원격 전기안전 교육을 위한 모바일 에이전트 시스템 설계

조현섭, 유인호*, 장성환**,유기수***

청운대학교 디지털방송공학과, *국립익산대학 전기과, **원광대학교 전기공학과, ***중소기업청
e-mail : chohs@chungwoon.ac.kr

Design of Mobile Agent System for Remote Electric Safety Education

Hyun-Seob Cho In-Ho Ryu* Sung-Whan Jang** Ki-Soo Rheu***
Dept of Digital Broadcast Engineering Chungwoon University, *Iksan National College
**Dept of Electrical Engineering Wonkwang University, ***Small and Medium Business Administration

Abstract -To effectively deal with video data, a semantic-based retrieval scheme that allows for processing diverse user queries and saving them on the database is required. In this regard, this paper proposes a semantic-based video retrieval system that allows the user to search diverse meanings of video data for electrical safetyrelated educational purposes by means of automatic annotation processing. If the user inputs a keyword to search video data for electrical safety-related educational purposes, the mobile agent of the proposed system extracts the features of the video data that are afterwards learned in a continuous manner, and detailed information on electrical safety education is saved on the database. The proposed system is designed to enhance video data retrieval efficiency for electrical safety-related educational purposes.

1. Introduction

With advancements in network and multimedia data compression technologies, rapid progress has recently been made in technologies that enable multimedia content services.

As a result, user demand for large-capacity video data in mobile environments is growing. To meet the diverse needs of different users, a vast amount of video data needs to be effectively managed [1]. The effective and efficient management of video data requires a technology that enables systematic classification and integration of large-capacity video data. In addition, a system allowing for effective retrieval and storage of video data should be in place to provide users with information on video data according to diverse user environments, for example, on mobile terminals [2]. However, compared with desktop computers, mobile terminals have many inherent restraints such as low CPU processing rate/bandwidth/battery capacity, and small screens[3,4]. In particular, low CPU processing rates and bandwidths are key inhibitors to servers that aim to provide seamless multimedia data. To ensure the effective retrieval and playout of video data on mobile terminals, CPU performance of terminals must be improved, and network technologies and systematic video data indexing technologies must be developed. Currently, studies that focus on addressing such restraints on mobile terminals and effectively indexing video data are being actively conducted [5,6]. However, such video indexing methods are based on simply classifying video genres or types, rather than on reflecting user needs. To attach various types of information to video data is difficult since video data contain no textualized information relative to typical text data. As such, there is a need for semantic-based retrieval using additional information such as frames, key frames, and annotations, in video. It is very important to make information on video data more systematic and concrete so that the user can perform content-based retrieval of such video data [7].

2. Video Retrieval Scheme

Studies currently being conducted on content-based video data retrieval are largely divided into the following: 1) a feature-based retrieval scheme that uses similarity by extracting features from key frames; and 2) an annotation-based retrieval scheme in which a comparative retrieval is performed on a user's annotation that has been inputted and saved for a key frame [8]. However, both content-based video data retrieval methods have some drawbacks. In the case of the annotation-based retrieval scheme, retrieval is performed such that, if the user attaches annotations using characters to the semantic information of individual video data whose automatic recognition is difficult, the video data are saved on the database and extracted by already attached annotations during retrieval[9].

3. Video Retrieval System Based on Mobile Agent
3.1 Automatical Indexing Processing

The indexing agent extracts keywords from the queries submitted from the user for ideo retrieval, and matches them with basic annotation information stored in the metadata. Then, it detects the keyframes that has the same keywords as annotation information, and sends them to the user.

Once entered, queries submitted from the user are analyzed, and keywords are extracted. The user's keywords extracted are matched with the annotation information of metadata stored in the annotation database. As a result of matching, the keyframes having exactly matched keywords are detected in the database, and then they are sent to the user. Additionally, the keywords that do not exactly match annotation information among queries received from the user are defined as potential keywords. If the mobile agent receives queries on retrieval of video contents from the user, it extracts keywords among query words and matches them to the basic annotation information of the metadata. It then detects key frames that have the same keywords as the annotation information and sends them to the user. Figure 1 shows an overall schematic of how annotations are automatically processed by the mobile agent. As illustrated in Figure 1, once a user's query is inputted, a corresponding keyword is extracted following an analysis of the user's query. The candidate key frames that are extracted according to the user's queries sent by the mobile agent are transmitted to the user in the form of exemplified image lists. They are displayed in descending order. If the user selects his/her desired key frame among key frame lists, the selected key frame becomes the secondary query image. Then, potential keywords are included in the annotation information of a corresponding key frame to make the meaning of the key frame more concrete. In addition, every time annotations are automatically processed by users' queries, the keywords extracted from individual users' queries and the selected example images are matched together. The similar weight values of accurately matched keywords are increased while those of unmatched keywords are decreased. This leads to making the meaning of annotations for corresponding key frames more concrete. Once the annotator has inputted basic annotation information, the annotation information on key frames is automatically updated whenever the user performs a retrieval of video contents. This significantly reduces the problems that may occur when all individual users do not use identical vocabularies to express video scenes.
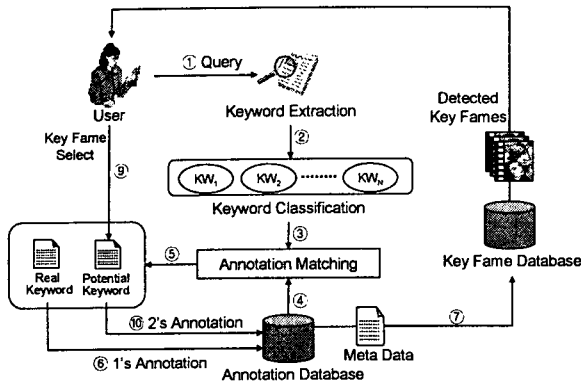
Figure 1. Automatical Annotation Processing

3.2 Automatic Annotation Processing Mechanism

Once a user's query consisting of one or more words is inputted, a corresponding keyword is extracted. The key frames containing users' key words are searched by the extracted keyword inputted by the user. The users' keywords are then classified into "real keywords" and "potential keywords". The accurately matched keywords among keywords in the annotation information are classified as "same keywords", while the accurately unmatched keywords are classified as "difference keywords". Figure 3 shows the classification of keywords which is a preprocessing step to automatic annotation processing.

The agent extracts key frames containing identical keywords and displays key frame lists to the user. If the user selects a specific key frame among the key frame lists, semantic weight values are calculated for each individual keyword that belongs to the specific key frame. Where annotation keywords in the keyframe are same keywords, new semantics weight is calculated as Equ (1):

$$W_{Keyword\_new} = W_{Keyword\_old} + \frac{1}{N_{Kframe\_SK}} \quad (1)$$

While $W_{keyword\_new}$ is the new semantics weight for annotation keywords and $W_{keyword\_old}$ is the previous semantics weight for annotation keywords. N_{Kframe_SK} is the number of keyframes with same keywords. In the meantime, where annotation keywords in the keyframe are difference keywords, new semantics weight is calculated as Equ (2):

$$W_{Keyword\_new} = W_{Keyword\_old} - \frac{1}{N_{Kframe\_SK}} \quad (2)$$

4. Implementation and Experimental Evaluation

4.1 Implementation

Figure 4 shows the interface that allows the user to perform semantic-based retrieval on a mobile terminal. The user can search key frames for his/her desired scenes by inputting several search words.
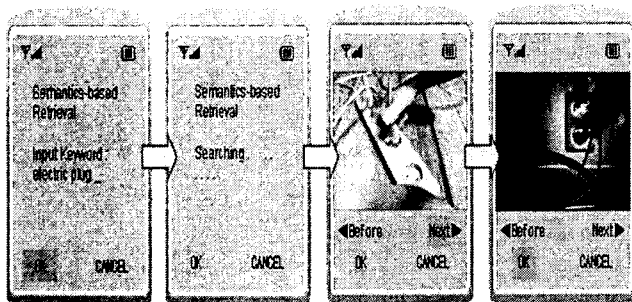

Fig.3. Mobile Client Interface for Annotation-based Retrieval

4.2 Experimental Evaluation

MPEG formatted movie video files were used as domains for video data in order to evaluate the proposed system. For the purpose of this study, we used some 20 movies for electric safety education that had corresponding videos totaling 38 video clip files, and detected a total of 20,292

keyframes. By default, a single annotation was attached to 3,714 keyframes, except that a keyframe has duplicate portions due to successive cuts or that individual objects of a keyframe are indistinguishable. In order to evaluate the proposed system for overall retrieval precision and recall, user queries were performed more than 500 times. Figure 5 illustrates the precision and recall of retrieval for this system.
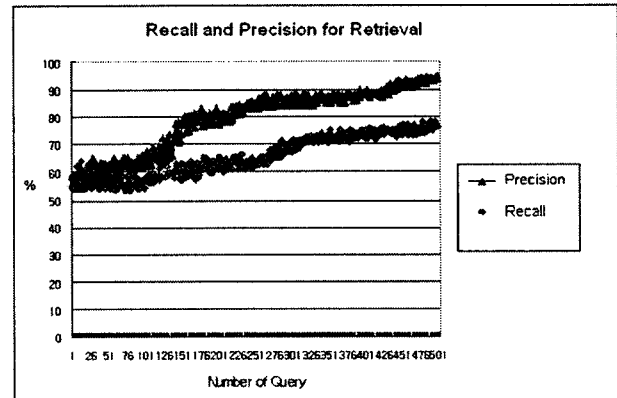

Fig. 5. Retrieval Recall & Precision of the Proposed System

5 Conclusions

This paper presents a video retrieval system that allows the user to perform various semantics retrievals using automatical indexing techniques for large capacity multimedia data in wireless network. As with experiment results, the proposed system was able to improve retrieval precision for semantics-based video data retrieval as well as produce high precision at an approximate rate of 95.9% as a result of testing and evaluating user queries. Proposed system showed precision that improve more using optimized comparison area detection and reduced overhead of system and retrieval time in mobile phone.

References

[1] Sibel Adali, et. al., "The Advanced Video Information System : data structure and query processing," *Multimedia System*, pp.172-186, 1996.
[2] B. Y. Ricardo and R. N. Berthier, Modern Information Retrieval, *ACM press*, 1999.
[3] T. Kamba, S. A. et al., "Using small screen space more efficiently," In *Proceedings of CHI'96*, ACM Press, 1996.
[4] O.Buyukkokten, H. et al., "Power Browser: Efficient Web Browsing for PDAs," In *Proceedings of CHI'2000*, ACM Press, 2000.
[5] Jiang Li. et al., "Scalable portrait video for mobile video communication," *IEEE Trans on CSVT*, Vol. 13, No. 5 pp.376-384, 2003.
[6] C. Dubuc. et al., "The design and simulated performance of a mobile video telephony application for satellite third-generation wireless systems," *IEEE Trans on Multimedia*, Vol. 3, No. 4 pp.424-431, 2001.
[7] Rune Hjelsvold, VideoSTAR - A Database for Video Information Sharing, Ph. D. Thesis, Norwegian Institute of Technology, Nov. 1995.
[8] H. Rehatsch[?], and H. Muller, "A Generic Annotation Model for Video Database," *The 3rd International Conference on VISUAL '99*, Amsterdam, Netherlands, pp.383-390, 1999.
[9] J. R. Smith and S. F. Chang, "VisualSEEK : a fully automated content-based image query system," *ACM Multimedia*, Boston, 1996.
[10] Chong-Wah Ngo, Ting-Chuen Pong, Hong-Jiang ZhangOn, "Clustering and retrieval of video shots through temporal slices analysis," *IEEE Trans on Multimedia*, Vol.04 No.04 pp.0446-0458, 2002.