A Study for Photo Album Retrieval System Based MPEG-7

Byeongtae Ahn*

Liberal & Arts College, Anyang University, Anyang-shi, South Korea. Email: ahnbt@anyang.ac.kr

Abstract. Semantic photo retrieval has been a crux to bridge "semantic gap" between the simple visual features and the abundant semantics delivered by a photo. Therefore, effective photo retrieval using semantics is one of the major challenges in photo retrieval. We propose a semantic retrieval method of photos using photo annotation user interface. A photo album management system that facilitates photo management and semantic retrieval, which fully relies on the MPEG-7 standard as information base, and using a native XML database has been designed and implemented.

Key Words *:* Semantic photo retrieval, MPEG-7 standard, Photo album management

1. Introduction

Photo retrieval research has been on-going for sometime. Most important barrier of multimedia retrieval is lack of a comprehensive, simple and flexible representation of multimedia data. MPEG-7 standard was proposed to solve this problem[1]. But semantic photo retrieval techniques are needed to overcome the drawbacks of conventional photo retrieval approach[2].

In this paper, we propose a MPEG-7 based Photo Album Management System(MPAMS), which supportssemantic photoretrieval using MPEG-7 documents stored in the native XML database system[3].

The work related to the present paper is provided in section 2. Event semantic of photo information is introduced in section 3. Section 4 presents our design process of MPAMS. The implementation of MPAMS is discussed in section 5. A comparison of MPAMS with other systems is given in section 6. Section 7 gives conclusions and future research.

2. Related Works

In the last several years, many endeavors have done for the photo retrieval system using MPEG-7 technologies. MIRRORis a platform for content-based image retrieval research and development [4]. CIRES is an image retrieval system, which enhances the performance of effective feature integration, and supports multiple image queries [5]. RPRWS is a robust retrieval mechanism for annotated photos by a keyword query [6].ASARSOCLT annotates documents to be retrieved with semantic tags that are defined and derived from a set of domain concepts or schemes called domain ontology and thesaurus [7].

As reported above most photo retrieval systems support content-based image retrieval based on the MPEG-7 standard, but not for the semantic photo retrieval. Only ASARSOCLT supports semantic photo retrieval based on ontology, but MPEG-7 descriptors for semantic annotation are not used. Our MPAMS system not only supports event semantic annotation, but also uses MPEG-7 documents stored in the native XML database system[8].

3. Event Semantic of Photo Information

Event semantic is one of the semantic objects of photos, which can be represented using MPEG-7. Event object is focusing on describing event information of photos. Event annotation on photos can be reused if some photos have the same event context.

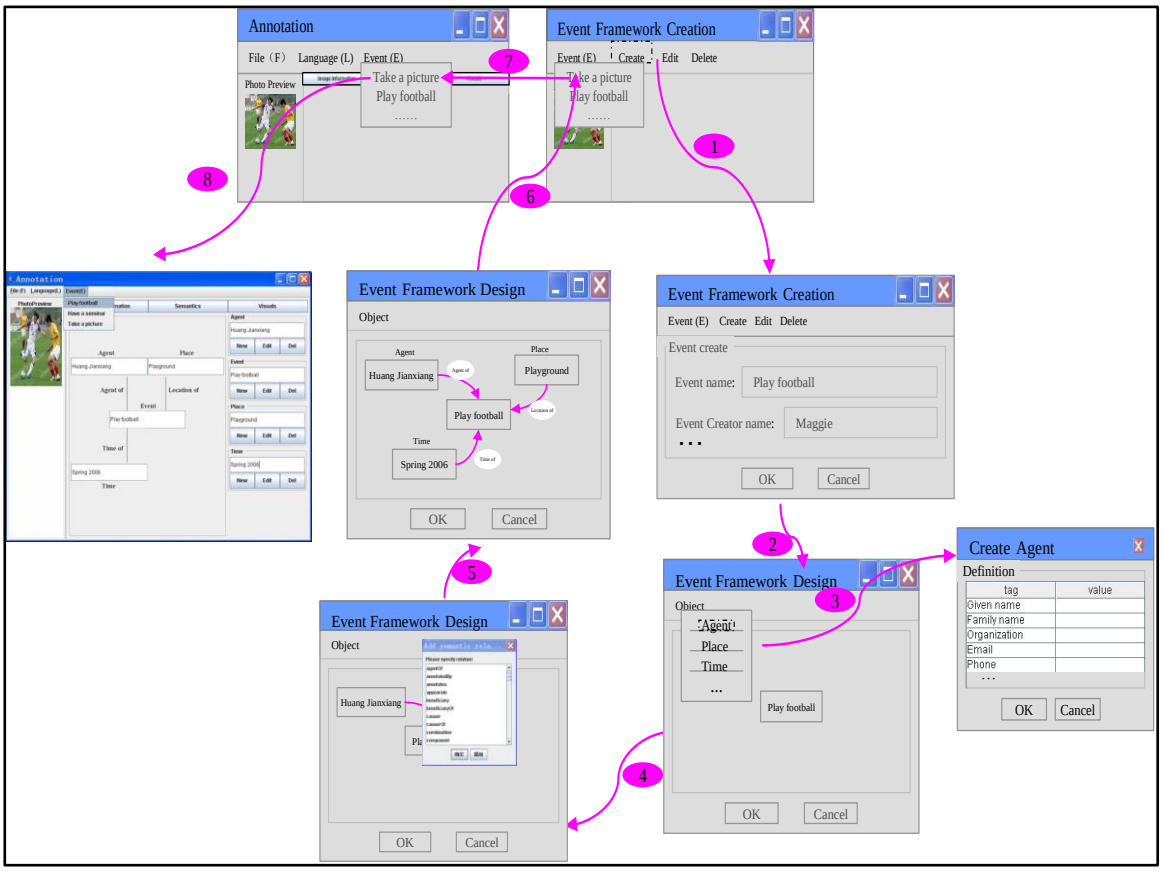

Fig. 1. Event Semantic Creation Processing

MPAMS offers event semantic creation function to support effective exploratory search on photos. The event semantic creation processing in our system is depicted in figure 1.The processing consists of two steps by using two subsystems: event framework creation (①-③), and event semantic creation (④-⑧). In first step, event framework creation subsystem allows event semantic designer to create new event framework of photos. Designer can enter event information in this subsystem manuallyto design event semantic frameworks, like event name, event creator name, and so on. In second step, event semantic creation subsystem allows the

end user to choose one of the semantic frameworks and to generate event semantic of his/her photo[9]. The photo information of event name is shown in the center of the event framework design window and this window allows user to define other semantic object like agents, places, and time. These semantic objects can be used for creating the description by dragging and dropping them onto the panel with the mouse. User can also interconnect these objects by drawing relationships between them using the middle mouse button. The event semantic graph can be saved as part of an MPEG-7 description. After this processing,a new kind of event information is displayed in the event menu on the top of event framework creation subsystem. It also displays on the top the annotation subsystem. End user can reuse this event semantic annotation if some photos have the same event context[10].

4. Design of MPAMS

MPAMS is a multimedia retrieval system focusing on photo objects, which fully relies on the MPEG-7 standard. It also supports the management of the multimedia content and metadata. Main characteristics of MPAMS are : 1. it supports event semantic annotationand semantic photo retrieval. 2. a native XML database system Berkeley DB XML for managing MPEG-7 data. 3. it offers multiple language interfaces toserve different national users. The characteristics of MPAMS are depicted in figure 2.

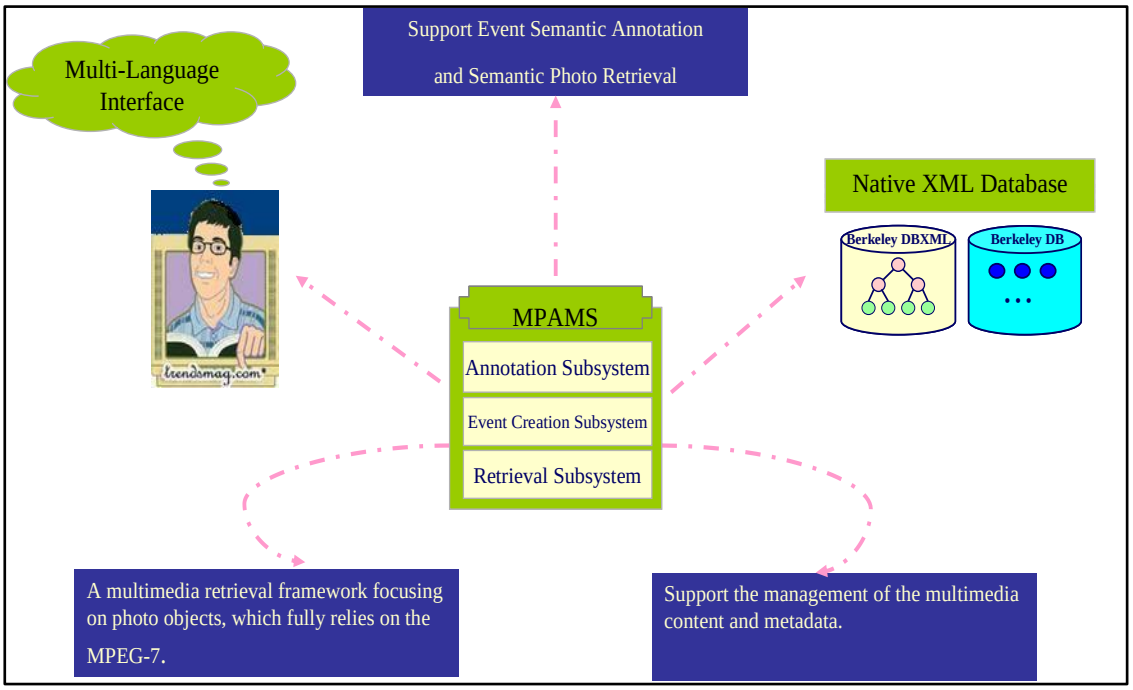

Fig. 2. Characteristics of MPAMS

4.1 System Architecture of MPAMS

MPAMS contains three main components: photo annotation subsystem, photo retrieval subsystem, and multimedia database. The general system architecture of MPAMS is depicted in figure 3.

Availability of high quality content descriptions is essential to provide well performing retrieval methods. This is achieved by content annotation subsystem. The annotation on photos is stored as MPEG-7 conforming descriptions[11].

Photos together with the according MPEG-7 documents are transferred into the multimedia database system. The raw photos database (Photo DB) and their corresponding MPEG-7/XML database (MPEG-7 DB)is separately stored in Berkeley DB and Berkeley DBXML[12].

Photo retrieval subsystem is also operated based on the multimedia database, which performs retrieval and displays the retrieval results.

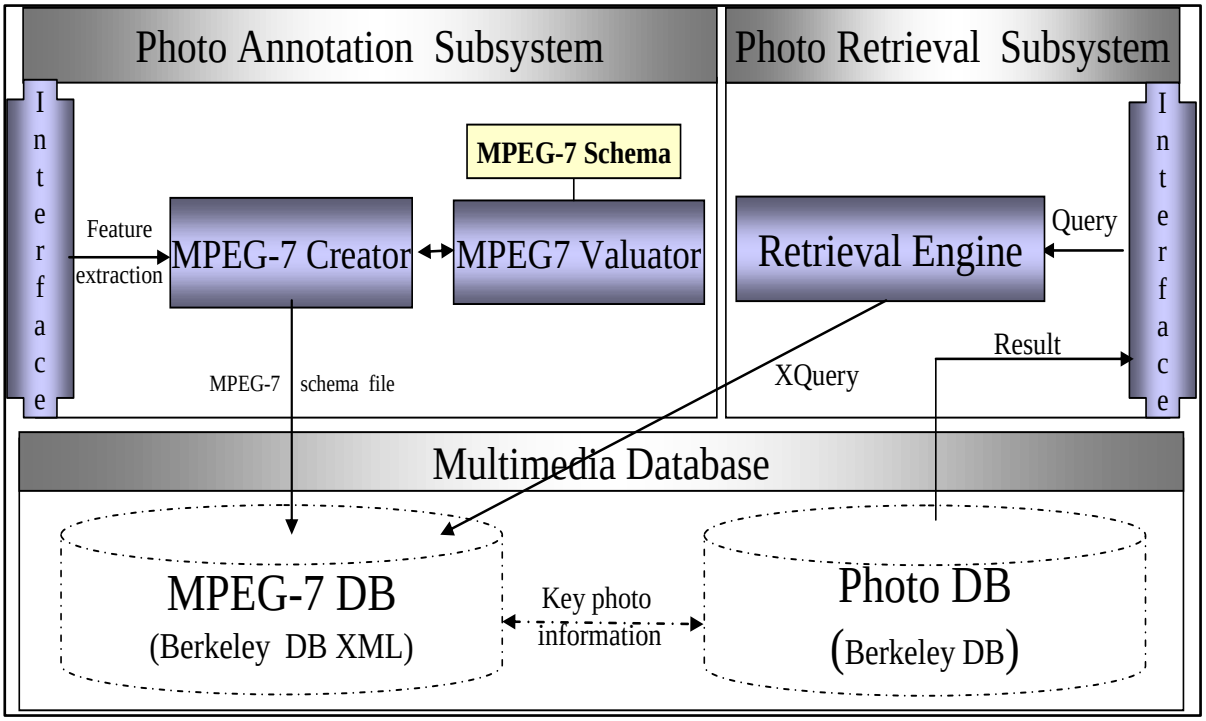

Fig. 3. System Architecture of MPAMS

4.2 Photo Annotation Subsystem

Photo annotation subsystem contains two main components: MPEG-7 Create Module and MPEG-7 Validate Module. Figure 4 presents a detailed architecture of photo annotation subsystem[13].

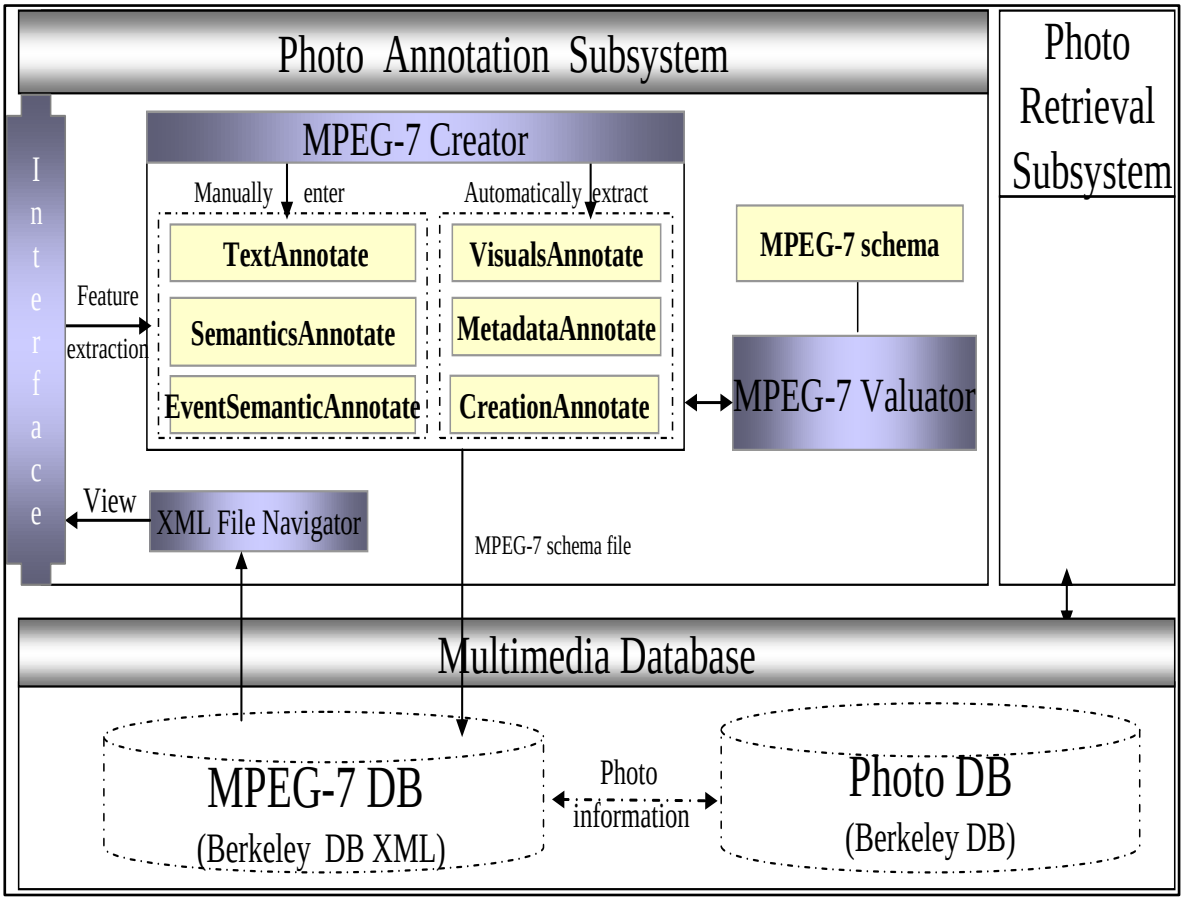

Fig. 4. Architecture of Photo Annotation Subsystem

MPEG-7 create module allows user to enter text, semantics, and event semantic annotation manually. It can also extract visuals, metadata and creation information automatically. This information through MPEG-7 create module creates their description into the corresponding part of the MPEG-7 document, whichis stored into Berkeley DBXML[14].

MPEG-7 valuator module offers validation to ensure the MPEG-7 documentconforms to the specified XML schema.The XML schema is defined according to the W3C XML Schema.

4.3 Photo Retrieval Subsystem

Photo retrieval subsystem contains three main componentsRetrieval Panel, Retrieval Engine Module and Result View. The general architecture of the photo retrievalsubsystem is depicted in figure 5.

Retrieval engine module receives and translates a user's query into its corresponding description in the form of XQuery to search information from Berkeley DB XML.Berkeley DB XML is a native XML database system, which allows XQuery-based access to documents stored in containers. The result of query processing is generated in XML form from the Berkeley DBXML, which is an XML fragment of MPEG-7 metadata. By the matching, the

corresponding photo results from the Berkeley DB are displayed on the result view panel to the user[15].

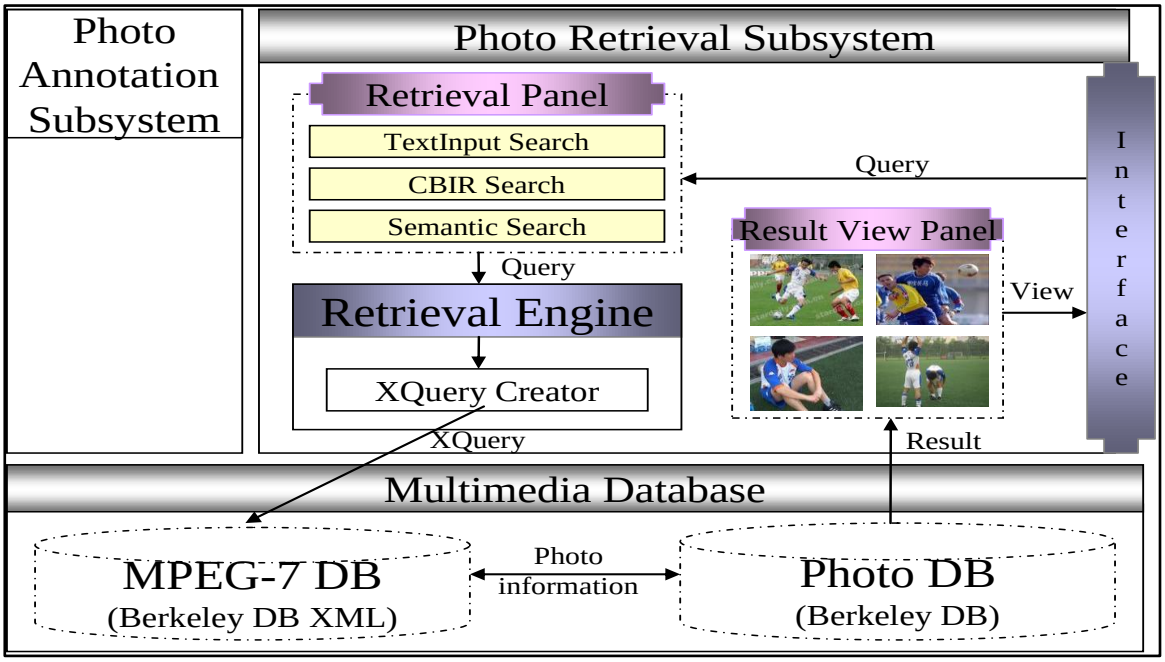

Fig. 5. Architecture of Photo Retrieval Subsystem Multimedia Database

Multimedia database consists of two main parts: Berkeley DB XML, which stores MPEG-7 document. Berkeley DB, which stores raw photos.Architecture of multimedia database is depicted in figure 6.

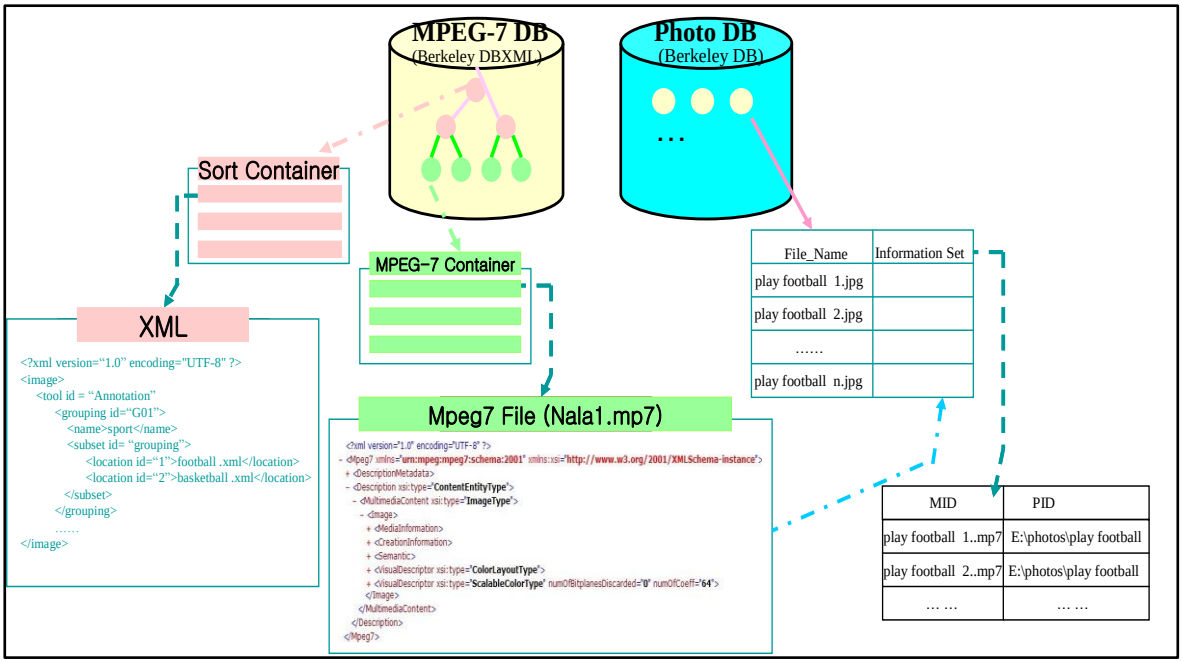

Fig. 6. Multimedia Database Architecture of MPAMS

On the left of multimedia database is Berkeley DB XML. MPEG-7 documents are stored in it. On the top of the Berkeley DB XML is sort container, which store XML file of uniform sort. MPEG-7 container is put underside of the Berkeley DB XML, where the detailed MPEG-7 file is stored. On the right of multimedia database is Berkeley DB, where raw photos are stored. An index with MPEG-7 file and photo is also made to implement the retrieval work.By virtue of the index with them, it is possible to get the correspondingphoto[16].

5. Conclusion and Future research

A combined semantic technology and MPEG-7 standard to supports effective photo retrieval to develop a photo album management system using MPEG-7 (MPAMS) is designed and implemented. An event semantic annotation based on MPEG-7 is proposed for efficient photo retrieval. This work is one of few efforts to manage photo information using MPEG-7 technology based on a native XML database system.

Combining event semantic technology with other semantic technology for more effective photo retrieval approach will be carried out as a future work.

References

1. A. Ono, M. Amano, M. Hakaridani, T. Satou, and M. Sakauchi, "A Flexible Content-Based Image Retrieval System with Combined Scene Description Keyword," IEEE, pp. 201-208, June 2019.
2. B. Bruegge, J. Blythe, J. Jackson and J. Shufelt, " Object-Oriented System Modeling with OMT," OOPSLA '92, 27(10), PP. 414-427, Oct. 2019.
3. D. Gall, "MPEG: A Video Compression Standard for Multimedia Applications," CACM, 34(4), pp. 47-58, April 2019.
4. E. Oomoto and K. Tanaka, "OVID: Design and Implementation of a Video-Object Database System," IEEE Trans. on Knowledge and Data Engineering, 5(4), pp. 62-72, 2019.
5. H. Aoki, S. Shimotsuji, and O. Hori, "A Shot Classification Method of Selecting Effective Key-Frames for Video Browsing," ACM Multimedia 96, pp. 1-10, 2019.
6. H. Frater and D. Paulissen, Multimedia Mania, Abacus, 2019.
7. I. Center, "Query by Image and Video Content: The QBIC System," IEEE Computer, 28(9), pp. 23-32, Sept. 2019.
8. J. Monaco, "How to Read a Film," The Art, Technology, Language, History and Theory of Film and Media, Oxford University Press, 2019.
9. J. Meng and S. Chang, "CVEPS - A Compressed Video Editing and Parsing System," ACM Multimedia 96, pp. 43-53, 2020.
10. J. Rumbaugh, M. Blaha, W. Premerlani, F. Eddy, and W. Lorensen, Object-Oriented Modeling and Design, Prentice-Hall, 2020.

11. J. Smith and S. Chang, "Searching for Images and Videos on the World-Wide Web," Technical Report #459-96-25, Center for Telecommunications Research, Columbia University, New York, August 2020.
12. J. Smith and S. Chang, "VisaulSEEK: a Fully Automated Content-based Image Query System," ACM Multimedia '96, Nov. 2020.
13. K. Hirata and T. Kato, "Query by Visual Example Content-based Image Retrieval," Advances in Database Technology(EDBT '92), pp. 56-71, 2020.
14. M. Davis, "Media Streams: Representing Video for Retrieval and Repurposing," Ph. D. Thesis, Massachusetts Institute of Technology, 2020.
15. R. and J. Gray, "Similar-shape Retrieval in Shape Data Management," IEEE Computer, pp. 57-62, Sept. 2020.
16. R. Hjelsvold, "Video Information Contents and Architecture," In Proceedings of the 4th International Conference on Extending Database Technology, pp. 28-31, March 2020.