Design of Collection and Semantic Annotation System for Web Images

Ruojuan Xue
LiShan College, Shandong Normal University, Jinan, China
Email: xuerj@sdnu.edu.cn

Wenpeng Lu
School of Science, Shandong Polytechnic University, Jinan, China
Email: lwp@spu.edu.cn

Jinyong Cheng
School of Information, Shandong Polytechnic University, Jinan, China
Email: cjy @spu.edu.cn

***Abstract*—In order to satisfy the requirement of educators to automatically construct educational resource database with Web images, the paper designs and implements a collection and semantic annotation system for Web images. With the help of meta-search engine, the system collects Web images and their associated texts. In order to annotate semantic keywords, semantic dictionaries are built. Based on semantic dictionaries, the system annotates semantic keywords for Web images. The system is composed with image collection module and semantic annotation module, which are respectively divided into five and three child modules. The paper introduces the design and implementation of the system in detail. The collection and semantic annotation system can automatically search interested images, download them, annotate them with semantic keywords and store them into educational resource database, which would relieve the workload of educators greatly. The experimental results demonstrate that the system is efficient for collection and annotation of Web images, which can basically satisfy the requirement of the construction of educational resource.**

***Index Terms*—educational resource, semantic annotation, Web image, image annotation**

I. Introduction

The construction of education resource is one of the core tasks of education informatization[1, 2]. There are lots of images on World Wide Web (WWW). If Web images can be automatically collected and annotated, they could be integrated into educational resource database. This would be helpful to implement the intelligent construction of educational resource and promote the development of education informatization[3].

In order to automatically construct educational resources with Web images, we have designed a collection and semantic annotation system (CSAS) for Web images in the paper. The software system mainly includes two functions: the function to collect Web images and the function to annotate semantic information of Web images. When users specify the list of keywords of Web images, CSAS would search and download the relevant Web images with the help of meta-search engine. For each Web images, the relevant texts would be collected at the same time. Based on the relevant texts and semantic dictionaries, CSAS can infer the semantic keywords of Web images and annotate them.

The software system can help users to quickly collect and annotate all kinds of interested Web images, which is easy to be operated with a good man-machine interface. The software system can automatically collect and annotate Web images, which would greatly reduce the workload of educators to construct educational resources.

In the paper, the detailed design and implementation of CSAS are described. The rest of the paper is organized as follow. Section II introduces semantic annotation knowledge to extract semantic keywords of Web images. The framework of CSAS is introduced briefly in Section III. The detailed implementation of each child module is described in Section IV. The experiments are introduced in Section V. As last, we give the conclusion and future work.

II. Semantic Annotation Knowledge

In order to annotate semantic keywords for collected images, semantic annotation knowledge is required. As is described in Ref.[4, 5], the semantic dictionaries should be constructed. Besides, in order to annotate semantic keywords, the associated text of Web image should be extracted. The semantic dictionaries are the basis of semantic keywords and the associated texts are the source of semantic keywords.

Project number: Shandong Province Higher Educational Science and Technology Program (J12LN09 and J10LG20), China, and Natural Science Foundation of Shandong Province (ZR2011FQ038), China
Corresponding author: Wenpeng Lu

doi:10.4304/jsw.8.3.694-700

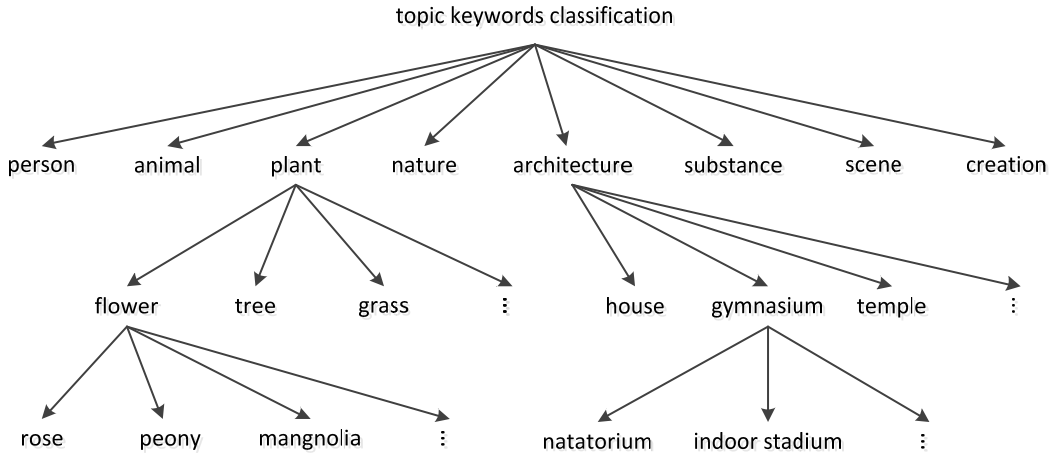

Figure 2. Topic Keywords Classification

A. *Semantic Dictionary*

1) Image Topic Dictionary

Topic keywords can express the main semantic information of an image. According to the research on Ref.[4], the construction of classification of image topic should consider the following factors: topic keyword hypernymy and hyponymy relation, image visual feature, cognitive habit. In order to satisfy the factors, the structure of tree form is selected as the classification structure of image topic. Topic keywords are divided into eight categories, which are person, animal, plant, nature, architecture, substance, scene and artistic creation. Each categories includes several sub-categories. The classification of topic keywords can be descripted as follow: first category name | second category name | third category name | …

- Person

Person category includes the keywords, such as name, occupation, nation or race in an image, which has three sub-categories: history, modern and myth person.

- Animal, plant, nature, architecture, substance

Animal category includes eight sub-categories: myth animal, ancient animal, mammals, reptile, amphibian, fish, bird and insect.

Plant category includes fourteen sub-categories, such as crop, vegetable, flower, tree, vine, bamboo, grass, leaf, pteridophyta, bryophyta, phycophyta, et al.

Nature category has fifty-seven sub-categories, such as mountain, peak, valley, ravine, mine, water, rain, snow, ice, fog, et al.

Architecture category has forty-eight sub-categories, such as city, town, wall, door, road, street, storied building, house, room, palace, temple, et al.

Substance category has six sub-categories, such as vessel, tool, equipment and material, et al.

- Scene

Scene category includes six sub-categories: landscape, night piece, street scene, vista, stage photo and occurrence.

- Creation

Creation category includes three sub-categories: paint, chart, art photo, which is used to describe the artificial images.

2) Image Object Dictionary

Based on object visual feature, image object dictionary is classified. Similar with topic dictionary, object dictionary also is constructed with tree form structure. the objects in the same category have same visual features. Image objects includes six categories: person, animal, plant, nature, creation and graphics. Each of them includes several sub-categories.

3) Image Attribute Dictionary

Attribute dictionary includes the keywords that are used to describe the visual feature of images, which includes: 262 color words, 82 shape words, 18 texture words, 15 posture words, 11 position words and 8 direction words.

B. *Image Associated Text*

The semantic keywords are extracted from image associated texts. In the paper, we thought that seven kinds of associated texts are necessary[4, 6, 7].

- Image file name

Image file name is usually English word, abbreviation, pinyin or Chinese word, which usually contains the topic or object name of Web image.

- Image file name or Web page title in a lower rank

If an image has a hyperlink which refers to an image or Web page in a lower rank. The file name of the target image or the title of the target Web page usually contains some semantic information of source image.

- Text of ALT tag of the image

The alternative text of an image often contains the simple description of the image.

- Text surrounding the image

The text surrounding the image usually contains most relevant semantic information of the image.

• Title of the Web page

The title of Web page describes the core information of the Web page, which has some semantic relation with the images located in the Web page.

• Section name of the Web page

Section name of the Web page usually expresses the classification of topic and object of the Web image.

• URLs

URL of the image, URL of the Web page and URL of the Web page linked by the image always includes part of sematic information of Web image.

III. FRAMEWORK OF CSAS

According to the function requirements of CSAS, as is shown in Fig.2, the software system is composed with eight modules, which can be divided into two categories, which are image collection module and semantic annotation module[8]. Imge colleciton module includes five child modules, which are the module of meta-search engine, the module of engine pages analysis, the module of download of Web pages and Web images, the module of extraction of associated texts. Semantic annotation module includes three child module, which are the module of annotation of attribute keywords, the module of annotation of object semantic keywords and the module of annotation of topic semantic keywords.

As in shown in Fig.2, firstly, according to the keywords specified by users, the module of meta-search engine is reponsible to access meta-search engine to get returned engine pages. Secondly, the engine pages are processed by the module of engine pages analysis to get URLs of Web pages and URLs of Web images. Thirdly, the URLs of Web pages and Web images are respectively sent into the module of download of Web pages and Web Images, which would download them. Fourly, the downloaded Web pages would be analyzed by the module of extraction of associated texts to get the associated texts of current Web image. Fively, based on attribute dictionary, the module of annotation of attribtue keywords extracts attribute keywords from the associated texts. Sixly, based on object dictionary, object keywords are extracted. Sevenly, based on topic dictionary, topic keywords are extracted. Lastly, Web image and its semantic keywords are stored into educational resource database.

A. Framework of Image Collection Module

When users specify a list of interested keywords for CSAS, this modules would retrieve relevant Web images with meta-search engine, such as Google and Baidu, download the images and extracted their associated texts. As is shown in Fig.2, this module is divided into five child modules, which are as follow.

1) Module of Meta-Search Engine

According to the list of keywords specified by users, the module would access meta-search engine to retrieve relevant Web images and download all of returned engine pages.

2) Module of Engine Pages Analysis

Web pages returned by meta-search engine contain URL information of relevant Web images. The module is

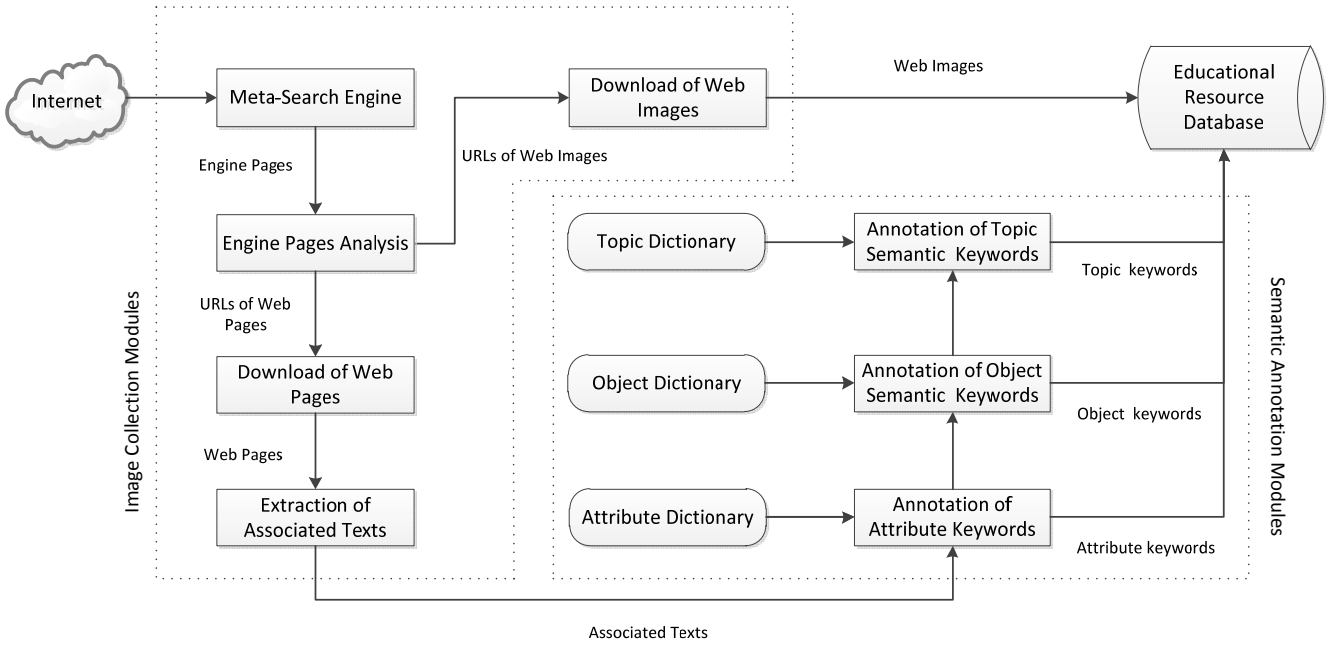

Figure 2. Framework of CSAS

responsible to analyze the engine pages to extract URLs of Web images and URLs of its Web pages.

3) Module of Download for Web Images

According to obtained URLs of Web images, the module would download Web images. When its semantic keywords are extracted, they would be stored into educational resource database together.

4) Module of Download for Web Pages

According to obtained URLs of Web pages, the module would download Web pages which Web images are located in. The Web pages would be used to extracted associated texts of Web images.

5) Module of Extraction of Associated Texts

As described in Section II, the module would extracted the seven kinds of associated texts which would be as the source of semantic keywords.

B. Framework of Semantic Annotation Module

In image collection modules, Web images and their associated texts have been downloaded and extracted. The associated texts contain important sematic information of Web images. According to semantic dictionaries, this module would extract the semantic keywords, such as topic, object and attribute keywords, and annotate Web images with them. This module is composed with three child modules, which are as follow.

1) Module of Annotation of Topic Keywords

Topic keywords represent topic information of the image, which can represent all or part of key semantic information. With the help of topic dictionary, this module is responsible to extract topic keywords from associated texts of Web images.

2) Module of Annotation of Object Keywords

Objects are the main objects which are outstandingly described in Web images. According to object dictionary, this module is responsible to extract object keywords from associated texts of Web images.

3) Module of Annotation of Attribute Keywords

Attribute dictionary lists descriptor words which describe visual features of images. According to attribute dictionary, this module is responsible to extract attribute keywords form associated texts of Web images.

IV. Implementation of Each Module of CSAS

A. Implementation of Image Collection Module

As described in Section III, this module is composed with five child modules: module of meta-search engine, module of engine pages analysis, module of download for Web images, module of download for Web pages and module of extraction of associated texts.

1) Implementation of Module of Meta-Search Engine

According to user-specified keywords, this module visits meta-search engine and downloads all of search engine pages. The detailed program flow is described as follow:

a) Select Google or Baidu as meta-search engine of Web images;

b) Obtain the user-specified list of keywords. Users can specify one or more keywords as the targets of image collection.

c) Take an image keyword as the object of this loop.

d) According to the keyword and query formula to generate the query address. For example, if Google is meta-search engine and apple is the keyword, then query address is "http://www.google.com.hk/search?q=apple%20fruit&um=1&ie=UTF-8&hl=zh-CN&tbm=isch&source=og&sa=N&tab=wi".

e) According to the query address, visit meta-search engine to request them to return relevant engine result page.

f) Read engine result page.

g) Save current engine result page.

h) Judge whether current engine result page is the last page of meta-search engine. If it isn't, then go to Step d); otherwise, go to Step i).

i) Judge whether current keyword is the last keyword. If it isn't, then go to Step c); otherwise, go to Step j).

j) Exit module of meta-search engine.

2) Implementation of Module of Engine Pages Analysis

After engine pages about a keyword have been downloaded, this module analysis them to get all of detailed URL information of Web images and their Web pages. The detailed program flow is described as follow:

a) Read an engine page.

b) Judge the source of the current engine page. If it is from Google, then get the tag-data of Google pages; if it is from Baidu, then get the tag-data of Baidu pages.

c) Initialize the variables to prepare for URL extraction.

d) According to the variables and tag-data of search engine, extract one pair of URLs (the URL of Web image and the URL of Web page) from current engine page.

e) Save the URLs into address file, and modify the value of variables.

f) Judge whether the URLs are the last one of current engine page. If they are, then go to Step g); otherwise, go to Step d).

g) Judge whether current engine page is the last page. If it is, then go to Step h); otherwise, go to Step a).

h) Exit module of engine pages analysis.

3) Implementation of Module of Download for Web Images

When address file of target Web images has been obtained, this module is responsible to download the Web images. The detailed program flow is described as follow:

a) Read one pair of URLs from address file.

b) Analysis the URL of Web image to get server address, port number, relative address, et al.

c) Establish Internet Connection with target server.

d) Judge whether the connection is succeeded. If it is, then go to Step e); otherwise, go to Step a).

e) Open a HTTP session and get its handle.

f) With the HTTP handle, Send Get command and obtain its return code.

g) Judge the return code. If it is address relocation, then go to Step b); if it is failure, then go to Step a); If it is OK, then go to Step h).

h) With a length of buffer, read the Web image and get return code.

i) According to the return code, judge whether the image has been completely downloaded. If it hasn't, then go to Step h) to read next part of images; otherwise, then go to Step j).

j) Save this image data into resource database.

k) Judge whether this pair of URLs is the last one of current address file. If it isn't, then go to Step a); Otherwise, then go to Step l).

l) Exit module of download for Web images.

4) Implementation of Module of Download for Web Pages

When address file of target Web pages has been obtained, this module is responsible to download the Web pages. The program flow is similar with the implementation of module of download for Web images.

5) Implementation of Module of Extraction of Associated Texts

After Web pages are downloaded , this module is responsible to analyze its HTML code to extract associated texts. We simply explain the method to extract each kind of associated text as follow:

a) Image file name

It is contained in the URL of the image, which is the last part of URL.

b) Image file name or Web page title in a lower rank

If the Web image has a hyperlink to another Web image or Web page, then extract the target URL of the hyperlink.

If the hyperlink is referred to an image, then extract its file name from the URL.

If the hyperlink is referred to a Web page, then download the Web page and extract the content between "<title>" and "</title>", which is the title of target Web page.

c) Text of ALT tag of the image

If the content of <alt> tag is not null, extract the text.

d) Text surrounding the image

The extraction of this kind of associated text is most complex. We need to analyze the layout of Web page, and select the text that are neighboring with Web image.

e) Title of Web Image

It is contained between "<title>" and "</title>", which is easy to be extracted.

f) Section name of the Web page

There are some markers to extract section name, such as "your position", "current position". For example, if there is "current position: homepage > instructional resource database > nature images > mammal" in Web page, we can extract "mammal" as the section name of the page.

g)URLs

URLs is contained between "" and "", which is easy to extract.

B. Implementation of Semantic Annotation Module

As described in Section III, when Web images and their associated texts have been collected, this module is responsible to extract and annotate semantic keywords. This module is composed with three child modules: module of annotation of topic sematic keywords, module of annotation of object semantic keywords and module of annotation of attribute keywords.

Semantic information of Web image is extracted based on relevant texts, which include image filename, filename of linked image, title of linked Web page, alternative text, surrounding text, title of located Web page, section name of located Web page, URL of image, URL of located Web page, URL of linked image or Web page. For ease of description, the relevant texts are denoted as T1~T10.

1) Implementation of Module of Annotation of Topic Keywords

According to image topic dictionary, based on relevant texts, this module extracts and annotates topic keywords for Web images[6, 7]. The detailed program flow is as follow:

a) Sort T1, T2, T3, T4, T5 (not more than a certain length) by descending priority order, then take each word in each text to match with the words of topic dictionary. If they are matched, all of matched words in the relevant text would be extracted as topic keywords and the rest would be discarded, then go to Step f). If all of them are not matched, then go to Step b).

b) Among T5, T6, T7, T8, T9, T10, if no less than two relevant texts have same words which can match with the words of topic dictionary, then extract the same words as topic keywords, go to Step f). Otherwise, go to Step c).

c) Among T6, if there are the words which can match with the words of topic dictionary, then extract the words as topic keywords, go to Step f). Otherwise, go to Step d).

d) Among T7, if there are the words which can match with the words of topic dictionary, then extract the words as topic keywords, go to Step f). Otherwise, go to Step e).

e) Extract T1 as topic keywords.

f) Annotate topic information of the image with extracted keywords.

g) Exit topic keywords annotation module.

2) Implementation of Module of Annotation of Object Keywords

According to image object dictionary, based on relevant texts, this module extracts and annotates object keywords for Web image. The detailed program flow is as follow:

a) Among T1, T2, T3, T4, extract the words which can match with the words of object dictionary as object keywords.

b) Among T5, T6, T7, T8, T9, T10, extract the words which are included in two or more relevant texts as object keywords.

c) Among T5, sort the words with the frequency by descending order, extract top-5 words which can match with the words of object dictionary as object words.

d) Annotate object information of the image with extracted keywords.

e) Exit object keywords annotation module.

3) Implementation of Annotation of Attribute Keywords

According to image attribute dictionary, based on relevant texts, this module extracts and annotates attribute keywords for Web image. As the difficulty of natural

language understanding, for the texts with simple grammatical structure, we analyze them to get attribute keywords. For the texts with complex grammatical structure, the keywords are extracted with statistical approach. The detailed program flow is as follow:

a) If there is an attribute word in image object name, or there is an attribute word before image object name, extract them as attribute keywords.

b) If a relevant text contains an attribute name, such as "color", "shape", "texture" et al., verb "be" follows the attribute name, and an attribute word follows verb "be", then extract the attribute word as attribute keyword.

c) Sort all of attribute words on their frequency by descending order, extract the word with highest frequency as attribute keywords.

d) Annotate attribute information of the image with extracted keywords.

e) Exit attribute keywords annotation module.

V. Experiments

After the collection and annotation system has been implemented as described in section III and section IV, we downloaded and processed 1,000 images from Internet with it.

In order to evaluate the effectiveness of the system, we defines three indicators: annotation coverage, accuracy, total accuracy. We mark the number of total images as N, the number of images which are annotated with more than one semantic keywords as M, the number of right semantic keywords which is annotated with the system automatically as R, the number of semantic keywords which is annotated manually as A, the number of semantic keywords with the system automatically as B. The three indicators are defined as follow:

$$Extraction\ rate = \frac{M}{N} \quad . \tag{1}$$

$$Accuracy = \frac{R}{A} . \tag{2}$$

TABLE I.
EXPERIMENTAL RESULTS

	Topic	Object	Color	Shape	Texture
Extraction Rate	80.1%	73.1%	4.2%	2.2%	0.5%
Accuracy	61.2%	64.3%	71.4%	50%	40%
Precision	54.2%	52.3%	61.9%	54.5%	20%

$$Precision = \frac{R}{B}. \tag{3}$$

The experimental results are shown in Table 1.

As shown in Table 1, on the one hand, the extraction rates of color, shape and texture keywords are rather low. This is caused with the reason : the three kinds of keywords appears with a low frequency, so their extraction rate are low. On the other hand, the extraction rate, accuracy and precision of topic and object keywords are better, which can basically satisfy the requirement of semantic annotation of Web images.

VI. Conclusions and Future Work

In order to satisfy the requirement of automatically construct educational resource database, the paper designs and implements a collection and annotation system for Web images. The software system can automatically search interested images, download them and annotate them with semantic keywords. This would relieve the workload of educators greatly and would be help to retrieve Web images[8].

Though the software can achieve the integrated operations of educational resource database construction, it is not be fully satisfied. The extraction rate, accuracy and precision of semantic keywords need to be improved in the further. In the next work, we would like to try more complex and effective method to improve its accuracy.

Acknowledgment

This work is supported by two projects of Shandong Province Higher Educational Science and Technology Program (J12LN09 and J10LG20), China, and by Natural Science Foundation of Shandong Province (ZR2011FQ038), China, and by grant 2011WSB11001 of the Shandong medical health science and technology development programs to Shao-qing Wang.

References

[1] Ling Zhao, Ya Liusan and Zhao Gang, "Construction of digital educational resources supermarket," *J. Harbin Inst. Technol.(New Series). Harbin*, vol. 15, no. SUPPL, pp. 57-61, 2008.

[2] Xiangzeng Meng and Lei Liu, "On retrieval of flash animations based on visual features," Lecture Notes in Computer Science, vol. 5093, no. 2008, pp. 270-277, 2008.

[3] Mesyura V.I., Khoshaba O.M. and Yukhimchuk S.V., "Effective decisions for constructing a model for information resources management in distance education," *Upravlyayushchie Sistemy i Mashiny*, no. 4, pp. 45-49, 2004.

[4] Xiangzeng Meng, Hua Zhang, Xiangying Wang, et al., "Semantic Extraction for WWW Images," Progress on Chinese artificial intelligence, People's Post and Telecommunication Publishing House, 2003.

[5] Xue Ruojuan, "An effective approach for instructional resource database construction with Web images," *Key Eng. Mater.*, vol. 439-440, no. 2010, pp. 1361-1366, 2010.

[6] Shen H.T., Ooi B.C. and K.-L. Tan, "Giving meanings to WWW images," In Proceedings of 8th ACM International Conference on Multimedia (ACM Multimedia 2000), pp. 39-47, 2000.

[7] Zhiguo Gong, Leong Hou U. and Chan Wa Cheang, "Web image semantic extractions from its associated texts," In Proceedings of Proceedings of the Eighth IASTED International Conference on Internet and Multimedia Systems and Applications, pp. 97-102, 2004.

[8] Wenpeng Lu, Ruojuan Xue, Haixia Li, et al., "A strategy of semantic information extraction for Web image," In Proceedings of 2009 International Conference on Business Intelligence and Financial Engineering, pp. 480-483, 2009.

Ruojuan Xue was born in Shandong, China. She received the bachelor and master degree in educational technology in July 2002 and July 2005 from Shandong Normal University. Her current research interests are supporting technology for cooperative learning, E-learning. Besides, she has begun to do some basic research on information security.

She worked at Shandong Normal University after her graduation. She has participated in multi researches on educational technology and development projects on monitor and control system.

Wenpeng Lu was born in Shandong, China. He received the bachelor and master degree in educational technology in July 2002 and July 2005 from Shandong Normal University.

He worked at Shandong Polytechnic University after his graduation. He participated in multi research and development projects on monitor and control system. His current research interests are in natural language processing and artificial intelligence.

Jinyong Cheng was born in ShanDong, China, on January 31,1981. He received his Bachelor of Science degree in applied chemistry from Shandong Normal University, Jinan, China, in 2002, and Master of Engineering degree in computer software and theory from Qingdao University, Qingdao, China, in 2005.

He is an instructor in Shandong Polytechnic University. His research interests include image processing, wavelet analysis, and biological information processing.