ICIC Express Letters ICIC International ©2017 ISSN 1881-803X
Volume **11**, Number **3**, March **2017** pp. **633–641**

FLCSE: A NOVEL FEATURE EXTRACTION METHOD FOR CONTENT STRUCTURES IN FLASH ANIMATION RESOURCES

Zhenguo Xu[1], Jing An[1], Xiangzeng Meng[1,*] and Xin Wang[2]

[1]School of Communication
[2]School of Information Science and Engineering
Shandong Normal University
No. 88, East Wenhua Road, Jinan 250014, P. R. China
[*]Corresponding author: mxz@sdnu.edu.cn

Received July 2016; accepted October 2016

Abstract. *With the developments of Internet and multimedia technologies, Flash animation as one of the most popular multimedia resources has presented explosive growth. More and more Flash resources can be downloaded via search engines on the Internet. However, there are difficulties in the quality and accuracy of retrieval, due to the complex structures and diverse content features. The content structure features which include metadata information, component elements, dynamic effects and interactive modes are important parts of the retrieval for Flash resources. They determine the efficiency and quality of retrieval. The shortages of research about feature extraction for content structures seriously impose restrictions on the dissemination and utilization of Flash resources. In order to solve this problem, this paper proposes a novel method called FLCSE (Flash content structure extraction) which is used for extracting content structure features. The method can extract the content structure features of Flash animation, such as metadata information, component elements, dynamic effects and interactive modes. The research has an important and practical significance on improving the efficiency and accuracy of the retrieval system for Flash resources. On the basis of optimizing the efficiency of FLCSE, we build an experimental prototype of FLCSE through C++. Several simulation experiments based on three parameters which include precision ratio, recall ratio and time efficiency, have been carried out respectively. Results show that FLCSE not only has good performances of recall ratio and precision ratio, but also has stable performances of time efficiency.*
Keywords: Retrieval of Flash resources, Content structure features, Component elements, Dynamic effects, Interactive modes

1. **Introduction.** With the leap-style development of multimedia technology, there are an increasing number of Flash animation resources that can be retrieved and downloaded. Flash can integrate various types of media and overcome monotonous display mode of single-media, so it is adored by professionals and amateurs. Relevant researches [1] show that more than 500 million devices are addressable today with Flash technology, and it is projected there will be over 1 billion addressable devices by the end of 2016. Flash is a vector-based movie format with a variety of interactive functions. It can be embedded with text, shapes, images, audio, video and other media elements. Many irregular effects can be generated through shape scaling, shape rotating, color changing and transparency changing. Movie clips, buttons and ActionScript can be used to generate flexible interactions and actions [2]. Compared with other multimedia, Flash has outstanding artistic expression. Due to the unique advantage of Flash resources, they can spread in the Internet rapidly. According to statistics, there are nearly 5 million Flash resources in Chinese Internet [3]. Therefore, how to retrieve Flash resources quickly and accurately has become a hot issue in current researches.

The feature extraction of content structures in Flash resources is the basis of content-based retrieval system. It is also the quantitative expression of content features. It can intuitively reflect the content features of Flash resources which can be directly applied to retrieval. It is an important part of the content feature extraction. Because the extraction process is complicated and involves a lot of contents, there are hardly any researches in this field. Aiming at this problem, this paper proposes a novel method for extracting structure feature of Flash based on the research of Flash file structure. The method extracts the structure feature of Flash resources from four aspects respectively, that is metadata information, component elements, dynamic effects and interactive modes. Results will be stored in the database and can be used to improve the performance of content-based retrieval system for Flash. As far as we know, this paper is the first work in the field of extracting content structure features. We have developed the prototype of FLCSE which includes the optimization of algorithm, practical application and experimental platform. This research has important influence on promoting performance and expanding functions of current search engines.

At present, the Flash resources can be retrieved mainly through the keyword search, such as Google, and Flash Kit [4]. Those contents that have been retrieved mainly involve text, images, audio, video, but without the visual features and interaction features. At present, researches about retrieval system for Flash resources mainly concentrate on the external characteristics of the animation and contextual information [5]. Adobe has made a statement that they will cooperate with Google, Yahoo and other search engines to optimize technology of Adobe Flash Player. This would enhance capabilities of search engines for the contents of Flash resources. However, there is not new research report or new functions in the retrieval system in this field, so far [6].

Content-based multimedia retrieval made up for the limitations of keywords-based retrieval. They have been studied in the field of database retrieval of image, audio, video, and have achieved remarkable results [7, 8, 9]. However, there are fewer researches about content-based retrieval due to the complexity of contents and the characteristic of nonlinearity structures. At present, Yang et al. have made some researches results in this field. Yang et al. have put forward a content-based retrieval method of Flash resources which is based on FLAME (Flash movies access and management) [10]. The method divided the content features of the Flash resource into three levels: objects, events and interactions. Based on the above research, Yang et al. proposed three-layer structures in the Flash resource retrieval system, presentation layer, index layer, and retrieval layer [11].

Ding et al. proposed the Flash resources retrieval method based on semantic expression. It takes ETS (expressive text sensing) as main framework [12]. The method extracts scene complexity, interactivity, movie mood, movie rhythm and movie category from flash file to bridge the feature gaps between low level and high level. Meng divided content features of Flash resources into four levels [13]. The first layer is the overall content features of Flash resources. The second layer is content features of logical scene. The third is the content features of visual scene. The fourth layer is content features for component elements. Based on the above study, Meng et al. proposed the content-based retrieval system [14]. Other researchers also proposed Flash resources retrieval method based on visual features [15], which splits a series of scenes on time line.

Above researches mainly focused on the theory and overview of the retrieval of content features. There are neither specific extraction algorithms, nor implementations of visual research platforms. They focused on the content feature extraction of component elements. However, there are not in-depth studies in interactive features and dynamic effects. Those works do not involve the content structure features. Aiming at shortages of the above research, this paper proposes a new feature extraction method for content structures, FLCSE.

The remainder of this paper is organized as follows. Firstly, the feature extraction of content structures is introduced in Section 2. Then, the algorithm of FLCSE is described in detail in Section 3. Next, the results of simulation experiments and their analysis are presented in Section 4. Finally, Section 5 will summarize the outlooks and conclude the paper.

2. **The Feature Extraction of Content Structures.** The feature extraction of content structures is in the second level of the content feature extraction. Due to the complexity and difficulty of extraction, there are fewer related researches in this field [15]. The results which are extracted from content structure feature would be expressed by formal description language which can be identified by computers and can be applied to retrieval of Flash resources. The efficiency and accuracy of the retrieval system based on content structure features are higher than the content-based retrieval system. According to the storage structure and generating principle, FLCSE analyzes whether there exist component elements, dynamic effects and interactive modes in the Flash resources. If any, FLCSE will analyze which kind of the component elements, the dynamic effects, the interactive modes it is and calculate the number of these objects. The component elements of Flash include text, shape, image, audio, video, button and movie clips mainly. The dynamic effects of Flash can be analyzed from its movement, color changing, shape changing, scaling and rotation. The analysis of Flash's interactive modes includes key up event, key down event, mouse up event, mouse down event, mouse move event, mouse drag over event, mouse rollout event, mouse press event, and clip unload event mainly. The extraction results are used to describe and retrieve Flash resources. FLCSE can enrich and refine the theory of content-based multimedia retrieval system, and can make up the blank in the research of the feature extractions of content structure.

3. **FLCSE.** FLCSE mainly includes four levels: the extraction of metadata information, the extraction of component elements, the extraction of dynamic effects and the extraction of interactive modes. The most importance is that different process levels are distinguished by a three-byte signature in the header of Flash files. The header begins with a three-byte signature either 0×46, 0×57, 0×53 (FWS) or 0×43, 0×57, 0×53 (CWS) [16]. FWS indicates that the Flash is uncompressed. CWS indicates that the entire file after the first 8 bytes was compressed by ZLIB [17]. Before analyzing file structures, Flash files must be uncompressed. Extracting the signature is the first step for the analysis of content structure features. The difficulties are identifying the types of signatures according to the main label of Flash files. Then the relative processing is carried out according to different types of signatures. The main label of Flash file is made up of different fields. Each field has different storage structures, so the analysis is relatively complicated. Flash resource file storage structure is shown in Figure 1 [16].

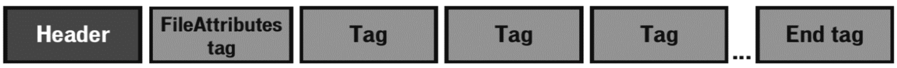

FIGURE 1. The Flash file structure

FLCSE divides the feature extraction of content structures into four levels. The first is the pretreatment of Flash resources. FLCSE decides whether to decompress Flash resources according to the signature, and then reads the header of flash files. The second is that FLCSE reads the main label of Flash resource file, and then determines whether it need further analysis according to the signatures or not. The third is to analyze storage structure of label specifically and to extract the content structure features. Finally, FLCSE stores the extracting results into a database. The algorithm of FLCSE is given as Algorithm 1. The details of FLCSE are presented in the following part.

Algorithm 1 The Algorithm of FLCSE

Input:
The set of experimental Flash resources, $F_n = \{F_i | i \in n\}$;
Output:
The set of content structure features, $SC_n = \{SC_i | i \in n\}$;

```
1:  if(Open(F_i))
2:      Res(header);
3:      if(F_i.Compress)
4:          F_i.Uncompress(), flag: C → F;
5:          reload Flash;
6:          label=true;
7:          while(label) do
8:              Lvalue=LabelRead()
9:              if(Lvalue==0)
10:                 label=false;
11:                 goto(7);
12:             for(k=0; k<t; k++)
13:                 if(Lvalue==CE[k])
14:                     ECE(), goto(7);
15:             if(Lvalue==26)
16:                 EPO2();
17:                 DPlaceObject2(), goto(7);
18:             if(Lvalue==70)
19:                 EPO3();
20:                 DPlaceObject3(), goto(7);
21:             if(Lvalue==34)
22:                 EDB2(), goto(7);
23:     else
24:         execute(7)-(22);
25: return SC_n;
```

Open(F_i): It represents an experimental flash object that can be opened. And its content structure information can be accessed. The mode of opening file is CFile::modeRead (a read-only pattern of opening files in the C++).

Res(header): It is the analysis process of header of experimental Flash object.

***F_i*.Compress**: It denotes whether the Flash object has been compressed. CWS indicates that the document was compressed by ZLIB. FWS indicates the opposite.

***F_i*.Uncompress()**: It denotes the uncompress operation, which can directly invoke the function Uncompress() provided by ZLIB. When the Flash has been decompressed, the first segment in the header will be changed from C to F. The formal representation of this process is: $C \to F$.

LabelRead(): FLCSE reads label to get value. Every label has the only identity value. We can use this only value to judge the label's type and function.

CE[k]: It is an array of identity values of defined labels which is used to define component elements.

ECE(): This operation is used to judge what types of component elements are according to the identity value of tag, and then extract content features of each component element.

EPO2(): It is a method for extracting dynamic effects of the control label, PlaceObject2. Compared with the segment PlaceFlagHasColorTransform by using the Boolean AND operator, if the result is true, there is a discoloration. Compared with the segment

PlaceFlagHasRatio by using the Boolean AND operator, if the result is true, there is a distortion. Compared with the segment PlaceFlagHasMatrix by using the Boolean AND operator, if the result is true, there is a matrix. FLCSE can analyze whether there are changes of movement, rotation and scaling through further analyzing the matrix.

EPO3(): It is a method for extracting dynamic effects of the control label, PlaceObject3. Compared with the segment PlaceFlagHasColorTransform by using the Boolean AND operator, if the result is true, there is a discoloration. Compared with the segment PlaceFlagHasRatio by using the Boolean AND operator, if the result is true, there is a distortion. Compared with the segment PlaceFlagHasMatrix by using the Boolean AND operator, if the result is true, there is a matrix. FLCSE can analyze whether there are changes of movement, rotation and scaling through further analyzing the matrix.

DPlaceObject2(): It is a method for extracting interaction features of the control label, PlaceObject2. Compared with the segment PlaceFlagHasClipActions by using the Boolean AND operator, if the result is true, there is a tag, CLIPEVENTFLAGS. Compared with corresponding field, FLCSE can further analyze the CLIPEVENTFLAGS by using the Boolean AND operator. If the result is true, there is a corresponding type of interaction.

DPlaceObject3(): It is a method for extracting interaction features of the control label, PlaceObject3. Compared with the segment PlaceFlagHasClipActions by using the Boolean AND operator, if the result is true, there is a tag, CLIPEVENTFLAGS. Compared with corresponding field, FLCSE can further analyze the CLIPEVENTFLAGS by using the Boolean AND operator. If the result is true, there is a corresponding type of interaction.

EDB2(): It is a method for extracting interaction features of defined label, DefineButton2. FLCSE reads the field of actions whose type is BUTTONCONDACTION. Compared with corresponding field of the tag of BUTTONCONDACTION, if the result is true, there will be a corresponding action of interaction.

4. **Simulation Experiment.**

4.1. **The experimental setup.**
We use C++ to develop the prototype experimental platform of FLCSE, as Figure 2. The experimental contract is the shockwave Flash 7.4 [18], which is a kind of professional decompiling software of Flash. The experimental resources are downloaded from Internet through the crawler developed C++ by ourselves. We divide Flash resources into eight types: 3D Flash resource, MTV, animation, advertisement, greeting cards, courseware, websites and games. In order to assure the accuracy of experimental results, we remove the invalid Flash resources that cannot play normally and select 1000 samples randomly from each type of Flash resources.

4.2. **Experimental parameters.**
In order to evaluate the performance of FLCSE, we set three evaluation parameters: recall ratio, precision ratio and time efficiency. The recall ratio and the precision ratio are important indexes to measure the performance of decompiling platform [19]. The time efficiency is the consumption function when FLCSE extracts different types of sample with numbers.

(1) Precision ratio is the specific value of the number of content structure features extracted correctly to the total number of the content structure features which is extracted by FLCSE. The formal representation is as follows:

$$QAR(N) = \frac{|\{flcse(O)\}| - |\{flcse(o)|o \in O\}|}{|\{flcse(O)\}|}$$

N represents all the Flash objects. *flcse*() denotes the extraction of content structure feature. o denotes the Flash object which is wrongly extracted by FLCSE. O denotes all Flash objects whose content structures have been extracted.

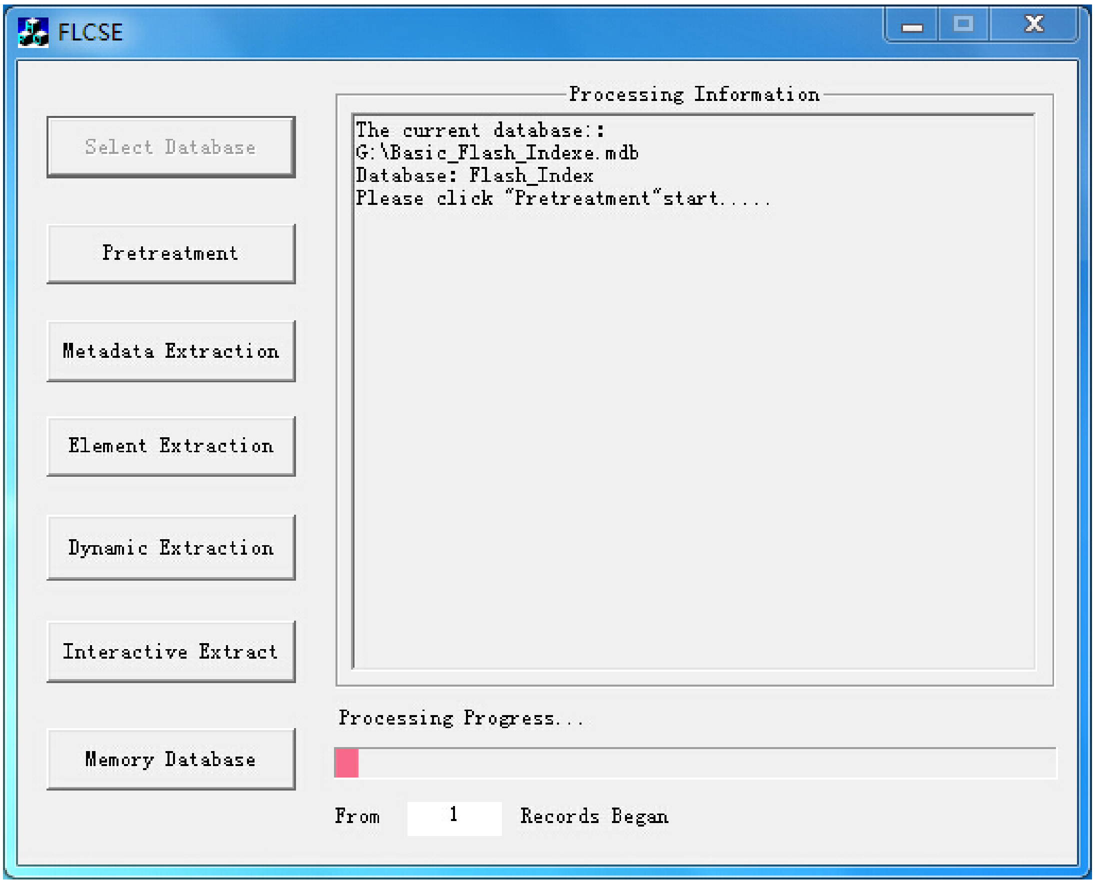

FIGURE 2. The experimental platform of FLCSE

(2) Recall ratio is the specific value of the number of content structure features extracted correctly to the total number of the content structure features which is extracted by the shockwave Flash. The formal representation is as follows:

$$QTR(N) = \frac{|\{flcse(O)\}| - |\{flcse(o)|o \in O\}|}{S(O)}$$

N represents all the Flash objects. $S(O)$ denotes that the total number of content structure features which is extracted by the shockwave Flash. *flcse*() denotes the extraction of content structure feature. o denotes the Flash object which is wrongly extracted by FLCSE. O denotes all Flash objects whose content structures have been extracted.

(3) Time efficiency is the time consumption which evaluates the performance of FLCSE with different numbers of samples. The formal representation is as follows:

$$T(N) = \sum_{1}^{n} (t_m + t_c + t_d + t_i)$$

n represents the number of one type of Flash resources. t_m indicates the time consumption of extracting the metadata information. t_c indicates the time consumption of extracting the component structures. t_d indicates the time consumption of extracting the dynamic effects. t_i indicates the time consumption of extracting the interactive features.

4.3. **Results analysis.** In this section, we present the experimental results and the performance analyses of FLCSE.

(1) In the aspect of precision ratio, the performances of FLCSE in types of advertisement and MTV are higher than other types of resource extraction, as Figure 3. Because the application of advertisement and MTV is relatively early on the network, meanwhile they are types of Flash resources which enjoy perfect technologies. Although the content

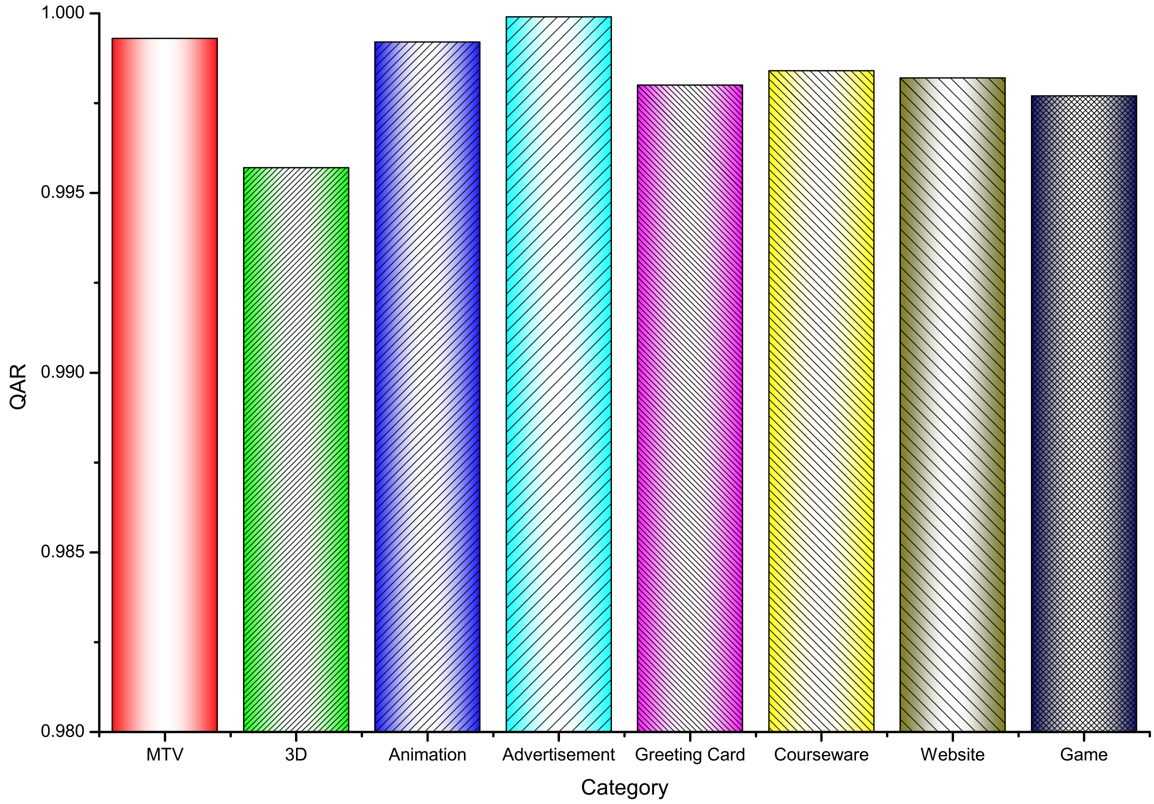

Figure 3. The precision ratio

structure features are complex, the extraction processes are relatively easy. Therefore, it owns a higher value of precision ratio. From an overall perspective, all the precision ratio of different types of Flash resources is over 99%. However, the performance in the 3D is lower than the average precision. This is related to characteristics of 3D Flash resources. The technology of this type is not perfect, and the extraction process is more complex.

(2) In the aspect of recall ratio, the performances of FLCSE in advertisement, greeting cards and MTV are better, but there are still a lot of deficiencies in 3D, as Figure 4. The technologies of advertisement, greeting cards and MTV are relatively mature; meanwhile definition methods of content structures are relatively fixed. Therefore, the recall ratio of these three types is higher than other types. The component elements of the 3D resources are simple, but the dynamic effects and interaction are more complicated, and dynamic effects and the interactions are implemented by ActionScript scripts. Therefore, the recall ratio of 3D is lower than other types.

(3) In the aspect of time efficiency, the time consumptions of FLCSE increase with the number of experimental samples, as Figure 5. Time Efficiency mainly depends on the number of experimental samples. The time efficiency in cartoon is higher than the time efficiency in greeting card. The main reason is that it has complex cartoon components, many dynamic effects and higher logical scenarios, so time consuming is higher than other types. The advertisement scenes are simple, and usually have only 2-3 scenarios. These scenarios are mainly including text, shapes and images, so the time efficiency of this type is lower than other types.

5. **Outlooks and Conclusion.** With the development of multimedia technology, Flash resources are becoming more and more abundant in the Internet. Therefore, the retrieval methods have drawn the attention of academic and search engine providers. After many years of study, researchers have been making remarkable achievements in the field of the extraction of content structures. However, the study about the feature extraction

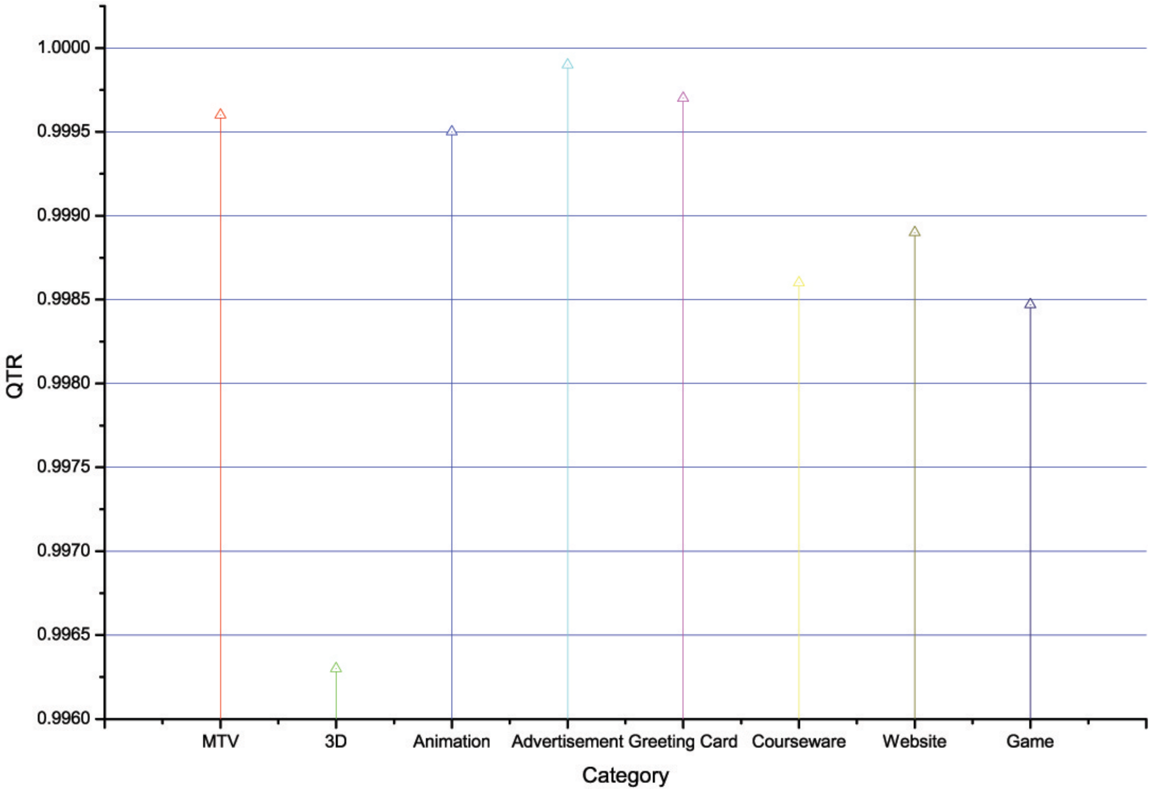

FIGURE 4. The recall ratio

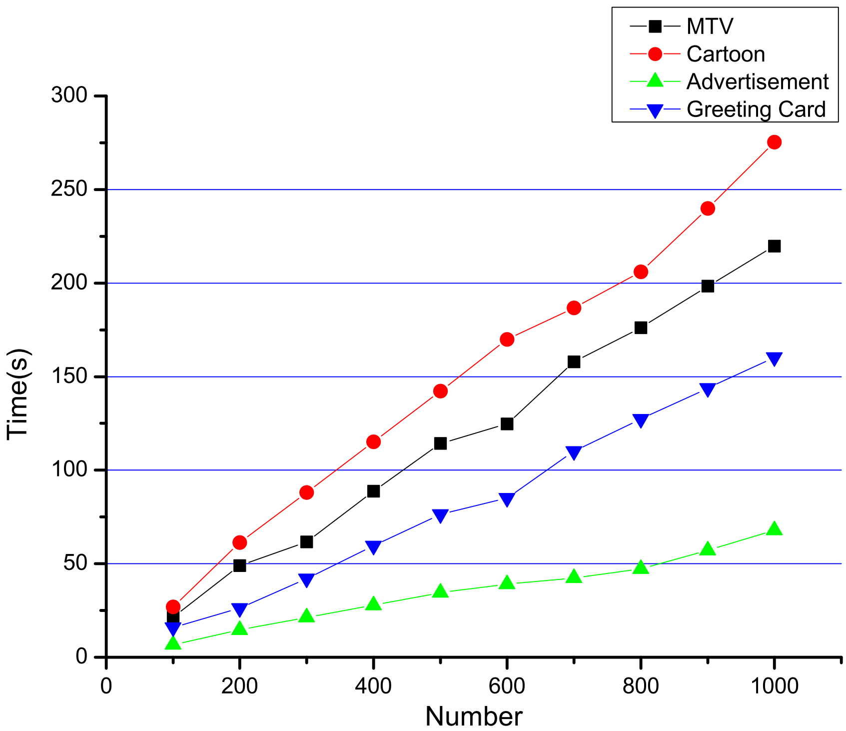

FIGURE 5. The time efficiency

of content structures is still in the initial phase. The component elements, dynamic effects and interactive features are important content structure features of Flash resources. They can make the retrieval systems more comprehensive and personalized, and improve the accuracy and efficiency of retrieval systems. Content structure features of Flash are complex. Different media elements of different types have their own characteristics, so improvements of precision ratio and recall ratio of the Flash retrieval system need further studies. With the continuous development of technology being used in the field of multimedia content analysis, the analysis and feature extraction of Flash resource content are gradually becoming mature. The classification and retrieval based on the content structure features of Flash would be a hot topic in the future. We will optimize the accuracy and efficiency of FLCSE ulteriorly in the following researches. And we will make up the shortages of FLCSE in the field of 3D and games.

Acknowledgment. This work is partially supported by the Science and Technology Development Project of Shandong Province (No. 2011GGX10112).

REFERENCES

[1] *Macromedia Flash Player Adoption Statistics*, http://www.macromedia.com/software/player_census/Flashplayer, 2016.
[2] S. Reimers and N. Stewart, Adobe Flash as a medium for online experimentation: A test of reaction time measurement capabilities, *Behavior Research Methods*, vol.39, no.3, pp.365-370, 2007.
[3] Z. Xu and X. Meng, The research of flash content feature extraction, *e-Education Research*, no.6, pp.55-60, 2015.
[4] *Flash Kit*, http://www.flashkit.com/index.shtml, 2016.
[5] J. Yang, Q. Li, W. Liu et al., Content-based retrieval of $Flash^{TM}$ movies: Research issues, generic framework, and future directions, *Multimedia Tools and Applications*, vol.34, no.1, pp.1-23, 2007.
[6] X. Meng, Analysis on content construction of teaching resources in flash web, *e-Education Research*, no.10, pp.81-85, 2010.
[7] J. T. Foote, Content-based retrieval of music and audio, *Voice, Video, and Data Communications*, pp.138-147, 1997.
[8] V. N. Gudivada and V. V. Raghavan, Content based image retrieval systems, *Computer*, vol.28, no.9, pp.18-22, 1995.
[9] H. J. Zhang, J. Wu, D. Zhong et al., An integrated system for content-based video retrieval and browsing, *Pattern Recognition*, vol.30, no.4, pp.643-658, 1997.
[10] J. Yang, Q. Li, W. Liu et al., FLAME: A generic framework for content-based flash retrieval, *The 4th International Workshop on Multimedia Information Retrieval, in Conjunction with ACM Multimedia*, 2002.
[11] J. Yang, Q. Li, W. Liu et al., Searching for flash movies on the web: A content and context based framework, *World Wide Web*, vol.8, no.4, pp.495-517, 2005.
[12] D. Ding, J. Yang, Q. Li et al., Towards a flash search engine based on expressive semantics, *Proc. of the 13th International World Wide Web Conference on Alternate Track Papers & Posters*, pp.472-473, 2004.
[13] X. Meng, On the content-based retrieval of flash educational resources on Internet, *e-Education Research*, no.9, pp.77-79, 2009.
[14] X. Meng, Z. Xu and R. Liu, A method of searching for flash movies on Internet based on component and structure, *Journal of Library Science in China*, vol.42, no.1, pp.83-95, 2016.
[15] X. Meng and L. Liu, On retrieval of flash animations based on visual features, *International Conference on Technologies for E-Learning and Digital Entertainment*, pp.270-277, 2008.
[16] Adobe Corporation, *SWF File Format Specification Version 10*, http://www.adobe.com/devnet/swf.html, 2016.
[17] P. Deutsch and J. L. Gailly, *ZLIB Compressed Data Format Specification Version*, 1996.
[18] *Shan Ke Jing Ling*, http://www.shankejingling.com, 2016.
[19] M. K. Buckland and F. C. Gey, The relationship between recall and precision, *JASIS*, vol.45, no.1, pp.12-19, 1994.