

장면전환검출과 사용자 프로파일을 이용한 비디오 학습 평가 시스템

신 성 윤[†]

요 약

본 논문에서는 사용자 프로파일을 기반으로 한 정보 필터링을 사용하여 학생 개인의 특성에 맞는 효율적인 원격 비디오 학습 평가 시스템을 제안한다. 비디오를 이용한 문제 출제를 위하여 위치, 크기, 그리고 컬러 정보를 기반으로 키 프레임 추출하고 그레이 레벨 히스토그램 차이와 시간 윈도우를 이용하여 문제 출제 구간을 추출한다. 또한 효율적인 평가를 위하여 카테고리 기반 시스템과 키워드 기반 시스템을 합성하여 문제를 출제하도록 한다. 따라서 학생들은 부족한 영역을 보충하고 관심 있는 영역을 유지하면서 학업 성취도를 향상시킬 수 있다.

Video Evaluation System Using Scene Change Detection and User Profile

Seong-Yoon Shin[†]

ABSTRACT

This paper proposes an efficient remote video evaluation system that is matched well with personalized characteristics of students using information filtering based on user profile. For making a question in forms of video, a key frame extraction method based on coordinate, size and color information is proposed. And Question-making intervals are extracted using gray-level histogram difference and time window. Also, question-making method that combined category-based system with keyword-based system is used for efficient evaluation. Therefore, students can enhance their study achievement through both supplementing their inferior area and preserving their interest area.

키워드 : 사용자 프로파일(User Profile), 비디오 학습 평가 시스템(Video Evaluation System), 키 프레임 추출(Key Frame Extraction), 정보 필터링(Information Filtering)

1. 서 론

정보처리(information processing)란 주어진 정보로부터 달성해야 할 목적에 도움이 되는 정보를 얻는 능동적인 과정을 의미한다. 인터넷상에서 정보를 처리하는 기능은 크게 정보 검색(information retrieval)과 정보 필터링(information filtering)의 두 가지 기능으로 분류할 수 있다.

정보 검색은 원하는 정보를 찾아주는 기능으로서, 수집한 정보나 자료의 내용을 분석한 후 알맞게 가공하여 축적해 놓은 데이터베이스로부터 사용자의 요구사항에 적합한 정보를 탐색하여 찾아내는 일련의 과정을 의미한다[1].

정보 필터링은 사용자의 요구사항에 맞도록 정보를 가공하고 걸러주는 기능으로서, 시스템에 의해 생성되고 유지되는 사용자의 관심에 따른 장기적인 프로파일(profile)에 기반을 둔 동적인 정보 스트림의 필터링을 말하는데, 대부분의

전용 필터링 시스템들은 학습 기술을 이용하여 사용자 관심 프로파일을 자동으로 생성하고 유지한다[2,3]. 정보 필터링은 정보의 개인화의 중요 과정으로서 전통적으로 내용-기반(content-based) 필터링, 사회적(social) 필터링, 그리고 경제적(economic) 필터링의 세 가지 형태로 분류되며, 이들을 서로 조합하여 사용하기도 한다[4].

내용-기반 필터링은 인식(cognitive) 필터링이라고도 하는데, 객체는 객체들의 내용과 사용자의 우선권 사이의 관계에 의하여 선택된다. 내용-기반 필터링의 대표적인 예로는 키워드-기반(keyword-based) 필터링이 있다[3,5]

사회적 필터링은 협동(collaborative) 필터링이라고도 불리는데, 객체들은 사용자에게 유사한 작업을 가진 다른 사람들의 관점에서 필터링 된다[6]. 사회적 필터링은 다수의 참여관계자와 객체들이 필요하다는 단점이 있지만 많은 시스템에서 이용하고 있는데, 대표적인 예는 다음과 같다. Tapes-try 시스템[7]에서는 사용자들이 직접 주석을 부여하고 관심 분야에 대한 판단을 결정하고 나면 프로그램 자체의 복

[†] 정 회 원 : 군산대학교 컴퓨터정보학과 IT 교수
논문접수 : 2003년 6월 16일, 심사완료 : 2003년 10월 20일

합 필터가 연속적으로 저장된 문서들에 대하여 필터링을 수행한다. Stanford Information Filtering Tool(SIFT)[8]은 웹상에서 필터링 서비스를 제공하는 시스템으로서, 사용자들은 매칭 전략을 사용한 키워드를 명시하는 하나 이상의 프로파일을 통하여 필터링 서비스를 제공받는다. GroupLens [9]는 Usenet News의 분산 협동 필터링을 위한 시스템으로서, 사용자들에게 읽고자 하는 기사에 대하여 자신들이 가중치를 부여할 수 있도록 하였다.

경제적 필터링은 비용 요소를 기반으로 정보를 필터링하는 방법이다[8]. 여기에서 사용되는 비용 요소들에는 사용되는 비용과 이윤 사이의 관계 또는 네트워크 대역과 객체 크기 사이의 관계 등이 있다.

위의 세 가지 필터링 방법들은 서로 조합하여 사용되기도 하는데, 대표적인 예로서 NewsWeeder 시스템[10]은 Usenet News를 위하여 내용-기반 필터링과 사회적 필터링을 조합한 방법이다.

최근 몇 년 동안, 사용자 프로파일을 이용한 많은 연구들이 수행되어 왔다. XML의 발전과 함께 내용뿐만 아니라 구조를 기반으로 한 웹 문서의 필터링이 폭넓게 수행되었는데, 대표적인 최신 웹 문서의 필터링으로 XFilter 시스템 [11]이 있다. Franklin 등[12]은 데이터 재충전을 위하여 사용자 프로파일을 이용하였다. 이들은 의미 있는 프로파일 언어로 작성된 사용자 프로파일을 기반으로 자동 데이터 재충전 방법을 제시하였다. Schwab 등[13]은 명확한 사용자 관찰만을 통하여 잠재적인 관심 사용자 프로파일 학습을 위한 방법을 제시하였다.

인터넷의 급속한 발달로 인하여 교육의 변화 가운데 가장 주목할만한 분야는 웹상에서의 원격 학습 평가(remote evaluation) 분야이다. 원격 학습 평가 시스템은 기존의 평가 시스템이 갖는 시공간적인 제약 및 평가 회수의 제한 등의 다양한 문제점을 극복하고 평가의 객관성, 투명성, 그리고 신속성을 보장하는데 중요한 역할을 수행하였다. 하지만 이 시스템은 학생 개인의 특성과 성향을 고려하여 평가를 수행하지 못하고 일률적이고 획일적인 평가만을 수행한다는 문제점을 갖고 있다.

이와 같은 문제점들은 개인화된 사용자 프로파일을 이용하면 효율적으로 해결할 수 있다. 즉, 학생들의 개인 사용자 프로파일을 이용하여 문제은행 데이터베이스로부터 개인의 특성 및 성향에 맞도록 문제를 필터링한 후 학생들에게 제공하고 문제를 풀도록 함으로서, 자신의 부족한 부분은 보충하고 우수한 부분은 더 개발할 수 있도록 유도할 수 있다[14]

본 논문에서는 개인화된 사용자 프로파일을 기반으로 한 정보 필터링을 이용하여 학생 개개인에 특성과 성향에 잘 부합되는 효율적인 비디오 학습 평가 시스템을 제시한다. 비디

오로 문제를 출제하기 위하여 EBS에서 방영된 장학퀴즈 프로그램을 이용한다. 문제 출제 장면의 키 프레임 추출하기 위하여 비디오 프레임에서 문제 번호 영역의 위치, 크기, 그리고 컬러 정보를 기반으로 한 장면 전환 검출 방법을 제시하고, 문제 출제 구간의 추출은 그레이 레벨 히스토그램 차이와 시간 윈도우를 이용하여 추출한다. 특히, 사용자 프로파일의 적용 과정에서 카테고리-기반 방법과 키워드-기반 방법을 병행한 문제 출제 방식을 적용하여 학생들의 취약한 영역을 보충하면서 관심 소재를 다루어 평가에 대한 흥미를 갖도록 한다.

본 논문의 구성은, 2장에서는 문제 출제 장면의 분할을 위한 장면 전환 검출에 대해 설명하고, 3장에서는 사용자 프로파일을 이용한 학습 평가를 설명한다. 4장에서는 시스템의 실질적인 구현과 활용을 설명하고, 5장에서는 결론을 맺도록 한다.

2. 장면 전환 검출

비디오의 장면 전환 검출(scene change detection)이란 비디오에서 연속적으로 주사되는 각각의 프레임에 담겨있는 정보를 사용하여 장면이 전환되는 지점을 찾아내는데 이용되는 기술이다. 디지털 비디오 정보를 데이터베이스로 구축하고자 할 때 선행되어야 할 것은 내용별로 색인화를 수행하는 것인데, 장면 전환 검출과 색인화를 통해서 사용자는 원하는 비디오 정보에 비순차적인 접근이 가능하고 내용-기반 검색이 가능해진다. 한 편의 비디오를 내용별로 분할하기 위한 최소 단위는 보통 장면이 전환되는 경계로 구분을 지을 수 있다. 따라서 비디오 정보의 데이터베이스 구축에 있어서 비디오의 장면 전환 검출은 필수적인 기술이다.

본 논문에서는 비디오를 통한 문제 출제를 위하여 장면 전환 검출을 이용하여 키 프레임과 문제 출제 구간을 추출하는 방법을 제시한다.

2.1 키 프레임 추출

본 논문에서는 비디오를 이용하여 문제를 출제하기 위하여 EBS에서 방영된 장학퀴즈 프로그램을 이용한다. 장학퀴즈 프로그램은 총 3라운드의 평가를 수행하지만 본 논문에서는 비디오를 분할하기 위하여 1라운드만을 이용한다. 분할된 문제 영역은 언어(Language), 수학(Mathematics), 사회(Society), 과학(Science), 그리고 외국어(Foreign Language)의 총 5개 영역으로 분류하여 문제를 출제하는데 이용한다.

장학퀴즈 프로그램의 1라운드 평가에 대한 구조적 특징을 요약하면 다음과 같다.

- ① 문제를 출제할 때는 항상 문제 번호와 문제 내용이 나타난다.

- ② 각각의 문제 번호와 문제 내용이 처음에 나타날 때에는 문제 번호가 깜박거리는 효과가 나타난다.
- ③ 문제 번호와 문제 내용은 각각 고정된 크기를 갖는다.
- ④ 문제 번호 영역은 일정한 컬러를 갖는다.
- ⑤ 문제 번호와 문제 내용은 문제 출제가 완료되면 사라진다.

위의 특징을 기반으로 문제 출제 장면의 키 프레임들은 문제 번호 영역의 위치, 크기, 그리고 컬러 정보를 이용한 유사성 측정을 통하여 추출한다. 유사성 측정 방법은 식 (1)과 같다.

$$Similarity = CI_i(p, s, c) - TI(p, s, c) \quad (1)$$

where $i = 1 \dots m$

식 (1)에서 $TI(p, s, c)$ 는 사전 지식을 기반으로 생성된 문제 번호 영역의 위치 (p), 크기 (s), 그리고 그레이 컬러 (c) 값들을 갖는 템플릿을 말한다. 그리고 $CI_i(p, s, c)$ 는 입력 프레임의 문제 번호 영역에 대한 각각의 정보를 갖는다.

키 프레임은 문제 번호 영역에 대한 유사성 측정을 통하여 임계값과의 비교를 수행함으로써 추출한다. 즉 결과값이 임계값 이내이면 키 프레임으로 추출하고 임계값 범위에 들지 않으면 일반 프레임으로 간주한다.

2.2 문제 출제 구간 추출

문제 출제 구간(QMI : question-making interval)은 문제 출제를 위하여 각각의 문제가 나타나서부터 사라질 때까지의 문제 번호와 문제 내용이 존재하는 구간을 말한다. 문제 출제 장면은 추출된 키 프레임을 바탕으로 문제 출제 구간의 추출을 통하여 얻어진 장면이다. 문제 출제 구간의 추출을 위하여 식 (2)와 같이 문제 번호 영역에 대한 그레이 레벨 히스토그램 차이(D_i)의 계산을 먼저 수행한다.

$$D_i = \sum_{j=1}^{Bins} |H_i(j) - H_{i-1}(j)| \quad (2)$$

식 (2)에서 $H_i(j)$ 는 프레임 i 에 대한 그레이 레벨 히스토그램의 j -번째 빈(bin)을 의미한다. D_i 는 현재 프레임 F_i 와 이전 프레임 F_{i-1} 사이의 히스토그램 차이값을 말한다.

그리고 각각의 문제들을 출제할 때 처음에 나타나는 깜박임 현상 때문에 문제 출제 구간이 잘못 추출되는 것을 방지하기 위하여 키 프레임으로부터의 시간 윈도우(W_i)를 부여하여 측정한다. 문제 출제 구간(QMI)은 그레이 레벨 히스토그램 차이값이 임계값(D_T) 이내에 존재할 경우 추출된다. 또한 히스토그램 차이값이 임계치보다 크더라도 키 프레임 시점(W_k)과 입력 프레임 시점(W_i) 사이의 프레임 차이값(W_f)이 시간 윈도우(W_T) 내에 존재할 경우 문제 출제

구간(QMI)으로 추출한다. 그렇지 않은 경우에는 일반 프레임 구간(NFI : Normal Frame Interval)으로 간주한다. 세부적인 문제 출제 구간 추출 알고리즘은 (그림 1)과 같다.

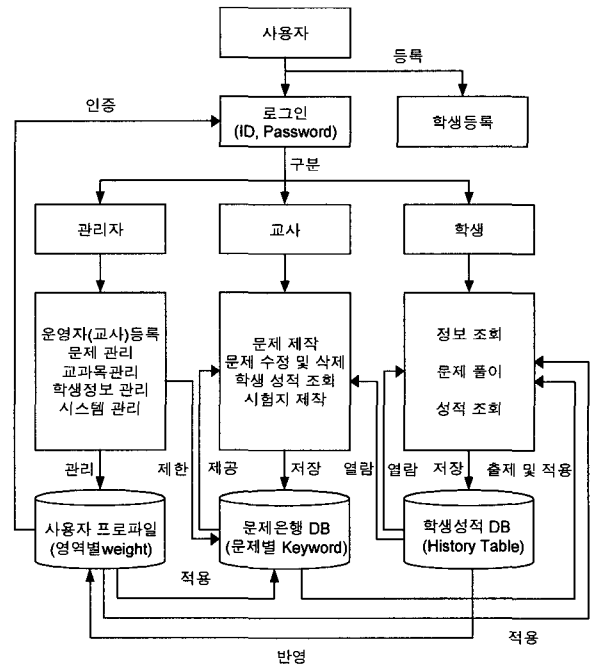
```

For (i = 2; i ≤ n; i++) {
    Di = gray-level histogram difference;
    Wf = |Wk - Wi|;
    If (Di ≤ DT) Fi = QMI;
    Else IF (Di > DT AND Wf ≤ WT)
        Fi = QMI;
    Else Fi = NFI;
}
    
```

(그림 1) QMI 추출 알고리즘

3. 사용자 프로파일을 이용한 학습 평가

전통적인 학습 평가 시스템들은 과거의 평가 제도에서 크게 벗어나지 못하고 시공간적인 제약과 개인의 특성과 성향 및 개인차를 고려하지 않은 획일적인 형태로서 평가의 형식, 방법, 그리고 처리에서 매우 소모적이고 비효율적이며 평가의 생산성과 능률성이 매우 저조하였다.



(그림 2) 전체 시스템 구조

본 논문에서는 사용자 프로파일을 이용하여 전통적인 평가 시스템이 갖는 문제점들을 해결하고 학생 개개인의 특성과 성향을 고려한 학습 평가 방법을 제시한다. 먼저, 문제은행 데이터베이스를 구축하여 웹상에서 시공간을 초월하여 평가를 수행할 수 있도록 한다. 다음으로, 사용자 프로파일을 이용하여 문제를 필터링한다. 마지막으로, 학생들의 개인차와 특성 및 성향을 고려한 문제 출제로 평가의

효율성을 향상시키도록 한다. 전체적인 시스템 구조는 (그림 2)와 같다.

(그림 2)와 같이 전체적인 시스템은 크게 관리자, 교사, 그리고 학생의 3개의 모듈로 구성된다. 관리자 모듈에서는 운영자가 되는 교사의 등록 및 관리, 새로운 교과목의 관리, 학생 정보 관리, 시스템 관리 등 시스템 전반을 관리하는 기능들로 구성된다. 교사 모듈은 영역별 문제 출제, 시험지 제작, 학생 성적 조회 등의 기능들로 구성된다. 학생 모듈은 학생 자신의 정보 조회, 문제풀이, 성적 조회 등의 기능들로 구성되며 문제 풀이를 통하여 데이터베이스에 관련 정보를 저장하고 프로파일을 적용할 수 있도록 한다.

3.1 시스템 세부 구조

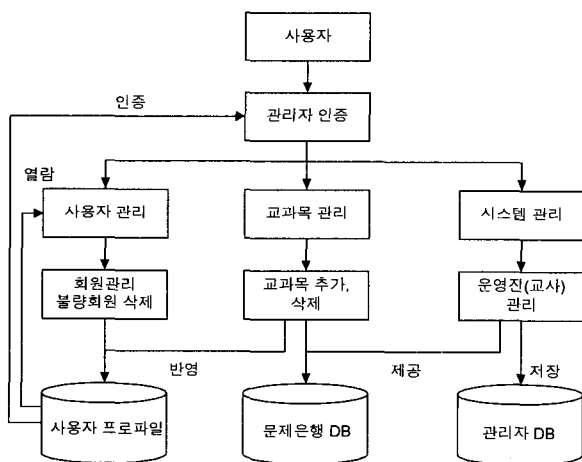
본 논문에서는 전통적인 평가 시스템이 갖는 문제점을 해결하기 위하여 카테고리-기반 사용자 프로파일과 키워드-기반 사용자 프로파일을 병합한 시스템을 제시한다. 본 논문에서 문제 출제는 정보 검색 에이전트(agent)[5]를 적용한다. 즉, 사용자 프로파일을 이용하여 문제은행에 저장된 문제들 중에서 사용자의 특성과 성향 및 개인차에 따라 알맞은 문제를 검색하고 필터링 하는 방법을 제시한다.

3.1.1 사용자 등록

사용자 등록은 웹 페이지를 통하여 사용자 아이디(ID)와 패스워드(Password)를 이용하여 로그인(Login)한 후 사용자 프로파일에서 사용자를 관리자, 교사, 또는 학생으로 구별하여 권한을 부여한다.

3.1.2 관리자 모듈

관리자는 (그림 3)과 같이 사용자 관리, 교과목 관리, 시스템 관리 등에 대한 모든 권한을 갖는다.



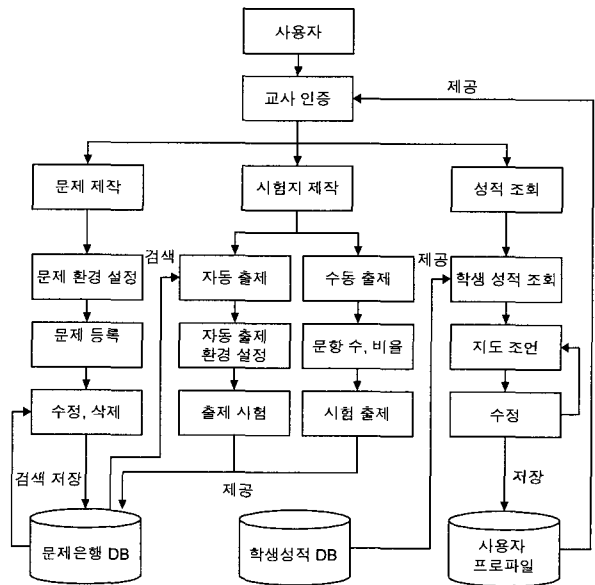
(그림 3) 관리자 모듈

사용자 관리의 사용자에 대한 정보 검색, 조회, 수정, 삭제 등을 수행하고, 교과목 관리의 교과목의 추가 및 삭제를

수행한다. 시스템 관리에서는 전체 데이터베이스의 관리와 미적용 프로파일 임계값($C_p(V)$)의 조정 등 전체적인 시스템을 관리한다.

3.1.3 교사 모듈

교사는 (그림 4)와 같이 문제의 제작, 문제 출제, 성적 조회 등의 권한을 갖는다. 문제가 제작되면 관리자에 의하여 문제은행 데이터베이스에 저장되는데, 문제를 제작할 때 도움말도 함께 출제하여 도움을 줄 수 있으며 영상, 음악 등의 멀티미디어를 활용한 문제도 출제할 수 있다.



(그림 4) 교사 모듈

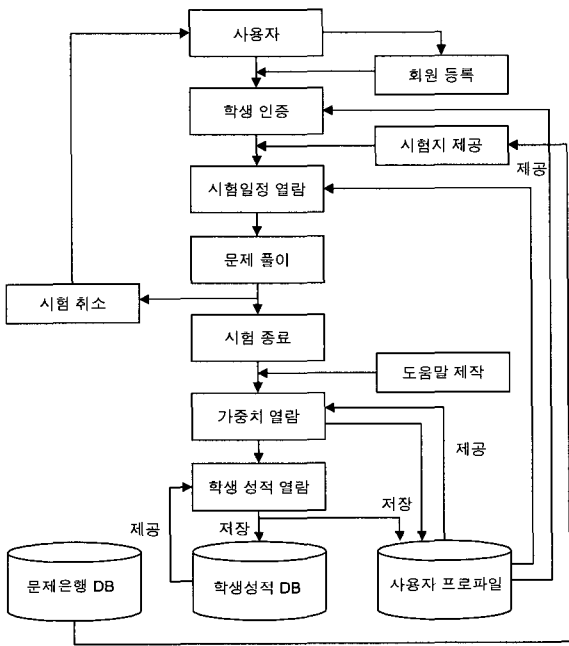
교사는 문제를 출제할 때 키워드-기반 사용자 프로파일 시스템을 지원하기 위한 키워드를 정해줘야 한다. 이는 문제를 제작할 때 해당 문제의 가장 중요한 주제를 키워드로 선정하여 입력함으로써 구현된다.

3.1.4 학생 모듈

학생 모듈은 사용자 프로파일의 실질적으로 적용되는 가장 중요한 부분이다. 학생들은 사용자 프로파일을 이용하여 최적화된 문제를 풀고 결과에 따라 다음 시험에 다시 새로운 프로파일이 반영된다. 로그인한 학생은 시스템으로부터 시험 회수, 영역별 문제수, 시간 등의 시험 일정을 제공받고 본격적인 시험을 치르게 된다.

학생에게 출제되는 문제는 사용자 프로파일에 의하여 학생의 개별적 특성과 성향에 최적화된 문제이어야 한다. 최적의 사용자 프로파일을 생성하기 위하여 관리자에 의하여 설정된 미적용 프로파일 임계값($C_p(V)$) 만큼의 시험에서는 모든 영역에서 사용자 프로파일을 반영하지 않는다. 이때는 초기화된 미적용 프로파일 문제수($iP_p(N)$) 만큼의 문

제를 부여받아 문제풀이를 통하여 사용자 프로파일을 축적한다. 이러한 과정을 통하여 $C_b(V)$ 가 만족되면, 시스템은 영역별 가중치($weight[i]$)와 사용자 관심 단어($I_u(W)$)를 획득하여 사용자 프로파일을 생성하고 유지한다.



(그림 5) 학생 모델

3.2 사용자 프로파일을 이용한 문제 필터링

학생이 시험을 치를 경우 시스템은 사용자 프로파일을 참고하여 사용자에게 대한 정보를 분석하고 확인하여 문제를 필터링한 다음 알맞은 문제를 제공한다. 그리고 영역별 문제수는 영역별 가중치의 계산에 의하여 결정된다. 본 논문에서는 카테고리-기반 방법과 키워드-기반 방법을 합성한 방법을 이용하여 문제를 필터링한다.

3.2.1 카테고리-기반 방법

영역별 점수는 각 영역에 대한 개별적 문제들의 총점을 구하여 얻을 최적화된 문제를 풀고 결과에 따라 다음 시험에 다시 새로운 프로파일이 반영된다. 로그인한 학생은 시스템으로부터 시험 회수, 영역별 문제수, 시간 등의 시험 일정을 제공받고 본격적인 시험을 치르게 된다. 사용자에게 취약한 영역에 대해 더 큰 비중을 부여함으로써 사용자의 흥미를 저하시킬 수 있다. 따라서 본 논문에서는 키워드-기반 시스템을 합성하여 수행함으로써 카테고리-기반 시스템의 단점을 보완한다.

3.2.2 키워드-기반 방법

키워드-기반 방법은 정성적 요소를 반영할 수 있는 방법이다. 카테고리-기반 시스템은 평가를 위하여 문제의 영역

을 나누고 영역별 가중치를 부여한다. 반면에 키워드-기반 시스템은 세부적인 요소인 개별 문제에서 가장 중요시되는 관심사인 주제를 찾아내어 이를 키워드로 하여 문제에 삽입한다. 다음으로, 채점 결과에서 사용자가 정답을 맞힌 문제를 대상으로 가장 빈번하게 나오는 키워드를 선별하여 기존의 키워드 데이터가 저장되어 있는 큐에 삽입하여 새로운 키워드 프로파일을 생성한다. 그 후, 다음 평가에서 카테고리-기반 시스템으로 할당받은 각 영역의 문제수에 다시 키워드 프로파일을 반영한 사용자 관심 단어를 가진 문제를 우선적으로 출제한다.

카테고리-기반 시스템에서 키워드를 선택하는 기준은 주관적인 소재이므로 이를 반영하기 위하여 문제은행 데이터베이스에서 문제를 추출할 때 검색기능을 사용하여야 한다. 본 논문에서는 이를 위하여 MySQL에서 제공하는 표준 SQL 규약의 LIKE 문을 이용하였다.

3.2.3 문제 필터링 과정

① 문제를 출제할 때 사용자 프로파일을 반영하여 출제할 것인지의 여부를 결정한다.

이 과정은 사용자 프로파일을 축적하기 위하여 필요한 과정이다. 출력은 Yes나 No의 형태로 나타나며, 이 과정을 통하여 불투명하고 불명확하며 함축적인 형태의 정보인 사용자 프로파일에 신뢰성을 부여할 수 있다.

본 논문에서는 사용자 프로파일의 축적을 위하여 미적용 프로파일 임계값($C_b(V)$)을 설정하였다. 아래와 같이 사용자의 전체 평가 계수($T_i(C)$)가 $C_b(V)$ 보다 작을 경우 사용자는 프로파일이 적용된 평가가 아닌 일반적인 평가를 받게 되며 이를 통하여 프로파일을 축적한다.

IF ($T_i(C) < C_b(V)$)

THEN FOR($i = 1 ; i \leq T_a(N) ; i++$)

$Q_a(N)[i] = iP_i(N)$

$C_b(V)$: 미적용 프로파일 임계값

$T_i(C)$: 전체 평가 계수

$T_a(N)$: 전체 영역수

$Q_a(N)$: 영역별 문제수

$iP_i(N)$: 미적용 프로파일 문제수

② 사용자에게 부여할 영역별 문제수를 계산한다.

데이터베이스에서 문제수 형태로 저장되어있는 사용자 프로파일을 가져온다. 이처럼 영역별로 상대적 학습수준에 맞는 문제를 제시함으로써 카테고리-기반 시스템이 구현되며, 이 루틴에서 각 영역별 출제 문제수가 결정된다.

FOR($i = 1 ; i \leq T_a(N) ; i++$)

$Q_a(N)[i] = \text{Query}(\text{Select Area}[i] \text{ From}$
 User_Profile_DB where ID = userid

③ 해당 영역에 $Q_a(N)$ 만큼의 문제를 할당한다.

이 루틴은 키워드-기반 시스템을 구현하는 루틴이다. 채점과 프로파일 반영을 통하여 생성되고 수정된 $I_u(W)$ 를 통하여 사용자 관심 단어와 관련된 문제를 우선적으로 선택한다.

FOR($i = 1 ; i \leq Q_a(N) ; i++$)

Question_Array += Query(Select Question_No From
 Question_DB
 where Question_Keyword Like
 % $I_u(W)$ %)

$I_u(W)$: 사용자 관심 단어

만일 $I_u(W)$ 값이 Null이거나 $I_u(W)$ 에 해당하는 문제 레코드 셋을 모두 추가하였다면 Random() 함수를 이용하여 무작위로 선택한다. 두 가지 경우 모두 선택된 레코드 셋을 사용자에게 제공될 Question_Array에 삽입할 때는 중복검사 과정을 거치고, 정렬 알고리즘을 적용하여 사용자에게 주어질 문제들을 선택한다. 본 논문의 경우 선택정렬을 이용하였다.

④ 시스템은 지문과 문제를 사용자에게 보여주고 사용자가 입력한 답안을 배열 형태로 받아들인다.

⑤ 문제은행에서 가져온 정답과 사용자가 입력한 답안을 비교하는 채점 루틴으로서, 영역별 총점과 전체 총점을 계산한다.

⑥ 카테고리 기반 시스템에 적용하기 위한 영역별 가중치를 계산한다.

영역별 가중치는 학생 개인의 특성과 성향을 판단하는 중요한 데이터이다. 가중치는 시험 영역별로 학생이 획득한 점수를 바탕으로 우수한 영역일수록 보다 작은 음수값을 갖으며 부족한 과목일수록 보다 큰 양수값을 갖는다. 이 가중치 부여를 통하여 학생이 관심이 있고 문제 해결 능력이 있는 영역은 출제 문제수를 줄이고, 학생이 부족한 영역을 출제 문제수를 늘림으로서 부족한 영역을 보충하고 실력을 향상시키도록 한다.

For($i = 1 ; i \leq T_a(N) ; i++$)

$weight[i] = \left\{ \left(\frac{T_i(S)}{T_i(N)} - \frac{T_a(S)[i]}{Q_a(S)[i]} \right) \cdot T_a(N) \right\}$

$T_i(N)$: 전체 만점 기준이 되는 총점

$T_i(S)$: $T_i(N)$ 을 만점 기준으로 얻은 전체 시험 점수

$Q_a(S)$: 영역별 만점 기준이 되는 총점

$T_a(S)$: $Q_a(S)$ 를 만점 기준으로 얻은 영역별 시험 점수

이 가중치는 선행 시험의 문제 비율(소수점 변환)과 문제수와 함께 곱해진다. 이 곱해진 결과는 반올림 되어 선행 시험 문제수와 더해지거나 빼져서 다음 시험 문제수를 산출한다.

⑦ 키워드-기반 시스템에 적용하기 위한 $I_u(W)$ 를 구한다.

For($i = 1 ; i \leq T_a(N) ; i++$)

$I_u(W) = \text{Query}(\text{Select Question_Keyword From Area}[i]$
 Where Question_No = Area[i][right_j]
 right_j : 해당 영역 중 정답집합의 문제번호

⑧ 카테고리-기반 시스템에서 획득한 가중치와 키워드-기반 시스템에서 획득한 $I_u(W)$ 를 사용자 프로파일 데이터 베이스에서 갱신한다. 가중치는 기존의 프로파일 데이터에 Plus 연산을 통하여, $I_u(W)$ 는 기존의 프로파일 데이터를 Queue에 삽입한 후 새롭게 획득한 데이터를 Push 연산을 통하여 갱신한다.

⑨ 채점을 통하여 얻어진 영역별 총점과 전체 총점을 사용자에게 제시하고 사용자 히스토리 데이터베이스에 저장한다.

4. 비디오 학습평가 시스템의 구현

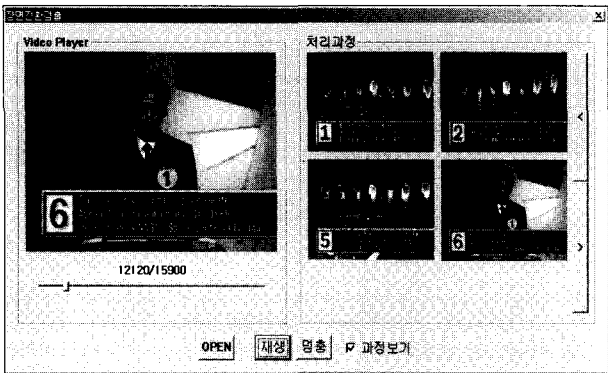
4.1 구현 환경

본 논문에서 제시한 사용자 프로파일을 이용한 비디오 학습 평가 시스템은 Pentium 4-1.4GHZ PC에서 TCP/IP 환경을 사용하기 위하여 Apache 웹 서버를 이용하였다. 장면 전환 검출은 EBS 장학퀴즈 비디오를 대상으로 Visual C++6.0을 이용하여 구현하였다. 또한 Windows 2000 Server에서 중형 데이터베이스 시스템을 구축하기 위하여 MySQL을 이용하였고 데이터베이스 액세스를 위하여 PHP3를 이용하였다.

4.2 장면 전환 검출 구현

장면 전환 검출을 위한 비디오 데이터는 EBS 장학퀴즈의 첫 번째 라운드를 이용하였다. 키 프레임은 비디오의 문제 번호 영역에 대한 구조적 특징(위치, 크기, 그리고 컬러 정보)을 이용하여 추출하였다. 문제 출제 구간은 그레이 레벨 히스토그램 차이와 시간 윈도우를 이용하여 추출하였다. 장면 전환 검출은 (그림 6)과 같이 수행된다.

다음 <표 1>은 장면 전환 검출을 위한 문제 번호 영역의 위치, 크기, 평균 그레이 컬러, 시간 윈도우(프레임), 그리고 그레이 레벨 히스토그램 차이에 대한 기준값과 임계값을 나타낸다.



(그림 6) 장면 전환 검출

<표 1>에서 기준값은 문제 번호 영역의 표준이 되는 값이고 임계값은 기준값에 대한 임계 영역에 해당되는 값을 의미한다. 만일 모든 특징들이 임계값 범위 이내에 존재하게 되면 키 프레임과 문제 출제 구간이 추출된다.

<표 1> 장면 전환 검출을 위한 임계값

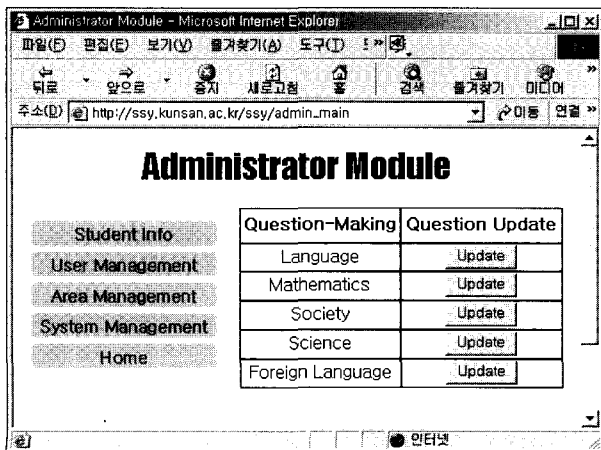
features	기 준 값	임 계 값
위치(x, y)	25, 171	기준값±1
크기(area)	1739	기준값±83
평균 그레이 컬러	170	기준값±10
W_T	60	기준값±5
D_i	100	기준값±10

4.3 사용자 모듈 구현

사용자는 관리자로부터 인증 받은 개인 ID와 Password를 이용하여 시스템에 접속할 수 있다.

4.4 관리자 모듈 구현

관리자는 시험 문제의 출제와 영역에 관련된 전반적인 사항들을 제어하고 관리한다. 또한 학생과 교사 정보의 관리

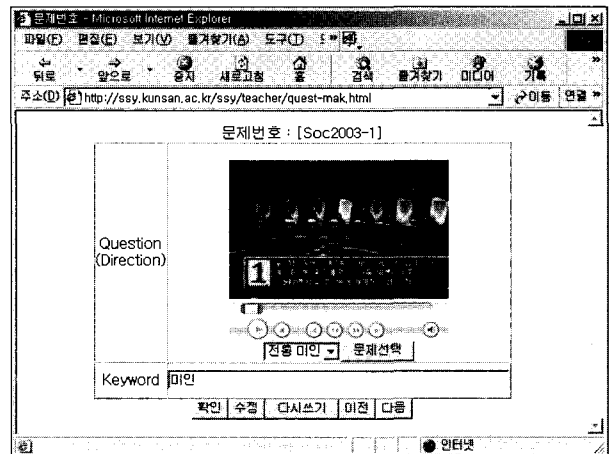


(그림 6) 관리자 모듈 화면

와 프로파일 축적을 위하여 필요한 $C_p(V)$ 와 $iP_q(N)$ 를 설정한다. (그림 6)은 관리자 모듈의 화면을 보여주고 있다.

4.5 교사 모듈 구현

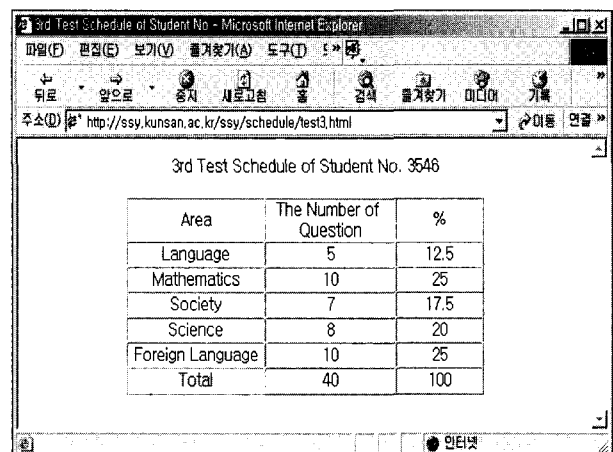
교사는 관리자의 기능 중에서 시스템 관리, 교과목 관리, 운영자 관리를 제외한 학생정보 관리와 문제 출제 및 수정 작업을 수행할 수 있다. 문제를 출제할 때 키워드-기반 방법을 구현하기 위하여 (그림 7)과 같이 비디오 형태의 문제(지문)와 키워드를 삽입할 수 있다.



(그림 7) 교사 모듈의 문제 출제 화면

4.6 학생 모듈 구현

비디오 학습 평가 시스템에 학생이 접속한 경우, 초기화면에는 (그림 8)과 같은 사용자 프로파일 분석에 의하여 학습 평가를 위한 영역별 문제수와 비율을 보여준다.

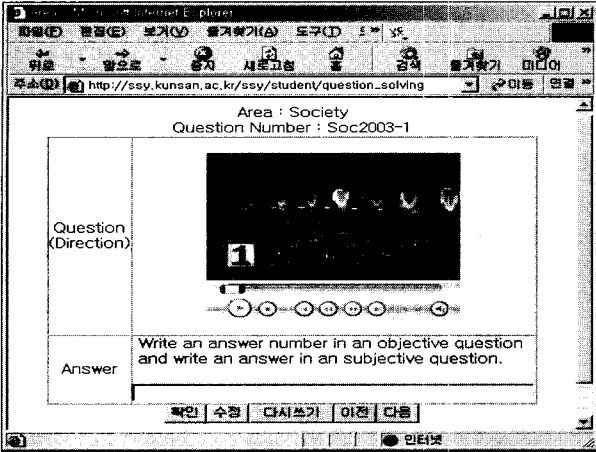


(그림 8) 학생 모듈의 시험 일정

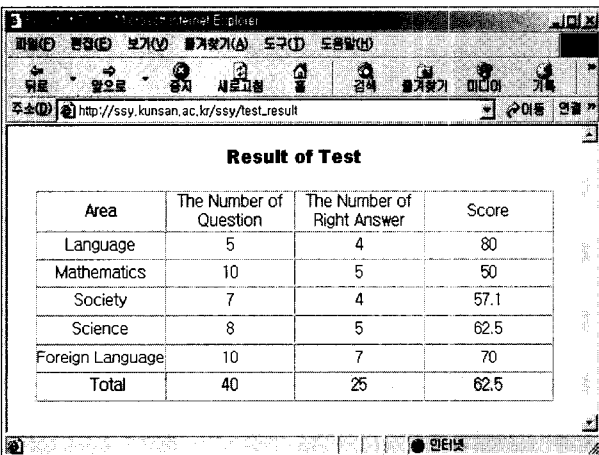
학생들은 (그림 8)의 시험 일정을 열람한 후 (그림 9)와 같이 연속적으로 시험 문제를 풀게 된다.

시험 문제 풀이에서는 각 문제에 대한 해당 영역과 문제 번호를 사용자에게 제시하고 사용자가 선택한 답을 입력받

는다. 부여된 문제 풀이를 완료한 경우 자동으로 (그림 10) 과 같은 시험 결과 확인 페이지로 이동하여 자신의 시험 결과를 확인한다.



(그림 9) 학생 모듈의 문제 풀이 화면



(그림 10) 학생 모듈의 시험 결과 화면

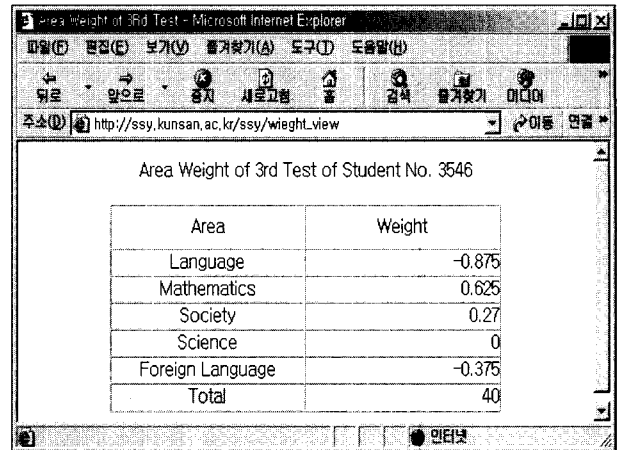
이렇게 채점된 시험 결과는 영역별 상대 점수에 의하여 카테고리-기반 시스템의 사용자 프로파일을 갱신한다. (그림 10)에서 수학(Mathematics) 영역을 살펴보면, $T_i(N)$ 은 전체를 100점 만점으로 하여 100을 갖고 $T_i(S)$ 는 $T_i(N)$ 을 만점 기준으로 얻은 시험의 총점 62.5를 갖는다. $Q_a(S)$ 는 영역별 총점을 100점 만점으로 하여 100을 갖고 $T_a(S)$ 는 $Q_a(S)$ 를 만점 기준으로 얻은 영역의 총점 50을 갖는다. 그리고 $T_a(N)$ 은 전체 영역수인 5를 갖는다.

(그림 11)은 (그림 10)의 시험 결과를 바탕으로 영역별 가중치를 부여한 결과를 나타낸다.

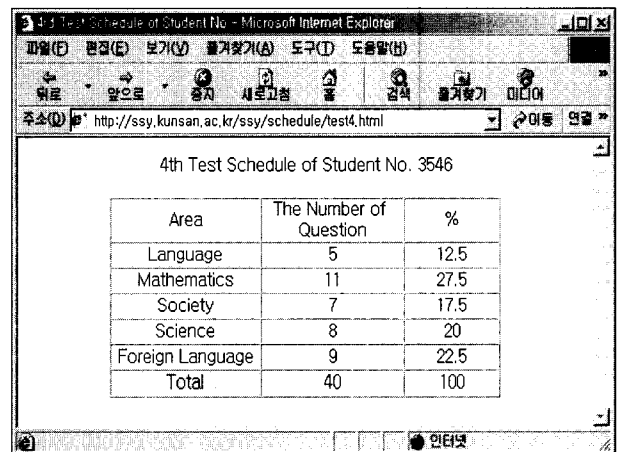
다음 시험에서는 선행 시험을 통하여 갱신된 사용자 프로파일이 적용되어 (그림 12)와 같이 문제수에서 변화를 갖는다.

<표 2>에서는 3545번 학생의 사용자 프로파일이 적용된

3번째 시험과 4번째 시험 사이의 문제수 변화를 비교하여 나타내었다.



(그림 11) 학생 모듈의 영역별 가중치



(그림 12) 학생 모듈의 다음 시험 일정

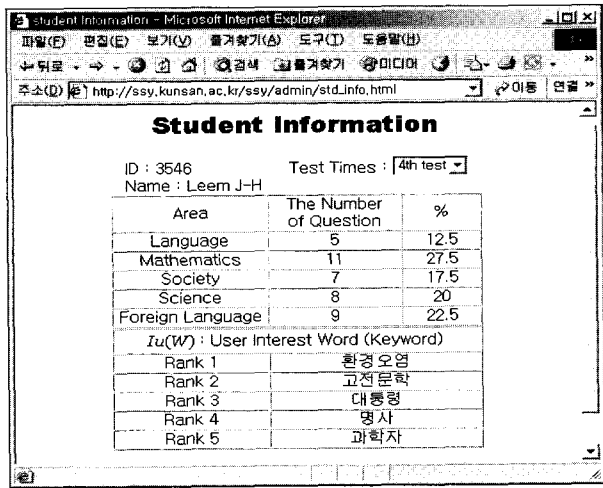
<표 2> 3번째 시험과 4번째 시험의 문제수 변화

Area	3번째 시험		4번째 시험	
	문제수	% (영역/전체)	문제수	% (영역/전체)
Language	5	12.5%	5	12.5%
Mathematics	10	25%	11	27.5%
Society	7	17.5%	7	17.5%
Science	8	20%	8	20%
Foreign Language	10	25%	9	22.5%

<표 2>를 통하여 알 수 있는 것처럼, 각 영역에 대한 개인화된 특성을 바탕으로 사용자 프로파일을 적용하고 문제수를 조정함으로써 학생들에게 보다 효율적인 학습 평가를 수행하였다.

(그림 13)은 관리자 모듈에서 3545번 학생의 시험 정보를 열람한 화면이다.

(그림 13)에서 $Iu(W)$ 는 키워드 기반 시스템에 적용되어 다음 시험 평가에 반영된다.



(그림 13) 관리자 모듈의 학생정보 화면

4.7 비교 및 평가

단순하고 확실적인 전통적인 학습 평가 시스템과 본 논문에서 제시한 새로운 학습 평가 시스템 사이의 특성을 <표 3>과 같이 비교 평가하였다.

<표 3> 시스템 비교 평가

구분	전통적 시스템	제안 시스템
문제 출제 및 평가 형태	단순, 전체적, 확실적	개인별로 다른 형태
영역별 문제수	전체 동일	개인별로 다름
문제 출제 및 평가 기준	출제 및 평가자 기준	개인의 특성과 성향 및 관심
문제 출제 및 평가 우선순위	우선순위 없이 동일	부족한 영역 우선
문제 출제 매체	텍스트	비디오
평가 목적	단순한 학습 성취도 평가	학습 성취도 향상

<표 3>의 비교 평가 내용과 같이 사용자 프로파일을 이용한 비디오 학습 평가 시스템의 가장 중요한 목적은 전통적인 평가 시스템이 갖는 단순하고 확실적인 학습 성취도 평가가 아니라 학생 개개인의 특성과 성향 및 관심을 고려하여 부족한 영역에 대한 보충과 우수한 영역에 대한 관심 유지를 통하여 학습 성취도를 향상시키는데 있다.

5. 결론

본 논문에서는 사용자 프로파일을 기반으로 한 정보 필터링을 이용하여 학생 개개인의 특성 및 성향과 잘 부합되

는 효율적인 비디오 학습 평가 시스템을 제시하였다. 평가에 사용자 프로파일을 적용하기 위하여 카테고리-기반 시스템과 키워드-기반 시스템을 합성한 문제 출제 방법을 이용하였다. 또한 본 논문에서는 비디오 형태의 문제 출제를 위하여 위치, 크기, 그리고 컬러 정보를 바탕으로 한 구조적 특징을 이용하여 문제 출제 장면의 키 프레임을 추출하였으며, 문제 출제 구간은 그레이 레벨 히스토그램 차이와 시간 윈도우를 이용하여 추출하였다. 학생들은 이 시스템을 이용하여 자신들의 부족한 영역을 보충하고 관심 영역을 유지하면서 학습 성취도를 향상시킬 수 있다. 특히, 사용자 프로파일을 이용하여 학습 평가에 대한 효율성을 매우 높였으며, 사용자 관심 단어를 이용하여 학습에 대한 흥미를 유도하였다. 또한 교사의 입장에서도 학생들을 지도하는데 커다란 도움을 주도록 하였다.

참고 문헌

- [1] D. Lefkovitz, "File Structure for Online Systems," Rochelle Park : Hayden, 1969.
- [2] Feynman, C., "Nearest neighbor and maximum likelihood methods for social information filtering," International Document, MIT Media Lab, Fall, 1993.
- [3] Sheth, B. D., "A Learning Approach to Personalized Information Filtering," SM Thesis, Department of EEVS, MID, Feb., 1994.
- [4] T. W. Malone, et al., "Intelligent Information Sharing System," Communications of the ACM, Vol.30, No.5, pp. 390-402, 1987.
- [5] Salton, G. and McGill, M. J., "Introduction to Modern Information Retrieval," McGraw-Hill, 1993.
- [6] Thomas Kahabka, Mari Korkea-aho and Günther Specht, "GRAS : An Adaptive Personalization Scheme for Hypermedia Databases," Proc. of the 2nd Conf. on Hypertext-Information Retrieval-Multimedia (HIM '97), pp.279-292, 1997.
- [7] Goldberg, D., Nicholas, D., Oki, B. and Terry, D., "Using Collaborative Filtering to Weave an Information Tapestry," CACM, Vol.35, No.12, pp.61-70, Dec., 1992.
- [8] Tak Y. Yan and Hector Garcia-Molina, "SIFT-A tool for wide-area information dissemination," In Proc. of the 1995 USENIX Technical Conf., pp.177-186, 1995.
- [9] Paul Resnick, Neophytos Iacovou, Mitesh Suchak, Peter Bergstrom and John Riedl, "GroupLens : An open architecture for collaborative filtering of netnews," In Proc. of ACM 1994 Conf. on Computer Supported Cooperative Work, pp.175-186, 1994.
- [10] Lang, K., "NewsWeeder : An Adaptive Multi User Text

Filter," Research Summary, Aug., 1994.

- [11] M. Altinel and M. J. Franklin, "Efficient Filtering of XML documents for Selective Dissemination of Information," Proc. VLDB Conf., Sep., 2000.
- [12] M. Cherniack, M. J. Franklin and S. Zdonik, "Expressing User Profiles for Data Recharging," IEEE Personal Communications, pp.6-13, 2001.
- [13] I. Schwab and A. Kobsa, "Adaptivity through Unobstrusive Learning," KI 3(2003), Special Issue on Adaptivity and User Modeling, pp.5-9, 2002.
- [14] Lee E. B., Kwak D. H. and Ryu K. H., "Understanding of Computer," KNOU press center, 1999.



신성운

e-mail : syshin@cs.kunsan.ac.kr

1994년 군산대학교 전산학과 졸업

1997년 군산대학교 대학원 정보통신공학과
(공학석사)

2003년 군산대학교 대학원 컴퓨터학과
(이학박사)

1998년~2002년 군장대학 컴퓨터응용학부 겸임교수

1999년~2002년 (주) 넷플러스 동여상연구팀 연구원

2003년~현재 군산대학교 컴퓨터정보과학과 IT 교수

관심분야 : 컴퓨터비전, 패턴인식, 멀티미디어 정보검색, 감성공학, 멀티미디어 데이터베이스