
그림책 콘텐츠를 위한 전자책 변환기 설계

E-Book Converter Design for the Picture Book Contents

김유선, YooSun Kim*, 김현진, Hyunjin Kim**,
김우철, Woo-Cheol Kim***, 박상현, Sanghyun Park****

요약~ 휴대기기의 보급이 보편화되면서 전자책에 대한 관심이 증대되고 있다. 휴대기기 소지 연령 또한 낮아지면서 그림책이 좋은 전자책 콘텐츠가 될 것으로 예상되고 있다. 하지만 그림책은 일반 책과는 달리 글뿐만 아니라 그림도 중요한 정보이므로 전자책으로 만들려면 글과 함께 그림도 보존되어야 한다. 그런데 그림책 영상 한 쪽을 그대로 전자책 한 쪽으로 변환하면 휴대기기의 작은 화면으로 인해 글자들을 읽기가 힘들게 된다. 이러한 문제를 해결하기 위하여 본 논문에서는 그림책 영상에서 문자를 인식해 문자 부분을 확대하여 콘텐츠로 만드는 전자책 변환기를 설계하고자 한다. 이를 위해 먼저 그림책 영상 한 쪽에서 그림영역과 문자 영역을 분리하여 추출한다. 추출된 그림영역은 축소해 JPEG 형식으로 저장하고, 문자영역은 OCR 모듈을 통해 문자인식을 하여 문자 파일을 만든다. 이렇게 만든 JPEG 형식의 그림과 문자 파일, 그림과 문자의 배치 정보를 입력으로 하여 전자책으로 변환한다.

Abstract~ E-books are getting more popular with growing popularity of mobile devices to use. As children use mobile devices, picture books are expected to be good e-book contents. Picture books are different from the general books ; text is not only important information but also picture is. When converting into an e-book, a picture have to be preserved with text together. Then if a page of a picture book is converted into the page of an e-book, the letters from e-book are hard to read due to the small screens of mobile devices. In this paper to solve this problem, e-book converter which recognizes and expands the texts in a picture book image has designed. Firstly, a text region and a picture region is extracted from a picture book image. Secondly, OCR module recognize texts from a text region and a extracted picture region is resized and changed into JPEG format. Finally, these texts from OCR module, a JPEG format picture, and the placement information of texts and picture make E-book.

핵심어: 전자책, 그림책, Epub 형식, 휴대용 기기, 문자 추출, 광학문자인식.

Keywords: E-Book, Picture Book, Epub format, Mobile Device, Text Extraction, OCR.

본 논문은 중소기업청에서 지원하는 2009년도 산학협력실 지원사업(00035990)의 연구수행으로 인한 결과물임을 밝힙니다.

*주저자 : 연세대학교 컴퓨터과학과 박사과정 e-mail: yoursun@cs.yonsei.ac.kr

**공동저자 : 연세대학교 컴퓨터과학과 석사과정 e-mail: chriskim@cs.yonsei.ac.kr

***공동저자 : 연세대학교 컴퓨터과학과 박사후과정 e-mail: twelvepp@cs.yonsei.ac.kr

****교신저자 : 연세대학교 컴퓨터과학과 교수 e-mail: sanghyun@cs.yonsei.ac.kr

1. 서론

휴대전화· PDA· 휴대용게임기와 같은 휴대용 기기의 보급이 보편화되고 기능이 향상되면서 전자책(E-Book)의 활용도가 점점 높아져 가고 있으며 다양한 콘텐츠에 대한 요구도 증대되고 있다. 특히 휴대용 기기의 소지 연령이 낮아지고 있어 유아 및 아동을 위한 그림책 또한 전자책으로 많이 활용될 것으로 예상되고 있다.

그림책은 그림과 글자가 모두 중요한 정보로 전자책으로 변환시 글자와 함께 그림도 보존 되어야 한다. 그렇지만 스캐된 그림책 한 페이지를 그대로 전자책으로 바꾸게 되면 휴대용 기기에서는 글자가 너무 작아 읽기 힘들게 된다. 이 문제를 해결하기 위하여 본 논문에서는 스캐된 그림책 각 페이지에서 그림부분을 잘라내고 문자 영역을 추출하여[8] 광학문자인식(OCR)을 이용하여 문자를 인식한 뒤[1] 글자 크기를 조정하여 그림과 글자를 조합하여 전자책으로 자동 전환하는 변환기를 설계한다.

본 논문에서는 먼저 그림책 영상에 대해 살펴보고 전자책 표준인 ePub 형식(format)에 대해 정리한 후 제안된 전자책 변환기의 전체적인 구조를 설명하고 마지막으로 결론 및 향후 계획을 기술한다.

2. 그림책 영상

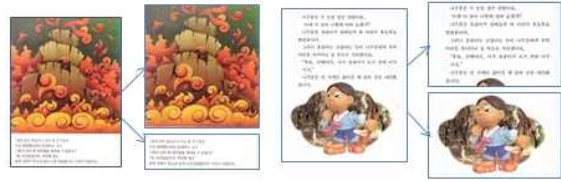
2.1 그림책 영상의 특성

그림책은 유아나 아동을 대상으로 하기 때문에 그림이 많은 부분을 차지할 뿐만 아니라 중요한 정보를 담고 있다. 이에 반해 글자 영역은 적은 부분을 차지하지만 일반적인 도큐먼트에 비해 글자가 크고 시각적으로 선명하게 분리되어 있으며, 많은 경우 회거나 옅은 색 바탕에 한 단어는 같은 색상을 사용하면서 글줄은 가로로 되어 있다. 독특한 디자인을 가지는 소수의 그림책을 제외한 대부분의 그림책들이 이와 같은 특징을 가진다.

2.2 그림책 영상의 분류

본 논문의 목적이 그림책 영상에서 그림영역과 문자영역을 잘라내어 재구성하는 것이므로 이러한 측면에서 그림책 영상을 분류해 보면 크게 그림 1과 같이 그림영역과 문자영역이 겹치는 경우와 그림 2에서 보듯이 그림영역과 문자영역이 겹치지 않는 경우로 나눌 수 있다. 그림 1은 그림과 글자를 상하 또는 좌우로 나누는 직선으로 구분 가능한 (a) Manhattan Layout과 이러한 직선으로 구분 불가능한 (b) Non-Manhattan Layout으로 구분할 수 있다[4].

그림영역과 문자영역이 겹치지 않는 경우는 그림부분과 문자 부분을 잘라낼 수 있으나 그림영역과 문자영역이 겹치는 경우는 문자영역을 추출할 수는 있으나 그림영역을 잘라내지는 못하고 전체영역을 그림영역으로 취급하여야 한다.



(ㄱ) 선으로 분리되는 경우 (ㄴ) 선으로 분리되지 않는 경우
(Manhattan Layout) (Non-Manhattan Layout)

그림 1. 그림영역과 글자영역이 겹치지 않는 경우



그림 2. 그림영역과 글자영역이 겹치는 경우

3. ePub Format

본 논문에서 설계한 전자책 변환기는 ePub 형식을 따르는 것으로 이는 IDPF(International Digital Publishing Forum)에서 제정한 표준으로 2.01 version까지 제정되어 있다[3]. ePub는 XHTML과 CSS, XML을 기반으로 한 Zip packaging format으로 Reflowable한 성질을 가지는데, 이는 다양한 사이즈의 휴대기기 화면에서 전자책 viewer를 통해 전자책을 볼 때 화면의 가로세로 비율 및 크기에 따라 자동적으로 화면에 맞추어 보여 주는 특성이다.

4. 그림책을 위한 E-Book 변환기

본 논문에서 설계한 E-Book 변환기의 개요는 그림 3과 같다.

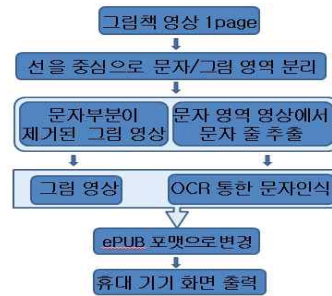


그림 3. 시스템 개요도

입력으로 받은 그림책 영상 한 쪽(page)에 대해 선을 중심으로 문자 영역과 그림 영역을 분리한다. 분리된 그림영역은 크기를 줄인 뒤 JPEG 형식으로 변환하고, 문자 영역은 줄 단위로 추출해 각각 영상을 만든 뒤 광학문자인식(OCR)을 통해 문자파일(text file)을 만든다. 이를 변환기에 입력으로 주어 ePub형식으로 변환하면, 휴대기기의 view가 세로 방향일 때와 가로 방향일 때, 화면 크기가 바뀔 때마다 배치가 달라지는 reflowable한 epub 파일이 만들어진다. 각 단계

별 결과는 그림 4에서 살펴볼 수 있다.

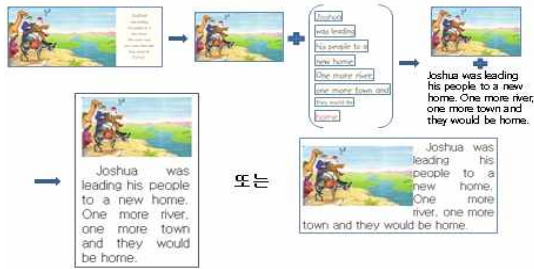


그림 4. 단계별 결과

4.1 문자 영역 추출

본 논문에서는 그림책 영상의 전자책 전환기의 설계에 중점을 두고 있으므로 문자 영역을 추출에 대해서는 가장 간단한 처리를 필요로 하는 그림책 종류인 선으로 그림영역과 문자영역이 분리되는 경우에 한정하여 문자 영역 추출한다. 이를 위해 몇 가지 전처리(preprocessing)를 해주어야 하는데, 먼저 칼라 영상을 그레이 영상으로 바꾸고 잡음을 제거하고 대조도를 증가(contrast stretch)시킨 뒤 반복적 이진화로 얻어진 문턱치(threshold)에 기반한 높은 문턱치를 이용하여 이진영상을 얻는다. 이제 (1),(2)와 같이 가로 투영(horizontal projection)과 세로 투영(vertical projection) [5]을 하고, 이 데이터를 분석하여 문자영역과 그림영역을 자를 수 있는 선을 찾고 이를 중심으로 그림영역과 문자영역을 분리한다.

$$\text{행 투영 } h_i = \text{행에 있는 255의 개수}, 0 \leq i \leq N-1 \quad (1)$$

$$\text{열 투영 } v_j = \text{열에 있는 255의 개수}, 0 \leq j \leq M-1 \quad (2)$$

그림영역으로 분리된 부분은 축소하여 JPEG 형식으로 저장하고, 문자영역으로 분리된 영역은 그 내부에서 다시 투영 정보를 이용하여 줄 단위로 문자부분을 영상으로 잘라내 저장한다[2][6].

4.2 광학문자인식(OCR)

줄 단위 문자영상들을 입력으로 받아 광학문자인식 모듈을 통해 하나의 문자 파일을 생성한다. 광학문자인식의 효율을 높이기 위해 문자영상들을 이진화해주고 광학문자인식 결과에 대해서도 띄어쓰기 보정과 같은 후처리를 해준다[7].

4.3 E-Book 변환기

위의 일련의 과정들로부터 생성된 JPEG 형식의 그림영상과 문자파일을 입력으로 받고 원래 그림책 영상에서의 그림영역과 문자영역의 배치관계를 입력으로 받아 epub형식의 전자책을 생성한다. 그림과 문자의 배치정보는 원래 그림책 형태를 가능한 보존하기 위해 포함하는 것으로 기본배치에는 영향을 줄 수 있으나 ePub형식 자체가 reflowable하므로 전자책을 보는 화면의 크기와 모양에 따라 화면배치가 자동으로 달라진다. 현재까지 사용되고 있는 전자책 변환기에서

는 글과 그림의 상하배치 정보는 줄 수 있었지만 좌우배치 정보는 줄 수 없게 되어 있는데, 제안된 전자책 변환기에서는 그림과 글의 좌우배치를 가능하게 하여 그림 4의 아래쪽처럼 휴대기기를 가로로 보느냐 세로로 보느냐에 따라 화면 활용도를 높이는 전자책이 생성된다.

5. 결론 및 향후 계획

본 논문에서는 그림책 영상에서 그림부분은 JPEG 형식으로 변환하고 문자 영역은 줄 단위로 추출하여 광학문자인식을 통해 문자파일을 만들어 ePub형태의 전자책 콘텐츠를 생성하는 전자책 변환기를 설계하였다. 이는 휴대용 기기에서도 그림이 보존되면서 문자를 읽는 데 불편하지 않을 뿐만 아니라 그림책의 좌우배치를 가능하게 하여 그림과 문자의 배치 정보를 주므로 그림책 원래의 형태를 가능한 보존할 수 있고 화면의 활용도를 높이는 전자책 콘텐츠를 생성할 수 있다. 하지만 문자 영역 추출은 가장 단순한 형태의 그림책 영상만을 대상으로 하여 기본적인 알고리즘으로 접근하였으므로 향후에는 다양한 그림책 영상에서도 문자 영역을 추출할 수 있는 알고리즘이 필요하다. 또한 제안된 변환기는 그림책 한 쪽을 대상으로 하였으므로 이를 여러 쪽으로 이루어진 책 한 권에 대해 수행하도록 확장하여야 한다.

6. 참고문헌

- [1] V. Wu, R. Manmatha and E.M. Riseman, Textfinder: an automatic system to detect and recognize text in images, IEEE Transactions on PAMI 20 (1999), pp. 1224-1229.
- [2] G. Nagy, S. Seth, "Hierarchical representation of optically scanned documents", Proceedings of the 7th International Conference on Pattern Recognition (ICPR), Montreal, Canada, 1984. pp. 347-349.
- [3] <http://www.idpf.org/>
- [4] Stefano Ferilli, Marenglen Biba, Floriana Esposito, Teresa M. A. Basile, A Distance-Based Technique for Non-Manhattan Layout Analysis, Proceedings of the 2009 10th International Conference on Document Analysis and Recognition, p.231-235, July 26-29, 2009
- [5] 오일석, 패턴인식, 교보문고, pp.266-268, 2009.
- [6] 박종경, 음봉규, 권용식, 진성아, "모바일 환경의 OCR Anyword", 한국콘텐츠학회 2006 춘계종합학술대회 논문집 Vol. 4 No. 1, pp.152-155, 2006
- [7] 강승식, "한글 문장의 자동 띄어쓰기", 한글 및 한국어 정보처리 학술대회, pp.137-142, 1998.
- [8] Y. J. Song, K. C. Kim, Y. W. Choi, H. R. Byun, S. H. Kim, S. Y. Chi, D. K. Jang, Y. K. Chung, "Text Region Extraction and Text Segmentation on Camera captured Document Style Images", ICARD '05, Vol.1, pp.172-176, 2005.