
MPEG-7 기반의 이벤트 의미 포토 검색 관리 시스템

Event Semantic Photo Retrieval Management System based on MPEG-7

안병태, 정범석, 이종하
유한대학 경영정보과

Byeong-Tae Ahn(ahnbt@yuhan.ac.kr), Bhum-Suk Chung(cbs622@yuhan.ac.kr),
Chong-Ha Lee(chlee@yuhan.ac.kr)

요약

의미 포토 검색은 포토의 간단한 시각화 특성과 적합한 의미를 분류하는데 있어서의 검색을 간소화 시키는 데 중요한 역할을 한다. 의미 검색을 이용한 효과적인 포토 검색은 포토 검색에 있어서 매우 중요한 과제 중의 하나이다. 따라서 우리는 사용자 인터페이스의 포토 주석을 이용한 새로운 이벤트 의미 포토 검색 기법을 제안한다. 본 논문에서는 순수 XML 데이터베이스와 MPEG-7 표준을 기반으로 포토 관리 및 의미 검색이 쉬운 포토 앨범 관리 시스템을 설계 및 구현하였다.

■ 중심어 : | 의미 포토 검색 | MPEG-7 표준 |

Abstract

Semantic photo retrieval has been an important role in reducing the semantic gap between the simple visual features and the abundant semantics delivered by a photo. Effective photo retrieval using semantics is one of the major challenges in photo retrieval. And we propose a new event semantic photo retrieval method by using photo annotation user interface. In this paper, A photo album management system that facilitates photo management and semantic retrieval, which fully relies on the MPEG-7 standard as an information base and a native XML database, has been designed and implemented.

■ keyword : | Semantic Photo Retrieval | MPEG-7 Standard |

1. Introduction

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]. In order to overcome the drawbacks of conventional photo retrieval approaches, Semantic photo retrieval techniques are needed. In this paper, we propose a MPEG-7 based Photo Album Management System

(MPAMS), which supports semantic photo retrieval for MPEG-7 documents stored in the native XML database system.

Related works are described in Section 2. Event semantic of photo information is introduced in Section 3. Section 4 presents the proposed MPAMS. The implementation of MPAMS is discussed in Section 5. A comparison of MPAMS with existing systems is given in Section 6. Section 7 gives conclusions and

future research.

II. Related Works

In the last several years, many endeavors have done for the photo retrieval system using MPEG-7 technologies. MIRROR is 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 described in [4-7], 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. The proposed MPAMS system not only supports event semantic annotation, but also uses MPEG-7 documents stored in the native XML database system.

III. 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.

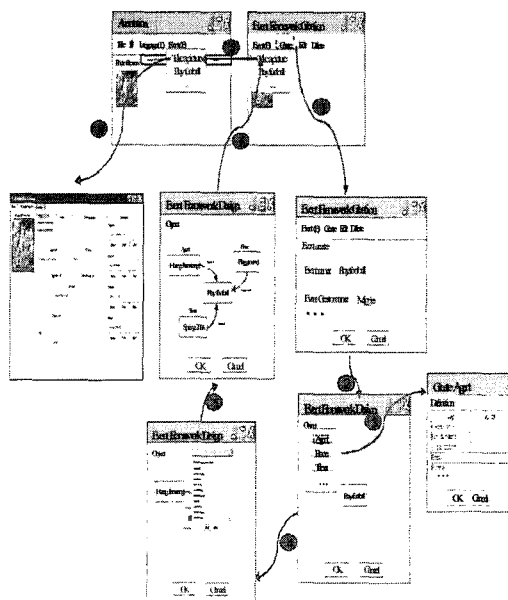


Fig. 1. Event Semantic Creation Processing

MPAMS offers event semantic creation function to support effective exploratory search on photos. The event semantic creation processing in the proposed system is depicted in [Fig. 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 manually to design event semantic frameworks. For instances, 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 photos. The photo information of event name is shown in the center of the event framework design window. This window allows user to define other semantic objects such as 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 by 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.

IV. 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 as follows: (1) It supports event semantic annotation and semantic photo retrieval. (2) A native XML database system is Berkeley DB XML for managing MPEG-7 data. (3) It offers multiple language interfaces to serve different national users. The characteristics of MPAMS are depicted in [Fig. 2].

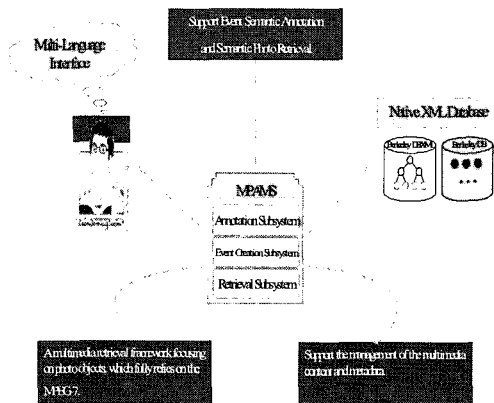


Fig. 2. Characteristics of MPAMS

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 [Fig. 3]. In order to provide well performing retrieval methods, availability of high quality content descriptions is essential. It is achieved by content annotation subsystem. The annotation on photos is stored as MPEG-7 conforming descriptions. Photos together with the 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) are separately stored in Berkeley DB and Berkeley DB XML. Photo retrieval subsystem is also operated based on the multimedia database, which performs retrieval and displays the retrieval results.

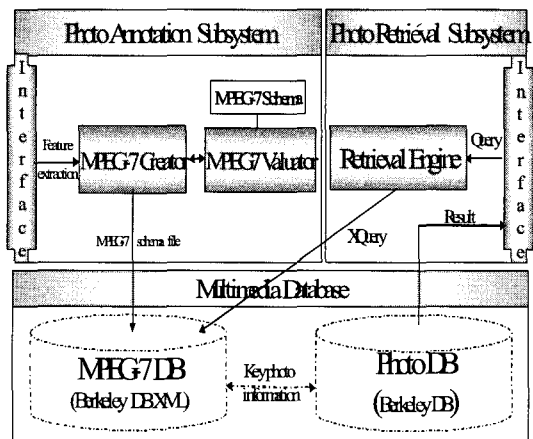


Fig. 3. System Architecture of MPAMS

4.2 Photo Annotation Subsystem

Photo annotation subsystem contains two main components: MPEG-7 Create Module and MPEG-7 Validate Module. [Fig. 4] shows an architecture of photo annotation subsystem in detail.

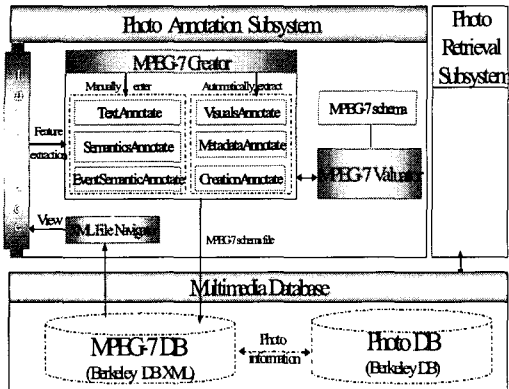


Fig. 4. Architecture of Photo Annotation Subsystem

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 makes module creates their description for the corresponding part of the MPEG-7 document, which is stored into Berkeley DB XML.

MPEG-7 valuator module offers validation to ensure the MPEG-7 document conforming 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 components: Retrieval Panel, Retrieval Engine Module and Result View. The general architecture of the photo retrieval subsystem is depicted in [Fig. 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[8]. 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 DB XML, 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 [2].

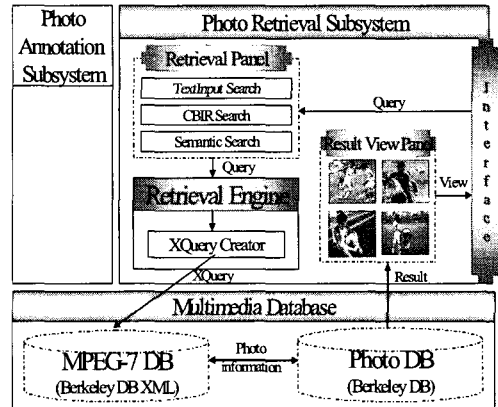


Fig. 5. Architecture of Photo Retrieval Subsystem

4.4 Multimedia Database

Multimedia database consists of two main parts: (1) Berkeley DB XML, which stores MPEG-7 document. (2) Berkeley DB, which stores raw photos. Architecture of multimedia database is depicted in [Fig. 6].

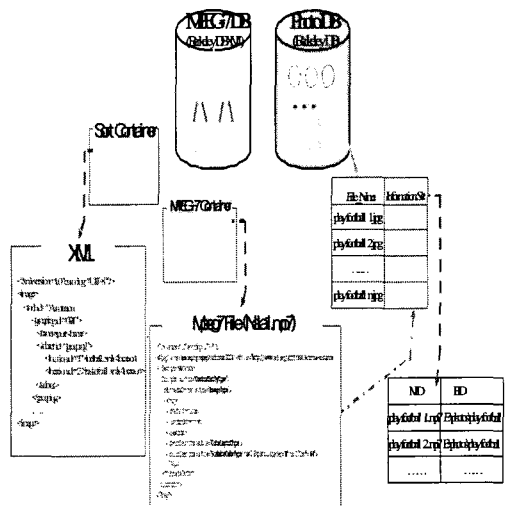


Fig. 6. Multimedia Database Architecture of MPAMS

On the left of multimedia database is Berkeley DB XML [9]. 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 puts the base 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 using the index with them, it is possible to retrieve the targeted photos [3].

V. Implementation of MPAMS

This section mainly focuses on semantic annotation and semantic retrieval implementation of MPAMS.

5.1 Annotation User Interface

MPAMS provides user with an interactive, easy-to-use annotation user interface. As mentioned above, end user can reuse the event semantic annotation if some photos have the same event context. The event annotation is shown on the semantic annotation panel, which allows user to edit semantic object except event object to make sure the annotation according to the current photo information.

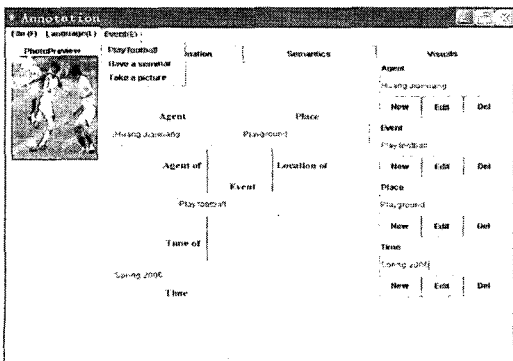


Fig. 7. Photo Annotation User Interface (Semantic)

5.2 MPEG-7 File of the Event Semantic

MPEG-7 description of the event semantic part is shown in [Fig. 8], which is created from the event semantic annotation of photo annotation subsystem. As shown in [Fig. 8], we know the description of the semantic objects and their relationships. After the MPEG-7 metadata instance file is created by the annotation subsystem, the file will be stored and will be available for photo retrieval.

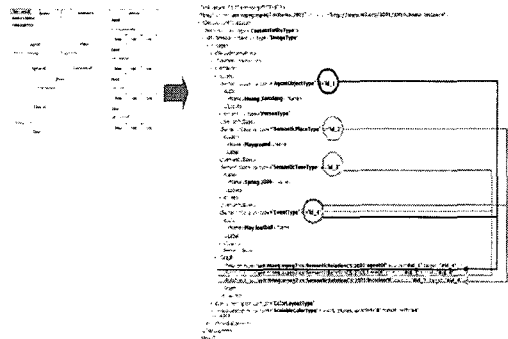


Fig. 8. MPEG-7 File for Semantic Information

5.3 Retrieval User Interface

MPAMS offers a retrieval mechanism for searching semantic descriptors. It allows user manually to enter information on the right hand of this panel as shown in [Fig. 9].

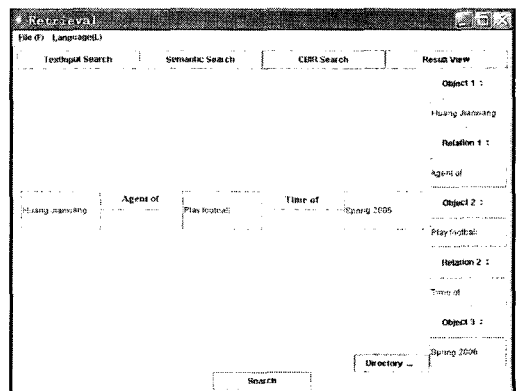


Fig. 9. Photo Retrieval User Interface (Semantic)

5.4 Photo Retrieval Processing

Photo retrieval processing works as shown in [Fig. 10]. The semantic retrieval process has five steps: (1) Define retrieval term, (2) Create XQuery, (3) Retrieve MPEG-7 file, (4) Match photo, and (5) View result.

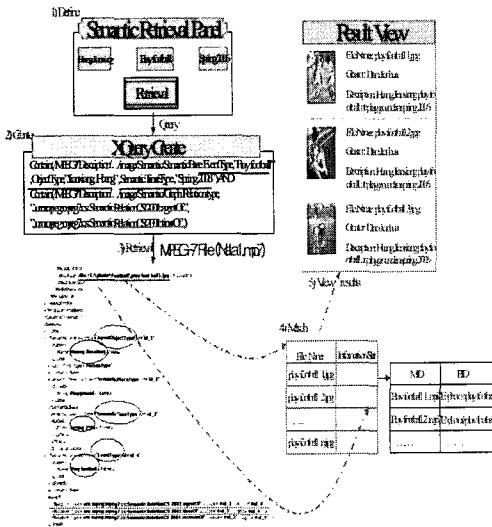


Fig. 10. Photo Retrieval Processing

1) Define Retrieval Term

Photo retrieval subsystem offers a good retrieval user interface to enable user specify an easy and institute query. Semantic retrieval mechanism allows user to define retrieval term manually.

2) Create XQuery

After the retrieval term is defined, the retrieval query is translated into its corresponding description in the form of XQuery by the XQuery creator module.

3) Retrieve MPEG-7 file

After the XQuery is created according to the retrieval term, the corresponding MPEG-7 file is retrieved from the Berkeley DB XML.

4) Match photos

When the corresponding MPEG-7 file is retrieved, the corresponding photo results obtained by matching the File_Name and the MPEG ID from Berkeley DB.

5) View result

User can view the result by the view panel one by one.

VI. Comparison of The Proposed MPAMS with Existing System

A comparison of the proposed MPAMS with existing systems is shown in [Table 1]. Photo Album Management System Using MPEG-7 (MPAMS) is an efficient photo album management and retrieval system, which overcomes the drawbacks of other photo management systems. In this section, we shall compare MPAMS with other systems to prove the capability of MPAMS.

MIRROR (MPEG-7 Image Retrieval Refinement based on Relevance feedback) is a platform for content-based image retrieval research and development using MPEG-7 technologies. The system core is based on MPEG-7 Experimentation Mode with web-based user interface for query by image example retrieval. A new merged color palette approach for DCD similarity measure and relevance feedback are also developed in this system. The system is highly modularized, new algorithms, new ground truth set, and even new image database can be added easily.

CIRES (Content based Image Retrieval System) is effective feature integration, and multiple image queries in enhancing the performance of an image retrieval system. CIRES weighted integration of structure, color and texture features is studied. In addition, it is a methodology of retrieval consisting of multiple query images, as opposed to the traditionally

used model of a single query image.

RPRWS (Robust Photo Retrieval Using World Semantics) is created by media Laboratory software agents group. It is a robust retrieval mechanism for annotated photos by a keyword query. By "annotated photos," it means a photo accompanied by some metadata about the photo, such as keywords and phrases describing people, things, places, and activities depicted in the photo. By "robust retrieval," it means that photos should be retrievable not just by the explicit keywords in the annotation, but also by other implicit keywords conceptually related to the event depicted in the photo.

ASARSOCLT (Automated Semantic Annotation and Retrieval Based on Sharable Ontology and Case-based Learning Techniques) is created by Von-Wun Soo and Chen- Yu Lee. 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 so that the information retrieval can be conducted to some extent at the abstract "semantic" level instead of at the purely syntactic keyword matching level. As shown in this table, we shall know that all of MPAMS, MIRROR, CIRES, RPRWS, and ASARSOCLT are effective photo retrieval mechanism. But MIRROR and CIRES are the platforms only for content-based image retrieval research. They can only implement color-based, shape-based, structure-based, and texture-based retrieval using the MPEG-7 technology. The low retrieval precision difficulty to formulate an exact feature query drawbacks of content-based retrieval technique. It need us offer a more effective and precise technology to overcome these drawbacks, that is semantic retrieval technology. Although RPRWS and ASARSOCLT deal with semantic knowledge about photo assets the MPEG-7 descriptors for semantic annotations are not used. MPAMS offers a

standardized approach to represent semantic objects and relationships to implements effective photo retrieval by using semantics, which is towards solving this problem. Semantic content-based photo retrieval techniques can overcome these drawbacks. It is one of the major challenges of MPAMS, which is more effective and precise than other systems. MPAMS also provides a rich set of automatic feature extraction modules, which fully rely on the MPEG-7 standard as information base. But MPAMS also need to be improved in some way. The summary of comparisons among systems is show in [Table 1].

Table 1. Comparison of the proposed MPAMS with existing systems

Function	MPAMS	MIRROR	CIRES	RPRWS	ASARSOCLT	
Annotation Module	Visuals	O	O	O	X	X
	Event Semantic	O	X	X	X	X
	Word Semantic	X	X	X	O	X
	Semantic base on Ontology	X	X	X	X	O
	Semantic base on MPEG 7	O	X	X	X	X
	Media information	O	X	X	X	X
	FreeText Annotation	O	X	X	X	X
	Structure/Text Annotation	O	X	X	X	X
Retrieval Module	CBR	O	O	O	X	X
	Textinput Search	O	X	X	X	X
	Semantic Search	O	X	X	O	X
Relevance feedback	X	O	X	X	X	
Remark	MPAMS: Photo Album Management System by using MPEG 7 MIRROR: MPEG 7 Image Retrieval Refinement based on Relevance feedback CIRES: Content based Image Retrieval System RPRWS: Robust Photo Retrieval using Word Semantics ASARSOCLT: Automated Semantic Annotation and Retrieval Based on Sharable Ontology and Case based Learning Techniques Support: O Nonsupport: X					

VII. Conclusion and Future Research

In this paper, we develop a photo album management system using MPEG-7 (MPAMS). Its main contribution is the integration of a semantic technology and MPEG-7 standard to support effective photo retrieval. An event semantic annotation based on MPEG-7 is proposed for an efficient photo retrieval. It can handle photo information using MPEG-7 technology based on a native XML database system.

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

Reference

- [1] J. Martinez, "Overview of the MPEG-7 Standard, ISO/IEC JTC1/SC29/WG1," N4980 (Klagenfurt Meeting), July 2002.
- [2] W3C, "XML Path Language," May 2000.
- [3] Sleepycat, "Berkeley DB XML, Berkeley DB,"
- [4] "MPEG-7 Image Retrieval Refinement based on Relevance feedback," Postgraduate and undergraduate research projects in the Department of Electronic Engineering, City University of Hong Kong, 2004.
- [5] A. Smeulders, "Content-Base Image Retrieval at the End of the Early Years," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.22, No.12, Dec. 2000.
- [6] H. Liu and H. Lieberman, "Robust Photo Retrieval Using World Semantic," MIT Media Laboratory Software Agents Group, 20 Ames St., E15-320G Cambridge, MA 02139.

- [7] V. W. Soo, C. Y. Lee, C. C. Li, S. L. Chen, and C. C. Chen, "Automated Semantic Annotation and Retrieval Based on Sharable Ontology and Case-based Learning Techniques," Department of Computer Science, National Tsing Hua University, HsinChu, 30043, Taiwan Graduate School of Library and Information Science, Simmons College, Boston, MA 02115, USA, 2003.
- [8] W3C, "XQuery 1.0: An XQuery Language," Working Draft, May 2003.
- [9] M. Olson, "Selecting and Implementing an Embedded Database System," IEEE Computer, Vol.33, No.9, pp.27-34, Sept. 2000.
- [10] 홍성용, 나연목, "의미 기반 전자 카탈로그 이미지 검색을 위한 XML 데이터베이스 시스템 구현", 한국멀티미디어학회 논문집, Vol.7, No.9, pp.74-83, 2004.

저자 소개

안 병 태(Byeong-Tae Ahn)

정희원



• He received the B.S in Computer Science from Kookmin University, Korea in 1999 and the M.E in Computer Engineering from Gyeongnam University, Korea in 2001.

He received Ph.D in Computer Science from Gyeongsang University, Korea in 2006. From 2001 to 2006, he has been engaged in business of Web portal. Since 2006, He is currently a professor in Dep. of MIS at Yuhan College. His main research interests include XML, Embedded DB, Multimedia Contents and MPEG-7.

정 범 석(Bhum-Suk Chung)

정회원



- He received the B.S, M.S and Ph. D in Management System from Kookmin University, Korea in 1982, 1984 and 1994 respectively. He is currently a professor in Dep. of MIS at Yuhan College.

His main research interests include Financial affairs, Multimedia Contents and semantic.

이 중 하(Chong-Ha Lee)

정회원



- He received the B.E. in Industrial Engineering from Korea University, Korea in 1983 and the M.S. in Industrial and Management systems Engineering from University of Nebraska, USA in 1987. He received Ph.D.

in Industrial and Management Systems Engineering from University of Nebraska, USA in 1996. Since 1996, he has been with Department of Management Information System, Yuhan College, Bucheon, Gyeonggi-do, Korea. His main research interests include Fuzzy Set and Artificial Intelligence Application, Design and Analysis of Manufacturing Systems, ERP.