Journal of Korea Multimedia Society Vol. 9, No. 5, May 2006(pp. 658-670)

모바일 교육 시스템을 위한 효율적인 영상 검색 구축

한은정[†], 박안진[††], 정기철[†††]

요 약

교육 매개체의 디지털화를 위해 정보 통신망과 기술을 본격적으로 교육에 도입함으로써 교육의 패러다임이 e-learning(electronics learning)으로 이동하고 있다. 그러나 e-learning 콘텐츠는 컴퓨터가 있는 장소에서만 제공받을 수 있기 때문에, 언제 어디서나 사용자가 원할 때 교육받기 힘든 단점이 있으며, 기존 오프라인 콘텐츠를 e-learning 콘텐츠로 재구성하기 위해 많은 비용과 시간이 소요된다. 본 논문에서는 정적인 2차원 시각 정보만을 제공하는 기존 오프라인 콘텐츠의 그림 영상에 모바일 기기를 이용하여 동적인 인터랙션과 다양한 멀티미디어 정보를 융합할 수 있는 교육용 모바일 콘텐츠 (education mobile contents: EMC) 시스템을 제안한다. 제안된 영어 교육 콘텐츠는 기존 오프라인 콘텐츠의 그림 영상을 저해상도 카메라가 장착된 모바일 기기로 인식하기 위해 모양 기반의 영상 검색 방법을 이용하며, 수행속도와 크기와 기울기 변화에 강건한 시작점 일치를 통한 차분 체인코드(differential chain code)를 사용한 DTW(dynamic time warping)를 이용한다. EMC는 모바일 기기를 이용하여 장소에 상관없이 오프라인 콘텐츠에 적합한 온라인 콘텐츠를 빠르고 정확하게 제공할 수 있다.

Effective Image Retrieval for the M-Learning System

Eunjung Han[†], Anjin Park[††], Keechul Jung[†††]

ABSTRACT

As the educational media tends to be more digitalized and individualized, the learning paradigm is dramatically changing into e-learning. Existing on-line courseware gives a learner more chances to learn when they are home with their own PCs. However, it is of little use when they are away from their digital media. Also, it is very labor-intensive to convert the original off-line contents to on-line contents. This paper proposes education mobile contents(EMC) that can supply the learners with dynamic interactions using various multimedia information by recognizing real images of off-line contents using mobile devices. Content-based image retrieval based on object shapes is used to recognize the real image, and shapes are represented by differential chain code with estimated new starting points to obtain rotation-invariant representation, which is fitted to computational resources of mobile devices with low resolution camera. Moreover we use a dynamic time warping method to recognize the object shape, which compensates scale variations of an object. The EMC can provide learners with quick and accurate on-line contents on off-line ones using mobile devices without limitations of space.

Key words: Contents Recycling(콘텐츠 재활용), Mobile Learning(모바일 학습), Multimedia Education Contents(멀티미디어 교육 콘텐츠), Shape-based Image Retrieval(모양 기반 검색), DTW (dynamic time warping)(동적타임와핑(DTW)), Mobile Vision(모바일 비전)

※ 교신저자(Corresponding Author) : 정기철, 주소 : 서울시 동작구 상도5동 1-1 숭실대학교(156-743), 전화 : 02) 828-7260, FAX : 02)822-3622, E-mail : kcjung@ssu.ac.kr
접수일 : 2006년 1월 17일, 완료일 : 2006년 4월 17일

[†] 준회원, 숭실대학교 IT대학 미디어학부 콘텐츠 공학(박사과정) (E-mail : hanej@ssu.ac.kr)
[††] 숭실대학교 IT대학 미디어학부 미디어 공학(박사과정) (E-mail : danny0411@nate.com)
[†††] 정회원, 숭실대학교 IT대학 미디어학부 교수

1. 서 론

정보통신의 급속한 발전은 글로벌화, 네트워크화, 디지털화로 지식·정보사회의 끊임없는 변화와 개발을 요구하고 있다. 교육환경도 이에 부응하여 다양한 교육 콘텐츠가 학습자 중심의 콘텐츠로 발전하고 있다[1-2]. 또한, 디지털 컨버전스의 핵심기기로 발전하는 휴대단말기를 통해 언제·어디서나 교육할 수 있는 유비쿼터스(Ubiquitous) 교육 환경이 구축을 위한 연구가 활발히 진행하고 있다[3].

전통적인 오프라인 콘텐츠는 다른 매체에 비해 기록성, 역사성이 보장되는 장점을 가진 반면 신속성이 떨어지며, 정적인 2차원 정보만을 가지고 있어 다양한 콘텐츠를 제공하기에는 어려움이 있다. 이에 따라, 국내의 경우 1999년 이후 새로운 교육 콘텐츠로써, 전통적인 교수자 중심의 면대면 교육체제 대신 테크놀로지를 활용하는 학습자 중심의 e-learning 학습체제가 확산되고 있으며, 다양한 교육현장에서 전통적 오프라인 학습의 한계를 극복할 수 있는 대안적 학습방법으로 자리잡아 가고 있다[1]. 그러나 e-learning 콘텐츠는 컴퓨터가 있는 장소에서만 제공받을 수 있기 때문에, 언제 어디서나 사용자가 원할 때 교육받기 어려우며, 기존 오프라인 콘텐츠를 e-learning 콘텐츠로 재구성하기 위해 많은 비용과 시간이 소요된다. 이런 단점을 해결하기 위해, '모바일 러닝(m-learning)'과 '블렌디드 러닝(blended learning)'에 대한 개념이 등장하였다. 모바일 러닝은 교육 환경에 이동성을 부가함으로써 학습자가 언제 어디서나 원하는 강의를 들을 수 있는 교육환경을 제공한다[2]. 이재석[4]은 모바일 장치를 이용하여 단어 및 숙어를 사전식 방법으로 접근하는 영어 어휘 학습 시스템을 개발하였으며, 이영석[5] 등은 학습 성취 속도가 빠른 학생과 뒤떨어지는 학생들의 수준별 학습을 위해, 모바일 장치를 통해 제공되는 지능형 교육시스템을 개발하였다. 또한 Sharples[6] 등은 학습자 스스로 학습할 수 있게 모바일 장치에 적합하고 직관적인 자기주도적 학습 콘텐츠를 개발하였다.

블렌디드 러닝은 유비쿼터스 시대에 대비하기 위해 다양한 매체를 통한 교육서비스를 준비함과 동시에 온·오프라인을 통합한 교육 환경을 제공한다[1)][7]. Billinghurst[8] 등은 학습자들이 오프라인 책 위에서 3D입체의 가상 환경을 경험할 수 있는 MagicBook을 개발하였으며, 이민경과 우운택[9]은 요리 강좌 교재 위에 실습하는 모습을 동영상 강의를 통하여 학습자에게 필요한 교육 효과를 주었다. 위의 두 시스템[8,9]은 카메라를 이용하여 실제 오프라인 콘텐츠를 입력받고, AR(augmented reality) 기술을 이용하여 온라인 정보를 제공한다. 대교에서는 P2I(Print to Internet) 제품에 삽입된 '아이콘'이라는 원형 코드에 PC 카메라를 비추면 관련된 온라인 콘텐츠로 연결되어 학습자에게 교육내용과 연계된 다양한 학습정보를 제공하는 기술을 제안하였다[10].

Chen[11] 등은 모바일 러닝과 블렌디드 러닝 두 개념의 장점을 포함하는 콘텐츠로써, 고해상도 카메라를 이용하여 나비영상을 입력받고, 인식된 나비영상과 그와 관련된 정보를 각 사용자에게 모바일 기기를 통해 제공하였다. KTF는 교육 콘텐츠는 아니지만, 핫코드(hot code) 서비스를 통해 신문, 잡지, 지하철, 광고 등에서 핫코드에 카메라폰을 비추면 관련된 모바일 콘텐츠나 모바일 커머스 사이트로 연결시켜주는 새로운 무선 인터넷 접속 서비스를 제공한다[12]. 앞에서 언급한 블렌디드 러닝을 이용한 학습방법[8,9]와 대교, KTF에서 제안한 방법은 오프라인 콘텐츠와 온라인 콘텐츠를 연결하기 위한 오프라인 콘텐츠의 인식방법으로 패턴 마커를 이용하기 때문에, 마커가 있는 장소에서만 콘텐츠를 제공받을 수 있는 단점이 있다.

본 논문에서는 모바일 러닝과 블렌디드 러닝의 개념을 모두 포함하는 교육용 모바일 콘텐츠 (education mobile contents: EMC) 시스템을 제안한다. 제안된 교육 콘텐츠는 첫째, 유비쿼터스 시대를 맞이 하여 언제·어디서나 공간적 제약없이 콘텐츠를 제공받을 수 있는 모바일 기기를 이용한다. 둘째, 오프라인 콘텐츠와 온라인 콘텐츠를 연결하기 위한 오프라인 콘텐츠의 인식 방법으로, 책의 각 페이지마다 사용해야 하는 패턴 마커 대신 오프라인 콘텐츠의 그림 영역을 이용하며, 모바일 기기의 저해상도 카메라에 의한 색상 검색 방법의 어려움을 해결하기 위해, 모양 기반 검색 방법(shape-based retrieval method)을 이용한다. 연산자원이 부족한 모바일 기기에서 실시간으로 수행하기 위해 외곽선에 대한 정보 손실이 적고, 저장 공간을 줄일 수 있는 체인 코드(chain

1) 오프라인 콘텐츠 활용은 온라인화 작업시에 소요되는 노동을 최소화하기 위한 것이다

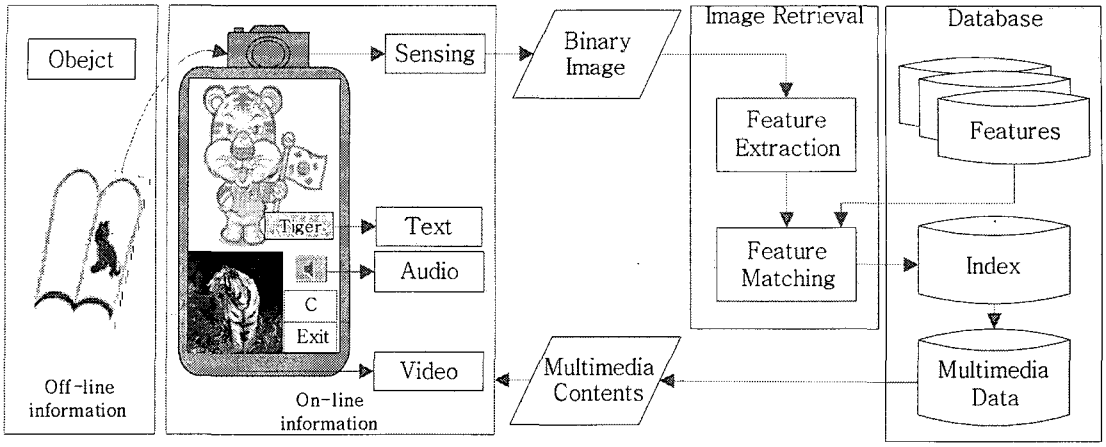


그림 1. EMC 시스템의 흐름도

code)를 이용한다. 오브젝트의 회전에 대한 문제를 해결하기 위해 간단한 연산을 가진 시작점 일치 방법과 차분 체인코드(differential chain code)를 이용하며, 인식단계에서 DTW(dynamic time warping)를 이용하여 크기 변화에 대한 문제를 해결한다. 결과적으로 EMC는 모바일 기기를 이용하여 공간적 제약없이 빠르고 정확하게 오프라인 콘텐츠와 인터랙션을 수행할 수 있으며, 다양한 멀티미디어 학습 콘텐츠(동영상, 텍스트, 원어민의 영어 발음)를 제공할 수 있는 장점을 가진다.

본 논문 구성은 다음과 같다. 2장에서 모바일 기기를 통한 영어 학습 콘텐츠의 전체적인 시스템 구조와 실행 과정을 보여주고, 모양 기반의 영상 검색 기법의 그림 추출, 특징값 추출 및 인식하는 방법에 대해 알아본다. 3장에서 디지털 카메라와 PDA에 장착된 카메라를 이용한 색상의 실험 결과를 보여주고, 카메라를 통해 영상을 제공할 때 회전 변환을 통한 수행 결과 분석과 사용자에게 직접 실험한 결과를 통해 학습에 대한 효과를 분석하며, 마지막으로 4장에서 결론을 맺는다.

2. EMC를 위한 영상 검색 구축

EMC는 기존의 다양한 오프라인 콘텐츠를 바탕으로, 카메라가 장착된 PDA를 통하여 동적인 인터랙션과 다양한 멀티미디어 정보를 융합함으로써, 정적인 2차원 시각 정보만을 제공하는 기존 오프라인 책의 활용범위를 넓히는 학습 시스템이다. 그림 1은 EMC의 전체 흐름도로써, 영상 검색 모듈과 데이터베이스(DB)로 이루어져 있으며, PDA 카메라를 통해 입력받은 유아용 영어 사전의 한 면에 대응하는 DB 내에 저장되어 있는 멀티미디어 학습 콘텐츠(영어 단어 철자, 동영상, 원어민의 영어 발음)를 제공한다.

기존 영상 검색 방법에 사용되는 색상 정보를 이용한 검색 방법은 영상의 크기에 무관하며, 다른 정보에 비해 계산이 간단하지만[13], 색상 정보를 제대로 입력받지 못하는 모바일 기기에 장착된 저해상도 카메라를 이용하면 제대로 수행되지 않는다. 질감(texture)을 이용한 검색 방법은 옷감의 무늬, 벽, 타일 등의 균질 또는 비 균질의 패턴에서 높은 인식을 보이지만, EMC 시스템의 오프라인 콘텐츠와 같은 일정한 패턴이 없는 그림 책에는 적합하지 않다. 우리는 저해상도 모바일 기기에 장착된 카메라를 위해 모양기반의 검색 방법을 이용하며, 다양한 회전이나 크기 변화에 대한 문제를 해결하고 실수연산기가 없는 모바일 기기에서 실시간으로 영상 검색을 수행하기 위해, 간단한 연산을 가진 시작점 일치 방법과 차분 체인코드를 이용하여 회전에 강건한 특징값을 추출하며, 크기 변화에 대한 문제는 인식 단계에서 널리 이용되는 DTW를 이용한다.

2.1 오브젝트 추출

유아용 영어 사전의 그림 영역[2]을 인식의 대상으

2) 실험에서 사용한 유아용 영어 사전은 그림, 글자, 배경으로 구성되어 있다(그림 2(a))

로 생각하고, 입력받은 영역 중 그림책의 내용(글자와 그림)과 배경(흰 바탕)을 분리하며, 내용에서 가장 넓은 영역을 차지하고 있는 부분을 그림 영역으로 추출한다. 그림 2에서 보는 바와 같이 흰색으로 되어 있는 그림책의 배경과 그 밖의 색상으로 되어 있는 그림책의 내용을 분리하기 위해, 임계화(thresholding)로 영상을 이진화하며, 영상의 외곽에서 발생하는 임계화의 오류는 영상의 중심영역과 연결 성분(connected component)의 면적을 이용하여 해결한다. 그림 2(b)는 영상 중심영역과 면적이 가장 큰 영역의 외곽선을 회색으로 표시한 영상이며, 그림 2(c)는 외곽선 내부의 오브젝트를 추출한 결과이다.

2.2 오브젝트 인식

추출된 오브젝트의 모양 정보를 이용하여 인식을 수행하며, 모양 정보를 표현하기 위해 오브젝트 외곽선의 체인코드를 이용한다. 체인코드는 외곽선에 있는 에지 포인터의 방향을 이용하여 오브젝트를 표현하는 방법이며, 에지 포인터의 방향은 그림 3(a)의 8방향 중 하나이다.

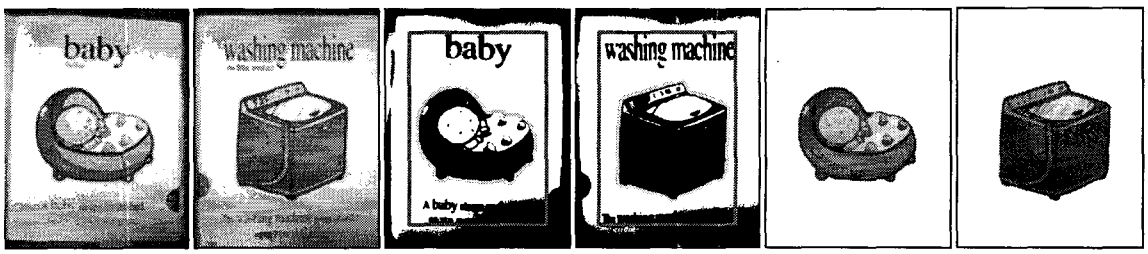
체인코드는 외곽선에 대한 정보 손실이 적고, 저장공간을 줄일 수 있는 장점이 있지만, 같은 오브젝트라도 회전과 크기에 따라 오브젝트의 표현이 변화하는 단점이 있다[13]. 오브젝트의 회전에 대한 문제를 해결하기 위해, 먼저 오브젝트 회전에 의해 변하는 시작점을 일치시키며, 획득된 시작점에서 회전을 고려하여 특징값을 획득하기 위해 차분 체인코드(그림 3(b))를 이용한다. 그리고 시작점(*StartPoint*)을 이차원 모멘트(moment)(식 1)를 이용하여 추출된 오브젝트(*object*)의 중심값(식 2)에서 가장 가까운 외곽점을 이용하여 구한다(식 3). X, Y는 외곽점의 좌표 집합이고, x_i, y_i는 i번째의 외곽점 나타낸다.

$$M_{pq} = \sum_x \sum_y x^p y^q object(x, y) \tag{1}$$

$$C_x = \frac{M_{10}}{M_{00}}, C_y = \frac{M_{01}}{M_{00}} \tag{2}$$

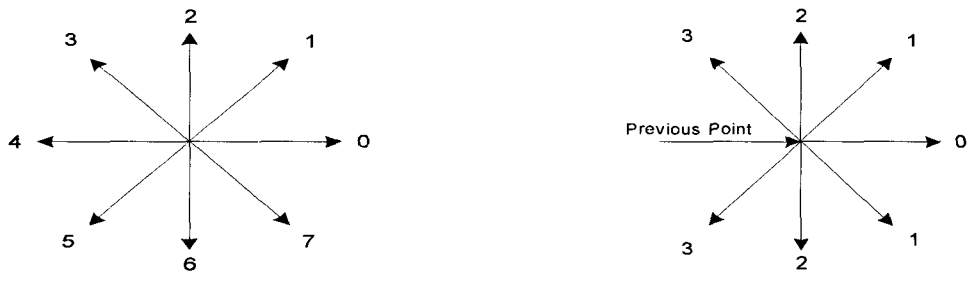
$$StartPoint = \arg\min_{x_i \in X, y_i \in Y} \sqrt{(x_i - C_x)^2 + (y_i - C_y)^2} \tag{3}$$

그림 4는 추출된 오브젝트(그림 4(a,e))와, 오브젝트의 외곽선, 중심값(검은색 사각형) 그리고, 시작위



(a) (b) (c)

그림 2. 오브젝트 추출: (a) PDA를 통해 입력받은 그림 책, (b) 오브젝트 외곽선과 영상의 중심영역을 표시한 이진 영상, (c) 추출된 오브젝트



(a) (b)

그림 3. 체인코드: (a) 8방향성 체인코드의 방향 벡터, (b) 차분 체인코드

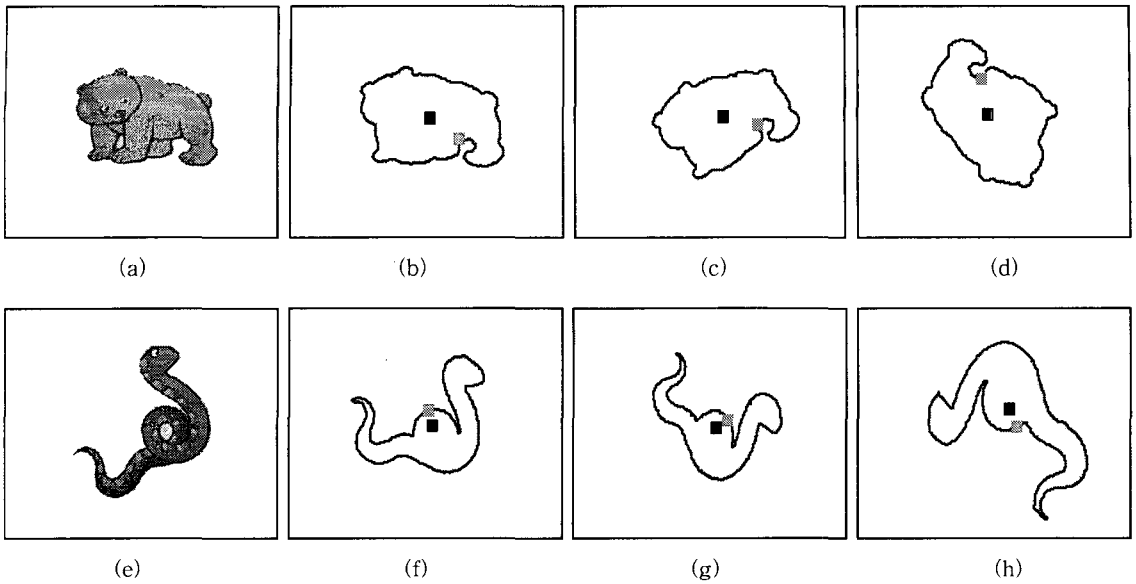


그림 4. 오브젝트의 중심점과 시작위치: (a,e) 입력영상의 오브젝트, (b-d,f-h) 오브젝트의 외곽선, 중심값, 시작위치를 나타낸 영상

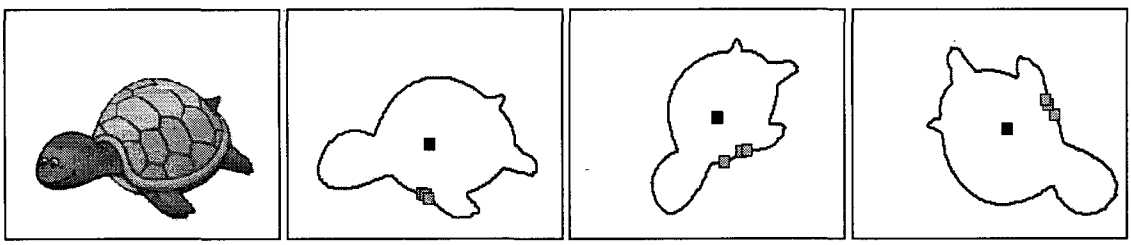


그림 5. 중심점에서 가장 가까운 3개의 점을 이용한 시작 위치

치(회색 사각형)를 표시한(그림 4(b-d),(f-h)) 예를 보여준다. 그림 4(e-h)에서 보는 바와 같이 중심점에서 외곽선까지 거리의 최소값이 원형 모양에 있을 때 시작점이 일치하지 않는 경우가 생기며, 이를 해결하기 위해 중심점에서 가장 가까운 3개의 점을 이용하여 시작위치를 구한다[3](그림 5).

실제 환경에서 입력받은 대부분의 영상에서 같은 오브젝트라도 외곽선의 길이(체인코드의 수)는 다른 경우가 대부분이다. 따라서 길이가 다른 두 패턴 사이의 유사도를 측정하기 위해서 동적 프로그래밍(dynamic programming: DP)을 기반으로 한 DTW [15]를 이용한다.

입력 영상의 특징값과 참조 영상의 특징값이 각각 $A=\{a_1,a_2,a_3,\cdots,a_I\}$와 $B=\{b_1,b_2,b_3,\cdots,b_J\}$일 때, 길이가 다른 두 특징값 사이의 유사도는 식 (4)를 이용하며, 입력패턴의 길이에 의해 유사도가 달라지는 문제는 정규화를 통하여 해결한다(식 (5))[15].

$$\gamma(i,j)=\min\{\gamma(i,j-1)+d(a_i,b_j),\gamma(i-1,j-1) +2\times d(a_i,b_j),\gamma(i-1,j)+d(a_i,b_j)\} \quad (4)$$

$$DTW\ (A,B)=\frac{\gamma(I,J)}{I+J} \quad (5)$$

$d(a_i,b_j)$는 두 특징값 a_i와 b_j 사이의 차이를 나타내며, 유클리디언 측도(Euclidean distance)를 이용하여 구한다.

3) 5개의 영상을 이용하여 테스트한 결과 3개의 영상이 가장 가까운 하나의 점에서 시작점이 일치하지 않았으나, 가장 가까운 3개의 점을 이용하면 모두 일치하였다.

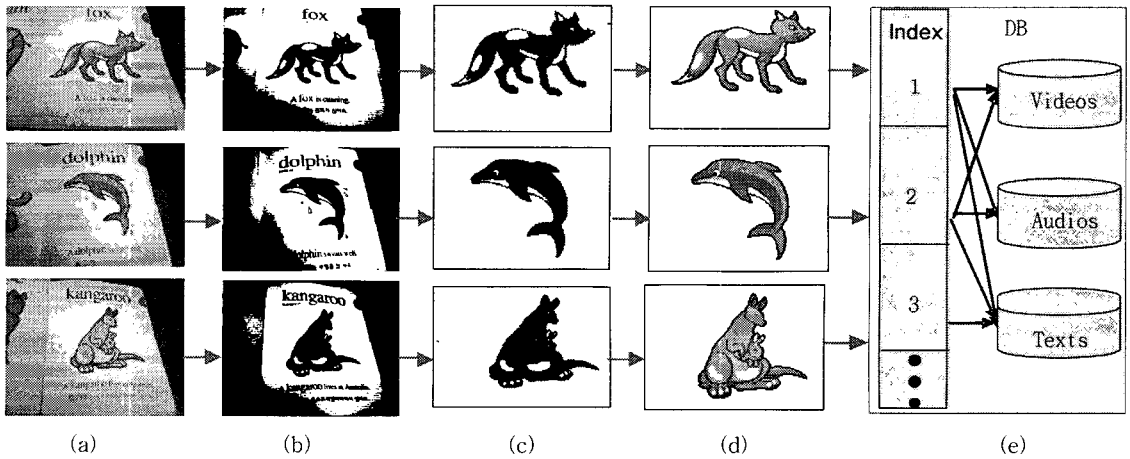


그림 6. DB 생성 과정: (a) PDA를 통해 입력받은 그림책, (b) 이진 영상, (c) 경계선을 표시한 오브젝트, (d) 추출된 오브젝트, (e) DB

DB에 있는 모든 특징값과 입력 영상에서 획득한 3개의 특징값을 DTW를 통해 비교하여 가장 유사한 결과값(index)을 획득한다(식 6).

$$index = \arg\min_{i\in\{1,2,\ldots,N\}} DTW(A_j, B_i)$$
$$j = \{1,2,3\} \quad (6)$$

B_i는 DB에 저장된 특징값, A_j는 입력 영상에서 획득한 특징값이며, N은 DB에 저장된 특징값의 개수이다.

2.3 DB 구성도

DB는 기존 오프라인 책의 각 그림에 대한 특징값과 인덱스(index) 그리고 멀티미디어 정보로 구성되어 있다(그림 6).

일반적인 모바일 환경에서 DB를 사용할 경우 파일의 크기를 고려해야 한다. 본 실험에서 사용한 PDA는 64MB Flash ROM과 128MB SDRAM의 메모리를 사용하고, DB에는 500KB 크기의 동영상 정보 50개, 10KB 크기의 문자 정보 50개, 그리고 700B 크기의 특징값 정보 150개가 저장되어있으며, 약 30MB의 공간을 필요로 한다[4].

DB의 구조는 Contents, Audio, Video와 Text 네 개의 원소로 그림 7과 같이 구성된다. Contents의 성분 요소는 주키(Content key) 와 content name, 외부키(video key, audio key, text key)로 구성된다. DB 구성은 각각의 요소를 만족시키며, 주키(video key, audio key, text key)와 파일 이름(video filename, audio filename, text filename) 그리고 텍스트 원소, 텍스트 내용의 넓이와 높이(text width, text height)를 포함한다.

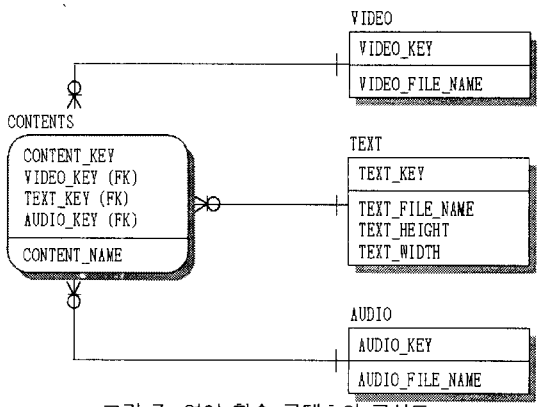


그림 7. 영어 학습 콘텐츠의 구성도

3. 실험 및 결과

실험 및 결과를 세 가지로 분류하여 기술한다. 첫째, 기존의 색상 기반 영상 검색방법을 모바일 기기에서 수행할 경우 생기는 문제점을 지적하기 위해

4) 본 논문에서 사용한 동영상은 16120 크기에 30초 길이, 음성은 10초의 길이, 문자는 640 크기를 사용한다.

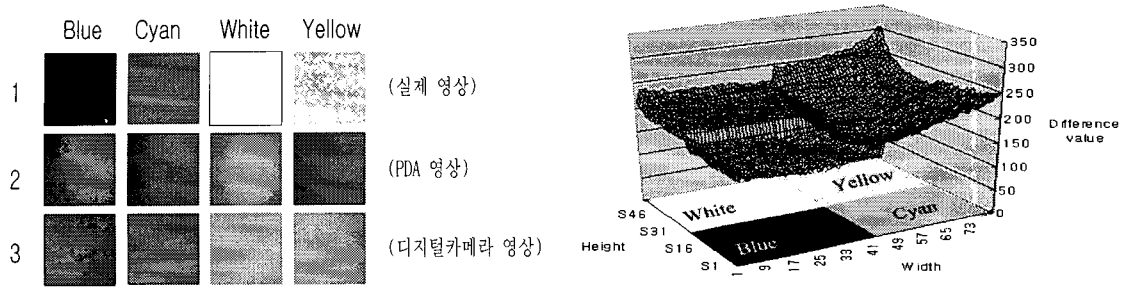


그림 8. 왜곡 분석을 위한 영상: (a) 파란색, 청록색, 노란색, 흰색에 대해 디지털 카메라와 PDA 카메라로 입력받은 영상, (b) 실제 영상과 PDA 카메라로 입력받은 영상의 차이값

모바일 기기에 장착된 30만 화소 카메라에 대한 색상 정보 추출의 취약점을 분석하고, 둘째, 모양 기반 영상 검색방법에서 발생할 수 있는 문제점으로, PDA에 장착된 카메라를 통하여 영상을 인식하는데 있어서 움직임 즉, 회전과 크기가 변화된 영상 정보를 얼마나 효율적으로 인식할 수 있는지에 대해 실험한다. 셋째, 사용자 측면에서의 EMC 학습 효과를 분석하기 위해, 기존 영어 책과 EMC 시스템을 통해 학습한 후 사용자 평가를 수행한다.

3.1 PDA 카메라의 색상 정보 분석

본 논문에서 일반 디지털 카메라와 PDA 카메라의 영상 획득을 통해 PDA 카메라의 색상 정보 추출의 취약점을 알아본다. 디지털 카메라는 800만 화소의 Olympus C-8080WZ 모델이며, PDA는 Pocket PC-2003 기반의 30만 화소 카메라가 장착된 POZ x301 모델이다. PDA를 위한 소프트웨어 개발도구로 embedded Visual C++ 4.0을 사용한다.

그림 8(a)는 파란색, 청록색, 노란색, 흰색에 대해 디지털 카메라와 PDA에서 입력받은 영상으로, PDA에서 입력받은 영상은 디지털 카메라에 비해 색상을 제대로 입력받지 못하며, 원래 색상에 비해 상당히 많은 차이점이 보인다. 그림 8(b)는 영상의 외곽에서 생기는 왜곡을 분석하기 위한 영상으로, 그림 8(a)의 4가지 색상을 그림 8(b)처럼 붙혀서 PDA 카메라를 통해 한번에 입력받고, 실제 색상간의 차이를 식(7)을 이용하여 나타낸 그래프이며, 파란색을 제외한 3가지 색의 외곽에서 영상의 왜곡이 더욱 심해짐을 알 수 있다. 이와 같은 이유로 본 논문에서는 PDA 카메라에서 입력받은 취약한 색상 정보 대신 모양 정보를 이용하여 인식을 수행하며, PDA에서의 수행 속도를 감안하여 이런 문제를 해결하기 위해 왜곡이 심한 영역을 제외한 영상의 중앙 260×180영역[5]에 대해서만 추출을 수행한다.

$$\text{DifferenceValue}=\sqrt{(r_1-r_2)^2+(g_1-g_2)^2+(b_1-b_2)^2} \quad (7)$$

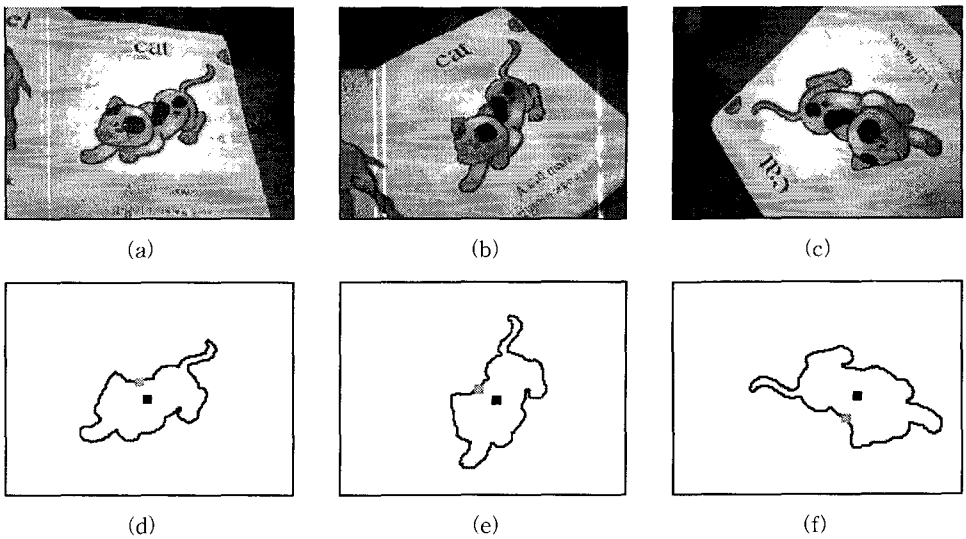
3.2 오브젝트의 회전과 크기 변화의 분석

모양 기반의 영상 검색방법은 다양한 회전과 크기 변화에서도 항상 높은 인식률을 보여야 한다[14].

그림 9는 다양한 회전에 대해 인식률을 평가하기 위한 데이터 중 하나의 예이며, 그림 9(d-f)에서 회전이 다양하게 변해도 중심점에서 가장 가까운 시작위치를 3개로 했을 경우 일치함을 알 수 있다[6]. 그림 10은 그림 9에 대한 차분 체인코드를 이용하여 획득한 특징값을 색상을 통해 표현한 영상으로, 그림 10(a-c)는 논문에서 제안한 시작 위치를 고려한 차분 체인코드의 영상이며, 그림 10(d-f)는 시작 위치를 고려하지 않고, 차분 체인코드를 이용한 영상이다. 그림 10(a-c)는 거의 유사한 특징값을 보이며, 그림 10(d)와 10(e)가 유사하게 표현되는 건 시작위치가 우연히 꼬리부분으로 같기 때문이다.

그림 11은 다양한 크기에 대해 인식률을 평가하기 위한 데이터 중 하나의 예로써, 그림 11(b)는 DB에 저장되어 있는 특징값을 추출한 일반적인 영상의 크기이며, 특징값의 길이는 175~200사이이다. 그림 12

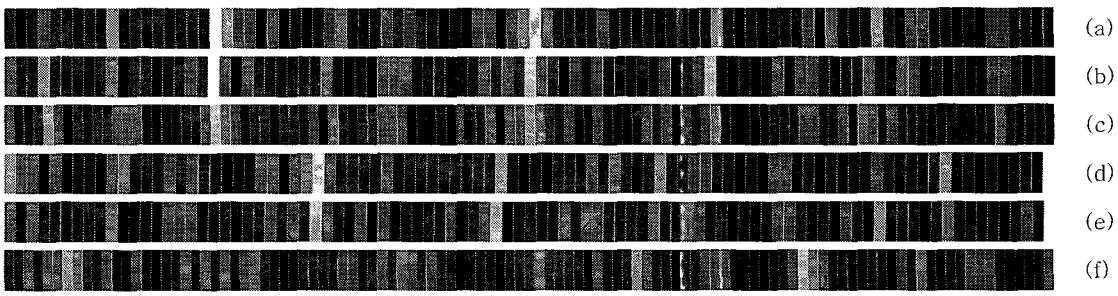
5) 실험에서 사용한 입력영상의 해상도는 320×240이다.
6) 그림 9에서 예를 든 고양이 그림은 다양한 회전에서 중심점에서 가장 가까운 외곽점이 일치하였기 때문에, 시작위치를 하나만 표시하였다.



(a) (b) (c)

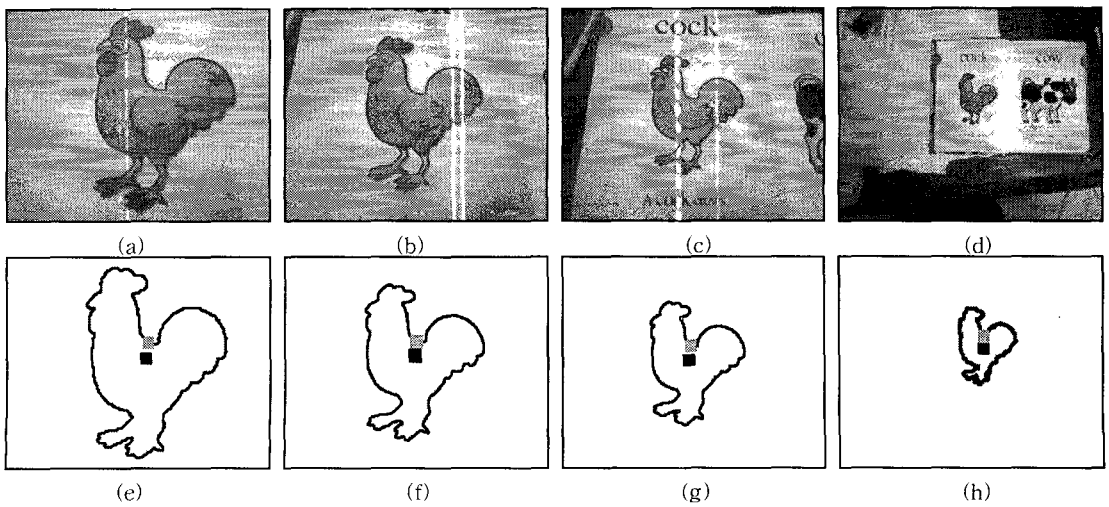
(d) (e) (f)

그림 9. 회전 변화에 대한 예: (a-c) 입력영상, (d-f) 오브젝트의 외곽선, 중심값, 시작위치를 표시한 영상



(a)
(b)
(c)
(d)
(e)
(f)

그림 10. 오브젝트의 특징값: (a-c) 그림 9(d-f)를 시작위치를 고려하여 특징값을 표현한 영상, (d-f) 시작위치를 고려하지 않고 특징값을 표현한 영상



(a) (b) (c) (d)

(e) (f) (g) (h)

그림 11. 크기 변화에 대한 예: (a-d) 입력영상, (e-h) 오브젝트의 외곽선, 중심값, 시작위치를 표시한 영상

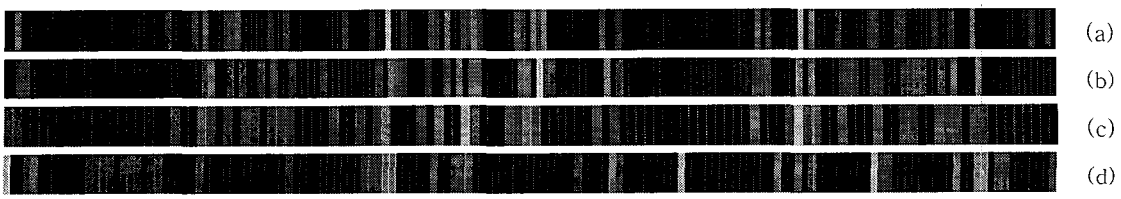


그림 12. 그림 11(e-h)에 대한 오브젝트 특징값

는 그림 11에 대한 차분 체인코드를 이용하여 색상으로 표현한 영상으로, 그림 11(e-h)와 같이 거의 유사한 특징값을 가지지만, 그림 11(h)는 오브젝트가 작기 때문에[7] 디지털화시 정보 손실이 있고, 오브젝트를 제대로 표현하지 못한다. 따라서 오브젝트의 특징값 길이를 100이상으로 제한하며, 작은 오브젝트는 인식의 대상에서 제외한다.

그림 13은 다양한 회전과 크기의 변화가 동시에 있을 때의 예를 보여주며, 그림 14는 그림 13(d-f)에 대한 특징값을 색상으로 보여주는 영상이다. 그림 14에서 시간 축에서 약간씩 어긋남을 볼 수 있으며, DTW를 이용한 인식단계에서 해결된다.

3.3 수행 결과 분석

그림 15는 PDA에 제공된 EMC 시스템을 실행하는 화면을 보여준다. 사용자가 PDA를 이용해서 원하는 영상을 선택하면(그림 15(b)), EMC 시스템은 영상을 인식하고, 인식된 결과를 기반으로 동영상과 단어를 사용자에게 제공하며(그림 15(c)), 사용자가 단어의 발음을 요청하면 오디오 파일을 통해 원어민 발음을 제공한다.

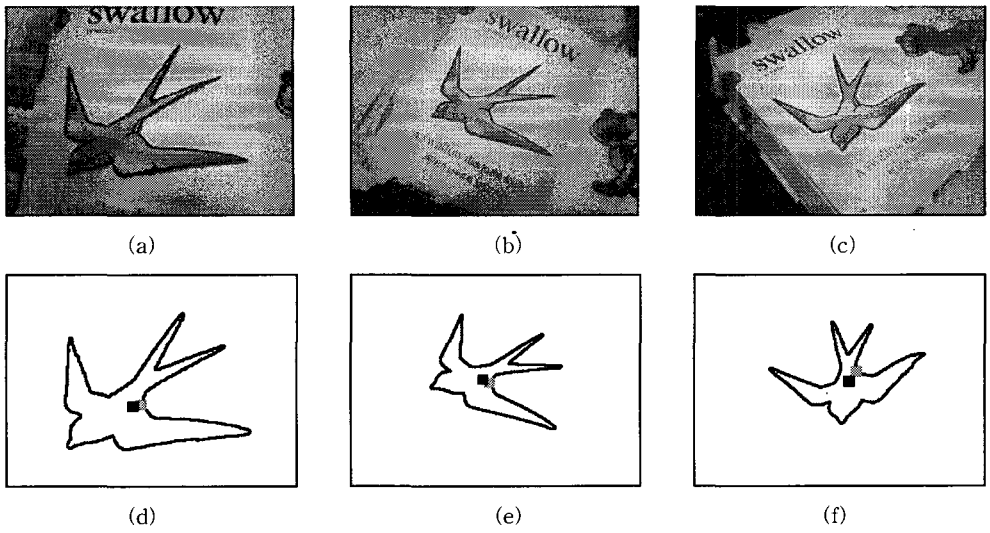


그림 13. 크기와 회전 변화에 대한 예: (a-c) 입력 영상, (d-f) 오브젝트의 외곽선, 중심값, 시작위치를 표시한 영상

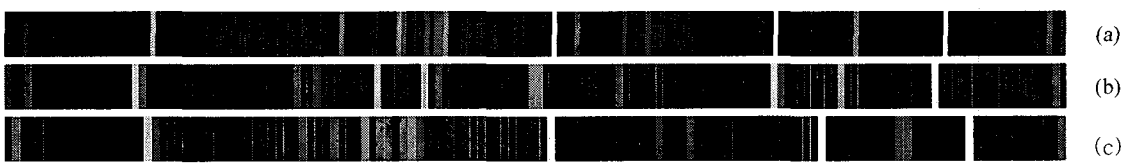


그림 14. 그림 13(d-f)에 대한 오브젝트 특징값

7) 그림 12(h)의 특징값 길이는 75이다.

(a) (b) (c)

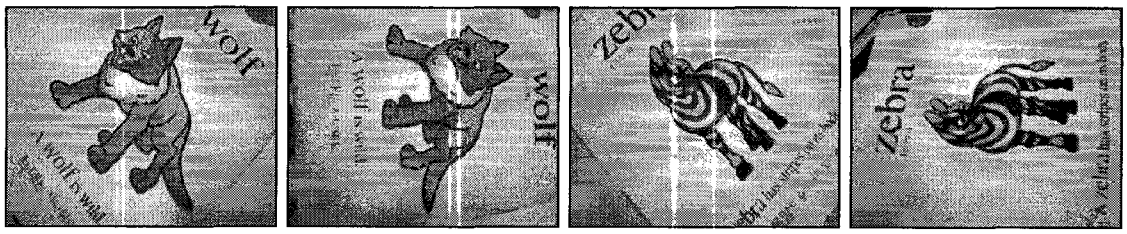
그림 15. 영상 검색 결과: (a) 수행 화면, (b) 영상 선택, (c) 멀티미디어 교육 콘텐츠

표 1은 한 영상에 대한 수행시간과 50장(논문에서 사용한 오프라인 콘텐츠의 총수)에 대한 인식률을 보인다. 그림 16은 오인식에 대한 첫 번째 예를 보여주며, 그림 17은 그림 16(e-h)에 대한 특징값을 색상으로 보여주는 영상이다. 그림 16에서 모양 정보의 외곽선만을 이용하기 때문에 유사한 특징값이 나올 때에는 오인식이 발생한다[13]. 그림 18은 오인식에 대한 또 다른 예를 보여주면, 오브젝트의 색상이 밝을 경우 오브젝트를 제대로 추출하기 못하는 경우가 생기며, 이로 인해 오인식이 발생한다.

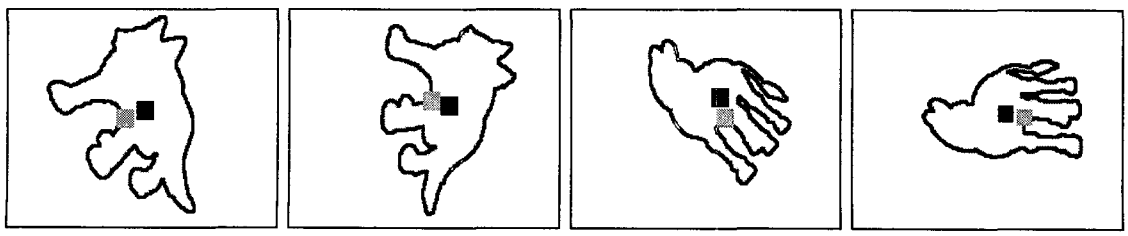
모바일 기기에서 기존의 오프라인 콘텐츠를 모두 DB에 저장하고, 손(또는 버튼)으로 검색하여 해당 멀티미디어 교육 콘텐츠를 제공하면 사용자가 직접 DB를 모두 검색해야 하기 때문에 오랜(약 5초) 시간

표 1. 수행 시간(ms)과 인식률(%)

수행시간(ms)				인식률 (%)
오브젝트 추출	특징값 추출	DTW를 이용한 오브젝트 인식	전체수행시간	
100	150	500	750	92



(a) (b) (c) (d)



(e) (f) (g) (h)

그림 16. 오인식에 대한 예: (a-d) 입력영상, (e-h) 오브젝트의 외곽선, 중심값, 시작위치를 표시한 영상

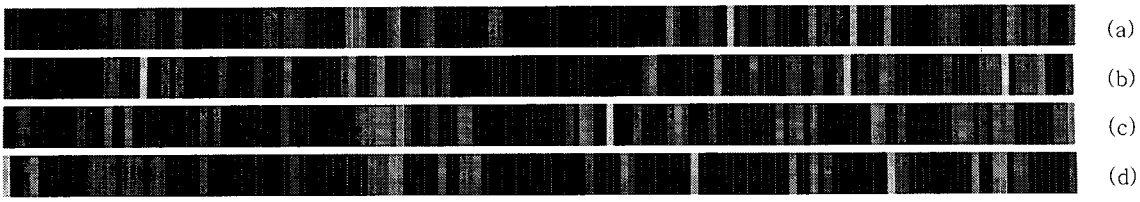
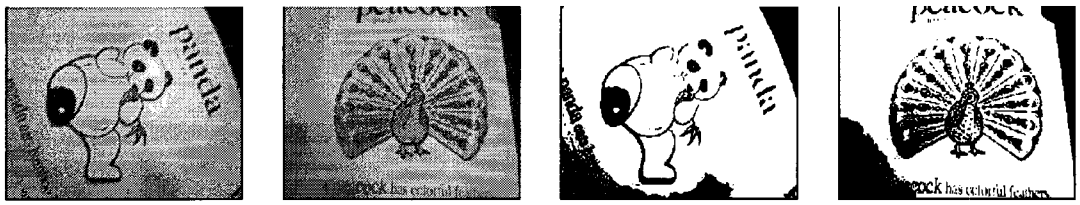


그림 17. 그림 15(e-h)에 대한 오브젝트 특징값



(a) (b)

그림 18. 오인식에 대한 예: (a) 입력영상, (b) 오브젝트 추출을 위한 이진영상

이 소요되지만, EMC 시스템은 모바일 기기의 카메라를 통해 빠르고(대략 0.75초) 정확하게 멀티미디어 콘텐츠를 제공받을 수 있으며, 빠르게 발전해가는 정보사회 교육 콘텐츠에 알맞다.

3.4 사용자에 의한 EMC 학습 효과 분석(User Study)

사용자로부터의 평가를 얻기 위해 수행시간과 인식률을 통해 실험하였던 50장을 이용하여 테스트를 실시하였다. 사용자 평가를 위해 50명이 참석하였으며, 이들은 8세에서 13세 사이의 남자 15명, 여자 35명으로, 학생들마다 실력차가 있는 것을 가만하여 영어 학습 성적의 수준이 비슷한 학생들로 구성하였다. 영어 책으로 학습을 하는 A그룹과 EMC 시스템이 들어있는 모바일 기기로 학습하는 B그룹으로 나누어 평가를 수행하였다. 학습의 효과를 알아보기 위해 읽기, 쓰기, 듣기 각각 10문제를 테스트하였고, 대체적으로 얼마나 오래 기억을 하고 있는지를 알아보기 위해 중간 학습을 하지 않은 제약을 두어, 21일 이후 다시 테스트를 하였으며 결과는 표 2와 같다. 표 2에서 EMC 시스템을 이용한 교육은 일반 오프라인 콘텐츠만을 이용한 교육(9.9)보다 높은 평균(13.6)이 나왔으며, 원어민의 발음을 듣고 따라하는 반복 학습을 통해 읽기(5.0)와 듣기(5.3)의 향상된 학습 점수가 나왔다. 또한 동영상을 통해 실감있는 학습을 제공함으로써 시간이 지나도 오래 기억함을 알 수 있다(A:9.9→4.3, B:13.6→12.6). 본 실험을 통해, 전통적인 학습 환경을 탈피하여 기존의 책에서 제공하지 않은 말하기/듣기 영역을 정확한 발음과 함께 학습할 수 있는 새로운 환경이 필요하다는 연구 결과를 얻을 수 있었다. 그러나 이 실험은 초기 실험 결과로써, 다수의 사람을 대상으로 하여 이루어진 실험 결과가 아니기 때문에, 학습 평가에 있어서 신빙성은 떨어진다. 따라서 우리는 추후에 보다 신빙성 있는 실험 결과를 얻기 위해 다수의 사람을 대상으로 실험을 실시할 계획이다.

4. 결 론

본 논문에서는 정적인 2차원 시각 정보만을 제공

표 2. 전체 평균값

콘텐츠	날짜	그룹	읽기	쓰기	듣기	전체성적평균
기존 영어 책	7월 1일	A그룹	2.9	4.3	2.7	9.9
	7월 22일	A그룹	1.6	1.7	1	4.3
EMC	7월 1일	B그룹	5.0	3.3	5.3	13.6
	7월 22일	B그룹	4.5	3.3	4.8	12.6

하는 기존 오프라인 콘텐츠의 그림 영상에 모바일 기기를 이용하여 동적인 인터랙션과 다양한 멀티미디어 정보를 융합할 수 있는 교육용 콘텐츠를 제공하였다. 오프라인 콘텐츠와 온라인 콘텐츠를 연결하기 위한 오프라인 콘텐츠의 인식 방법으로, 책의 각 페이지마다 사용해야 하는 패턴 마커 대신 오프라인 콘텐츠의 그림 영역을 이용하였으며, 모바일 기기의 저해상도 카메라에 의한 색상 검색 방법의 어려움을 해결하기 위해, 모양 기반 검색 방법을 이용하였다. 모양 정보를 표현하기 위해 체인코드를 이용하였고, 회전에 대한 문제를 해결하기 위해 시작점 일치와 차분 체인코드를, 크기에 대한 문제를 해결하기 위해 DTW를 이용하였으며, 실험에서 다양한 회전과 크기에서 높은 인식률을 보였다. 그러나 우리 시스템은 크게 세 가지 문제점이 있다. 첫째, 간단한 이진 영상만으로 그림 영역을 추출하기 때문에 빛과 같은 주변 환경에 민감하게 반응한다. 둘째, 모양 정보의 외곽선만을 이용하기 때문에 유사한 특징값이 나올 때에는 오인식이 발생한다. 셋째, PDA의 제한된 연산 자원 때문에, 이진영상, 체인코드와 같은 간단한 알고리즘을 이용하지만, 실시간으로 수행하는데 어려움이 있다. 이런 문제점들을 해결하기 위해 복잡한 알고리즘을 이용하여 그림 영역을 정확하게 찾고, 모양 정보의 외곽선 외 추가적인 정보를 이용하여 더욱 정확하게 인식하며, 실시간으로 수행하기 위해 Client/Server 구조를 이용할 것이다.

Acknowledgement

"본 연구는 숭실대학교 교내연구비 지원으로 이루어졌음"

참 고 문 헌

[1] 한은정, 정기철, 임충재, "멀티미디어를 이용한 e-learning 자기주도적 학습 콘텐츠 개발", 정보처리학회, 춘계학술발표대회 논문집, 제12권, 제1호, pp. 1019-1022, 2005.

[2] ERICSSON, "From e-learning to m-learning," http://learning.ericsson.net/leonardo, 2002.

[3] Eija Kaasinen, "User needs for location-aware mobile services," *Personal and Ubiquitous Computing*, Vol.7, Issue 1, pp. 70-79, 2003.

[4] 이재석, "영어 어휘 학습을 위한 모바일 콘텐츠의 설계 및 구현," 대구가톨릭대학교 석사학위논문, 2003.

[5] 이영석, 조성원, 최병욱, "모바일 영어 학습을 위한 지능형 교육 시스템의 설계 및 구현", 정보처리학회 논문지, 제10권, 제5호, pp. 539-550, 2003.

[6] M. Sharples, D. Corlett, and O. Westmancott, "The Design and Implementation of a Mobile Learning Resource," *Personal and Ubiquitous Computing*, Vol. 6, Issue 3, pp. 220-234, 2002.

[7] http://www.contentsmedia.com/

[8] M. Billinghurst, H. Karo, and I. Poupyrev, "The MagicBook: A Transitional AR Interface," *Computer Graphics*, pp. 745-753. 2001.

[9] 이민경, 우운택, "u-learning 환경을 위한 요리 강습 시스템," 한국 HCI 학회 논문집, 제1권, 제1호, pp. 585-589, 2004.

[10] 대교 교육 서비스 기법, http://www.daekyo.co.kr.

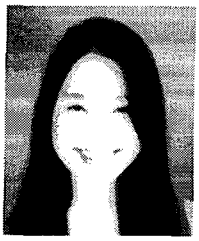
[11] Y. S. Chen, T. C. Kao, G. J. Yu, and J. P. Sheu, "A Mobile Butterfly-Watching Learning System for Supporting Independent Learning," *Proceedings of International Workshop on Wireless and Mobile Technologies in Education*, pp. 11-18, 2004.

[12] KTF, http://www.ktf.co.kr.

[13] L. G. Shapiro and G. C. Stockman, *Computer Vision*, Prentice Hall, 2000.

[14] S. Li, M. C. Lee, and D. Adjeroh, "Effective Invariant Features for Shape-based Image Retrieval," *Journal of the American Society for Information Science and Technology*, Vol. 56, Issue 7, pp. 729-740, 2005.

[15] H. Sakoe and S. Chiba, "Dynamic Programming Algorithm Optimization for Spoken Word Recognition," *IEEE Transactions on Acoustic, Speech, and Signal Processing*, Vol. 26, No. 1, pp. 43-49, 1978.



한 은 정

2003년 신라대학교 국어국문학과(인문학사)
2005년 동서대학교 소프트웨어대학원 소프트웨어학과(공학석사)
2005년~현재 숭실대학교 IT대학 미디어학부 콘텐츠공학(박사과정)
관심분야 : 모바일 콘텐츠, E-Learning, M-Learning, HCI, CBIR, 온톨로지

박 안 진

2003년 인제대학교 컴퓨터공학과(공학사)
2004년 숭실대학교 IT대학 미디어학부 미디어 공학(공학석사)
2006년~현재 숭실대학교 IT대학 미디어학부 미디어공학(박사과정)
관심분야 : 모바일 콘텐츠, 컴퓨터 비전, 인터랙티브 게임

정 기 철

1994년 경북대학교 컴퓨터공학과(공학사)
1996년 경북대학교 대학원 컴퓨터 공학과(공학석사)
1999년 Intelligent User Interfaces group at DFKI (The German Research Center for Artificial Intelligence GmbH), Germany, 방문연구원
2000년 Machine Understanding Division, Electro Technical Laboratory in Japan, 방문연구원
2000년 경북대학교 대학원 컴퓨터공학과(공학박사)
2000년~2002년 PRIP lab, Michigan State University Postdoc
2003년~현재 숭실대학교 IT대학 미디어학부 교수
관심분야 : HCI, 콘텐츠공학, 인터랙티브 게임, 영상처리/컴퓨터 비전, 증강현실, 인공지능