Journal of Applied Sciences

ISSN 1812-5654

science alert

ANSInet
an open access publisher
http://ansinet.com

Journal of Applied Sciences 13 (12): 2276-2281, 2013
ISSN 1812-5654/DOI: 10.3923/jas.2013.2276-2281

On Semantic Retrieval Oriented to Ietm Software

[1]Li Kairong, [1]Song Heng, [1,2]Zhu Junwu, [3]Chang Baoxian and [2]Li Xining
[1]College of Information Engineering, Yangzhou University, Yangzhou City Jiangsu Province, China
[2]Department of Computing and Information Science, University of Guelph, Guelph n1g2k8, Canada
[3]College of Sciences, Nanjing University of Technology, Nanjing, China

Abstract: Complex system is equipped with large-scale technical data, which can implement the task of the operation maintenance, repair of equipment and so on. With the increasing amount and complexity of the technical information, the accuracy and efficiency of knowledge retrieval with positioning has decreased significantly. This study presents an interactive electronic technical manual method of creation and implementation which based on semantic retrieval, starts from the image semantic annotation and the RDF cutting of text and then realizes the semantic retrieval of image and text. This study created the knowledge base about the instructions for use of Samsung Galaxy III, taking C# language as the development language, Visual Studio as integrated development environment, SQL Server as database in implementing the semantic retrieval. Compared with current searching for relevant information manually, the system's efficiency and accuracy of knowledge retrieval has improved greatly. Finally, some processes are presented in the software engineering, such as the demand for knowledge retrieval system analysis, algorithm design, coding, testing and maintenance.

Key words: Interactive electronic technical manual, semantic retrieval, knowledge base

INTRODUCTION

As the constant improvement of civil equipment, the application of the Interactive Electronic Technical Manual (IETM) is broadening, which no longer restricted to military equipment (Ventura, 1988; Greenough and Tjahjono, 2007). However, the increase in the number of technical data and the diversity of information carrier caused a sharp drop in the efficiency and accuracy of information retrieval. On the contrary, the use of technical information demand has strong timeliness; people hope to find the exact information in the very short time. Existing information organization forms including "chunks of text" and "image title" and so on, which unable to realize the accurate positioning and reorganize the information or provide personalized information display (Zhu *et al*., 2004; Luo *et al*., 2009; Xia and Wang, 2006). Taking example for electronic devices, when a user uses an electronic equipment at the first time, it is always hard to begin within the face of a thick manual and other technical data; For professional users, it is an urgent-need to find and understand part of information technology, full-text keyword retrieval certainly reduces the experience degrees of their users.

Therefore, the demand of the accuracy and efficiency of information retrieval become more and more obvious (Huang and Li, 2008a). The main reasons follow: First, accessing to information has obvious selectivity, users are only interested in certain pieces of information; Second, larger complex system of technical data, it is difficult to realize effective pieces of information location in the current information organization mode. However, as the popularization and application of theory and technology for knowledge representation (such as the ontology), using the semantic retrieval technology to enhance the quality and accuracy of information retrieval has become possible (Huang and Li, 2008b; Shi *et al*., 2009; Jiang *et al*., 2008). The semantic annotation of domain ontology for a piece of information (text, images and video, etc.,) provides the basic terminology and vocabulary, through multiple pieces of information granularity of annotation, which can maintain its complete semantic; Especially for image and video, semantic tags are mostly the description of its contents, which reduce the amount of calculation of image pattern recognition and there is hope to improve the retrieval speed and accuracy at the same time (Drori and Rinott, 2007; Bay *et al*., 2008; Shaer and Hornecker, 2010). After editing

Corresponding Author: Li Kairong, College of Information Engineering, Yangzhou University, Yangzhou City, Jiangsu Province, China

the technical information database, the semantic retrieval of its content is no longer matching the content for all the key words, only for images and text labels, computing the semantic distance to get the candidate collection of a set of search results.

A kind of creation and implementation methods of interactive electronic technical manual based on semantic retrieval is proposed in this paper, respectively from the semantic annotation image and the RDF cutting of text and then achieve both images and text semantic retrieval; Creating the knowledge base about the instructions for use of Samsung Galaxy 3, taking C# language as the development of language and Visual Studio as integrated development environment, implementing the semantic retrieval which takes SQL Server as database. Considering the user might enter synonyms or even the wrong keywords, the system provides the function of matching synonyms and recorrecting wrong characters.

OVERALL DESIGN OF IETM SYSTEM

System architecture of IETM system: This system is mainly divided into two modules: Image semantics and text retrieval management and text management all have functions semantic retrieval and management. Image semantic retrieval and management module is mainly search for related image information according to the key words that users input, while the text semantic management module is mainly to search for related text information according to the key words that users input. Image semantic of wrong character recorrection and synonyms matching. Finally processing and combining the images and text information searched and display the processed results to the users (Fig. 1).

SEMANTIC MATCHING DEGREE CALCULATION OF THE FINAL SEARCH RESULTS

Final matching degrees of search results mainly depends on the matching degrees of image search results and the matching degrees of text search results match in this study, we define the final search results matching degree is the sum of 30% of the matching degree of image search results and 70% the matching degree of text search results.

Final results match calculation is mainly the calculation about the matching degree of searching related text and image information in view of the related keywords user input:

- **Step 1:** Calculating the matching degree of image search results, defined as matchPer1
- **Step 2:** Calculating the matching degree of text search results, defined as matchPer2
- **Step 3:** Defining the matching degree of final search results as mathchPer then match Per = 0.3* match Per1+0.7* matchPer2.//Calculating the matching degree of final search results public static void MatchPercent CalZong(image ig) {if both the matching degree of image and text are zero, the matching degree of final search resultsif (ig.matchPer < = 0&&ig.matchPerText< = 0) ig.matchPer = 0; else {ig.matchPer = 0.3* ig.matchPer + 0.7 *ig.matchPer; }}

IMAGE SEMANTIC RETRIEVAL AND MANAGEMENT

Inputting related content in the search text box of the main interface to search the related image information about the keyword of the text box and calculating the matching degree of related information! The functions of wrong character correct and the synonyms for matching are given.

In order to be able to find the related content about the instructions of Samsung GalaxyIII, the related image information about Samsung GalaxyIII needs to be stored in the database. In order to be able to correct the typing errors inputted by users, wrong character library need to be stored in the database. what's more, synonyms thesaurus are also need to be stored to match the synonyms about the keywords inputted by users.

In Fig. 2, the mode of image semantic retrieval and management are divided into the knowledge base editing and knowledge base searching. Knowledge base edited is divided into synonyms for editing, words of similar pronunciation for editing, words of similar composition for editing and knowledge base editing. Knowledge base searched is divided into original word searched and error correction searched words of similar pronunciation for correcting, words of similar composition for correcting and synonyms for matching. The ultimate aim of original word searched and error correction searched is to get the related image information.

TEXT SEMANTIC MANAGEMENT

Input related content in the search text box of the main interface to search the related text information about the keyword of the text box and calculating the matching degree of related information. The functions of wrong character correct and synonyms for matching are given:

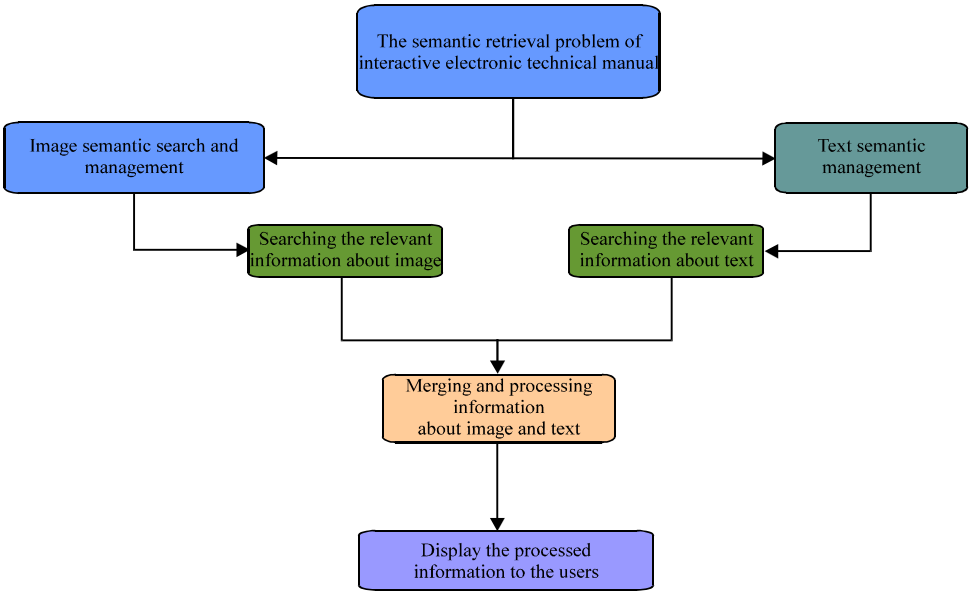

Fig. 1: IETM system which support semantic retrieval

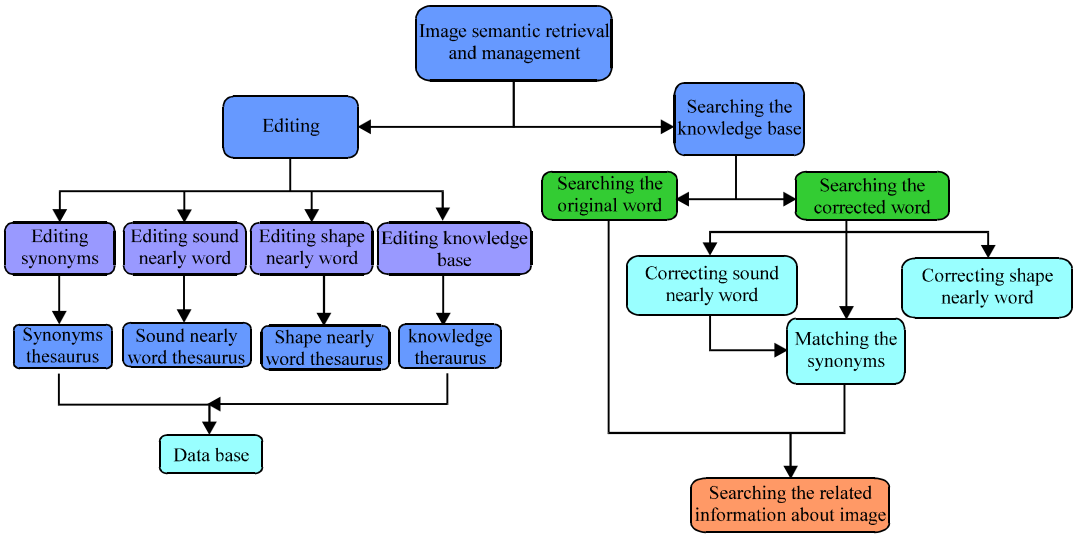

Fig. 2: System structure of image semantic management

- The system framework is as follows:

The calculation for the semantic matching degree of image searched is mainly the calculation about the matching degree of searching related image information in view of the user input related keywords (Fig. 3):

- **Step 1:** There is no words corrected and no synonyms matched, then matching degree matchPer is 1.
- **Step 2:** If there exists synonyms match, then matching degree match Per = 0.95*0.90*0.85*…*(1-0.05 n), where n is the number of words corrected and n<20
- **Step 3:** If there exists words recorrected, then the matching degree match Per = 0.8*0.75*0.70*…*(0.85-0.05n), where n is the number of words matched and n<17. //Calculating the matching degree of words corrected and synonyms matched, public static void Match Percent Cal (image ig, int matchWords, int wordsNum,int type){double match Per = 1; //type indicates whether the predicate and the object is complete: 1

In order to be able to find the related text information Samsung Galaxy 3, the related text information about Samsung Galaxy 3 needs to be stored in the database in the format of RDF. We know that there are relations between some of the sentences in the mobile phone

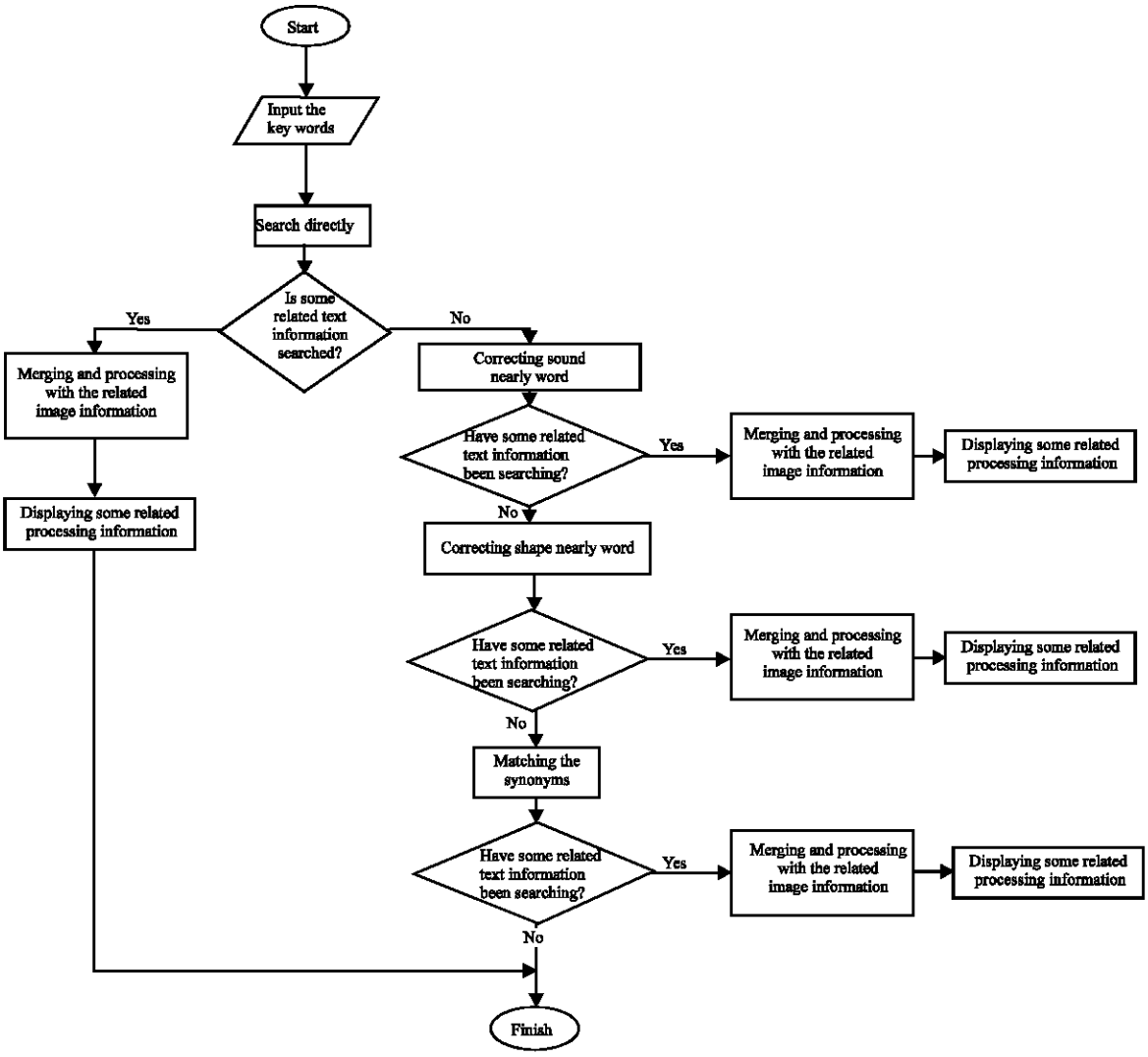

Fig. 3: Flow chart of image repository searching

represents predicate only, 2 represents object only, 3 represents both predicate and object. if (type == 1|| type = 2) {matchPer = 0.5;} else if (type == 3) {matchPer = 1;} else {ig.matchPer=-1;//Indicating error happened when calculating matching degree return; }for (int i = 0; i < matchWords; i++) {matchPer *= (0.95- 0.05 *i);}for (int i = 1; i <= wordsNum; i++) {matchPer *= (0.85-0.05*i);} ig.matchPer = matchPer;}

manual, such as cause and effect relationship. In order to string the sentences in the RDF table to form the rational relationship, the linkStruct table need to be built.

In Fig. 4, the text semantic management are divided into the text library edited and text library searched, text library edited is divided into RDF for editing, statements in the relationship for editing. Knowledge base searched is divided into original word searched and error correction searched, words of similar pronunciation for correcting, words of similar composition for correcting and synonyms for matching. The ultimate aim of original word searched and error correction searched is to get the related text information.

- Flow chart of the text search

According to text content input in the search text box by the user, search out the related information about image in the knowledge database. Firstly, if there are no results searched, perform error correction module, perform error correction module. Specific methods to correct mistakes and match the synonyms as well as the methods

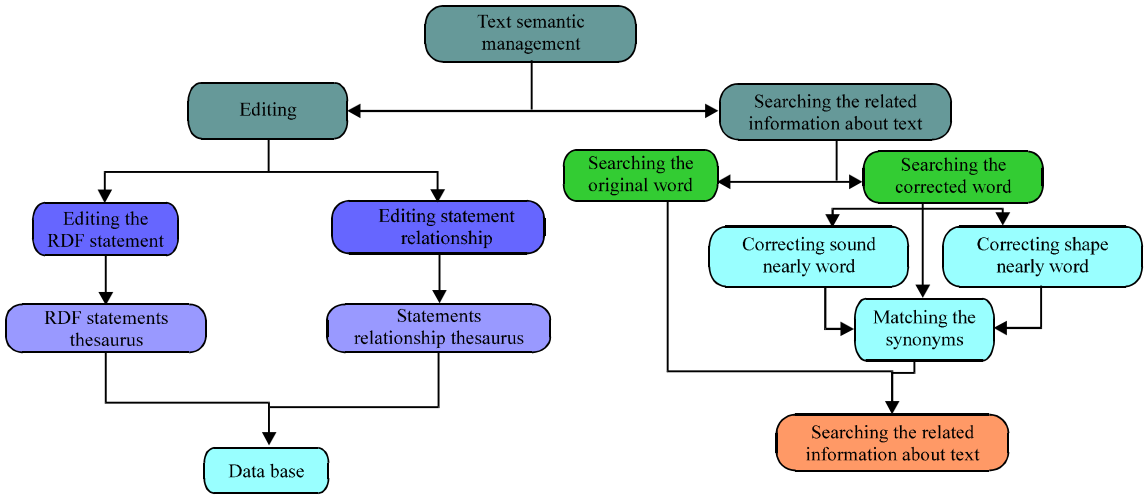

Fig. 4: System architecture of text semantic management

mentioned in the image semantic retrieval and management. Finally, search the information with the corrected text for keywords. If we can search out the relevant contents, querying the indirect relationship between this statement and other statements in the linkStruct database according to the id of a statement and calculating the layers of indirect relationship

- The matching degree algorithm of text searching results

```
else
{ig.matchPerText=-1;return;}
for(int i=0;i< matchWords;i++)
{ matchPer *= (0.95- 0.05 * i); }
for (int i =1; i<= wordsNum; i++)
{ matchPer *= (0.85-0.05 * i); }
for (int i = 1; i <= linkNum; i++)
{
// the matching degree will become 95% of its original matching degree when a layer //of indirect relationship is added
matchPer*=0.95; }
ig.matchPerText = matchPer;}
```

CONCLUSION

A kind of creation and implementation methods of interactive electronic technical manual based on semantic retrieval is proposed in this study, respectively from the semantic annotation image and the RDF cutting of text and then achieve both images and text semantic retrieval; The system is a knowledge retrieval system application which takes Object-oriented as design ideas, takes C # language as the development language, takes Visual Studio as integrated development environment and takes SQL Server as database, this system has simple application interface, the system searches some related information in the knowledge base according to the input information. The main function is to intelligently correct the wrong characters and match the synonyms according to the related keywords input by users to search the exact information users need.

ACKNOWLEDGMENT

This research is supported by the National Nature Science Foundation of China (No.~61170201) , (No.~61070133) and Science and Technology Project of Yangzhou City (No. YZ2011098).

REFERENCES

Bay, H., A. Ess, T. Tuytelaars and L.V. Gool, 2008. Speeded-Up Robust Features (SURF). Comput. Vision Image Understand., 110: 346-359.

Drori, T. and M. Rinott, 2007. Pixel materiali: A system for creating and understanding pixel animations. Proceedings of the 6th International Conference on Interaction Design and Children, June 6-8, 2007, Aalborg, Denmark, pp: 157-160.

Greenough, R.M. and B. Tjahjono, 2007. An interactive electronic technical manual for an advanced aerospace assembly machine. Int. J. Adv. Manuf. Technol., 33: 1045-1055.

Huang, J.Z. and F. Li, 2008a. Research and discussion on IETM for complex weapon system. Microcomput. Appl., 8: 28-30.

Huang, J.Z. and F. Li, 2008b. Research on IETM for complex weapon system. Ordnance Ind. Autom., 8: 22-22.

Jiang, Y., X. Wu and Y. Fu, 2008. Design and realization of the maintenance directed system of ship equipment on IETM. Ship Electron. Eng., 7: 131-133.

Luo, Z., H. Chang and X. Zhang, 2009. IETM in warship electromechanical equipment informationization support. Ship Boat, 4: 55-58.

Shaer, O. and E. Hornecker, 2010. Tangible user interfaces: Past, present and future directions. Found. Trends Hum. Comput. Interact., 3: 1-137.

Shi, H., X. Zhang, Z. Luo and Y. Liu, 2009. Application of IETM in warship equipment informationization support. Ship Electron. Eng., 8: 147-150.

Ventura, C.A., 1988. Why switch from paper to electronic manuals? Proceedings of the ACM Conference on Document Processing Systems, December 5-9, 1988, Santa Fe, NM., USA., pp: 111-116.

Xia, F. and W.G. Wang, 2006. Research on the technological scheme of military equipment training-oriented IETM. Sichuan Ordnance J., 3: 22-25.

Zhu, J., T. Chen and H. Xie, 2004. Study on IETM of class 5 based on ontology. J. Jiangsu Univ. Sci. Technol., 6: 90-96.