

스캐닝 된 교육 저작물에서 영상 저작물 검색 기술 연구

정다미, 최영주, 김병규*

숙명여자대학교 IT 공학과

e-mail : {dm.jeong, yj.choi}@ivpl.sookmyung.ac.kr, bg.kim@sm.ac.kr*

A Study on Image Retrieval Algorithm for Scanned Education Books

Da-Mi Jeong, Young-Ju Choi, Byung-Gyu Kim*

Dept of IT Engineering, Sookmyung Women's University

e-mail : {dm.jeong, yj.choi}@ivpl.sookmyung.ac.kr, bg.kim@sm.ac.kr*

요 약

저작물 소유권 및 저작권 문제를 해결하기 위한 방안으로 원 저작자에게 이용 허락을 받기 위한 기술적인 장치가 필요하다는 의견이 대두했다. 특히 교육 저작물은 그 특성상 이해를 돕기 위해 다양한 형태로 변형 후 사용되어지는 경우가 많다. 본 논문에서는 ORB 알고리즘의 강점을 이용하면서도 회전 왜곡을 복구하는 전처리 알고리즘과 색상 기반 특징 기술자를 이용한 매칭 알고리즘을 제안하여 스캐닝 된 교육 저작물에서 사용되는 이미지의 원본을 찾는 검색 기술을 제안하였다. 교육 저작물에서 흔히 볼 수 있는 그림, 사진 등 저작물의 변형에 대해 제안된 알고리즘은 평균 96.8%의 매칭 정확도를 보여준다.

1. 서론

인터넷의 보편화로 이미지를 얻을 수 있는 경로가 늘어남에 따라 무분별한 저작물의 사용은 창작자에게 정상적으로 수익이 지급되지 못하는 구조적 문제를 낳았다. 특히 교과서로 대표되는 교육 저작물의 경우 그 특성상 다양한 저작권을 가진 그림, 사진, 지도 등의 이미지를 사용해야 하므로 원 저작자에게 이용 허락을 받기 위한 기술적 장치를 필요로 한다[1].

본 논문에서 제안하는 교육 저작물 매칭 기술은 교육 저작물의 사용 목적에 맞게 변형된 이미지로부터 원본 이미지를 찾아내는 것을 목적으로 한다. 이 기술은 특허권에 관계없이 자유롭게 사용할 수 있는 ORB(Oriented FAST and Rotated BRIEF)[2] 알고리즘을 기반으로 회전 복구 알고리즘, 색상 기반 유사도 매칭 기법을 적용하여 기존 ORB 알고리즘 대비 최대 10.7% 성능을 향상시켰다. 제안하는 기술을 통해 교육저작물 검색 시스템을 구축하면 사용자 관점에서는 정당한 이용 허락을 받은 저작권 작업물을 사용할 수 있고, 창작자 관점에서는 자신의 저작물이 이용되는 현황과 그에 따른 수익을 정상적으로 제공받는 효과를 기대할 수 있다.

2. 관련연구

교육 저작물에서 사용되는 저작권 이미지는 교육 저작물의 특성에 맞게 편집, 변형, 변환되어지는 것이 일반적이다. 따라서 이미지 매칭 및 검색 작업 수행 시 다양한 왜곡에도 강인한 특성을 가지는 키포인트 추출 기법이 필요하다. 키포인트 추출과 관련된 연구는 과거로부터 활발하게 이루어지고 있는데 가장 대표적인 것이 Lowe[3]가 제안한 SIFT(Scale Invariant Feature Transform)

알고리즘이다. SIFT 알고리즘은 크기 변화에 불변하는 키포인트의 특성을 찾기 위해 피라미드 형식으로 키포인트 후보군을 구성한다. SIFT는 가능한 모든 조건을 탐색 후 키포인트를 결정하기 때문에 지역적 키포인트 추출 방법 중에는 정확도가 매우 높지만 그만큼 수행 시간도 길다는 단점이 있다. 이런 단점을 개선하기 위해 SURF(Speed Up Robust Feature)[4], SICK(Scale Invariant Corner Keypoints)[5] 같은 개선된 알고리즘들이 제안되었다. 그러나 SIFT와 SURF 알고리즘은 특허권이 있어 응용과 개발에 있어 한계가 있다. 이를 대체하기 위하여 OpenCV Labs에서는 키포인트 검출 알고리즘 FAST(Features from Accelerated Segment Test)[6]와 특징 기술자 추출 알고리즘 BRIEF(D)[7]를 각각 수정 및 혼합 적용하여 ORB 알고리즘을 개발하였다. ORB 알고리즘은 키포인트를 찾기 위해 다양한 피라미드 스케일에서 FAST를 적용하며 회전에 불변한 특성을 추가한 조정된 BRIEF 특징 기술자를 적용한다. 이러한 알고리즘에서 추출된 특징 기술자는 이미지 매칭 시 키포인트를 비교하는 기준으로 적용된다.

3. 왜곡에 강인한 키포인트 추출 및 매칭 기법

본 논문에서 제안하는 키포인트 추출 및 매칭 기법은 4가지의 세부적인 과정으로 나눌 수 있다.

3.1 회전 왜곡 복구 알고리즘

회전 왜곡은 픽셀의 기하학적 위치를 변화시키기 때문에 일반적인 키포인트 매칭 기술을 적용할 시 에러율이 매우 높다. 따라서 이미지가 회전된 각도를 감지해 원본과 유사하도록 복구하면 정확도를 높일 수 있다. 그림 1은 회전 왜곡에 대처하기 위한 복구

과정을 나타내고 있다. 먼저 팽창 연산을 통해 덩어리화된 이미지를 원래 이미지와의 차연산을 통해 윤곽을 드러낸다. 비트를 반전시켜 윤곽에 외접하면서 면적이 가장 작은 직사각형의 좌상단 꼭지점 좌표를 계산한다. 이 좌표로부터 구해진 각도만큼 회전을 시키고 테두리 영역을 삭제하면 회전되기 이전보다 크기가 작은 이미지를 얻을 수 있다. 크기가 작은 이미지는 원본 이미지와 같은 지역적 특징을 가지고 있기 때문에 이미지 매칭 및 검색에 용이한 형태가 된다.

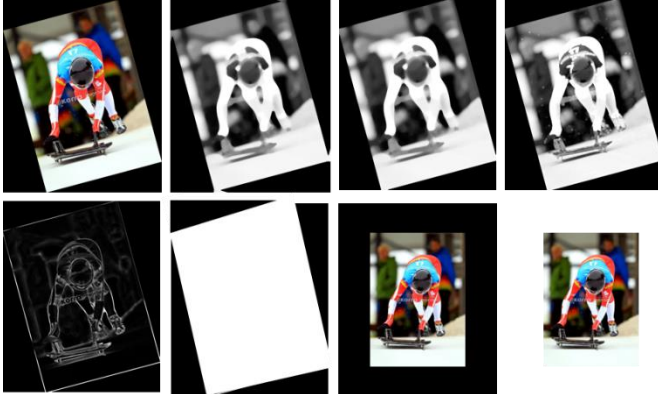


그림 1. 회전 왜곡 복구 과정.

3.2 ORB 특징 검출기를 이용한 키포인트 검출

전처리를 마친 이미지에서 키포인트 및 특징 기술자를 검출하기 위해 ORB 알고리즘이 적용된다. ORB 알고리즘은 키포인트 추출을 위해 키포인트 후보 픽셀을 중심으로 반지름이 9인 위치에 원형으로 놓인 픽셀들과의 차이를 비교하여 키포인트를 추출하는 FAST(Features from Accelerated Segment Test)-9 기법을 적용한다. 이후에 방향성을 고려하기 위해 Intensity Centroid 라는 독자적인 기법을 추가하여 키포인트에 방향성을 부여한다. rBRIEF(rotated BRIEF) 알고리즘은 이렇게 추출된 키포인트를 설명하는 특징 기술자를 추출한다. BRIEF(Binary Robust Independent Elementary Features) 알고리즘은 패치 내에서 가우시안 분포를 따르는 쌍들의 픽셀 밝기값을 비교하여 특징 기술자를 추출하는데 rBRIEF 알고리즘은 여기에 앞서 추출된 키포인트의 각도에 대해 회전 행렬을 적용한 알고리즘으로 왜곡에 더 강인하다. 수식 (1)은 이미지 하나에서 추출된 특징 기술자 벡터 결과이다.

$$D_i = \begin{bmatrix} v_{00} & \cdots & v_{031} \\ \vdots & \ddots & \vdots \\ v_{(k-1)0} & \cdots & v_{(k-1)31} \end{bmatrix} \quad (1)$$

여기서 i 는 이미지 번호를, k 는 키포인트의 개수를 의미한다. 최종적으로 키포인트 하나 당 32 차원의 특징 기술자가 추출된다.

3.3 CAD 를 이용한 1차 매칭

매칭 요청 시 먼저 유사도 개념으로 매칭이 수행된다. 1차 매칭은 입력된 이미지의 키포인트,

특징 기술자와 데이터 베이스에 있는 이미지들의 키포인트, 특징 기술자 정보를 CAD(Cosine Angle Distance) 계산으로 거리가 가까운 것부터 높은 유사도를 가지는 이미지를 찾아내는 방식으로 수행된다. 수식 2는 특징 기술자 벡터들 간의 거리를 비교하기 위한 CAD 수식이다.

$$\begin{aligned} CAD(X,Y) &= \frac{A \cdot B}{|A||B|} \\ &= \frac{\sum_{j=1}^{32} A_j \times B_j}{\sqrt{\sum_{j=1}^{32} (A_j)^2} \times \sqrt{\sum_{j=1}^{32} (B_j)^2}} \end{aligned} \quad (2)$$

여기서 j 는 32 차원의 특징 기술자 벡터에 하나씩 접근하여 순차적으로 곱할 수 있도록 주어진 인덱스이다. 이렇게 계산하여 $\cos(\theta)$ 의 값이 1에 가까우면 유사도가 높고, 0에 가까우면 유사도가 낮은 것으로 판단할 수 있다.

3.4 색상 특징 기술자를 이용한 2차 매칭

본 논문에서 제안하는 방법은 1차 매칭을 통해 선별된 상위 50개 검색 결과에 대해 색상 기반 유사도 매칭을 2차적으로 수행함으로써 최종 검색 결과의 인지적 유사성을 높이고자 한다. 색상 기반 유사도 매칭 기법은 지역적 색상 정보를 추출하여 인덱스로 활용하는 기법으로 그림 2와 같이 입력 이미지에 5개 영역으로 나뉜 마스크를 적용한다. 마스크는 각 영역별로 색상 특징 기술자를 추출하여 이미지의 지역적 특성을 강인하게 만들어준다. 마스크의 각 영역별로 특징 기술자를 추출하는 방법은 다음과 같다. 먼저 인간이 색상을 습득하는 방식과 유사하게 색상 정보를 추출하기 위하여 RGB 채널의 이미지를 HSV 채널 이미지로 변환한다. 실험적으로 H, S, V 채널이 사용할 히스토그램 bin의 개수를 각각 8개, 12개, 3개로 하여 각 채널의 정보를 누적한다. 이렇게 하면 각 영역의 색상 정보는 288개의 번호 목록을 가지고 표현될 수 있고, 한 이미지는 1440개의 색상 정보를 색상 특징 기술자로 사용할 수 있다. 그림 3은 색상 기반 유사도 매칭 기법 적용 전과 후의 결과 차이를 보여준다.



그림 2. 색상 특징 기술자 매칭 마스크.



그림 3. (a) 1차 매칭만 적용했을 때의 검색 결과, (b) 1차 매칭 후 2차 매칭까지 적용했을 때의 검색 결과.

4. 실험 결과

4.1 실험 데이터셋

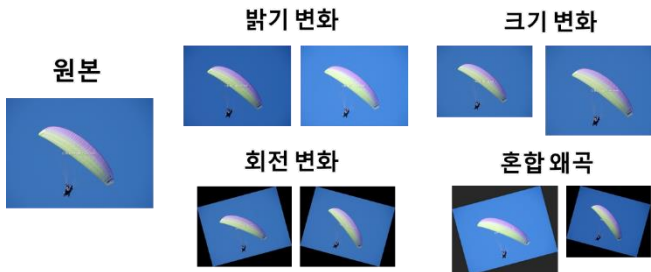


그림 4. 실험 데이터셋 예시.

본 논문에서 제안한 키포인트 추출 및 매칭 기법이 교육 저작물에서 흔히 볼 수 있는 밝기, 회전, 크기 및 이를 혼합한 왜곡에 강인하다는 것 증명하기 위해 6 가지 왜곡(Brighter, Darker, Bigger, Smaller, Bigger-brighter, Smaller-darker)을 적용한 변형 데이터셋을 사용하여 실험하였다. Brighter 데이터셋은 HSV 채널로 변환 뒤 V 채널의 값을 40 증가시킨 데이터셋이고, Darker 데이터셋은 감마 값을 0.85로 한 감마보정을 적용한 데이터셋이다. Bigger, Smaller 데이터셋 각각은 이미지 사이즈를 10% 증가, 감소 시킨 데이터셋이며 마지막으로 밝기와 크기 왜곡을 혼합한 데이터셋에 시계방향, 반시계방향으로 회전 왜곡을 더한 데이터셋이 Bigger-brighter, Smaller-darker 데이터셋이다. 그림 4는 각 왜곡 데이터셋에 대한 예시이다.

표 1. 검색 성능 비교: 정확도

Precision (%)		ORB[2]	Proposed
단일왜곡	Brighter	92.5	93.8
	Darker	95.0	98.5
	Bigger	93.6	99.8
	Smaller	96.0	99.2
혼합왜곡	Bigger-brighter	86.2	92.0
	Smaller-darker	88.2	97.7
평균		91.9	96.8

4.2 정밀도 검증

표 1에서 기존 ORB 알고리즘과 제안한 방법의 성능을 비교하고 있다. 성능 비교를 위해 Precision을 정확도 지표로 사용하였다. 기존 ORB 알고리즘과 제안한 방법은 회전 왜곡이 더해졌을 때 최대 10.7% 정밀도가 증가한 것을 확인할 수 있다.

4.3 색상 기반 유사도 매칭 결과

본 논문에서 제안하는 색상 기반 유사도 매칭 기법을 통해 추출된 검색 결과는 그림 5와 같이 시각적 유사도가 높다는 것을 알 수 있다.

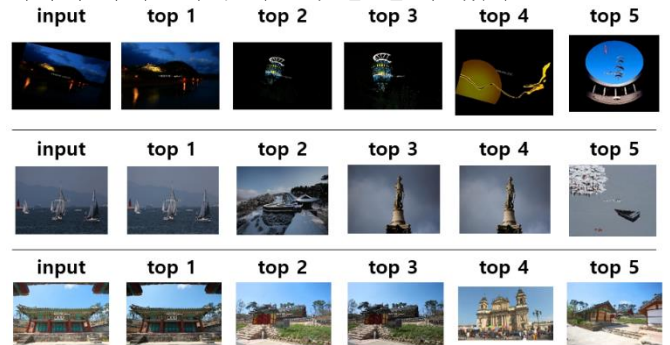


그림 5. 색상 기반 유사도 매칭 결과.

6. 결론

본 논문에서는 교육저작물에서 사용되는 저작권 저작물이 보다 정확하게 검색될 수 있도록 키포인트 추출 및 매칭 기법을 제안하였다. 제안한 기법은 기존에 오류가 많이 발생하던 원인인 회전 왜곡을 복구하는 전처리 기법을 도입하여 평균 정밀도를 5.3% 증가시켜 96.8%의 정밀도를 달성하였다. 또한, 상위 검색 결과로 유사한 이미지가 검색될 수 있도록 색상 특징 기술자를 용한 색상 기반 유사도 매칭 기법을 적용하여 검색 결과의 질을 높였다. 이러한 기법은 교육저작물에서 사용된 원본 저작물 이미지를 검색하고 찾아내는 데 사용될 것이다.

Acknowledgement

본 연구는 문화체육관광부 및 한국저작권위원회의 2019년도 저작권기술개발사업의 연구결과로 수행되었음.

참고문헌

- [1] 최영주, 김지해, 이영운, 이종혁, 홍광수, 김병규. "교육용 도서 영상을 위한 효과적인 객체 자동 분류 기술." 한국디지털콘텐츠학회 논문지, 18.7, pp. 1323-1331. Nov. 2017.
- [2] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," in Proc. IEEE Int. Conf. Comput. Vision, Barcelona, Spain, pp. 2564-2571, Nov. 2011.
- [3] D.G. Lowe, "Distinctive image features from scale-invariant keypoints." International Journal of Computer Vision 60 (2), pp. 91-110, 2004.

- [4] H. Bay, A. Ess, T. Tuytelaars, V. Gool, "Speeded-up robust features (surf)." Computer Vision and Image Understanding 110 (3), pp. 346-359, 2008.
- [5] L. BO, L. HAIBO, S. ULRIK, "Scale-invariant corner keypoints. In International Conference on Image Processing" Paris, France, pp. 5741-5745, 2014.
- [6] E. Rosten, T. Drummond. "Machine learning for high speed corner detection." in 9th European Conference on Computer Vision, vol. 1, 2006.
- [7] M. Calonder, V. Lepetit, C. Strecha, and P. Fua. "Brief: Binary robust independent elementary features." Computer Vision-ECCV 2010, pp. 778-792, 2010.