

SIFT 알고리즘을 이용한 혼합형 모바일 교육 시스템[†]

(Mixed Mobile Education System using SIFT Algorithm)

홍광진*, 정기철**, 한은정***, 양종열****

(Kwang-jin Hong, Kee-chul Jung, Eun-jung Han, Jong-yeol Yang)

요 약 무선 인터넷과 모바일 기기의 보급으로 언제 어디서나 원하는 정보를 얻을 수 있는 유비쿼터스 환경을 위한 인프라가 구축되면서 교육을 포함한 다양한 분야에서 오프라인과 온라인 컨텐츠를 동시에 이용함으로써 정보 전달의 효율성을 높일 수 있는 방법에 대한 연구가 활발하게 이루어지고 있다. 본 논문은 사용자에게 모바일 기기를 이용하여 오프라인과 온라인 컨텐츠를 함께 제공하여 교육의 효과를 높일 수 있는 혼합형 모바일 교육 시스템(Mixed Mobile Education System: MME)을 제안한다. 제안된 시스템은 기존의 연구와 달리 사용자에게 자연스러운 환경을 제공하기 위해서 부가적인 태그를 사용하지 않는다. 태그를 사용하는 시스템의 경우 새로운 데이터의 등록이 어렵고, 유사한 내용의 오프라인 컨텐츠 사용이 불가능하기 때문이다. 본 논문에서 우리는 저화질의 카메라를 통해 입력받은 영상에서 잡음, 색상 왜곡, 크기 및 기울기 변화에 영향을 적게 받는 특징점을 추출하고 오프라인 컨텐츠를 인식하기 위해 Scale Invariant Feature Transform(SIFT) 알고리즘을 사용하였다. 또한 클라이언트-서버 구조를 사용함으로써 모바일 장치의 제한적인 저장 능력 문제를 해결하고, 데이터의 등록 및 수정이 용이하도록 하였다. 실험을 통해 기존의 혼합형 교육 시스템과의 성능을 비교하고 제안된 시스템의 장단점을 확인하였으며, 시스템을 실생활에 적용하였을 경우 다양한 상황에서도 사용자에게 만족할만한 성능을 제공함을 확인하였다.

핵심주제어 : MME 시스템, 모바일 교육 컨텐츠, 컨텐츠 기반 정보 검색, SIFT 알고리즘

Abstract Due to popularization of the wireless Internet and mobile devices, the infrastructure of the ubiquitous environment, where users can get information whatever they want, anytime and anywhere, is created. Therefore, a variety of fields including the education studies methods for efficiency of information transmission using on-line and off-line contents. In this paper, we propose the Mixed Mobile Education system(MME) that improves educational efficiency using on-line and off-line contents on mobile devices. Because it is hard to input new data and cannot use similar off-line contents in systems used additional tags, the proposed system does not use additional tags but recognizes off-line contents as we extract feature points in the input image using the mobile camera. We use the Scale Invariant Feature Transform(SIFT) algorithm to extract feature points, which are not affected by noise, color distortion, size and rotation in the input image captured by the low resolution camera. And we use the client-server architecture for solving the limited storage size of the mobile devices and for easily registration and modification of data. Experimental results show that compared with previous work, the proposed system has some advantages and disadvantages and that the proposed system has good efficiency on various environments.

Key Words : MME system, mobile learning contents, contents-based information retrieval, SIFT algorithm

* 본 연구는 숭실대학교 교내 연구비 지원으로 이루어졌다.
** 숭실대학교 미디어학과

** 숭실대학교 미디어학과 조교수
*** University of British Columbia Postdoc
**** (주)에이티솔루션

1. 서 론

초고속 유무선 인터넷의 보급, 모바일 기기의 성능 향상 등의 정보 통신 기술 인프라 구축으로 언제, 어디서나 사용자가 원하는 정보를 개인 단말기를 통해 얻을 수 있는 유비쿼터스 환경이 보편화됨에 따라 다양한 분야에서 오프라인과 온라인 환경을 연결하기 위한 연구가 이루어지고 있다. 교육 분야에서도 기존의 오프라인 학습 자료와 멀티미디어 자료를 함께 이용함으로써 학습자에 대한 교육 효과를 높이기 위한 다양한 연구가 진행되고 있다. 특히 Smith가 제안한 블렌디드 러닝 [2]은 기존의 e-learning과 mobile learning(m-learning)의 장점을 극대화하기 위해 오프라인 교육 환경을 전통적인 교육 방법까지 포함함으로써 보다 효율적이고 개인 맞춤형 교육이 가능하도록 한다.

본 논문에서 우리는 기존의 온라인 교육 컨텐츠에 블렌디드 러닝 개념을 적용하여 오프라인과 온라인 정보를 동시에 이용하고, 모바일 기기를 사용하여 시간과 장소에 상관없이 정보를 제공받을 수 있도록 함으로써 보다 효율적인 교육이 가능한 혼합형 모바일 교육 시스템(mixed mobile education system: MME)을 제안한다. 일반적으로 오프라인과 온라인 컨텐츠의 연결을 위해 센서나 마커와 같은 부가적인 태그를 사용한다. 특정한 태그에 최적화된 패턴 인식 알고리즘을 적용할 경우 높은 인식률과 빠른 응답 속도를 제공할 수 있기 때문이다. 그러나 제안된 시스템은 사용자에게 자연스러운 학습 환경의 제공과 데이터베이스에 저장되어있는 데이터의 수정 및 사용자에 의한 새로운 데이터의 등록이 가능하도록 하기 위해서, 태그를 부착하지 않고, 카메라를 통해 입력된 영상을 분석하여 오프라인 컨텐츠를 인식하는 방법을 사용한다. 이를 위해 Scale Invariant Feature Transform(SIFT) 알고리즘 [1]을 이용하여, 낮은 화질의 카메라 입력 영상에서 발생하는 조명, 크기 및 회전 변화의 영향을 적게 받는 특징점을 추출함으로써 오브젝트 인식률을 향상시킨다. 하지만 SIFT 알고리즘은 태그를 이용하는 방법에 비해 상당한 계산량을 필요로 하며 모바일 기기에서 사용되는 CPU가 처리하기에는 많은 무리가 따른

다. 또한 데이터베이스에 저장된 정보에 대한 자유로운 수정 및 추가 역시 모바일 기기에 내장된 메모리 용량으로는 큰 효과를 얻을 수 없다. 따라서 제안된 시스템은 클라이언트-서버 구조를 적용하여 이와 같은 문제를 해결한다. 실험 결과를 통해 우리는 제안된 MME 시스템이 카메라 입력 영상에 대한 높은 인식률을 보이며, 시간과 공간의 제약 없이 학습자가 원하는 정보를 정확하게 제공해줄 수 있음을 확인하였다.

논문의 구성은 다음과 같다. 2장은 기존의 오프라인과 온라인 컨텐츠를 활용하여 효율적인 교육이 가능하도록 하기 위한 연구들의 장단점을 설명하고, 3장은 MME 시스템의 구성과 각 모듈의 기능을 설명한다. 4장은 기존의 혼합형 교육 시스템과 다양한 환경에서의 입력 영상에 대한 인식 성능을 비교하고, 5장에서 결론과 향후 연구 방향에 대해 이야기한다.

2. 관련 연구

기존의 오프라인 컨텐츠는 포함하는 정보가 정적이라는 특징을 가지며 컨텐츠에 대한 집중도가 높다는 장점을 가지고 있다. 그러나 해당 컨텐츠에 대한 보다 많은 정보를 얻기 위해서는 다른 다수의 컨텐츠 또는 미리 지식을 습득한 교육자를 통한 학습이 필요하다.

반면에 온라인 컨텐츠의 경우 포함하는 정보가 동적이고 학습자 개개인의 수준에 맞는 교육이 가능하기 때문에 오프라인 컨텐츠에 비해 효율적인 정보 전달이 가능하다는 장점을 가지고 있다.

이처럼 오프라인과 온라인 컨텐츠는 서로의 단점을 보완하여 보다 효율적인 정보 전달이 가능하도록 한다. 따라서 다양한 분야에서 오프라인과 온라인 컨텐츠 모두의 장점을 활용하기 위해 다양한 시스템과 컨텐츠가 연구되고 있다. 특히 교육 분야의 경우 e-learning, m-learning 등과 같은 온라인 교육 컨텐츠 개발을 통해 전통적인 교육 방법의 단점을 보완하는 연구를 진행하면서 오프라인과 온라인 컨텐츠의 통합에 대한 연구가 활발하게 진행되고 있다(표 1).

<표 1> 온라인/오프라인 컨텐츠의 장점을 활용한 새로운 교육 환경에 대한 연구

교육 기법	사용 컨텐츠	온라인-오프라인 컨텐츠 연결 방법	데이터 추가 및 수정 가능 여부	저자[논문]
m-learning	온라인 형태로 변환된 오프라인 컨텐츠	오프라인 컨텐츠 사용안함	불가능	Ericsson [4]
				이재석 [5]
				이영석 등 [6]
blended learning	온라인 컨텐츠와 오프라인 컨텐츠 모두 사용	태그 사용	불가능	M. Sharples et al. [7]
				M. Billinghurst et al. [8]
		영상처리 기법 사용	가능	이민경 등 [9] 한은정 등 [3] L. Paletta et al. [10]

기존의 e-learning은 오프라인 컨텐츠를 온라인의 형태로 바꾸어 학습자가 방대한 양의 정보를 쉽게 접근할 수 있도록 한다는 장점을 가진다. 그러나 학습을 하기 위해 교육 컨텐츠에 접근할 수 있는 컴퓨터를 이용하기 때문에 컨텐츠에 접근할 수 있는 장소가 제한적이고, 일방적으로 지식을 제공받는다는 단점이 있다.

m-learning [4]은 이러한 단점을 해결하기 위해 모바일 기기를 적용하여 시간과 장소에 상관없이 교육 컨텐츠에 접근할 수 있도록 ERICSSON사가 제안한 교육 개념이다. 그러나 ERICSSON사는 단순히 컨텐츠의 접근성에 대한 해결책만 제시하고 있다. 이재석이 제안한 영어 어휘 학습 시스템 [5], 이영석 등이 제안한 지능형 교육시스템 [6], Sharples 등이 제안한 자기주도적 학습 컨텐츠 [7]는 m-learning 방식에 학습자 개개인의 수준을 판단하여 적절한 컨텐츠를 제공할 수 있는 지능형 학습 컨텐츠를 적용함으로써 다른 사람의 도움 없이 학습자가 스스로 학습할 수 있도록 한다. 그러나 m-learning 방식의 학습 컨텐츠는 온라인 형태의 컨텐츠만 사용하기 때문에 기존의 오프라인 컨텐츠를 변환하기 위한 비용이 많이 들고, 컨텐츠 제작자에 의해 미리 저장된 정보만 습득할 수 있다는 단점이 있다.

이러한 단점을 보완하기 위하여 온라인과 오프라인을 통합한 교육 환경을 제공하는 '블렌디드 러닝 [2]'이라는 개념이 제안되었다. 특히 블렌디드 러닝에서의 오프라인 교육 환경은 단순히 컨텐츠의 종류에 국한되지 않고 교육자가 학습자에게 직접 지식을 전달하는 전통적인 교육 방법까지 포함한다. 따라서 보다 효율적이고 개인의 수준에 맞는 교육이 가능하다는 장점이 있다. Billinghurst

등의 MagicBook[8], 이민경과 우운택의 AR 요리 강좌 교재 [9], 한은정 등의 실감형 교육 컨텐츠 (Augmented Learning Contents: ALC) [3], Paletta 등의 Ambient Learning 시스템 [10]은 온라인/오프라인 통합 환경을 통해 학습자에게 보다 다양한 정보를 제공한다. MagicBook과 AR 요리 강좌 교재의 경우 학습자가 Head-Mounted Display(HMD)를 착용하고, 오프라인 책에 연결된 AR 객체를 함께 보여주고, ALC와 Ambient Learning 시스템의 경우 PDA를 사용하여 가상 컨텐츠와 실제 환경을 연결함으로써 보다 쉽고 효율적으로 정보를 제공받도록 한다. 따라서 이들 컨텐츠는 교육 효과를 높이는 것이 가능하고, 간단한 정보는 오프라인 컨텐츠를, 상세한 설명이 필요한 정보는 온라인 컨텐츠를 사용함으로써 컨텐츠 개발 비용을 줄일 수 있다는 장점이 있다.

블렌디드 러닝 환경은 카메라를 통해 실제의 오프라인 컨텐츠를 입력받아 그에 해당하는 온라인 정보를 검색하여 사용자에게 제공한다. 이때 오프라인 컨텐츠를 인식하는 방법에 따라 오프라인 컨텐츠에 부착된 RFID, 바코드, 마커 등의 부가적인 태그를 사용하여 인식하는 방법 [8, 9]과 영상처리 기법을 사용하여 카메라 입력 영상을 분석하고 인식하는 방법 [3]으로 구분할 수 있다. 태그를 사용하는 MagicBook과 AR 요리 강좌의 경우, 오프라인 컨텐츠에 대한 인식 속도가 빠르고 인식률이 높은 반면, 정보를 제공하기 위한 새로운 오프라인 컨텐츠의 등록이 어렵다는 단점이 있다. 영상을 분석하는 방법을 사용하는 ALC와 Ambient Learning 시스템은 카메라 입력 영상에서 추출한 오브젝트를 이용하기 때문에 사용자에게 자연스러운 환경을 제공하지만, 카메라 입력 영상의 화질이

떨어져 주변 조명의 밝기와 입력 영상 내의 오브젝트 크기 변화의 영향을 많이 받고, 추출된 오브젝트의 형태가 유사한 서로 다른 컨텐츠에 대한 오인식이 발생할 확률이 높다는 단점이 있다.

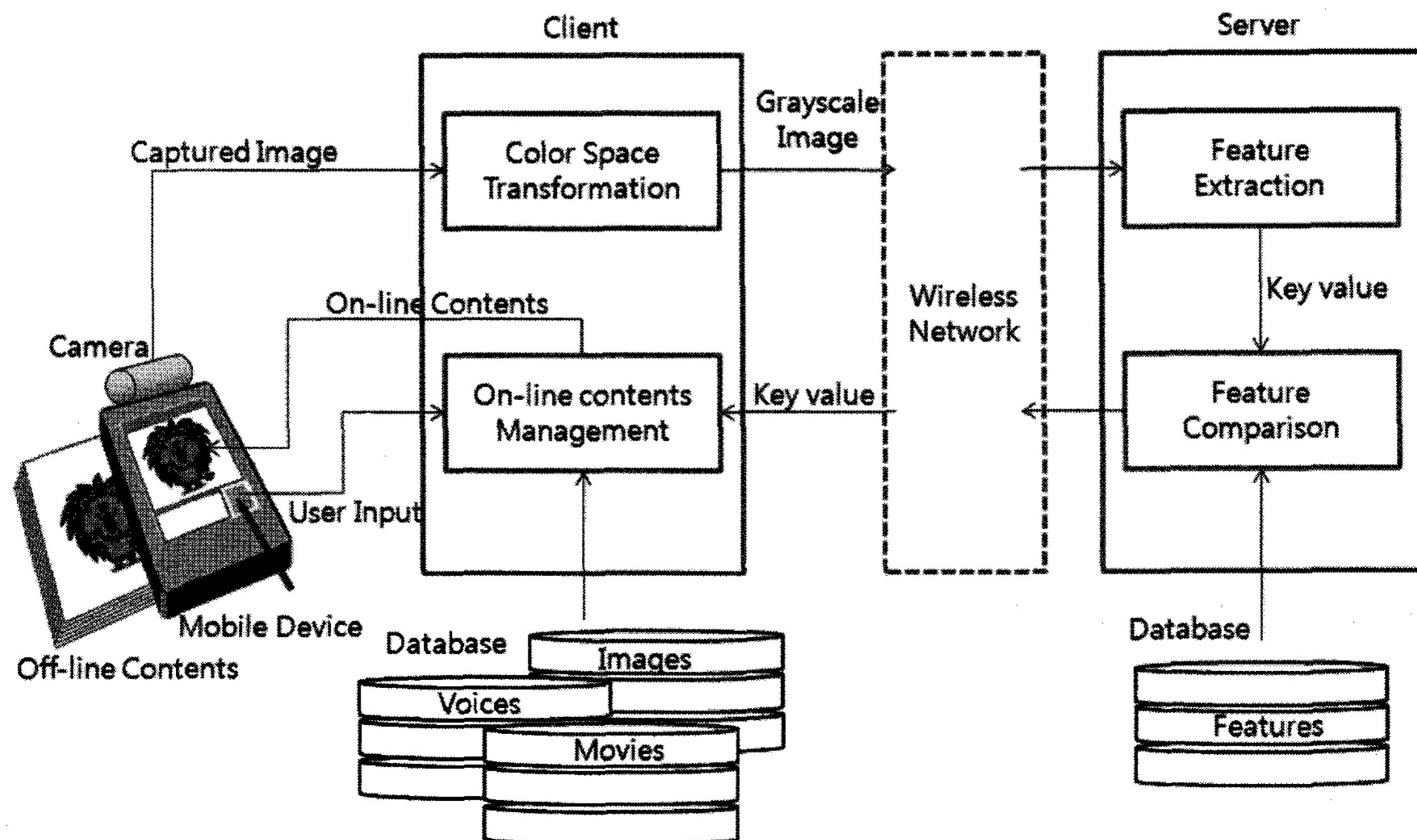
3. MME 시스템 구성과 SIFT 알고리즘을 이용한 컨텐츠 인식 과정

MME는 기존의 교육용 오프라인 컨텐츠를 기반으로, 모바일 장치를 통해 학습자와 온라인 컨텐츠 사이의 상호작용이 가능도록 함으로써 학습 효과를 높이는 시스템이다. 기존의 e-learning, m-learning, 블렌디드 러닝을 기반으로 제작된 컨텐츠들은 하나의 플랫폼에서 모든 과정을 수행하는 단일 구조로 이루어져 있다. 그러나 컨텐츠의 플랫폼이 모바일로 옮겨가면서 CPU 성능과 데이터 저장 공간의 제약으로 오프라인 컨텐츠의 인식을 위한 복잡한 연산의 수행이나, 다양한 멀티미디어 정보의 저장이 어려워지고 있다. 따라서, 제안하는 시스템은 이러한 모바일 기기의 한계를 보완하기 위해서 클라이언트-서버 구조로 이루어진다(그림1).

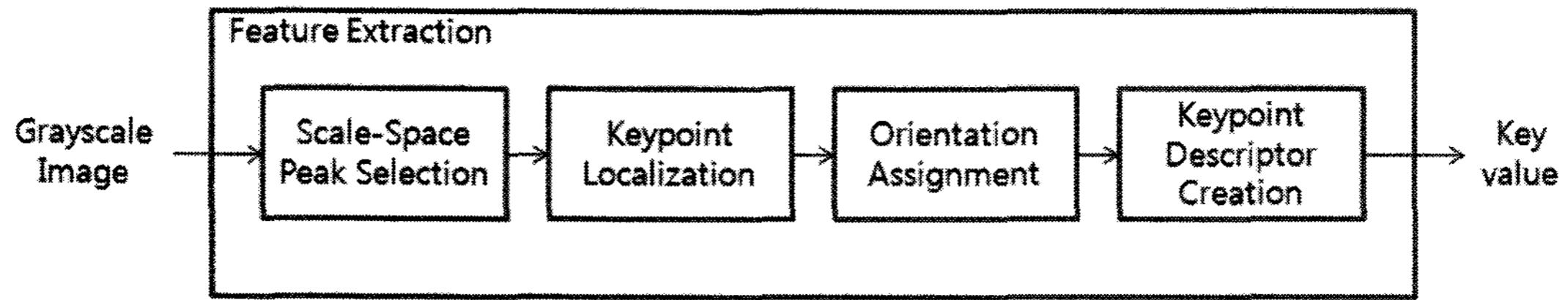
클라이언트 부분은 카메라가 부착된 모바일 기기로 구성되며, 오프라인 컨텐츠에 대한 영상 획득과 온라인 정보 출력, 사용자와 컨텐츠 사이의 상호작용과 관련된 기능을 수행한다. 서버 부분은 클라이언트에서 전송된 영상을 이용하여 오프라인 컨텐츠를 인식하고 그 결과를 클라이언트에 전송하는 기능을 한다.

3.1 오프라인 컨텐츠 획득

오프라인과 온라인 컨텐츠의 연결 및 사용자와 컨텐츠 사이의 상호작용을 위해 우리는 모바일 기기에 장착된 카메라를 이용한다. 카메라를 통해 입력받은 오프라인 컨텐츠 영상은 무선 네트워크를 이용하여 서버로 전송되는데, 이때 컨텐츠 인식에 불필요한 정보를 클라이언트에서 미리 제거하고 전송함으로써 전송 속도를 향상 시킨다. 입력 영상의 크기는 160×110 픽셀의 RGB 영상이므로 52.8($160 \times 110 \times 3$)Kbyte가 된다. 그러나 서버에서 입력 영상에 대한 특징값을 추출할 때 영상의 명암 정보만을 사용하므로 RGB 색상 공간으로 표현된 입력 영상에서 밝기(I)값만을 계산하여 전송함으로써 데이터의 양을 1/3로 줄인다(식(1)).



(그림 1) 시스템 구성도



(그림 2) 특징값 추출 단계에 대한 상세 구성도

$$I = \frac{1}{3}(R + G + B) \quad (1)$$

$$N(x, y, k\sigma) = \{d(x, y, k\sigma), 8 \neq \text{neighbors of } d(x, y, k\sigma)\} \quad (2)$$

$$E(x, y, k\sigma) = \begin{cases} d(x, y, k\sigma) & \text{if } d(x, y, k\sigma) = \max(N(x, y, k\sigma), N(x, y, k^2\sigma), N(x, y, \sigma)) \text{ or} \\ & \min(N(x, y, k\sigma), N(x, y, k^2\sigma), N(x, y, \sigma)) \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

3.2 오프라인 컨텐츠 인식

서버는 오프라인 컨텐츠의 인식을 위해 클라이언트로부터 전송된 입력 영상의 특징값을 추출한다. 그러나 모바일 기기에 장착된 카메라는 일반적인 디지털 카메라에 비해 CCD크기가 작아 동일한 화소의 이미지를 표현할 때 해상도가 낮고, 조명에 민감하여 정확한 색을 표현하는 것이 어렵다 [3]. 또한 모바일 기기의 특징인 이동성으로 인해 입력 영상 내의 오브젝트의 크기와 방향이 일정하지 않다. 따라서 우리는 오브젝트의 크기, 회전, 조명 변화의 영향을 적게 받는 SIFT 알고리즘을 사용하여 특징점을 추출하고, DB 내의 특징값들과 비교하여 입력 영상을 인식한다.

그림 2는 그림 1의 시스템 구성도 중에서 특징점 추출 단계에 대한 상세 구성도를 보여주며, SIFT 알고리즘을 이용한 특징값 추출은 다음의 4개 세부 단계로 구성된다[1].

첫 번째 단계에서는 크기 변화의 영향을 적게 받는 특징점 후보를 구하기 위해 입력 영상의 공간을 피라미드 형태의 스케일 공간으로 변환하고, Different of Gaussian(DoG)가 콘볼루션된 영상 D 내에서 지역적 극값(local extrema)을 찾는다. 지역적인 극값은 영상 D 의 한 픽셀 d 를 동일 스케일 레벨 내의 주변 픽셀과 위아래 레벨에서 d 에 대응하는 주변 픽셀 26개를 비교하여 결정된다. 각 스케일 레벨에서 결정된 극값으로 구성된 영상 E

는 식(2)로 표현되는 9개 픽셀 N 과 N 에 대응하는 상하 스케일의 픽셀을 통해 식(3)으로 표현된다. 이때 영상 D 는 식(4)로 표현되며, 사용되는 G , I , L 은 각각 가우시안 함수와 입력 영상(식(5)), 가우시안이 콘볼루션된 영상(식(6))을 의미한다. 또한, x , y 는 영상 내의 좌표를, σ 는 가우시안 함수의 분산을, k 는 스케일을 의미한다.

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \quad (4)$$

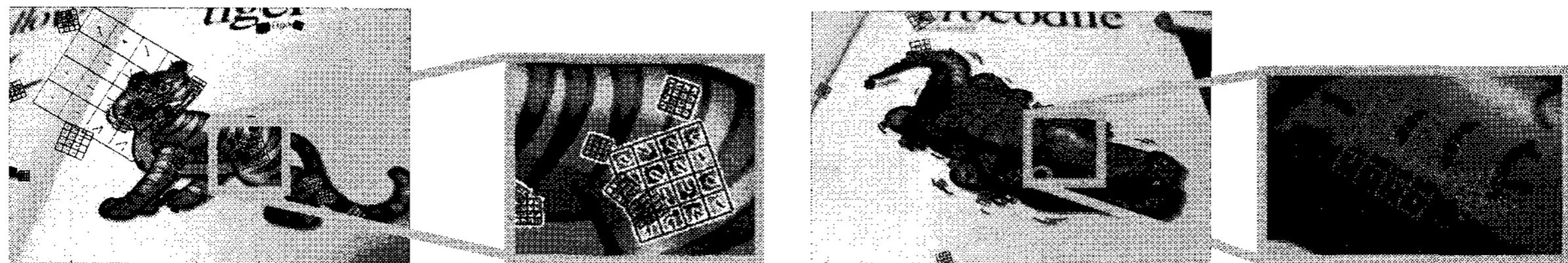
$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (5)$$

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}} \quad (6)$$

두 번째 단계에서는 이전 단계에서 구해진 극값 중에 큰 의미를 가지지 못하는 후보들을 제외시키고 영상 내에서 실제 특징값으로 사용될 후보를 결정한다. 첫 번째 단계에서 구해진 픽셀의 밝기값과 픽셀이 구해진 영역의 모서리 영역(corner-region) 여부를 이용하여, 잡음을 제거하고 모서리 영역일 경우 해당 픽셀을 최종 특징점 후보로 결정한다. 세 번째 단계에서는 회전 변화의 영향을 적게 받는 특징점 추출을 위해서 결정된 특징점의 주방향을 계산한다. 피라미드의 각 단계에서 결정된 특징점의 주방향을 계산하기 위해서, 각각의 특징점을 중심으로 일정한 크기의 윈도우(16×16)를 씌우고 윈도우 내의 각 픽셀의 크기와 방향을 계산한다(식(7)).

$$m(x,y) = \sqrt{(L(x+1,y) - L(x-1,y))^2 + (L(x,y+1) - L(x,y-1))^2}$$

$$\theta(x,y) = \tan^{-1} \frac{L(x,y+1) - L(x,y-1)}{L(x+1,y) - L(x-1,y)}$$
(7)



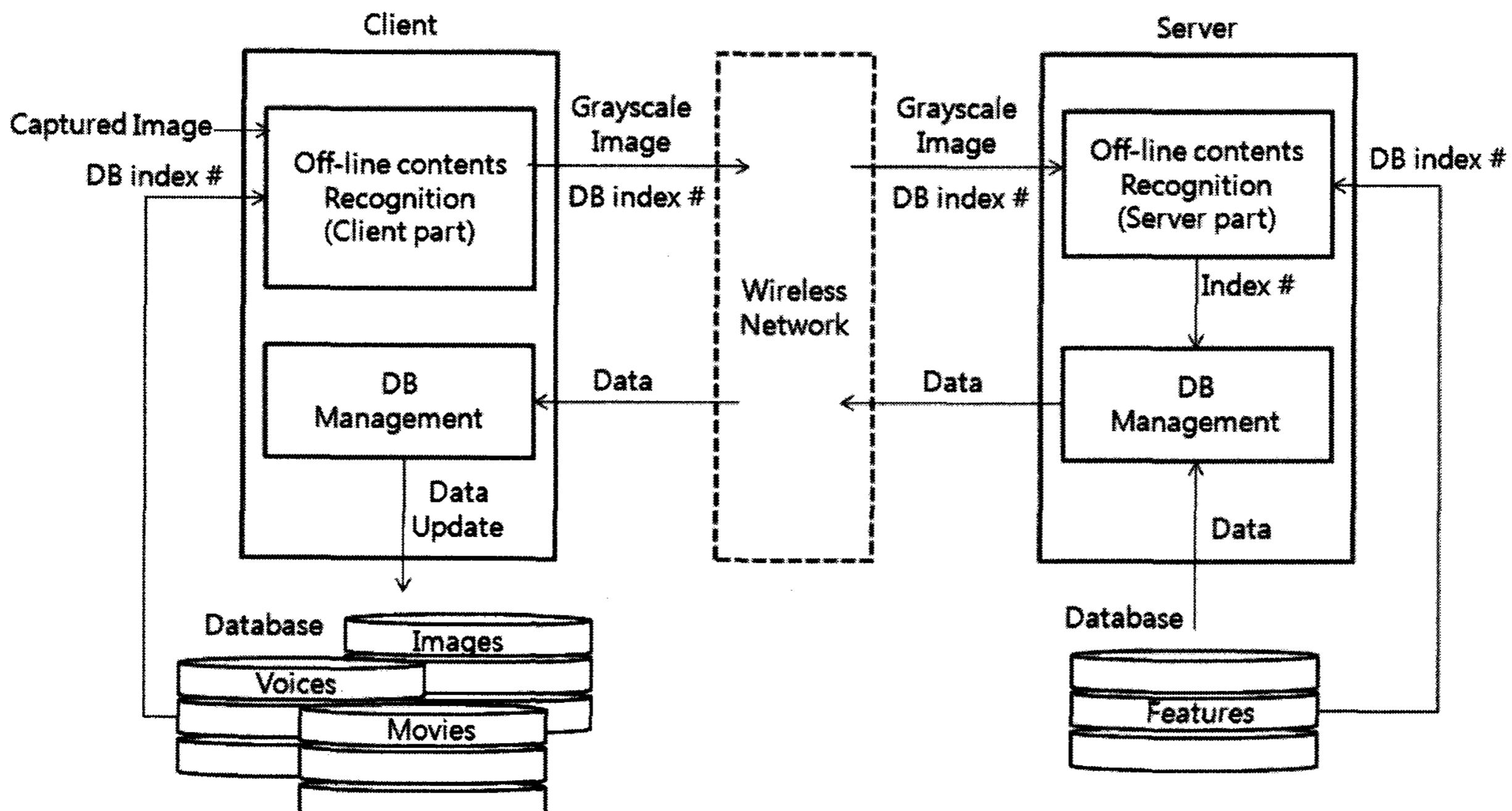
(그림 3) SIFT 알고리즘을 이용한 특징점 추출 결과.

최종적으로 하나의 특징점이 128차원의 벡터($4 \times 4 \times 8$)의 서술자로 표현되고 영상 인식에 사용된다(그림 3). 이렇게 구해진 특징값은 DB안에 저장되어 있는 영상들의 특징값들과의 매칭을 통해 입력 영상이 무엇인지를 인식하게 된다. 특징값의 매칭은 Square distance를 사용하며, 잘못된 매칭을 막기 위해 distance가 가장 적은 특징값과 두 번째로 적은 특징값 사이의 비율이 0.8 보다 작은 경우에 DB 내의 특징값과 매칭된다고 결정하게 된다.

3. 3 데이터 관리

서버에서 전송된 결과값을 이용하여 클라이언트는 일치하는 DB 내의 음성, 동영상, 사진으로 구성된 온라인 정보를 화면에 보여주고, 학습자는 원하는 형태의 정보를 선택할 수 있다.

본 논문에서 제안하는 MME 시스템은 서버와 클라이언트에 각각 DB를 가지고 있다. 제안하는 시스템의 가장 큰 장점은 교육 서비스를 제공하기 위한 오프라인 컨텐츠와 관련된 모든 온라인 컨텐-



(그림 4) 오프라인 컨텐츠 인식을 통한 클라이언트와 서버 사이의 데이터 동기화

츠를 서버에 저장하고, 필요한 컨텐츠를 모바일 기기에 다운받아 사용함으로써 모바일 기기의 저장 공간의 한계를 해결하고 온라인 컨텐츠의 업데이트가 용이하다는 것이다. 또한, 서버는 오프라인 컨텐츠 인식을 위한 특징값을 저장함으로써 DB의 크기를 줄이고, 검색의 효율을 높일 수 있다는 장점을 가진다.

그림 4는 클라이언트와 서버 사이의 데이터 동기화를 위한 과정을 보여준다. 사용자가 학습을 시작하기에 앞서 모바일 카메라를 통해 입력받은 오프라인 컨텐츠의 표지 영상을 서버로 전송하면, 서버는 오프라인 컨텐츠 인식 결과와 해당하는 온라인 컨텐츠의 업데이트 여부를 클라이언트로 전송한다. 클라이언트는 DB의 온라인 컨텐츠 인덱스 번호가 전송받은 인식 결과와 일치하는지를 확인한 뒤, 일치하지 않거나, 인덱스 번호는 일치하지만 서버의 컨텐츠가 업데이트되어 내용이 바뀐 경우 서버에 DB 업데이트를 요청하여 새로운 컨텐츠를 다운받는다.

4. 실험 및 결과

서버는 듀얼로 구성된 AMD Opteron 246 CPU 와 1Gbyte RAM, NVIDIA GeForce 5750으로 구성되어있다. 실험에 사용된 시스템은 Microsoft Visual C++ 6.0과 Microsoft eMbedded Visual C++ 4.0, PDA 카메라 SDK를 사용하여 구현되었다. 제안된 MME 시스템은 ALC 시스템과 동일한 컨텐츠를 사용하기 때문에, 본 논문에서는 제공하는 컨텐츠에 대한 학습 효과 분석에 관련한 실험은 실시하지 않고, SIFT 알고리즘을 사용함으로써 얻게되는 효과를 중심으로 실험하였다. 따라서 우리는 클라이언트-서버 구조 내에서 SIFT 알고리즘을 사용하는 MME 시스템과 기존의 모바일 기

기 상에서 chain-code 알고리즘 [11]을 사용하는 ALC 시스템의 정확도와 수행 속도 비교를 통해, 제안된 MME 시스템의 효용성을 평가한다.

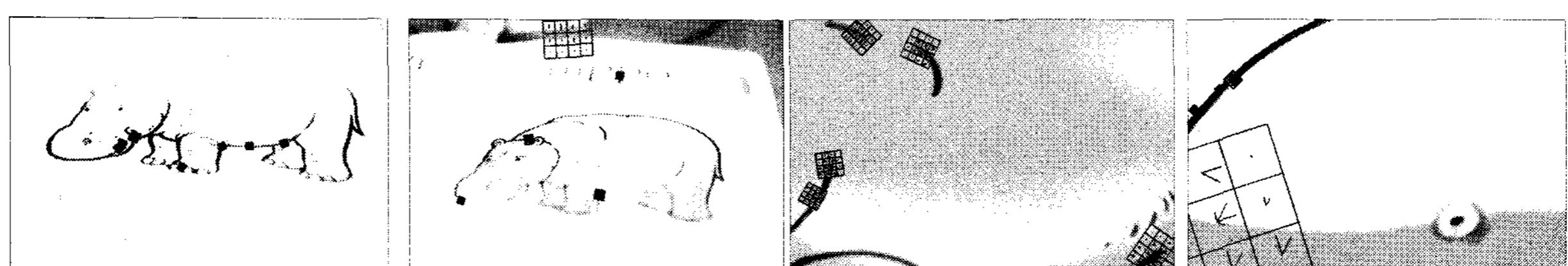
4.1 수행결과 및 분석

2장에서 설명한 것처럼, SIFT 알고리즘은 영상을 크기 공간으로 변환하여 표현하고 각 크기 단계 별로 지역적인 극한값을 추출하기 때문에 크기 변화의 영향을 적게 받고, 추출된 극한값은 주방향 계산을 통해 특징점 주변의 픽셀 방향을 수정하기 때문에 카메라 수평 방향의 회전에 영향을 적게 받는다. 이러한 장점 때문에 SIFT 알고리즘은 부가적인 태그를 사용하지 않는 모바일 학습 시스템에 매우 적합하다고 할 수 있다.

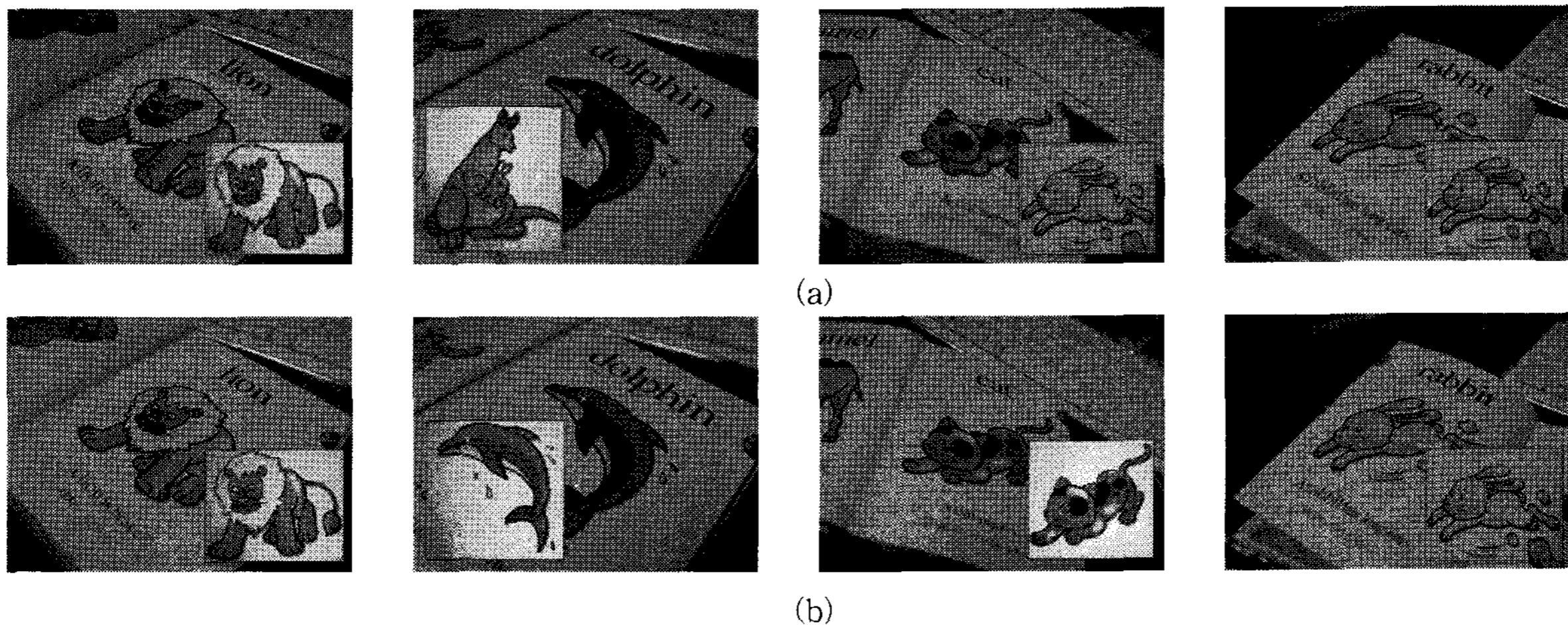
그러나 지나치게 밝은 조명 아래에서나 지역적인 극한값이 존재하지 않는 오브젝트의 일부분만을 입력 받을 경우 오브젝트 인식을 위한 충분한 개수의 오브젝트 특징점 추출이 어렵고, 이로 인해 오인식이 발생한다(그림 5). 그림 5에서 첫 번째와 두 번째 그림은 동일한 오브젝트에 지나치게 밝은 조명이 비춰졌을 경우의 특징점 추출 결과이고, 세 번째와 네 번째 그림은 오브젝트의 일부분을 촬영했을 때 촬영된 부분이 지역적 극한값이 존재할만한 텍스쳐를 가지지 않아서 비교를 위한 충분한 개수의 특징점을 추출하지 못한 경우를 보여준다.

4.2 효용성 평가

제안된 MME 시스템과 같은 블렌디드 러닝 기반의 학습 시스템, 특히 부가적인 태그를 사용하지 않는 시스템에서는 카메라의 위치가 일정하지 않기 때문에 동일한 오브젝트의 회전, 크기 및 조명에 의한 밝기 변화에 대해 영향을 적게 받는 영상 인식 알고리즘을 필요로 한다.



(그림 5) 오인식의 예.



(그림 6) 인식 결과 비교: (a) ALC 시스템, (b) MME 시스템.

그림 6은 각각 ALC 시스템과 MME 시스템에서 다양한 입력 영상(큰 영상)과 인식한 결과(사각형 안의 작은 영상)를 보여준다. 동일한 입력 영상에 대해 MME 시스템은 대부분 정확한 결과를 보인 반면, ALC 시스템은 유사한 형태를 가지는 입력 영상에 대해 부정확한 결과를 보이는 것을 확인할 수 있다.

그림 6의 결과를 바탕으로 오프라인 컨텐츠 인식에 영향을 미치는 조건을 이미지 크기와 회전으로 가정하여 실험하였다. 실험 중 조명의 밝기를 제외한 이유는 그림 5와 같이 오인식이 발생하는 극한 상황에서는 실제 학습 역시 불가능하기 때문이다.

그림 7은 입력 영상 내 오프라인 컨텐츠의 다양한 크기 변화에 대한 인식 결과이다. MME 시스템은 크기의 변화에 큰 영향을 받지 않고 정확한 인식 결과를 보였으나 ALC 시스템은 오프라인 컨텐츠가 일정 크기(40%) 이하로 줄어들 경우 인식률이 급격히 낮아진다. 본 실험에서 오프라인 컨텐츠의 크기는 B6 사이즈($130\text{mm} \times 185\text{mm}$)를 기준으로 하였으며 그림 7의 결과는 제안된 시스템이 일반적으로 사용되는 오프라인 컨텐츠를 등록하고 온라인 컨텐츠를 연결하여 사용할 수 있음을 의미한다고 할 수 있다.

오프라인 컨텐츠의 회전 변화에 대한 비교는 ALC 시스템과 MME 시스템 모두 만족할 만한 결과를 보였다. 회전 변화 비교는 오브젝트의 크기를 변화시키지 않고 오브젝트에 대해 촬영 방향을 달-

리하여 입력된 영상을 이용하여 비교하였기 때문에 검색 결과가 대부분 정확하게 나타난 것으로 보인다(그림 8(a)). 반면에 그림 8(b)에서 보는 것처럼 카메라 수직 방향의 회전에는 결과의 차이가 있음을 볼 수 있는데, 그 이유는 카메라 수평 방향의 회전과 달리 카메라 수직 방향의 회전은 오브젝트의 형태를 변화시키기 때문이다. 그림 8의 결과는 그림 7과 달리 실제 모바일 기기를 사용함에 있어서 가장 많이 발생하는 상황을 실험한 것으로 써, 일반적으로 모바일 카메라를 통해 영상을 촬영할 경우 대부분 카메라 수직방향으로 일정 각도 이상 기울어지게 된다. 그림 8(b)에서 45도를 기준으로 두 시스템 모두 정확도가 급격하게 떨어짐을 볼 수 있으나 45도 이상의 기울기에서는 시스템을 통한 학습이 어려운 환경이다. 따라서 제안된 시스템이 기존 시스템에 비해 실생활에서 보다 효율적으로 사용 될 수 있음을 확인할 수 있다.

표 2는 시스템 각 단계별로 걸리는 시간과 전체 수행 시간 및 오브젝트에 대한 인식률을 비교한 결과이다¹⁾.

1) 표 2는 시스템의 실제 사용 가능한 범위인 인식 대상의 크기변화 60% 이하, x-z 평면에 대한 회전 변화 45도 이내에서 30개 오브젝트를 대상으로 각각 100회씩 실험한 결과들의 평균으로 얻어졌다.

<표 2> 각 단계별 수행 시간과 인식률 비교

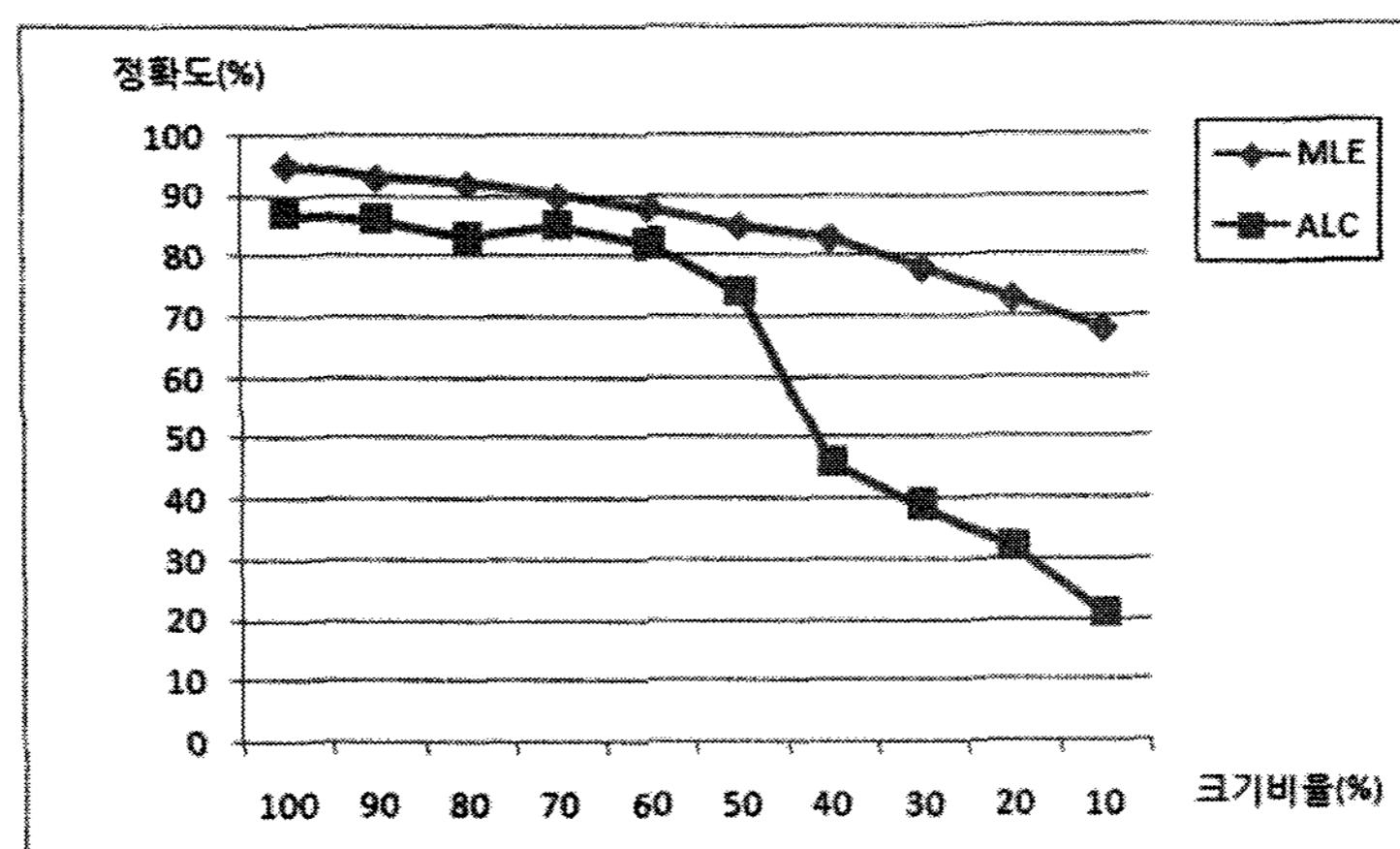
구분	수행시간(ms)				인식률 (%)
	물체 추출	특징값 추출	물체 인식	총수행 시간	
MME	800	178	380	1358	94
ALC	100	145	513	758	84

비교 결과 MME 시스템이 속도는 더 느리나 인식률은 더 높다는 것을 알 수 있는데, 그 이유는 ALC와 달리 클라이언트-서버 구조를 사용하기 때문에 PDA 카메라를 이용하여 입력받은 영상을 서버로 전송하는 시간이 필요하나, 서버 컴퓨터의 높은 계산 능력을 바탕으로 크기, 회전, 색상 왜곡에 강한 SIFT 알고리즘을 사용하였기 때문이다²⁾.

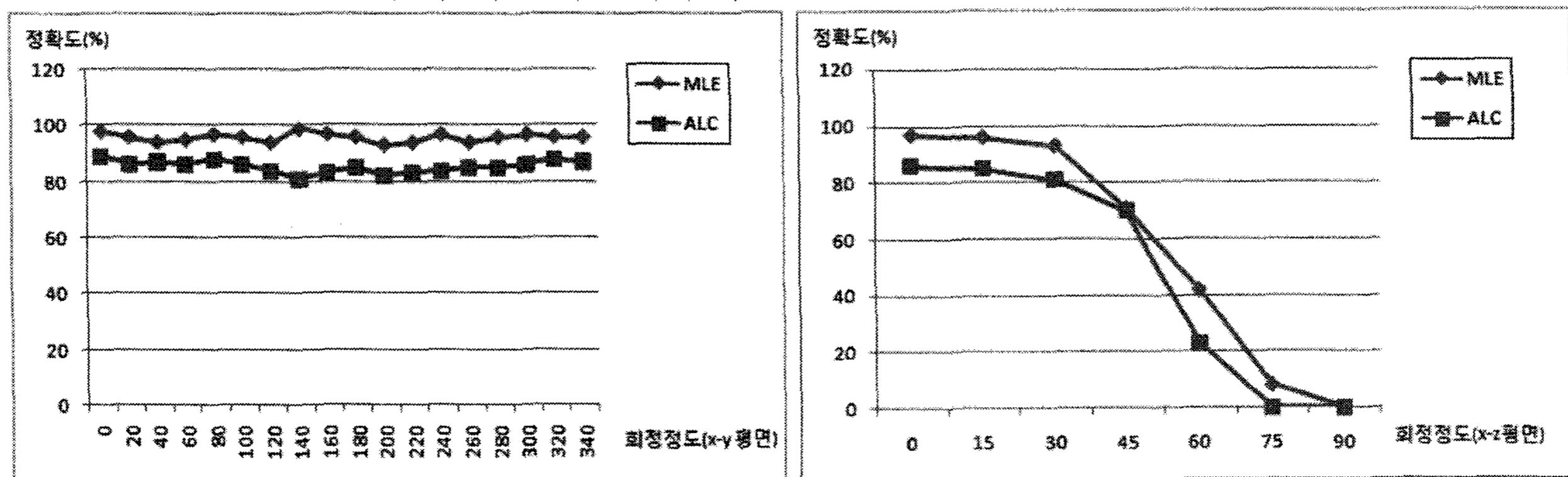
또한 클라이언트와 서버 사이의 데이터 전송 시간은 최근 점진적으로 보급되고 있는 802.11g 규격(초당 6.8Mbyte 전송)을 사용할 경우 이론적으로 약 5배의 이득을 얻을 수 있다.

5. 결 론

본 논문에서 우리는 오프라인 컨텐츠와 온라인 컨텐츠를 동시에 사용할 수 있는 혼합형 모바일 교육 시스템인 MME 시스템을 제안하였다. 제안된 시스템은 사용자에게 기존의 오프라인 교육 환경과 유사한 자연스러운 교육 환경을 제공하기 위해 오프라인 컨텐츠에 부가적인 태그를 부착하지 않고 모바일 장치에 장착된 카메라를 이용하여 오



(그림 7) 크기 변화에 따른 오프라인 컨텐츠 인식 결과.



(그림 8) 회전 변화에 따른 오프라인 컨텐츠 인식 결과: (a)카메라 수평 방향, (b)카메라 수직 방향.

2) 실험에 사용된 무선 네트워크 환경은 IEEE 802.11b 규격을 지원하기 때문에 전송속도는 11Mbps이고, 초당 1.4Mbyte의

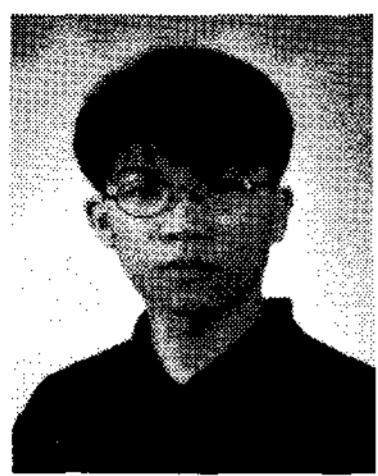
데이터 전송이 가능하다. MME 시스템은 입력 영상의 크기가 160×110 이고, 실제 데이터 전송양은 약 500Kbyte이다.

프라인과 온라인 컨텐츠를 통합하는 환경을 제공하였다. 또한 기존의 모바일 카메라를 이용하는 혼합형 컨텐츠의 단점인 다양한 왜곡을 가지는 입력 영상에 대한 부정확한 인식과 모바일 기기의 낮은 연산 능력 및 메모리의 한계로 인한 복잡한 알고리즘 사용의 어려움 등을 해결하기 위해 크기 및 회전 변화의 영향을 적게 받는 SIFT 알고리즘을 이용하고, 무선 네트워크 기반의 클라이언트/서버 구조를 이용하였다. 실험을 통해 우리는 제안된 시스템이 기존의 시스템에 비해 조명 변화의 영향을 적게 받고, 동일한 오프라인 컨텐츠의 크기가 변화하거나 카메라로 촬영한 위치가 바뀌어도 높은 인식률을 보임을 확인하였다. 또한 오프라인 컨텐츠에 대한 온라인 컨텐츠를 서버를 통해 전송받음으로써 모바일 기기의 DB를 최신의 상태로 유지할 수 있고, 모바일 기기의 제한된 메모리 크기 내에서 효율적으로 정보를 제공할 수 있음을 확인할 수 있었다.

현재 우리는 낮은 속도의 무선 네트워크 환경에서도 실시간으로 온라인 정보를 제공하기 위해 클라이언트와 서버 간의 작업 분배와 전송 데이터의 양을 줄이기 위한 연구를 수행 중에 있다. 특히 SIFT 알고리즘을 이용한 특징값 추출 결과의 차원을 낮추면서 각 특징값의 성질은 그대로 유지하기 위한 연구를 지속적으로 수행함으로써 다양한 분야에 개선된 알고리즘을 적용할 것이다.

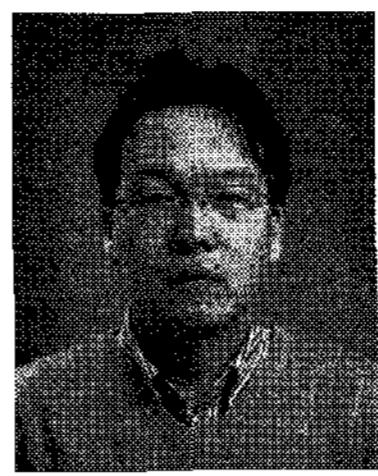
참 고 문 헌

- [1] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision, Vol. 60, No. 2, pp. 91-110, 2004.
- [2] P. J. Smith, "Modern learning methods: rhetoric and reality," Personnel Review, Vol. 31, No. 1, pp. 103-113, 2001.
- [3] 한은정, 박안진, 정기철, "모바일 교육 시스템을 위한 효율적인 영상 검색 구축", 멀티미디어학회 논문지, 제9권 제5호, pp. 101-113, 2006.
- [4] ERICSSON, From e-learning to m-learning, <http://learning.ericsson.net/leonardo>, 2002.
- [5] 이재석, "영어 어휘 학습을 위한 모바일 컨텐츠의 설계 및 구현", 대구가톨릭대학교 석사학위논문, 2003.
- [6] 이영석, 조성원, 최병욱, "모바일 영어 학습을 위한 지능형 교육 시스템의 설계 및 구현", 정보처리학회 논문지, 제10권, 제5호, pp. 539-550, 2003.
- [7] 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.
- [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] L. Paletta, C. Seifert, P. Luley and A. Almer, "Mobile Vision for Ambient Learning Mobile Vision for Ambient Learning in Urban Environments," 3rd European conference on mobile learning, 2004.
- [11] Z. Razak, O. A. Rahim, M. Y. Idna, N. M. Noor and M. Yaacob, "VHDL Implementation of JAWI Character Recognition via Chain Code Algorithm," 3rd International Symposium on Multispectral Image Processing and Pattern Recognition, Vol. 5286, pp. 457-460, 2003.



홍 광 진 (Kwang-jin Hong)

- 학생회원
- 2004년 2월 : 숭실대학교 컴퓨터 학과 (공학학사)
- 2006년 2월 : 숭실대학교 미디어 학과 (공학석사)
- 2006년 3월 ~ 현재 : 숭실대학교 미디어학과 박사 과정
- 관심분야 : HCI, 증강현실, 카메라 기반 3차원 물체 모델링



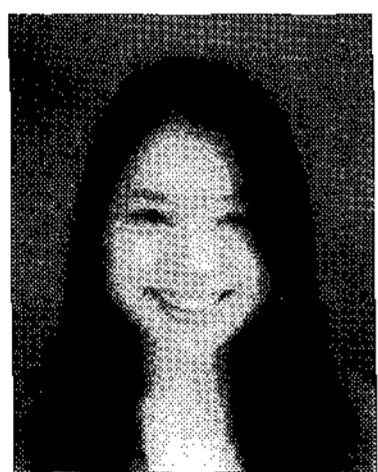
양 종 열 (Jong-yeol Yang)

- 1995년 3월~1997년 12월 : 포항 공대 정보통신연구소 연구원
- 1998년 1월~2000년 5월 : (주)다 우기술 대리
- 2000년 6월~2004년 12월 : (주) 일렉트릭아일랜드 개발이사
- 2008년 2월 : 숭실대학교 미디어학과 공학박사
- 2005년 3월~현재 : (주)에이티솔루션 재직중
- 관심분야 : HCI, 컴퓨터비전, 모바일비전 게임



정 기 철 (Kee-chul Jung)

- 정회원
- 2000년 2월 : 경북대학교 컴퓨터공학과 (공학박사)
- 2003년 3월 : 숭실대학교 미디어학과 전임강사
- 2005년 3월 ~ 현재 : 숭실대학교 미디어학과 조교수
- 관심분야 : HCI, 컴퓨터비전, 인공지능



한 은 정 (Eun-jung Han)

- 2005년 2월 : 동서대학교 소프트 웨어학과 (공학석사)
- 2007년 2월 : 숭실대학교 미디어학부 (공학박사)
- 2007년 2월-현재 : University of British Columbia Postdoc.
- 관심분야 : 모바일 콘텐츠, HCI, 온톨로지, Interactive Art and technology