

# 교육용 가상환경을 위한 이미지 선택 알고리즘 개발에 관한 연구

권수영\*, 김민영\*, 조용주\*\*, 박경신\*\*\*

\*상명대학교 컴퓨터과학부

\*\*상명대학교 디지털미디어학부

\*\*\*단국대학교 컴퓨터학부 멀티미디어전공

e-mail : [kpark@dankook.ac.kr](mailto:kpark@dankook.ac.kr)

## A Study on the Development of Image Selection Algorithm for Educational Virtual Environment

Sooyoung Kwon\*, Minyoung Kim\*, Yongjoo Cho\*\*, Kyoung Shin Park\*\*\*

\* Department of Computer Science, Sangmyung University

\*\* Department of Digital Media, Sangmyung University

\*\*\* Computer Division, Multimedia Engineering Major, Dankook University

### 요 약

본 논문에서는 교육용 가상환경에서 학습자들에게 좀 더 효과적인 학습 효과를 주기 위해 학습 중에 보고 촬영했던 사진들을 자동으로 정리해서 사용자가 교육용 가상환경에서 체험했던 학습 내용을 사진을 통해서 복습할 수 있도록 해주는 알고리즘을 소개한다. 기존의 날짜, 장소, 키워드 등의 정보를 이용하여 사진을 정리하는 알고리즘과는 달리, 본 논문에서는 사용자가 학습을 하면서 기억해야 할 중요한 내용이나 사용자의 관심도에 의해 사진 정리를 함으로써 사용자의 학습 효과를 높이는 것을 목적으로 하는 사진 정리 알고리즘을 소개한다. 이에 따라 알고리즘에서 학습적으로 중요한 사진을 뽑는 기준과 사용자의 관심도, 인지율 계산에 대해 설명하고 이 알고리즘을 기반으로 구현한 시스템을 설명한다. 또한 사용자 실험 분석을 하고 향후 연구 방향에 대해 논한다.

### 1. 서론

최근 각종 멀티미디어 기기의 대중화로 일반 사용자들이 관리하고 소유하는 디지털 콘텐츠가 풍부하게 되었다. 특히 이런 디지털 콘텐츠 중에서도 가장 많이 다루는 디지털 사진들을 관리하기 위해 사진을 찍은 날짜나 장소 등을 기준으로 사진을 정리해주는 알고리즘들이 많이 나왔다[1~4]. 이런 알고리즘들은 주로 사진을 찍은 장소나, 날짜, 사진에 있는 사람 등과 같은 관련 데이터나 키워드를 이용하여 사진을 분류하고 관리한다. 이런 알고리즘들은 궁극적으로 모아놓은 사진들 중에서 원하는 메타 정보를 이용하여 사진을 쉽게 찾을 수 있도록 하는데 그 목적을 두고 있다.

본 논문에서는 기존의 사진 정리 알고리즘들이 검색을 목적으로 하는 것과는 달리, 가상환경에서 학습 과정 중에 찍은 사진들을 메타 정보로 활용하여 정리하고 추출해서 다시 보여줌으로써 복습을 도와주는 목적으로 개발된 이미지 정리 알고리즘을 소개한다. 이런 알고리즘의 프로토타입 구현을 목적으로, 교육용 가상환경에서 사용자들이 학습 중에 직접 촬영한

디지털 사진들을 바탕으로 학습적 가치가 있는 사진들을 자동으로 선택, 정리하여 제공한다. 그리고 가상환경에서의 학습이 끝난 후 사용자에게 사진을 다시 한번 보여줌으로써 사용자의 학습 효과를 높일도록 하였다.

본 논문에서는 먼저 관련 연구를 살펴보고, 복습을 위한 이미지 선택 알고리즘에 대해 살펴본다. 그리고 알고리즘을 구현한 가상현실 시스템과 복습에 유용하도록 사진을 보여주는 인터페이스 시스템에 대해 설명한다. 그리고 이 시스템에 대한 사용자의 실험과 결과를 설명한다.

### 2. 관련 연구

일반적으로 방대한 양의 사진을 정리하는 시스템으로 PhotoCompas 가 있다[1]. 사람들은 사진을 정리할 때 생일이나 결혼과 같은 특별한 이벤트나 특정 장소를 기준으로 구분하는데, PhotoCompas 는 이런 성향을 시스템에 도입하여 사진 정리를 효과적으로 도와주고 있다. 또한 이 시스템은 사진에 포함되어 있는 시간, 위치 정보, 외부 온도와 이벤트 등과 같은 추가적인

메타 데이터를 활용하여 개인이 찍은 사진들을 자동으로 위치와 이벤트 중심의 계층 구조를 만들어 주고 그 계층구조를 활용해서 사진을 정리 할 수 있도록 해 준다.

TagMaps 는 지리정보 태그(Tag)가 포함된 사진을 사용자가 지도 위에서 쉽게 검색할 수 있도록 하는 시스템이다. 일반적으로 많은 양의 사진을 지도 위에 다 표시하기에는 지도의 크기가 제한적이다. 이를 해결하기 위해 TagMaps 는 사진의 위치 정보를 바탕으로 비슷한 지역끼리 묶는 클러스터링 (Clustering) 작업을 한다. 그 후, 각각의 클러스터 안에 있는 사진의 순위를 정하여 그 중 가장 높은 순위를 갖는 대표 사진을 추출하여 보여 준다[2].

야후 리서치 (Yahoo! Research Berkeley)에서는 대표 사진 추출법으로 온라인에서 사용자들의 피드백을 받는 방식을 사용했다. 이를 위해서 인터랙티브 웹 어플리케이션의 지도 위에 특정 지역의 랜드마크를 활용한 태그를 보여준다. 이 태그는 TF-IDF (Term Frequency Inverse Document Frequency) 방식에 따라 중요도가 결정되고 그 결과에 따라 지도 위에서 보이는 글자의 크기가 정해진다. 각각의 태그에는 그와 관련 있는 사진들이 보여지며 이들 이미지 중에서 사용자의 의견을 받아 지도 위의 특정 지역을 가장 잘 나타내는 사진을 결정한다[3].

최근엔 사진을 찍는 사람의 위치와 방향 정보를 이용하여 여러 장의 사진 중에 대표 사진을 추출한 연구도 있다[4]. 이 연구에서는 개인마다 대표 사진을 선택하는 기준이 다르기 때문에 사진의 위치 정보와 방향 정보가 그 사람의 관심을 반영한다고 가정했다. 그리고 사진을 먼저 건물 또는 풍경을 대표하는 사진으로 분류하였고, 그 분류된 사진 안에 포함된 대상물의 비율이나 방향의 정확도에 따라 점수가 부여되고 가장 높은 점수를 갖는 것이 대표 사진으로 선정된다.

### 3. 자동화된 사진 정리 요약 알고리즘

우리가 제안하는 사진 정리 요약 알고리즘은 기존의 사진 정리 알고리즘이 많은 양의 사진을 정리하고 대표가 되는 사진을 추출한다는 점에서 공통점이 있다. 그러나 이러한 연구들은 대부분 제한된 크기를 갖는 디스플레이 장치에 여러 장의 사진을 보여주는 데 따르는 어려움을 해결하거나 사용자가 자신의 사진 컬렉션을 쉽게 검색하도록 하는 데에 목적을 가지고 있다. 이에 반해 본 논문에서 제안하는 알고리즘은 가상환경에서 사용자가 관심을 가지고 촬영한 사진들 중 학습 목표를 바탕으로 교육적으로 의미 있는 내용을 포함하는 사진을 추출하여 요약해 준다.

이를 위해서 어떤 사진이 교육적으로 의미 있는 사진으로 간주할 지에 관한 기준을 다음과 같이 설정했다. 1. 가상현실 환경에서 사용자가 꼭 학습 해야 할 중요한 사물은 정해진 것이어야 하고, 그 사물에 대한 정보(예를 들어, 3 차원 위치, 크기, 정면방향)를 바탕으로 이미지를 선택 한다. 2. 교육하고자 하는 가상환경의 아이템 (즉, 학습내용)은 미리 정해져 있어야

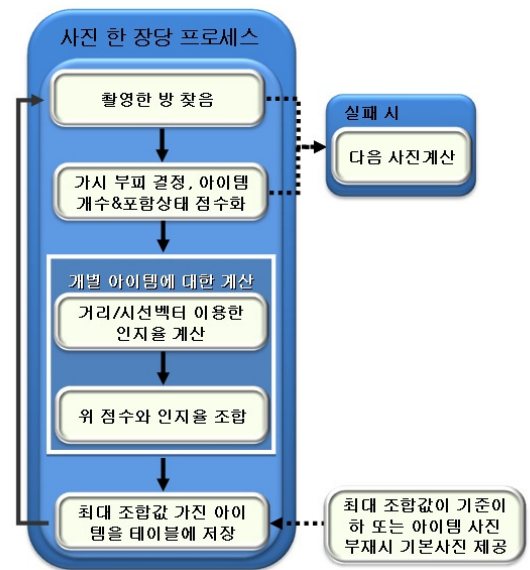


그림 1 사진 정리 요약 알고리즘 순서도

하며, 이 아이템들을 기준으로 하여 이들 위주로 촬영된 사진을 추출한다. 3. 사진 속 아이템들은 사용자가 적당한 거리에서 촬영한 것이어야 하며 그 크기가 어느 정도 식별 가능해야 한다. 4. 촬영된 사진 안에 보다 많은 아이템이 포함되어 있는 것일수록 추출될 가능성이 높다. 5. 위의 모든 조건이 충족되었다면 아이템의 모습을 가장 잘 표현할 수 있는 방향 (즉, 정면 방향에 가까운 것)에서 촬영한 사진일수록 추출될 가능성이 높다.

그림 1 은 교육용 가상환경에서 복습을 위해 자동화된 사진 정리 요약 알고리즘 순서도를 보여주고 있다. 먼저 사용자가 촬영한 위치와 바라보는 방향의 정보를 바탕으로, 근처에 있는 학습용 아이템들을 찾아낸다. 그리고 가상환경 내의 모든 아이템 크기는 존재하는 위치를 중심으로 하고 미리 지정된 크기의 반지름을 갖는 구의 형태로 간주한다. 그리고 사용자가 바라보는 방향을 중심으로 그림 2 에서 보인 것처럼 3 차원 가시 부피(View Volume)를 계산하고, 그 안에 포함된 아이템의 개수와 상태에 따라 차등적으로 점수를 부여한다.

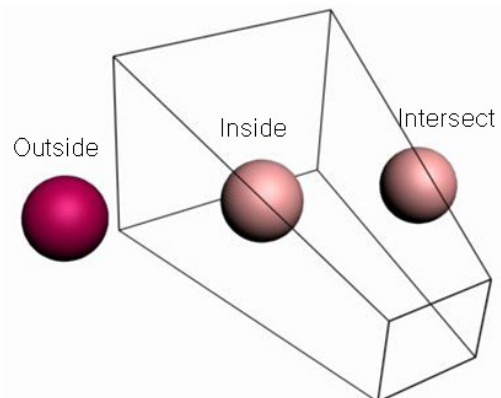


그림 2 가시 부피 내 아이템의 포함 여부 검사

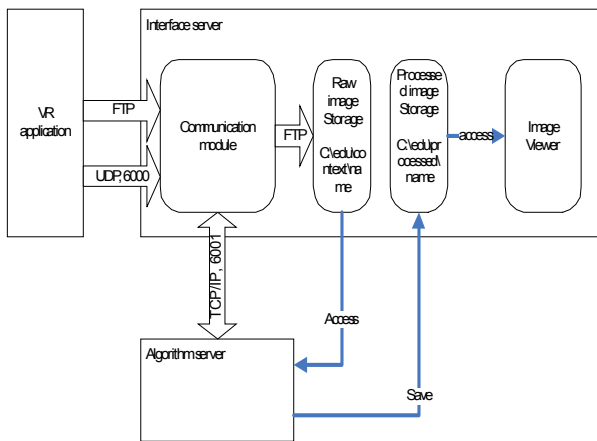


그림 1 전체 시스템 구성도

하지만 이렇게 가시 부피 안에 많은 아이템이 포함되어 있다고 해도, 아이템으로부터 지나치게 먼 곳에서 촬영하여 인지율이 낮은 아이템들이 찍힌 사진은 의미가 없기 때문에 사용자와 아이템간의 거리를 고려하여 사진 속에 있는 개별 아이템에 대한 인지율을 사용자와 아이템 간의 거리와 사용자의 시선 벡터를 사용하여 계산한다. 그리고 이렇게 계산된 인지율은 특정 비율로 가중치를 두어 조합된다.

이렇게 가시 부피 안에 많이 포함되고, 또 사용자가 아이템과 가까운 거리에서 촬영한 사진에 점수를 주게 된다면, 설사 인지율이 좋게 나오더라도, 옆이나 뒤에서 촬영한 사진들은 아이템을 제대로 보여줄 수 없기 때문에, 복습을 위한 사진으로 적절하지 못하다. 즉 복습을 위해서 가장 좋은 사진은 사용자가 가까운 거리에서 촬영하고 가시 부피 안에 모든 아이템들이 들어가 있으면서도 정면에서 찍은 것이다. 이런 사진들을 추출해낼 수 있도록 사용자의 시선 벡터를 이용하여 각 이미지에 대한 인지율을 다시 계산한다.

마지막으로 이렇게 구해진 사진들의 인지율들을 이용해서 최종 점수를 계산한다. 그리고 결과적으로 얻어진 값이 복습을 위해 사용될 수 있는 지 확인한다. 즉 학습용 가상환경을 만드는 사람들이 각 아이템 별로 찍어놓은 복습용 사진들의 값과 비교해서 그 값보다 점수가 좋은 사진들은 추출한다. 만약 제대로 찍힌 사진이 없다면 미리 찍어둔 사진으로 대체해서 보여지게 된다.

#### 4. 알고리즘을 적용한 가상환경 시스템 구성 및 동작

그림 3에 보이듯이 자동화된 사진 알고리즘을 적용한 가상환경 시스템은, 사용자가 학습을 하면서 상호작용을 할 수 있는 가상현실 시스템과 사용자가 가상현실에서 촬영한 사진을 알고리즘에 적용하여 선택 및 관리 할 수 있는 알고리즘 서버, 가상현실 시스템과 알고리즘 서버를 컨트롤하여 복습용 사진을 보여주고 복습을 할 수 있도록 도와주는 인터페이스 서버

(Interface Server, IS)로 이루어져 있다. 각 시스템은 각각 개별적으로 실행되며, 가상현실 시스템과 인터페이스 서버는 UDP 와 FTP 프로토콜을 이용해서 데이터를 주고 받고, 인터페이스 서버와 알고리즘 서버는 TCP 와 파일 공유 네트워크로 구성되어 있다.

복습을 위한 사진 정리 알고리즘을 구현하고 평가하기 위해 고구려 안악 3 호분을 가상현실로 재현한 디지털 고구려 가상환경[5]을 사용하였다. 디지털 고구려에는 사용자가 학습을 통해 알아야 하는 10 개의 아이템이 있다. 처음 학습이 시작되면, 사용자는 조이스틱 등을 이용해 고구려 안악 3 호분을 체험하면서 관심을 가진 사물이나 상황에 대해 가상 카메라로 촬영을 하도록 한다. 사용자가 가상현실을 체험 하면서 학습용 사물을 만나게 되면 해당 사물에 대한 쿼즈를 풀고 또한 그 사물을 촬영하기도 한다. 이렇게 학습이 끝나고 나면, 사용자가 촬영한 모든 사진들에 앞에서 설명한 사진 정리 알고리즘을 적용하여 복습에 필요한 사진들을 추출하게 된다. 그리고 체험 학습이 끝난 후에 추출된 사진들을 다시 한번 보여주어 복습시킴으로써 학습에 대한 효과를 높일 수 있도록 한다.

디지털 고구려 가상환경은 가상현실 저작도구인 Ygdrasil[6]로 구현되었다. 현재 가상환경은 리눅스 운영체제에서 구동되고 있으며, 알고리즘 서버와 인터페이스 서버는 윈도우 운영체제에서 동작한다. 가상환경에서 사용자가 사진을 찍게 되면, 가상현실 시스템에서는 UDP 로 인터페이스 서버에 이를 알리게 되고, 인터페이스 서버에서는 FTP 를 이용해서 가상환경으로부터 사진을 전송 받는다.

디지털 고구려는 좌측실의 왕과 왕비의 의관, 우측실의 부역 항아리, 회랑 대행렬도의 병사 등 역사학자의 자문을 받아, 고구려 안악 3 호분에 있는 벽화에서 유물로 가치있는 것들을 골라 모두 10 개의 아이템으로 배치시켰다. 그리고 사용자가 이 아이템들의 근처에 가면 내레이션으로 질문이 나오게 되고 사용자는 질문에 답을 하면서 고구려의 유물을 좀 더 자세히 살펴보고 학습할 수 있도록 꾸며져 있다. 특히 이 때 즉석 가상사진을 찍을 수 있도록 개선해서 본 연구에서 활용하였다.

알고리즘 서버에서는 인터페이스 서버가 가상환경으로부터 받은 사진들을 넘겨 받아 앞서 설명된 사진 추출 알고리즘을 적용시키고 복습을 위한 사진들을 추출한다. 그리고 사용자가 사진을 찍지 않아서 누락된 아이템들이나 사진을 찍기는 했지만 원하는 수준에 미치지 못하는 사진들은 미리 만들어두었던 복습용 사진으로 대체한다.

사용자의 학습 세션이 끝나게 되면 인터페이스 서버에 있는 최종 요약 사진 뷰어 (Final Summarized Image Viewer) 프로그램이 작동되면서 바로 복습 세션을 시작하게 된다. 지정된 시간 동안 사용자에게 복습용 사진들을 보여주고 자동으로 종료하고, 사용자 디렉토리로 이미지 파일을 이동시키고 다음 세션을 준비한다.

## 5. 실험 및 결과 분석

이렇게 계산한 비율이 사용자가 실제로 사용할 때 얼마만큼의 정확도를 가지고 적용 되는지 알아보기 위해 간단한 사용자 평가를 실시 하였다. 23~28 세 대학생 5 명을 상대로 학습용 가상현실 환경인 디지털 고구려를 사용하면서 가상 카메라를 이용하여 사용자가 관심을 보이는 사물에 대해서 촬영을 하도록 하였다. 이때 사용자가 촬영하는 사진의 개수에 대해서는 제한을 두지 않고 자유롭게 찍도록 하였다.

알고리즘의 정확도는 사용자들이 찍은 사진들과 알고리즘을 통해서 추출된 대표 사진들을 비교하는 것으로 확인하였다. 사용자들은 체험학습 시간 동안에 평균 30.8 장의 이미지를 촬영하였고, 그 중에서 중요 아이템을 찍은 사진의 비율은 약 82% 정도로 나타났다. 하지만, 알고리즘을 적용해서 선택된 사진들은 평균적으로 5.6 개 정도로 나타나서 사용자들이 별로 중요하게 생각하지 않은 부분에 대해서도 많은 사진을 찍었음을 확인할 수 있었다.

또 이렇게 알고리즘을 적용했을 때, 사용자들이 촬영한 이미지 중 중요하지 않거나, 사물이 제대로 인식되지 않은 사진 들에 대해서는 알고리즘에 의해 걸러진 것을 알 수 있었다. 특히 실험에서 사용자들이 촬영한 중요한 아이템들의 사진들은 이미 대표 사진으로 지정해놓았던 사진들과 비교했을 때 큰 차이가 나타나지 않는 것으로 보아, 알고리즘이 효율적으로 적용된 것을 알 수 있었다.

## 6. 결론 및 향후 연구 방향

본 논문에서는 교육용 가상환경에서 사용자가 상호작용적인 체험을 하고 학습 효과를 향상 시키기 위해 체험 학습 후에 복습할 수 있는 사진을 추출하는 사진 정리 알고리즘을 소개하였다. 따라서 교육용 가상환경에서 학습자들이 기억해야 할 의미있는 사진들에 대한 기준을 먼저 세우고, 이 기준을 만족시키는 사진을 추출할 수 있는 알고리즘을 구현하였다. 이렇게 구현된 알고리즘에서 쓰이는 사진의 인지율과 전반적인 알고리즘의 동작 방법을 시스템에 구현하여 실제 사용자 평가를 진행시켰다. 평가 결과 알고리즘을 통해서 실제 중요하지 않은 사진들이 많이 걸러졌음을 확인하여 이미지 정리 알고리즘의 효용성을 알 수 있었다.

본 논문에서 소개한 사진 정리 알고리즘은 아직은 실내에서 사용자의 시선 벡터나 위치 등을 파악하는 장치들이 많이 보급되어 있지 않지만, 알고리즘은 가상환경뿐만 아니라 실제 박물관 등의 실내 환경을 대상으로 적용이 가능하다. 향후 연구 방향은 좀더 다양한 교육용 가상환경 콘텐츠를 제작하고 증강현실 환경에서도 실용성 있는 사진 정리 알고리즘이 되기 위한 연구를 할 예정이다.

## 참고문헌

[1] M. Naaman, Y. J. Song, A. Paepcke, H. Garcia-Molina, "Automatic Organization for Digital Photographs with

Geographic Coordinates," Proceedings of the 4th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL), pp.53-62, 2004.

[2] A. Jaffe, M. Naaman, T. Tassa, M. Davis, "Generating Summaries and Visualization for Large Collections of Geo-Referenced Photographs," Proceedings of the 8th ACM International Workshop on Multimedia Information Retrieval (MIR), pp.89-98, 2006.

[3] S. Ahern, S. King, M. Naaman, R. Nair, "Summarization of Online Image Collections via Implicit Feedback," Proceedings of the 16th International Conference on World Wide Web, pp.1325-1326, 2007.

[4] S. Han, 'A Photo Summarization Technique based on Location and Orientation with Geographic Information', Master's Thesis, Information and Communications University, 2007.

[5] Y. Cho, K. Park, S. Park, H. Moon, "Designing Virtual Reality Reconstruction of the Koguryo Mural, Cooperative Design, Visualization and Engineering," Lecture Notes in Computer Science Vol. 3675, pp.194-201, Sep. 2005.

[6] D. Pape, 'Composing Networked Virtual Environments', Ph. D. Dissertation, University of Illinois at Chicago, 2001.