
디지로그 북 추적을 위한 레이아웃 기반 페이지 인식 기법

Layout Based Page Recognition for Tracking the Digilog Book

박종희, Jonghee Park*, 우운택, Woontack Woo†

요약 본 논문에서는 디지로그 북 추적을 위한 레이아웃 기반 페이지 인식 기법을 제안한다. 기존 디지로그 북에서 적용된 마커 기반 인식 기법은 페이지에 특정 패턴의 마커를 부착해야 하며, 텍스처 기반의 페이지 인식 기법은 문자로만 구성되어진 페이지를 인식하는데 어려움이 있다. 이를 극복하기 위해, 제안된 기법은 인쇄된 페이지의 레이아웃 정보를 활용하여 영역에 따라 적합한 알고리즘을 적용한다. 학습 단계에서 입력 영상은 페이지 레이아웃에 따라 그림 영역과 문자 영역의 영상으로 분리된다. 각 영상은 영역에 따라 적합한 특징점이 추출되고, 기술자(Descriptor)를 생성하여 해당 페이지를 학습한다. 카메라 입력 영상은 학습된 영상과의 특징점 매칭(Matching)을 통해 인식된다. 제안된 기법은 기존에 출판된 책이나 문서를 변형하지 않고 적용 가능하기 때문에 교육, 출판 업계 등 여러 응용분야에서 널리 활용될 수 있을 것이다.

Abstract In this paper, we propose a layout based page recognition method for tracking the Digilog Book. In case of marker based method which applied method to previous Digilog Book, specific patterns of marker should be attached to the book. In addition, texture based method cannot recognize text-only pages. In order to solve those problems, the proposed method applies different algorithms to each area of page layout. In learning phase, input images divides into picture area and text area. According to area's type, it learns pages by extracting proper feature points and generating descriptors. Images from camera are recognized by matching feature points of each area. We expect that proposed method would be applicable to many fields such as education and publication.

핵심어: *Digilog Book, Layout, Page Recognition*

본 연구는 문화관광부 및 한국문화콘텐츠진흥원의 문화콘텐츠기술연구소육성사업의 연구결과로 수행되었음.

*주저자: 광주과학기술원 U-VR 연구실; e-mail: jpark@gist.ac.kr

†교신저자: 광주과학기술원 U-VR 연구실; e-mail: wwoo@gist.ac.kr

1. 서론

디지로그 북은 증강현실 환경에서 출판물 (종이책) 에 인간의 시각, 청각, 촉각을 자극하는 멀티미디어 콘텐츠를 융합시켜 종이책에서 제공할 수 없는 부가적인 정보를 제공하는 책이다 [1]. 디지로그 북은 시각적 정보를 증강현실 기술을 이용하여 제공하는데, 이를 위한 페이지의 인식 및 추적기법은 디지로그 북을 제작하는데 가장 중요한 부분이다.

기존 디지로그 북들은 페이지를 인식 및 추적하기 위해 주로 마커와 텍스처 기반의 자연 특징점 매칭 기법이 적용되었다. Magic book [2]은 첫 번째 디지로그 북으로 마커 기반의 페이지 인식 및 추적기법을 이용하였다. 마커 기반 기법은 강건하게 페이지 인식 및 추적이 가능하지만, 인쇄된 출판물에 특정패턴의 마커를 부착해야 할 뿐만 아니라 마커의 일부가 가려질 경우 콘텐츠가 증강되지 않는 문제점이 있다. 이러한 문제점을 해결하기 위해 Haunted Book [3]은 그림영역의 자연 특징점 기반 인식 및 추적 기법을 적용하였다. 하지만, 그림 1과 같이 텍스처의 특징점 기반 인식 기법은 문자로 각 특징점 기술자의 높은 유사성 때문에 문자로만 구성된 페이지에는 적용되기 힘들다 [4].

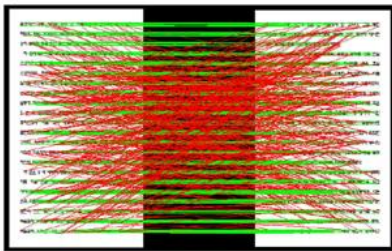


그림 1. 텍스트 영상에서 FAST [5]+SURF [6]로 매칭한 결과

제안된 기법은 페이지 레이아웃 정보를 이용하여, 문자나 텍스트로 이루어진 영역을 모두 인식할 수 있다. 또한, 문자와 그림이 함께 인쇄된 페이지의 경우 기존 기법보다 높은 매칭율을 보인다.

2. 레이아웃 기반 페이지 인식 기법

기존 출판된 책을 변형없이 적용하기 위해서, 제안된 기법은 자연 특징점 기반으로 페이지를 인식한다. 그림 2은 제안된 기법의 흐름도를 나타낸다. 카메라로부터 입력된 영상은 레이아웃에 따라 두개의 영상(그림, 문자 영역)으로 분리되어, 각 영역에 따라 적합한 알고리즘이 적용된다.

자연 특징점 기반 인식 기법은 정확한 페이지 레이아웃 추출기법보다 빠른 수행속도의 레이아웃 분석과정이 필요하다. 레이아웃을 분석하기 위해, 제안된 기법은 CRLA (Constraint Run Length Algorithm)을 수직, 수평방향으로 실

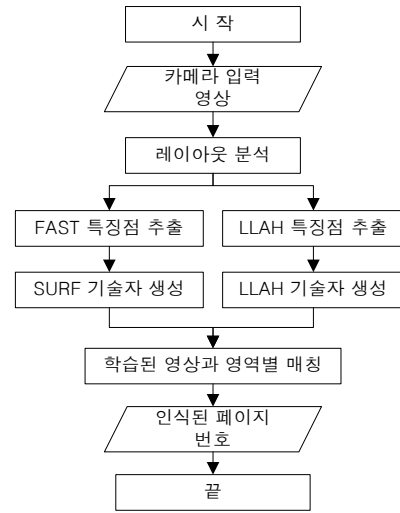


그림 2. 흐름도

행하여 공통된 영역을 찾는다. 각 영역들을 그림 영역과 문자 영역으로 구분하기 위해, 해당 영역의 히스토그램을 사용한다. 문자 영역의 경우, 보통 문자와 배경간의 색 차이가 크기 때문에 히스토그램에서 현저한 정점이 검출된다. 따라서, 해당 정점이 특정 값을 넘어간 영역은 문자영역으로, 아닌 경우에는 그림영역으로 간주한다.

레이아웃이 추출되면, 각 영역에 대한 특징점과 기술자를 생성한다. 그림영역 영상에서는 FAST 코너를 추출하여 SURF 기술자를 생성하고, 문자영역 영상은 LLAH [7]와 같이 각 단어의 중간점으로 추출하여 특징점으로 사용하고, 주변 특징점과의 Affine Invariant 값을 이용하여 기술자를 생성한다.

그림 3은 x 축으로 10° 만큼 회전된 영상과 학습된 영상과의 레이아웃 기반 매칭 결과를 보여준다. 인쇄된 콘텐츠의 종류에 따라 다른 알고리즘이 적용된 것을 확인할 수 있다. 해당 그림에서 빨간색 선은 매칭 결과의 outlier를 나타내고, 녹색 선은 inlier를 나타낸다.

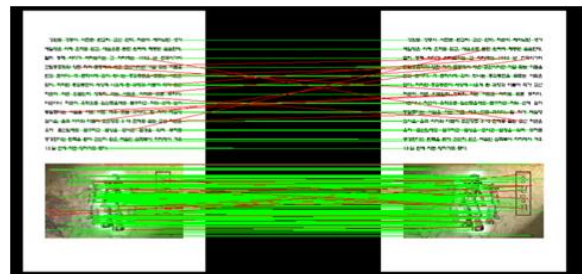


그림 3. 레이아웃 기반 매칭 결과

영역별로 생성된 기술자와 사전에 학습한 영상간의 매칭되는 점의 수에 따라 페이지가 인식된다. 이 때, 영역별로 페이지를 중복해서 검색하지 않고 각 영역의 넓이에 따라 좁은 영역의 매칭 결과로 후보군을 생성하고 넓은 영역의 매칭

결과로 페이지를 인식한다.

3. 실험

실험을 위한 하드웨어 구성은 Intel Xeon 2.66Ghz CPU와 3GB RAM을 사용하였다. 각 알고리즘은 OpenCV를 기반으로 구현되었다. 그림 4은 제안된 방법의 매칭율을 검증하기 위한 실험 영상 집합이다. 모든 영상은 640×480의 해상도이며, 각 영상을 x축으로 10°씩 회전하여 사용하였다. 매칭율은 수식 1와 같이 계산되었으며, n 과 T 는 각각 inlier의 갯수와 전체 매칭셋의 갯수이다. 해당 실험에서 호모그래피 에러가 3 픽셀 이상되는 매칭은 outlier로 가정하였다.

$$MatchingRate = n/T \quad (1)$$

그림 5는 해당 영상 집합에서 LLAH와 phonySURF, 그리고 제안된 방법의 각도별 평균 매칭율을 비교한 그래프이다. 제안된 방법은 해당 영상 집합에서 phonySURF 보다 평균 16.1%, LLAH 보다 평균 19.6% 높은 매칭율을 보였다.



그림 4. 테스트 영상 집합

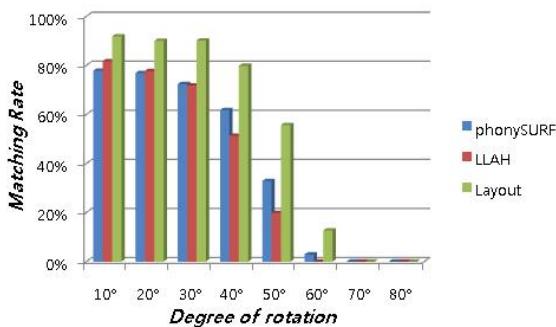


그림 5. 페이지 매칭율

4. 결론 및 향후연구

본 논문에서는 페이지의 고유정보인 레이아웃을 활용하여 페이지를 인식하는 기법을 제안하였다. 제안된 기법은 기존 디지털로 북에서 적용되는 마커나 텍스트 기반 인식기법보다 높은 특징점 매칭율을 보였다. 하지만, 증강현실 어플리케이션으로 활용되기 위해서는 좀 더 빠른 수행시간이 요구된다. 향후에는 실시간 인식이 가능한 기술자 생성 및 매칭 분야로 연구할 것이다.

참고 문헌

- [1] 이영호, 하태진, 이형목, 김기영, and 우운택, “디지털로 북 - 아나로그 책과 디지털 콘텐츠의 융합,” in 정보통신 분야학회 합동학술대회, vol. 14, pp. 186–189, 2007.
- [2] M. Billingham, H. Kato, and I. Poupyrev, “The MagicBook—moving seamlessly between reality and virtuality,” IEEE Computer Graphics and Applications, pp. 6–8, 2001.
- [3] C. Scherrer, J. Pilet, P. Fua, and V. Lepetit, “The haunted book,” in Proc. 7th IEEE/ACM International Symposium on Mixed and Augmented Reality ISMAR 2008, pp. 163–164, 2008.
- [4] H. Uchiyama and H. Saito, “Augmenting Text Document by On-Line Learning of Local Arrangement of Keypoints,” in Proc. 8th IEEE/ACM International Symposium on Mixed and Augmented Reality ISMAR 2009, pp. 95–98, 2009.
- [5] E. Rosten and T. Drummond, “Machine learning for high-speed corner detection,” European Conference on Computer Vision, vol. 1, pp. 430–443, 2006.
- [6] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, “Speeded-up robust features (SURF),” Computer Vision and Image Understanding, vol. 110, no. 3, pp. 346–359, 2008.
- [7] T. Nakai, K. Kise, and M. Iwamura, “Use of affine invariants in locally likely arrangement hashing for camera-based document image retrieval,” Lecture Notes in Computer Science, vol. 3872, p. 541, 2006.