image Process. 19 (1) (2010) 253–263.
[14] Li Kuang, Zhifang Liao, Wentao Feng, Haoneng He, Bei Zhang, Multimedia services quality prediction based on the association mining between context and qos properties, Signal Process. (2015).
[15] Zhi-fang Liao, Fei Cai, Miao Zhang, Zhi-ning Liao, Yan Zhang, A tripartite tensor decomposition fold-in for social tagging, J. Appl. Sci. Eng. 17 (4) (2014) 363370.
[16] Ying Liu, Dengsheng Zhang, Guojun Lu, Wei-Ying Ma, A survey of content-based image retrieval with high-level semantics, Pattern Recognit. 40 (1) (2007) 262–282.
[17] Fuhui Long, Hongjiang Zhang, David Dagan Feng, Fundamentals of content-based image retrieval. In Multimedia Information Retrieval and Management, Springer (2003), p. 1–26.
[18] Bo Lu, Guo-Ren Wang, Ye Yuan, A novel approach towards large scale cross-media retrieval, J. Comput. Sci. Technol. 27 (6) (2012) 1140–1149.
[19] Sudipta Mukhopadhyay, Jatindra Kumar Dash, Rahul Das Gupta, Content-based texture image retrieval using fuzzy class membership, Pattern Recognit. Lett. 34 (6) (2013) 646–654.
[20] Liqiang Nie, Meng Wang, Yue Gao, Zheng-Jun Zha, Tat-Seng Chua, Beyond text qa: Multimedia answer generation by harvesting web information, IEEE Trans. Multimed. 15 (2) (2013) 426–441.
[21] Liqiang Nie, Shuicheng Yan, Meng Wang, Richang Hong, and Tat-Seng Chua. Harvesting visual concepts for image search with complex queries, in: Proceedings of the 20th ACM international conference on Multimedia, ACM, 2012, pp. 59–68.
[22] Chunhua Shen, Junae Kim, Lei Wang, Scalable large-margin mahalanobis distance metric learning, IEEE Trans. Neural Netw. 21 (9) (2010) 1524–1530.
[23] Chung-Chih Tsai, Heng-Yi Lin, Jinshiuh. Taur, Chin-Wang Tao, Iris recognition using possibilistic fuzzy matching on local features, IEEE Trans. Syst. Man Cybern. Part B: Cybern. on 42 (1) (2012) 150–162.
[24] Guojun Wang, Md Zakirul Alam Bhuiyan, Jiannong Cao, Jie Wu, Detecting movements of a target using face tracking in wireless sensor networks, IEEE Trans. Parallel Distrib. Syst. on 25 (4) (2014) 939–949.
[25] Guojun Wang, Qiushuang Du, Wei Zhou, Qin Liu, A scalable encryption scheme for multi-privileged group communications, J. Supercomput. 64 (3) (2013) 1075–1091.
[26] Jianxin Wang, Zhixiong Liu, Shigeng Zhang, Xi Zhang, Defending collaborative false data injection attacks in wireless sensor networks, Inf. Sci. 254 (2014) 39–53.
[27] Shuhui Wang, Qingming Huang, Shuqiang Jiang, Qi Tian, Lei Qin, Nearest-neighbor method using multiple neighborhood similarities for social media data mining, Neurocomputing 95 (2012) 105–116.
[28] Roger CF Wong, HC Clement Leung, Automatic semantic annotation of real-world web images, IEEE Trans. Pattern Anal. Mach. Intell. on 30 (11) (2008) 1933–1944.
[29] Yingjie Xia, Xiumei Li, Zhenyu Shan, Parallelized fusion on multisensor transportation data: a case study in cyberits, Int. J. Intell. Syst. 28 (6) (2013) 540–564.
[30] Yingjie Xia, Yuncai Liu, Zhoumin Ye, Wei Wu, and Mingzhe Zhu. Quadtree-based domain decomposition for parallel map-matching on gps data, in: Proceedings of the 15th IEEE International IEEE Conference on Intelligent Transportation Systems (ITSC), 2012, pp. 808–813.
[31] Wankou Yang, Jianguo Wang, Mingwu Ren, Jingyu Yang, Lei Zhang, Guanghai Liu, Feature extraction based on laplacian bidirectional maximum margin criterion, Pattern Recognit. 42 (11) (2009) 2327–2334.
[32] Yi Yang, Feiping Nie, Dong Xu, Jiebo Luo, Yueting Zhuang, Yunhe Pan, A multimedia retrieval framework based on semi-supervised ranking and relevance feedback, IEEE Trans. Pattern Anal. Mach. Intell. on 34 (4) (2012) 723–742.
[33] Yi Yang, Fei Wu, Dong Xu, Yueting Zhuang, Liang-Tien Chia, Cross-media retrieval using query dependent search methods, Pattern Recognit. 43 (8) (2010) 2927–2936.
[34] Hong Zhang, Yun Liu, Zhigang Ma, Fusing inherent and external knowledge with nonlinear learning for cross-media retrieval, Neurocomputing 119 (2013) 10–16.
[35] Lining Zhang, Lipo Wang, Weisi Lin, Generalized biased discriminant analysis for content-based image retrieval, IEEE Trans. Syst. Man Cybern. Part B: Cybern. 42 (1) (2012) 282–290.
[36] Luming Zhang, Yue Gao, Rongrong Ji, Yingjie Xia, Qionghai Dai, Xuelong Li, Actively learning human gaze shifting paths for semantics-aware photo cropping, IEEE Trans. Image Process. 23 (5) (2014) 2235–2245.
[37] Luming Zhang, Yue Gao, Yingjie Xia, Qionghai Dai, Xuelong Li, A fine-grained image categorization system by cellet-encoded spatial pyramid modeling, IEEE Trans. Ind. Electron. 62 (1) (2015) 564–571.
[38] Luming Zhang, Yue Gao, Yingjie Xia, Ke Lu, Jialie Shen, Rongrong Ji, Representative discovery of structure cues for weakly-supervised image segmentation, IEEE Trans. Multimed. 16 (2) (2014) 470–479.
[39] Luming Zhang, Yingjie Xia, Rongrong Ji, Xuelong Li, Spatial-aware object-level saliency prediction by learning graphlet hierarchies, IEEE Trans. Ind. Electron. 62 (2) (2015) 1301–1308.
[40] Luming Zhang, Yingjie Xia, Kuang Mao, He Ma, Zhenyu Shan, An effective video summarization framework toward handheld devices, IEEE Trans. Ind. Electron. 62 (2) (2015) 1309–1316.
[41] Guang-Tong Zhou, Kai Ming Ting, Fei Tony Liu, Yilong Yin, Relevance feature mapping for content-based multimedia information retrieval, Pattern Recognit. 45 (4) (2012) 1707–1720.

Kehua Guo received his Ph.D. in Computer Science and Technology from the Nanjing University of Science and Technology in 2008. He is currently an Associate Professor in the School of Information Science and Engineering, Central South University. His research interests include social computing, ubiquitous computing, big data and image retrieval.

Ruifang Zhang received her bachelor degree in Computer Science and Technology from Central South University in 2015. She is currently a graduate student in the School of Information Science and Engineering, Central South University. Her research interests include green computing and big data.

Li Kuang received her Ph.D. in Computer Science and Technology from Zhejiang University in 2009. She is currently an Associate Professor in the School of Software, Central South University. Her research interests include social computing, intelligent transportation and big data.