

---

## 디지로그 북을 위한 레이아웃 기반 다감각 콘텐츠 저작 도구

### Layout Based Multimodal Contents Authoring Tool for Digilog Book

박종희, Jonghee Park\*, 우운택, Woontack Woo†

---

**요약** 본 논문에서는 디지로그 북 (Digilog Book) 저작 환경에서 레이아웃 기반 다감각 콘텐츠 저작 도구를 제안한다. 저작 과정에 사용자는 마우스나 펜 타입 장치를 통해 종이책 위에 반복적으로 가상 영역을 생성하고, 해당 영역에 증강될 콘텐츠의 위치, 사이즈, 파일 선택, 이벤트 처리 등의 속성을 지정한다. 저작이 완료되면 시스템은 기존에 인쇄된 페이지 번호를 인식하고, 해당 페이지에 생성된 영역과 속성 정보를 포함하는 레이아웃을 생성한다. 페이지 레이아웃은 증강현실 환경에서 렌더링 하기 적합한 씬 그래프 (Scene Graph) 형태로 표현 되고 XML 형태로 저장된다. 디지로그 북 뷰어는 저장된 레이아웃을 로딩하고 저작된 속성을 분석하여 해당 페이지에 영역별로 증강 및 기능을 실행한다. 제안된 저작 도구를 통해 사용자는 혼합 인터페이스를 통해 디지로그 북 환경에서 관심영역에 시각, 청각적인 다감각 콘텐츠를 손쉽게 저작할 수 있다. 증강현실 환경에서 사용자가 레이아웃을 보다 쉽게 생성하기 위하여 영역 템플릿을 제공한다. 또한, 제안된 저작 도구는 페이지 인식부와 페이지 추적부를 독립적으로 구성함에 따라 단일 마커만으로 다수 페이지 저작이 가능하다. 실험 결과, 제안된 저작도구는 증강현실 환경에서 적절한 수행시간을 나타냄을 보였다. 제안된 저작 도구는 교육, 출판 업계 등 여러 응용분야에서 널리 활용될 수 있을 것이다.

**Abstract** In this paper, we propose layout based multimodal contents authoring tool for Digilog Book. In authoring step, users create a virtual area using mouse or pen-type device and select property of the area repetitively. After finishing authoring step, system recognizes printed page number and generate page layout including areas and property information. Page layout is represented as a scene graph and stored as XML format. Digilog Book viewer loads stored page layout and analyze properties then augment virtual contents or execute functions based on area. Users can author visual and auditory contents easily by using hybrid interface. In AR environment, system provides area templates in order to help creating area. In addition, proposed authoring tool separates page recognition module from page tracking module. So, it is possible to author many pages using only single marker. As a result of experiment, we showed proposed authoring tool has reasonable performance time in AR environment. We expect that proposed authoring tool would be applicable to many fields such as education and publication.

**핵심어:** *Digilog Book, Authoring, Multimodal Contents, Layout*

---

본 연구는 문화관광부 및 한국문화콘텐츠진흥원의 문화콘텐츠기술연구소육성사업의 연구결과로 수행되었음.

\*주저자: 광주과학기술원 U-VR 연구실; e-mail: [jpark@gist.ac.kr](mailto:jpark@gist.ac.kr)

†교신저자: 광주과학기술원 U-VR 연구실; e-mail: [wwoo@gist.ac.kr](mailto:wwoo@gist.ac.kr)

## 1. 서론

디지털로 북은 증강현실 환경에서 출판물 (종이책) 에 인간의 시각, 청각, 촉각을 자극하는 멀티미디어 콘텐츠를 융합시켜 종이책에서 제공할 수 없는 부가적인 정보를 제공하는 책이다 [1]. 디지털로 북 저작도구로는 다음과 같은 연구들이 진행되었다. ARToolkit [2]은 사용자가 제공되는 라이브러리를 기반으로 프로그래밍 언어를 통해 다양한 확장이 가능한 장점이 있다. APRIL [3]은 GUI (Graphic User Interface) 인터페이스를 제공하고 재사용성이 뛰어난 컴포넌트 기반 저작을 지원한다. 하지만 ARToolkit, APRIL과 같은 저수준의 저작 도구는 프로그래밍에 익숙하지 않은 사용자가 저작하는 데에 어려움이 있다 [4]. 이를 극복하기 위한 고수준의 저작도구인 ARTalet [4]은 일반 사용자가 쉽게 3D 콘텐츠의 저작이 가능하다. 하지만 이는 시각적인 콘텐츠의 저작에 중점을 두고 있다.

본 논문에서는 디지털로 북 저작 환경에서 레이아웃 기반 다감각 콘텐츠 저작 도구를 제안한다. 제안된 저작 도구는 GUI 와 TUI (Tangible User Interface)을 혼합한 인터페이스를 제공한다. 사용자는 제공되는 인터페이스를 통해 가상 영역을 생성하고, 해당 영역에 속성 (시각, 청각 콘텐츠) 을 지정하는 작업을 반복적으로 수행함으로써 디지털로 북을 저작한다. 제안된 저작 도구는 증강현실 환경에서 사용자가 가상 영역을 쉽게 생성하기 위한 영역 템플릿을 제공한다. 또한, 제안된 저작 도구는 페이지 추적부와 페이지 인식부로를 독립적으로 구성 함으로써, 사용자가 단일 마커로 다수의 페이지를 저작 가능하게 한다.

혼합 인터페이스는 GUI 에 익숙한 사용자가 증강현실에서의 새로운 인터페이스를 습득해야하는 부담을 덜어준다. 사용자는 영역 생성 시 템플릿을 통한 빠른 영역 생성이 가능하다. 또한, 기존의 각 페이지에 마커가 인쇄되어 있는 디지털로 북의 시각적인 부자연스러움을 해결하였다.

## 2. 레이아웃 기반 디지털로 북 저작

디지털로 북 환경에서 페이지 레이아웃은 페이지에 증강될 콘텐츠의 배치와 양식을 나타낸다. 그림 1은 전체 시스템의 흐름을 나타낸다 [5]. 다감각 콘텐츠 저작 도구를 통해 사용자는 각 페이지에 해당하는 페이지 레이아웃을 생성하고, 생성된 페이지 레이아웃은 XML 형태로 저장된다. 저장된 페이지 레이아웃은 디지털로 북 뷰어에서 로딩 및 분석 되어 저작된 속성에 기반하여 영역별 증강 및 상호작용에 활용된다.

영역 생성부는 펜타입 장치, 키보드 및 마우스를 통해 입력으로 페이지 위에 2차원 가상 영역을 생성한다. 3차원 증강현실 환경에서 사용자 입력은 카메라 시점이 변하게 되

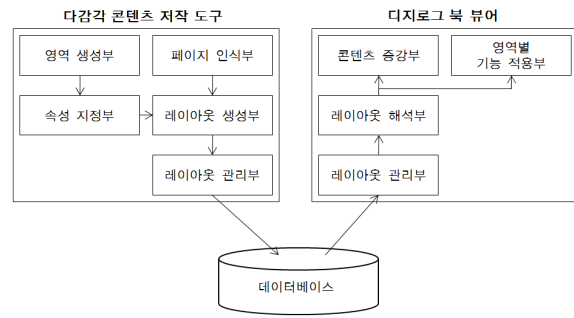


그림 1. 시스템 구조

므로 정확하게 입력하기 어렵다. 이를 보완하기 위해 제안된 저작도구는 자유 입력뿐만 아니라 영역 템플릿을 이용한 입력을 제공한다. 페이지는 마커와 평행하다는 가정 하에 영역 템플릿은 페이지의 X,Y 축에 평행하게 가상 영역을 생성한다.

속성 지정부는 생성된 영역에 영역 종류, 위치, 파일 위치, 이벤트 처리 등의 속성을 지정하는 GUI를 제공한다. 레이아웃을 식별하기 위해 페이지 인식부는 페이지 번호가 인쇄된 영역을 분리한다. 분리된 영역 영상은 문자인식 엔진을 통해 숫자로 인식되어 해당 페이지의 식별자로 사용된다. 각 영역은 식별자, 종류, 위치, 크기, 속성, 기능 등의 정보를 포함한다. 레이아웃 생성부는 한 페이지에서 저작된 영역들은 동일 레이아웃에 포함되며 계층 구조인 씬 그래프 (Scene Graph) 구조로 표현한다. 해당 씬 그래프는 레이아웃 관리부에서 XML 형태로 저장된다. 그림 2는 2개의 비디오 영역과 1개의 이미지 영역으로 이루어진 레이아웃과 해당하는 XML 파일을 나타낸다.

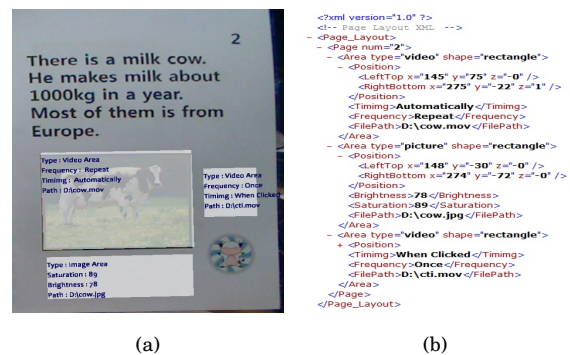


그림 2. 생성된 레이아웃과 해당 XML 구조

### 2.1 다감각 콘텐츠 저작 도구

사용자는 다감각 콘텐츠 저작 도구에서 제공되는 혼합 인터페이스를 통해 페이지 별로 레이아웃을 생성한다. GUI는 기존 사용자에게 이미 친숙하고, 2 자유도로 상호작용을 하기 때문에 조작하기 쉽다. TUI는 사용자에게 몰입감을

더해줄 수 있는 장점이 있다. 따라서 본 저작 도구는 GUI와 TUI를 혼합한 인터페이스를 제공한다. 영역 생성 및 선택, 콘텐츠 제어 등의 작업은 속성 지정 작업에 비해 조작이 간단하기 때문에 펜 타입 장치를 사용하여 몰입감을 더한다. 메뉴 조작, 이벤트 처리 등의 상대적으로 복잡한 속성 지정 작업에는 사용자에게 친숙한 GUI를 제공한다. 이는 사용자에게 새로운 UI의 조작 방법을 습득하는 부담을 덜어준다. 그림 3(a)는 사용자가 가상 영역을 TUI 기반으로 생성한 결과이다. 증강된 시각형은 관심 영역을 나타낸다. 그림 3(b)는 GUI 기반의 콘텐츠 선택 및 속성 설정을 나타낸다.

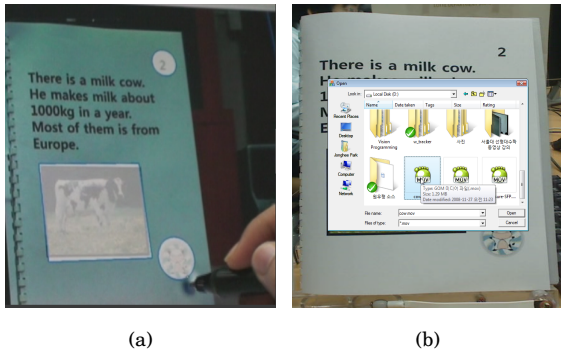


그림 3. 혼합 인터페이스 (a) TUI 기반 가상 영역 생성 (b) GUI 기반 콘텐츠 선택 및 속성 설정

디지털로그 북은 다수의 페이지로 이루어져 있으며, 사용자는 그림 4과 같이 페이지에 증강될 영역을 생성하고 속성을 지정하는 작업을 반복적으로 수행한다. 각 레이아웃은 사용자가 저작한 다수의 영역 및 해당 속성으로 이루어진다.

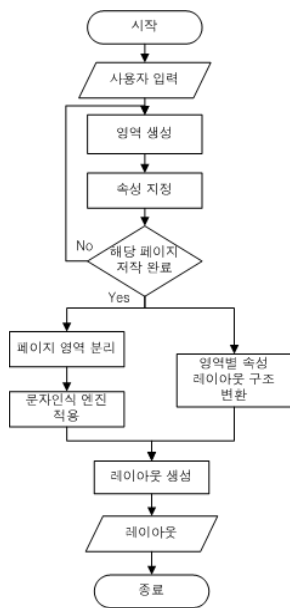


그림 4. 저작 흐름도

각 영역은 그림, 비디오, 문자, 소리 영역으로 나뉜다. 표 1은 선택할 수 있는 영역의 종류와 속성을 나타낸다.

영역의 종류에 따라 사용자가 선택할 수 있는 속성은 다르며, 지정된 속성은 디지털로그 북 뷰어에서 해석되어 영역별로 실행된다.

표 1. 영역별 속성 및 기능

영역 종류	속성	기능
그림	명도, 채도	증강
동영상	재생 시점	재생
	재생 횟수	
문자	스피치 구간	인쇄 문자 스피치
	속도	
	시작 시점	
소리	폰트, 사이즈, 색	증강
	재생 시점	재생

기존 디지털로그 북은 페이지 당 마커를 부착하여 페이지를 인식하거나 선 처리 과정을 통해 페이지 영상 특징점을 미리 학습하고, 입력 영상에서 특징점을 추출 및 매칭하여 페이지를 인식한다. 하지만 제안된 저작도구는 페이지 추적부와 페이지 인식부를 독립적으로 구성한다. 페이지의 식별자로 인쇄된 정보를 사용함으로써 단일 마커로 다수의 페이지 저작이 가능하다. 페이지 인식부가 페이지 추적부와 독립적인 경우, 저작 도구는 사용자가 페이지를 언제 넘겼는지에 대한 이벤트가 필요하다. 해당 이벤트를 처리하기 위하여 그림 영역의 색 변화량을 사용한다. 수식 1에서  $k$  번째 입력 영상과  $k+1$  번째 입력영상의 그림 영역을 각각  $A, B$  라 하고,  $n \times m$  크기의 영상이라 한다. 이때,  $w$  는 각각  $A, B$  의  $i, j$  번째의 색깔 값을 의미한다. 수식 1에서 얻어진 그림 영역 간 유사도가 특정 임계값을 넘을 경우, 사용자는 페이지를 넘겼다고 간주한다. 디지털로그 북 뷰어는 매프레임에 사용자가 페이지를 넘기는 이벤트를 체크하여 해당 레이아웃을 로딩 한다. 제안된 기법은 영상 특징점 기반의 디지털로그 북에서도 가능하다.

$$Similarity = \sum_{i=0}^n \sum_{j=0}^m (a_{ij} - b_{ij})^2 \quad (1)$$

## 2.2 레이아웃 기반 디지털로그 북 뷰어

디지털로그 북은 다각각 콘텐츠에서 생성된 XML 파일을 로딩하고 분석하여, 쉐인 그래프 구조로 변환한다. 우선, 초기 단계에 입력 영상으로부터 페이지 번호를 분리하고 문자인식 엔진을 적용한다. 해당하는 페이지 레이아웃을 로딩한 후, 사용자가 저작한 속성에 맞게 분석을 한다. 디지털로그 북 뷰어는 영역별로 다른 스펙트가 생성되며 해당하는 기능이나 증강을 한다. 그림 5는 페이지 레이아웃을 분석한 후, 영역별 저작된 콘텐츠를 증강시킨 결과이다.

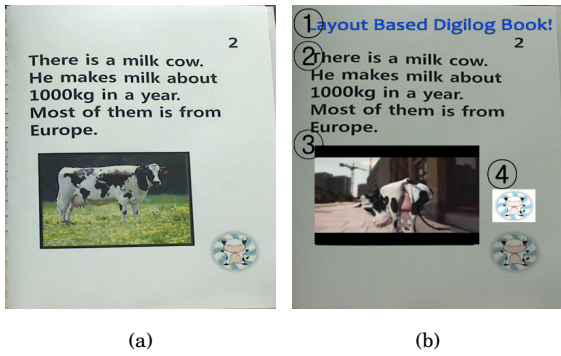


그림 5. 디지털로 북 뷰어를 통해 본 결과 (a) 원본 페이지 (b) 디지털로 북 뷰어를 통해 본 페이지 : ① 문자 ② 스피치 ③ 동영상 ④ 이미지 증강

### 3. 실험 및 분석

개발 환경은 Windows Vista, Intel Centrino T2400 1.83 GHz, 1.5 GB RAM, Visual Studio 2005 에서 개발 되었다. 라이브러리는 OpenCV (이미지 프로세싱) [6], OpenSceneGraph (렌더링) [7], Tesseract (문자인식 엔진) [8] 를 사용하였다.

사용자가 디지털로 북을 자연스럽게 보기 위해서 디지털로 북 뷰어는 빠른 수행시간이 요구된다. 그림 6은 페이지 수에 따른 페이지 넘김 이벤트가 발생한 시점부터 해당 페이지의 콘텐츠가 증강될 때까지의 수행 시간이다. 1 페이지 부터 20 페이지로 구성된 디지털로 북을 대상으로 실험한 결과, 최대 408ms 최소 320ms, 평균 361.53ms의 수행 속도를 보였다. 수행 속도의 주된 원인은 문자인식 엔진으로 나타났다. 제안된 저작도구는 페이지 수에 따라 수행시간이 선형적으로 증가하는 기존 디지털로 북과 달리 페이지 수가 증가되어도 일정 값을 유지한다. 이는 페이지 추적부와 페이지 인식부를 독립적으로 구성함에 따른 것이다.

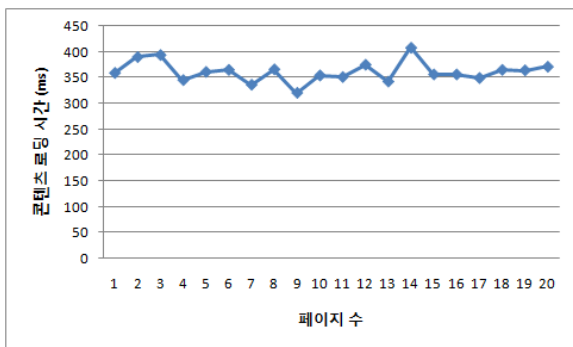


그림 6. 페이지수에 따른 콘텐츠 증강 수행시간

### 4. 결론 및 향후연구

본 논문에서는 프로그래밍에 익숙하지 않은 사용자가 디

지로그 북을 쉽게 저작할 수 있는 저작도구를 제안하였다. 제안된 저작 도구는 증강현실 환경에서 혼합 인터페이스를 제공하여 몰입 감 있는 저작이 가능하다. 사용자 입력 시에 영역 템플릿을 제공하여 정확한 영역 생성을 가능하도록 지원한다. 또한, 기존 디지털로 북과는 다르게 종이책의 번호 영역을 분리해서 페이지를 인식함으로써 하나의 마커로 다수의 페이지의 저작이 가능하다. 향후에는 종이책의 레이아웃을 분석하여 시각적으로 증강된 콘텐츠와 종이책의 콘텐츠와의 시각적으로 조화로운 증강에 대해 연구할 것이다.

### 참고 문헌

[1] 이영호, 하태진, 이형묵, 김기영, 우운택, “디지털로 북 - 아나로그 책과 디지털 콘텐츠의 융합,” 정보통신분야학회 합동학술대회, vol. 14, pp. 186-189, 2007.

[2] H. Kato and M. Billinghurst, “Marker tracking and hmd calibration for a video-based augmented reality conferencing system,” in Proc. 2nd IEEE and ACM International Workshop on Augmented Reality (IWAR '99), pp. 85-94, 20-21 Oct. 1999.

[3] F. Ledermann and D. Schmalstieg, “April: a high-level framework for creating augmented reality presentations,” in Proc. IEEE Virtual Reality VR 2005, pp. 187-194, 2005.

[4] 하태진, 이영호, 우운택, “디지털로 북 저작도구 artalet - 3 차원 객체 속성 저작,” 한국HCI 학회 학술지, vol. 1, pp. 314-318, 2008.

[5] 박중희, 김기영, 우운택, “디지털로 북 (digilog book) 을 위한 페이지 레이아웃 생성기법 및 응용,” 대한전자공학회 추계학술대회, vol. 31, pp. 961-962, 2008.

[6] OpenCV, <http://sourceforge.net/projects/opencvlibrary/>

[7] OSG, <http://www.openscenegraph.org/>

[8] Tesseract, <http://code.google.com/p/tesseract-ocr/>