

(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국

(43) 국제공개일
2012년 10월 11일 (11.10.2012)



(10) 국제공개번호
WO 2012/138005 A1

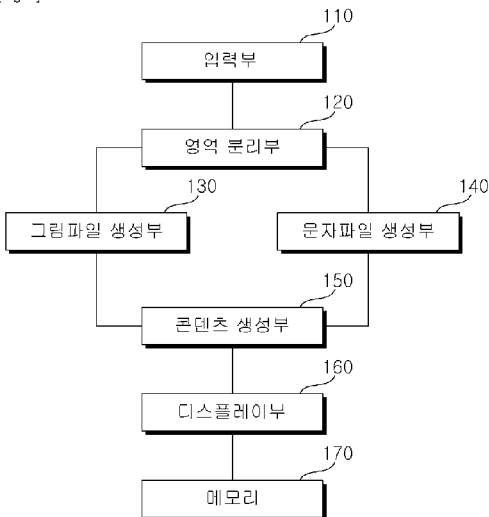
- (51) 국제특허분류: G06Q 50/00 (2006.01)
- (21) 국제출원번호: PCT/KR2011/002722
- (22) 국제출원일: 2011년 4월 15일 (15.04.2011)
- (25) 출원언어: 한국어
- (26) 공개언어: 한국어
- (30) 우선권정보: 10-2011-0030826 2011년 4월 4일 (04.04.2011) KR
- (71) 출원인 (US 을(를) 제외한 모든 지정국에 대하여): 연세대학교 산학협력단 (INDUSTRY-ACADEMIC CO-OPERATION FOUNDATION, YONSEI UNIVERSITY) [KR/KR]; 서울특별시 서대문구 신촌동 134, 120-749 Seoul (KR).
- (72) 발명자; 겸
- (75) 발명자/출원인 (US 에 한하여): 박상현 (PARK, Sang Hyun) [KR/KR]; 서울특별시 강남구 개포동 주공 1단지 120-503, 135-240 Seoul (KR). 김유선 (KIM, Yoo Sun) [KR/KR]; 서울특별시 중구 신당 4동 약수하이츠 아파트 111-301, 100-761 Seoul (KR). 안재균 (Ahn, Jae Gyoon) [KR/KR]; 서울특별시 서초구 서초동 1312-3 25/6 롯데캐슬클래식아파트 107-901, 137-778 Seoul (KR). 신은지 (SHIN, Eun Ji) [KR/KR]; 서울특별시 성북구 안암동 3가 133-48 도양빌 105, 136-073 Seoul (KR). 김현진 (KIM, Hyun Jin) [KR/KR]; 서울특별시 서대문구 홍은 1동 벽산아파트 103-303, 120-771 Seoul (KR).
- (74) 대리인: 특허법인 우인 (WOOIN, PATENT & LAW FIRM); 서울특별시 강남구 역삼동 648-15 신원빌딩 3층, 135-911 Seoul (KR).
- (81) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 국내 권리의 보호를 위하여): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 역내 권리의 보호를 위하여): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), 유라시아 (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), 유럽 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE,

[다음 쪽 계속]

(54) Title: METHOD AND APPARATUS FOR GENERATING E-BOOK CONTENTS

(54) 발명의 명칭 : 전자책 콘텐츠를 생성하기 위한 장치 및 그 방법

[Fig. 1]



- 110 ... Input unit
- 120 ... Area divider unit
- 130 ... Image file generation unit
- 140 ... Text file generation unit
- 150 ... Content generation unit
- 160 ... Display unit
- 170 ... Memory

(57) Abstract: According to the present invention, disclosed are a method and apparatus for generating e-book contents. According to the present invention, an apparatus for generating e-book contents comprises: an area divider unit dividing the input of the image page into an image area and a text area if a page image of an e-book is input; an image file generation unit generating an image file from the divided image area; a text file generation unit generating a text file from the divided text area through a text recognition; and a content generation unit combining the generated image file and the generated text file to generate one set of e-book content, wherein the content generation unit may arrange text included in the text file to generate the e-book content according to the format of the image file. Therefore, according to the present invention, better use can be made of a terminal screen displaying a page image of an image book, and reading text from a user's point of view can be made easier.

(57) 요약서: 본 발명에 의한 전자책(e-book) 콘텐츠를 생성하기 위한 장치 및 그 방법이 개시된다. 본 발명에 따른 전자책 콘텐츠를 생성하기 위한 장치는 그림책의 페이지 이미지를 입력받으면, 입력된 상기 페이지 이미지를 그림 영역과 문자 영역으로 분리하는 영역 분리부; 분리된 상기 그림 영역을 그림 파일로 생성하는 그림파일 생성부; 분리된 상기 문자 영역을 문자 인식을 통해 문자 파일로 생성하는 문자 파일 생성부; 및 생성된 상기 그림 파일과 상기 문자 파일을 결합하여 하나의 전자책 콘텐츠를 생성하는 콘텐츠 생성부를 포함하고, 상기 콘텐츠 생성부는 상기 그림 파일의 형태에 따라 상기 문자 파일에 포함된 문자를 배치하여 상기 전자책 콘텐츠를 생성할 수 있다. 이를 통해, 본 발명은 사용자 입장에서 문자 읽기가 수월해질 수 있을 뿐 아니라, 그림책의 페이지 이미지를 디스플레이하는 단말기 화면의 활용도를 높일 수 있다.

WO 2012/138005 A1



ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, **공개:**
MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK,
SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ,
GW, ML, MR, NE, SN, TD, TG).

— 국제조사보고서와 함께 (조약 제 21 조(3))

규칙 4.17 에 의한 선언서:

- 신규성을 해치지 아니하는 개시 또는 신규성 상실의 예외에 관한 선언 (규칙 4.17(v))

명세서

발명의 명칭: 전자책 콘텐츠를 생성하기 위한 장치 및 그 방법 기술분야

- [1] 본 발명은 전자책(e-book)에 관한 것으로, 특히, 그림책의 페이지 이미지에서 그림 영역과 문자 영역으로 분리하고 분리된 그림 영역과 문자 영역으로부터 그림 파일과 문자 파일을 생성하되, 문자 인식을 통한 문자 파일을 생성하여 그림 파일과 재배치하여 전자책 콘텐츠를 생성하기 위한 장치 및 그 방법에 관한 것이다.

배경기술

- [2] 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.
- [3] 그림책은 유아나 아동을 대상으로 하기 때문에 그림이 많은 부분을 차지할 뿐만 아니라 중요한 정보를 담고 있다. 이에 반해 글자 영역은 적은 부분을 차지하지만 일반적인 문서에 비해 글자가 크고 시각적으로 선명하게 분리되어 있으며, 많은 경우 회거나 옅은 색 바탕에 한 단어는 같은 색상을 사용하면서 글줄은 가로로 되어 있다.
- [4] 이러한 그림책은 그림과 글자가 모두 중요한 정보로 전자책으로 변환하는 경우에 글자와 함께 그림도 보존되어야 한다. 그렇지만 스캔된 그림책 한 페이지를 그대로 전자책으로 바꾸게 되면 휴대용 기기에서는 글자가 너무 작아 읽기 힘들게 된다.

발명의 상세한 설명

기술적 과제

- [5] 따라서 이러한 종래 기술의 문제점을 해결하기 위한 것으로, 본 발명의 목적은 그림책의 페이지 이미지에서 그림 영역과 문자 영역으로 분리하고 분리된 그림 영역과 문자 영역으로부터 그림 파일과 문자 파일을 생성하되, 문자 인식을 통한 문자 파일을 생성하여 그림 파일과 재배치하여 전자책 콘텐츠를 생성하기 위한 장치 및 그 방법을 제공하는데 있다.
- [6] 그러나 본 발명의 목적은 상기에 언급된 사항으로 제한되지 않으며, 언급되지 않은 또 다른 목적들은 아래의 기재로부터 당업자에게 명확하게 이해될 수 있을 것이다.

과제 해결 수단

- [7] 상기 목적들을 달성하기 위하여, 본 발명의 한 관점에 따른 전자책 콘텐츠를 생성하기 위한 장치는 그림책의 페이지 이미지를 입력받으면, 입력된 상기

페이지 이미지를 그림 영역과 문자 영역으로 분리하는 영역 분리부; 분리된 상기 그림 영역을 그림 파일로 생성하는 그림파일 생성부; 분리된 상기 문자 영역을 문자 인식을 통해 문자 파일로 생성하는 문자파일 생성부; 및 생성된 상기 그림 파일과 상기 문자 파일을 결합하여 하나의 전자책 콘텐츠를 생성하는 콘텐츠 생성부를 포함하고, 상기 콘텐츠 생성부는 상기 그림 파일의 형태에 따라 상기 문자 파일에 포함된 문자를 배치하여 상기 전자책 콘텐츠를 생성할 수 있다.

- [8] 바람직하게, 상기 영역 분리부는 상기 그림책의 페이지 이미지가 컬러이면, 컬러의 페이지 이미지를 그레이의 페이지 이미지로 변환하고, 변환된 상기 그레이의 페이지 이미지에 비등방성 확산필터를 적용하여 노이즈를 제거하며, 노이즈가 제거된 상기 페이지 이미지의 대조도를 증가시킨 후에 이진화를 통해 얻어 이진 데이터에 대해 행투영 데이터와 열투영 데이터를 획득하며, 획득한 상기 행투영 데이터와 상기 열투영 데이터를 분석하여 그림 영역과 그림 영역의 나눌 수 있는 기준선을 구하고, 구한 기준선을 기반으로 상기 그림 영역과 상기 문자 영역을 분리할 수 있다.
- [9] 이때, 상기 이진 데이터는 데이터가 존재하는 영역에서 1로 설정되고, 데이터가 존재하지 않는 영역에서 0으로 설정될 수 있다.
- [10] 이때, 상기 행투영 데이터는 가로축으로 이진 데이터가 1로 설정된 영역의 개수를 나타내고, 상기 열투영 데이터는 세로축으로 이진 데이터가 1로 설정된 영역의 개수를 나타낼 수 있다.
- [11] 필요에 따라, 상기 영역 분리부는 상기 그림 영역과 상기 문자 영역이 일부 겹쳐져 있는 경우에는 각기 다른 기준선을 기반으로 그림 영역과 문자 영역을 분리할 수 있다.
- [12] 필요에 따라, 상기 영역 분리부는 상기 그림 영역과 상기 문자 영역이 전부 겹쳐져 있는 경우에는 문자 영역을 분리한 후 전체 영역을 그림 영역으로 취할 수 있다.
- [13] 바람직하게, 상기 문자파일 생성부는 상기 문자 영역으로부터 줄 단위로 문자 영역을 추출하고, 그 추출된 문자 영역 각각을 OCR(Optical Character Recognition) 기법을 통해 문자를 인식하여 하나의 문자 파일로 생성할 수 있다.
- [14] 바람직하게, 상기 콘텐츠 생성부는 상기 전자책 콘텐츠를 화면의 가로세로 비율 및 크기에 따라 자동적으로 화면에 맞추어 보여주는 자동공간조정이 가능한 특성을 갖는 ePub(electronic Publication) 포맷으로 생성할 수 있다.
- [15]
- [16] 본 발명의 다른 한 관점에 따른 전자책 콘텐츠를 생성하기 위한 방법은 (a) 그림책의 페이지 이미지를 입력받으면, 입력된 상기 페이지 이미지를 그림 영역과 문자 영역으로 분리하는 단계; (b) 분리된 상기 그림 영역을 그림 파일로 생성하는 단계; (c) 분리된 상기 문자 영역을 문자 인식을 통해 문자 파일로 생성하는 단계; 및 (d) 생성된 상기 그림 파일과 상기 문자 파일을 결합하여 하나의 전자책 콘텐츠를 생성하는 단계를 포함하고, 상기 (d) 단계는 상기 그림

파일의 형태에 따라 상기 문자 파일에 포함된 문자를 배치하여 상기 전자책 콘텐츠를 생성할 수 있다.

- [17] 바람직하게, 상기 (a) 단계는 상기 그림책의 페이지 이미지가 컬러이면, 컬러의 페이지 이미지를 그레이의 페이지 이미지로 변환하고, 변환된 상기 그레이의 페이지 이미지에 비등방성 확산필터를 적용하여 노이즈를 제거하며, 노이즈가 제거된 상기 페이지 이미지의 대조도를 증가시킨 후에 이진화를 통해 얻어 이진 데이터에 대해 행투영 데이터와 열투영 데이터를 획득하며, 획득한 상기 행투영 데이터와 상기 열투영 데이터를 분석하여 그림 영역과 그림 영역의 나눌 수 있는 기준선을 구하고, 구한 기준선을 기반으로 상기 그림 영역과 상기 문자 영역을 분리할 수 있다.
- [18] 이때, 상기 이진 데이터는 데이터가 존재하는 영역에서 1로 설정되고, 데이터가 존재하지 않는 영역에서 0으로 설정될 수 있다.
- [19] 이때, 상기 행투영 데이터는 가로축으로 이진 데이터가 1로 설정된 영역의 개수를 나타내고, 상기 열투영 데이터는 세로축으로 이진 데이터가 1로 설정된 영역의 개수를 나타낼 수 있다.
- [20] 필요에 따라, 상기 (a) 단계는 상기 그림 영역과 상기 문자 영역이 일부 겹쳐져 있는 경우에는 각기 다른 기준선을 기반으로 그림 영역과 문자 영역을 분리할 수 있다.
- [21] 필요에 따라, 상기 (a) 단계는 상기 그림 영역과 상기 문자 영역이 전부 겹쳐져 있는 경우에는 문자 영역을 분리한 후 전체 영역을 그림 영역으로 취할 수 있다.
- [22] 바람직하게, 상기 (c) 단계는 상기 문자 영역으로부터 줄 단위로 문자 영역을 추출하고, 그 추출된 문자 영역 각각을 OCR(Optical Character Recognition) 기법을 통해 문자를 인식하여 하나의 문자 파일로 생성할 수 있다.
- [23] 바람직하게, 상기 (d) 단계는 상기 전자책 콘텐츠를 화면의 가로세로 비율 및 크기에 따라 자동적으로 화면에 맞추어 보여주는 자동공간조정이 가능한 특성을 갖는 ePub(electronic Publication) 포맷으로 생성할 수 있다.

발명의 효과

- [24] 이를 통해, 본 발명은 그림책의 페이지 이미지에서 그림 영역과 문자 영역으로 분리하고 분리된 그림 영역과 문자 영역으로부터 그림 파일과 문자 파일을 생성하되, 문자 인식을 통한 문자 파일을 생성하여 그림 파일과 재배치함으로써, 사용자 입장에서 문자 읽기가 수월해질 수 있는 효과가 있다.
- [25] 또한, 본 발명은 그림책의 페이지 이미지에서 그림 영역과 문자 영역으로 분리하고 분리된 그림 영역과 문자 영역으로부터 그림 파일과 문자 파일을 생성하되, 문자 인식을 통한 문자 파일을 생성하여 그림 파일과 재배치함으로써, 그림책의 페이지 이미지를 디스플레이하는 단말기 화면의 활용도를 높일 수 있는 효과가 있다.

도면의 간단한 설명

- [26] 도 1은 본 발명의 실시예에 따른 전자책 콘텐츠를 생성하기 위한 장치를 나타내는 예시도이다.
- [27] 도 2는 본 발명의 실시예에 따른 그림 영역과 문자 영역을 분리하는 원리를 설명하기 위한 예시도이다.
- [28] 도 3은 본 발명에 따른 문자 파일을 생성하는 원리를 설명하기 위한 예시도이다.
- [29] 도 4는 본 발명의 실시예에 따른 전자책 콘텐츠의 유형을 나타내는 예시도이다.
- [30] 도 5는 본 발명의 실시예에 따른 전자책 콘텐츠를 생성하는 원리를 설명하기 위한 예시도이다.
- [31] 도 6은 본 발명의 실시예에 따른 전자책 콘텐츠를 생성하기 위한 방법을 나타내는 예시도이다.

발명의 실시를 위한 형태

- [32] 이하에서는, 본 발명의 실시예에 따른 전자책(e-book) 콘텐츠를 생성하기 위한 장치 및 그 방법을 첨부한 도 1 내지 도 6을 참조하여 설명한다. 본 발명에 따른 동작 및 작용을 이해하는데 필요한 부분을 중심으로 상세히 설명한다. 명세서 전체를 통하여 각 도면에서 제시된 동일한 참조 부호는 동일한 구성 요소를 나타낸다.
- [33] 본 발명에서는 그림책의 페이지 이미지에서 그림 영역과 문자 영역으로 분리하고 분리된 그림 영역과 문자 영역으로부터 그림 파일과 문자 파일을 생성하되, 문자 인식을 통한 문자 파일을 생성하여 그림 파일과 재배치하여 전자책 콘텐츠를 생성할 수 있는 방안을 제안한다. 여기서, 그림책의 페이지 이미지는 그림책의 각 페이지를 스캔하거나 캡처하여 생성된 이미지를 의미할 수 있다.
- [34] 여기서 전자책(e-book)이란 지식과 정보의 콘텐츠를 종이 매체를 이용하지 않고 디지털 형태로 출간된 책을 가리키는 것으로, 이와 같은 전자책은 전자책을 읽을 수도 있도록 하는 프로그램이 탑재된 PC나 전용 휴대용 단말기 등을 통해 볼 수 있으며, 이와 같은 장치를 통칭해 전자책으로 칭하기도 한다. 즉, 넓은 의미에서 콘텐츠가 포함된 모든 소프트웨어적, 하드웨어적인 미디어물 전체를 말할 수 있는 것이다.
- [35]
- [36] 도 1은 본 발명의 실시예에 따른 전자책 콘텐츠를 생성하기 위한 장치를 나타내는 예시도이다.
- [37] 도 1에 도시한 바와 같이, 본 발명에 따른 전자책 콘텐츠를 생성하기 위한 장치는 입력부(110), 영역 분리부(120), 그림파일 생성부(130), 문자파일 생성부(140), 콘텐츠 생성부(150), 디스플레이부(160), 및 메모리(170) 등을 포함하여 구성될 수 있다. 이렇게 구성된 장치는 휴대폰, PDA(Personal Digital Assistants), 및 노트북 등의 휴대 기기에 적용될 수 있다.

- [38] 입력부(110)는 유선/무선 인터넷 망을 통해 그림책의 페이지 이미지를 입력받을 수 있다. 여기서, 그림책의 페이지 이미지는 그림책의 각 페이지마다 스캔되거나 캡처된 이미지를 의미하는데, 칼라(color) 이미지이거나 그레이(gray) 이미지일 수 있다.
- [39] 이렇게 입력받은 그림책마다 그 페이지 이미지는 메모리(170)에 저장될 수 있다.
- [40] 영역 추출부(120)는 입력된 그림책의 페이지 이미지를 그림 영역과 문자 영역으로 분리할 수 있다. 이를 상세히 설명하면 다음과 같다.
- [41] 영역 추출부(120)는 그림책의 페이지 이미지가 컬러이면, 컬러의 페이지 이미지를 그레이의 페이지 이미지로 변환할 수 있다. 영역 추출부(120)는 변환된 그레이의 페이지 이미지에 비등방성 확산필터(anisotropic diffusion filter)를 적용하여 노이즈를 제거할 수 있다. 이때, 노이즈는 원래의 그림책 각 페이지에 있던 노이즈일 수 있고, 그림책을 스캔 또는 캡처하는 과정에서 생긴 노이즈일 수도 있다.
- [42] 영역 추출부(120)는 노이즈가 제거된 페이지 이미지의 대조도(contrast)를 증가시킨 후에 문턱치에 기반한 반복적 이진화로 이진 데이터를 얻을 수 있다. 예컨대, 데이터가 존재하는 영역은 '1', 데이터가 존재하지 않는 영역은 '0'으로 설정될 수 있다.
- [43] 영역 추출부(120)는 이렇게 얻은 이진 데이터에 대해 가로 또는 행투영 데이터와 세로 또는 열투영 데이터를 다음의 [수학식 1]을 이용하여 획득할 수 있다.
- [44] [수학식 1]
- [45] 행투영 데이터 h_i = 행 i 에 있는 255의 개수, $0 \leq i \leq N-1$
- [46] 열투영 데이터 v_j = 열 j 에 있는 255의 개수, $0 \leq j \leq M-1$
- [47] 여기서, 행투영 데이터는 가로축으로 이진 데이터가 '1'로 설정된 영역의 개수를 나타내고, 열투영 데이터는 세로축으로 이진 데이터가 '1'로 설정된 영역의 개수를 나타낼 수 있다.
- [48] 영역 추출부(120)는 획득한 행투영 데이터와 열투영 데이터를 분석하여 그림 영역과 그림 영역의 나눌 수 있는 기준선 즉, 그림 영역의 경계선을 구할 수 있다. 즉, 연속된 흰부분이 차지하는 비율에 따라 가로축 또는 세로축의 기준선을 구할 수 있다.
- [49] 영역 추출부(120)는 구한 기준선을 기반으로 그림 영역과 문자 영역을 분리할 수 있다.
- [50]
- [51] 이때, 그림 영역과 문자 영역이 확연히 구분되어 겹치지 않는 경우도 있을 수 있지만 그림 영역과 문자 영역이 일부 겹치는 경우도 있을 수 있다. 이를 도 2를 참조하여 설명한다.
- [52] 도 2는 본 발명의 실시예에 따른 그림 영역과 문자 영역을 분리하는 원리를

설명하기 위한 예시도이다.

[53] 도 2에 도시한 바와 같이, 그림 (a)에서는 그림 영역과 문자 영역이 겹치지 않는 경우에 기준선을 기반으로 그림 영역과 문자 영역을 분리할 수 있다. 그림 (b)에서는 그림 영역과 문자 영역이 일부 겹쳐져 있는 경우에는 각기 다른 기준선을 기반으로 그림 영역과 문자 영역을 분리할 수 있다.

[54] 또한, 그림 (c) 내지 그림 (d)에서는 그림 영역과 문자 영역이 전부 겹쳐져 있기 때문에 문자 영역만을 분리한 후에 전체 영역을 그림 영역으로 취급할 수 있다.

[55]

[56] 그림파일 생성부(130)는 분리된 그림 영역을 JPEG(Joint Photographic coding Experts Group), GIF(Graphics Interchange Format), BMP(Bitmap image), PNG(Portable Network Graphics) 등의 포맷으로 변환하여 그림 파일로 생성할 수 있다.

[57] 문자파일 생성부(140)는 분리된 문자 영역을 OCR(Optical Character Recognition) 기법을 통해 문자 파일 예컨대, 텍스트 파일로 생성할 수 있다. 이를 도 3을 참조하여 설명한다.

[58]

[59] 도 3은 본 발명에 따른 문자 파일을 생성하는 원리를 설명하기 위한 예시도이다.

[60] 도 3에 도시한 바와 같이, 문자파일 생성부(140)는 분리된 문자 영역으로부터 줄(row) 단위로 문자 영역을 추출하고, 그 줄 단위로 추출된 문자 영역 각각을 OCR 기법을 통해 문자를 인식하여 하나의 문자 파일로 생성할 수 있다.

[61]

[62] 콘텐츠 생성부(150)는 생성된 그림 파일과 문자 파일을 결합하여 하나의 전자책 콘텐츠를 생성할 수 있다. 즉, 콘텐츠 생성부(150)는 그림 파일과 문자 파일을 결합하되, 그림 파일의 형태에 따라 문자 파일에 포함된 문자를 배치하여 전자책 콘텐츠를 생성할 수 있다. 여기서, 전자책 콘텐츠는 ePub(electronic Publication), pdf, azw 등의 포맷일 수 있는데, ePub 포맷을 사용하는 것이 바람직하다.

[63] ePub 포맷은 IDPF(International Digital Publishing Forum)에서 제정한 표준으로, XHTML, CSS, XML을 기반으로 한 Zip packaging format으로 자동공간조정이 가능한 성질을 가지는데, 이는 다양한 사이즈의 휴대 기기 화면에서 전자책 뷰어(viewer)를 통해 전자책을 볼 때 화면의 가로세로 비율 및 크기에 따라 자동적으로 화면에 맞추어 보여주는 특성이다.

[64] 도 4는 본 발명의 실시예에 따른 전자책 콘텐츠의 유형을 나타내는 예시도이다.

[65] 도 4에 도시한 바와 같이, 그림 (a)에서는 문자가 위쪽, 그림이 아래쪽에 배치되는 형태를 나타내고, 그림 (b)에서는 그림이 위쪽, 문자가 아래쪽에 배치되는 형태를 나타내며, 그림 (c)에서는 문자가 왼쪽, 그림이 오른쪽에 배치되는 형태를 나타내며, 그림 (d)에서는 그림이 왼쪽, 문자가 오른쪽에

- 배치되는 형태를 나타낼 수 있다.
- [66] 예컨대, 단말기의 화면 크기가 $7 \times 13 \text{ cm}^2$ 인 경우에는 그림은 7×7 의 크기로 하단에 배치되고, 문자는 7×6 의 크기로 상단에 배치될 수 있다.
- [67] 또한, 콘텐츠 생성부(150)는 생성된 그림 파일과 문자 파일을 결합하여 하나의 전자책 콘텐츠를 생성하되, 모든 유형의 전자책 콘텐츠를 생성하게 된다. 그래서 사용자가 단말기의 화면을 세로로 보게 되면 그림 (a)나 (b)처럼 화면이 디스플레이 되고, 가로로 보게 되면 그림 (c)나 (d)처럼 화면이 디스플레이 될 수 있다.
- [68] 이때, 실제 그림책의 문자와 그림보다 단말기의 화면이 작기 때문에 스크롤을 추가하여 그림은 고정 되고 문자는 그 내용에 따라 스크롤을 통해 볼 수 있게 된다.
- [69]
- [70] 도 5는 본 발명의 실시예에 따른 전자책 콘텐츠를 생성하는 원리를 설명하기 위한 예시도이다.
- [71] 도 5에 도시한 바와 같이, 콘텐츠 생성부(150)는 그림 파일과 문자 파일을 결합하되, 그림 파일의 형태에 따라 문자 파일에 포함된 문자를 배치하여 전자책 콘텐츠를 생성할 수 있다.
- [72]
- [73] 도 6은 본 발명의 실시예에 따른 전자책 콘텐츠를 생성하기 위한 방법을 나타내는 예시도이다.
- [74] 도 6에 도시한 바와 같이, 본 발명에 따른 전자책 콘텐츠 생성 장치는 그림책의 다수의 페이지 이미지를 입력받으면(S610), 입력된 상기 페이지 이미지를 그림 영역과 문자 영역으로 분리할 수 있다(S611).
- [75] 이를 상세히 설명하면, 전자책 콘텐츠 생성 장치는 그림책의 페이지 이미지가 컬러이면, 컬러의 페이지 이미지를 그레이의 페이지 이미지로 변환할 수 있다. 전자책 콘텐츠 생성 장치는 변환된 그레이의 페이지 이미지에 비등방성 확산필터(anisotropic diffusion filter)를 적용하여 노이즈를 제거하고, 노이즈가 제거된 페이지 이미지의 대조도(contrast)를 증가시킨 후에 문턱치에 기반한 반복적 이진화로 이진 데이터를 얻을 수 있다. 전자책 콘텐츠 생성 장치는 이렇게 얻은 이진 데이터에 대해 가로 또는 행투영 데이터와 세로 또는 열투영 데이터를 획득할 수 있다. 전자책 콘텐츠 생성 장치는 획득한 행투영 데이터와 열투영 데이터를 분석하여 그림 영역과 그림 영역의 나눌 수 있는 기준선 즉, 그림 영역의 경계선을 구하고, 구한 기준선을 기반으로 그림 영역과 문자 영역을 분리할 수 있다.
- [76] 다음으로, 전자책 콘텐츠 생성 장치는 분리된 상기 그림 영역을 그림 파일로 생성하고(S612), 분리된 문자 영역을 OCR(Optical Character Recognition) 기법을 통해 문자 파일로 생성할 수 있다.
- [77] 이때, 전자책 콘텐츠 생성 장치는 분리된 문자 영역으로부터 줄 단위로 문자

영역을 추출하고(S613), 그 줄 단위로 추출된 문자 영역 각각을 OCR 기법을 통해 문자를 인식하여(S614) 하나의 문자 파일로 생성할 수 있다(S615).

[78] 다음으로, 전자책 콘텐츠 생성 장치는 이렇게 생성된 그림 파일과 문자 파일을 미리 설정된 결합 형태에 따라 결합하여 하나의 전자책 콘텐츠를 생성할 수 있다(S616).

[79]

[80] 본 발명에 의한 전자책 콘텐츠를 생성하기 위한 장치 및 그 방법이 속하는 기술 분야에서 통상의 지식을 가진 자라면 본 발명의 본질적인 특성에서 벗어나지 않는 범위에서 다양한 수정 및 변형이 가능할 것이다. 따라서, 본 발명에 개시된 실시예들은 본 발명의 기술 사상을 한정하기 위한 것이 아니라 설명하기 위한 것이고, 이러한 실시예에 의하여 본 발명의 기술 사상의 범위가 한정되는 것은 아니다. 본 발명의 보호 범위는 아래의 청구범위에 의하여 해석되어야 하며, 그와 동등한 범위 내에 있는 모든 기술 사상은 본 발명의 권리범위에 포함되는 것으로 해석되어야 할 것이다.

[81]

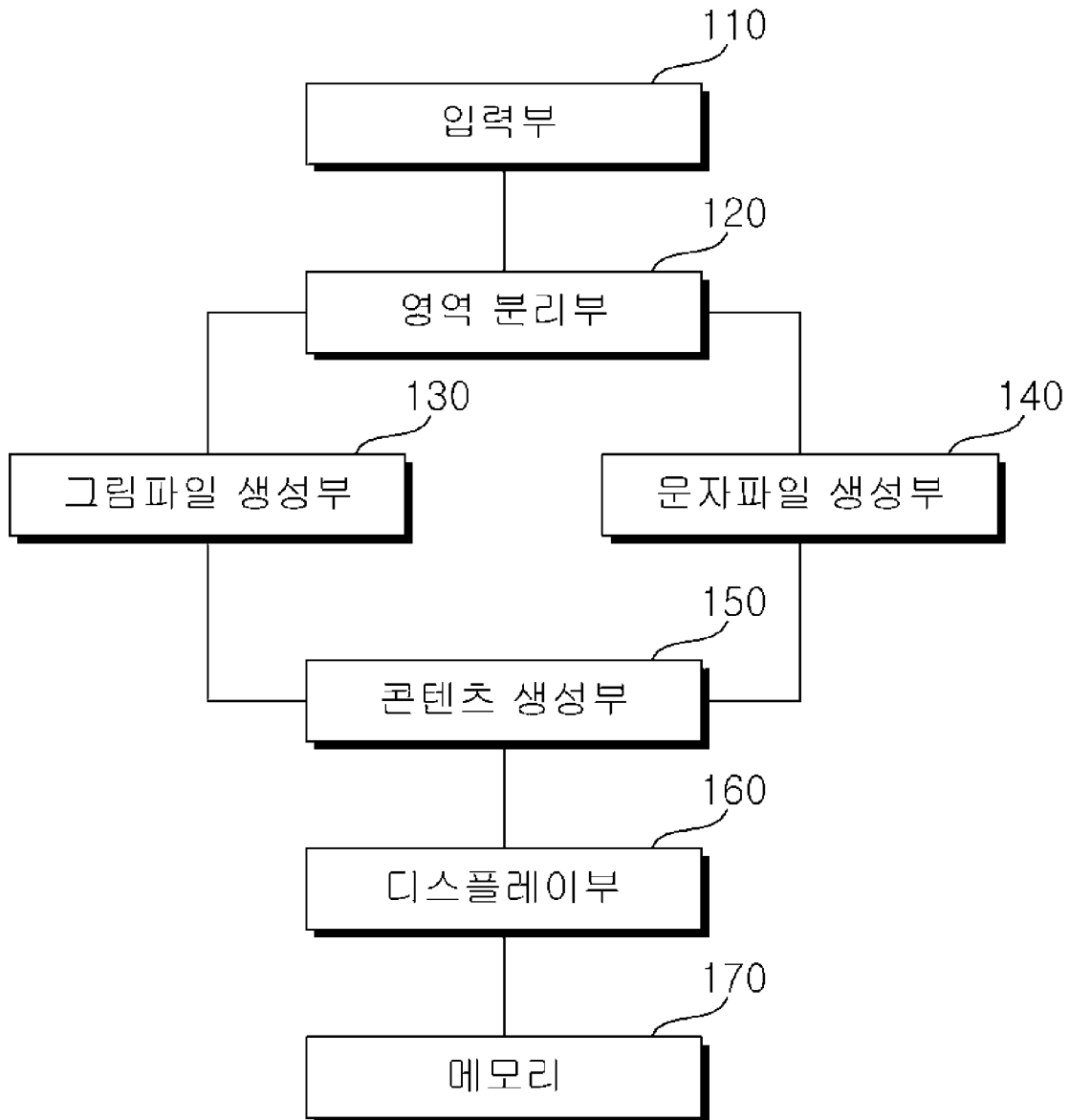
청구범위

- [청구항 1] 그림책의 페이지 이미지를 입력받으면, 입력된 상기 페이지 이미지를 그림 영역과 문자 영역으로 분리하는 영역 분리부; 분리된 상기 그림 영역을 그림 파일로 생성하는 그림파일 생성부; 분리된 상기 문자 영역을 문자 인식을 통해 문자 파일로 생성하는 문자파일 생성부; 및 생성된 상기 그림 파일과 상기 문자 파일을 결합하여 하나의 전자책 콘텐츠를 생성하는 콘텐츠 생성부; 를 포함하고, 상기 콘텐츠 생성부는 상기 그림 파일의 형태에 따라 상기 문자 파일에 포함된 문자를 배치하여 상기 전자책 콘텐츠를 생성하는 전자책 콘텐츠를 생성하기 위한 장치.
- [청구항 2] 제1 항에 있어서, 상기 영역 분리부는, 상기 그림책의 페이지 이미지가 컬러이면, 컬러의 페이지 이미지를 그레이의 페이지 이미지로 변환하고, 변환된 상기 그레이의 페이지 이미지에 비등방성 확산필터를 적용하여 노이즈를 제거하며, 노이즈가 제거된 상기 페이지 이미지의 대조도를 증가시킨 후에 이진화를 통해 얻어 이진 데이터에 대해 행투영 데이터와 열투영 데이터를 획득하며, 획득한 상기 행투영 데이터와 상기 열투영 데이터를 분석하여 그림 영역과 그림 영역의 나눌 수 있는 기준선을 구하고, 구한 기준선을 기반으로 상기 그림 영역과 상기 문자 영역을 분리하는 것을 특징으로 하는 전자책 콘텐츠를 생성하기 위한 장치.
- [청구항 3] 제2 항에 있어서, 상기 이진 데이터는, 데이터가 존재하는 영역에서 1로 설정되고, 데이터가 존재하지 않는 영역에서 0으로 설정되는 것을 특징으로 하는 전자책 콘텐츠를 생성하기 위한 장치.
- [청구항 4] 제3 항에 있어서, 상기 행투영 데이터는 가로축으로 이진 데이터가 1로 설정된 영역의 개수를 나타내고, 상기 열투영 데이터는 세로축으로 이진 데이터가 1로 설정된 영역의 개수를 나타내는 것을 특징으로 하는 전자책 콘텐츠를 생성하기 위한 장치.
- [청구항 5] 제2 항에 있어서, 상기 영역 분리부는, 상기 그림 영역과 상기 문자 영역이 일부 겹쳐져 있는 경우에는

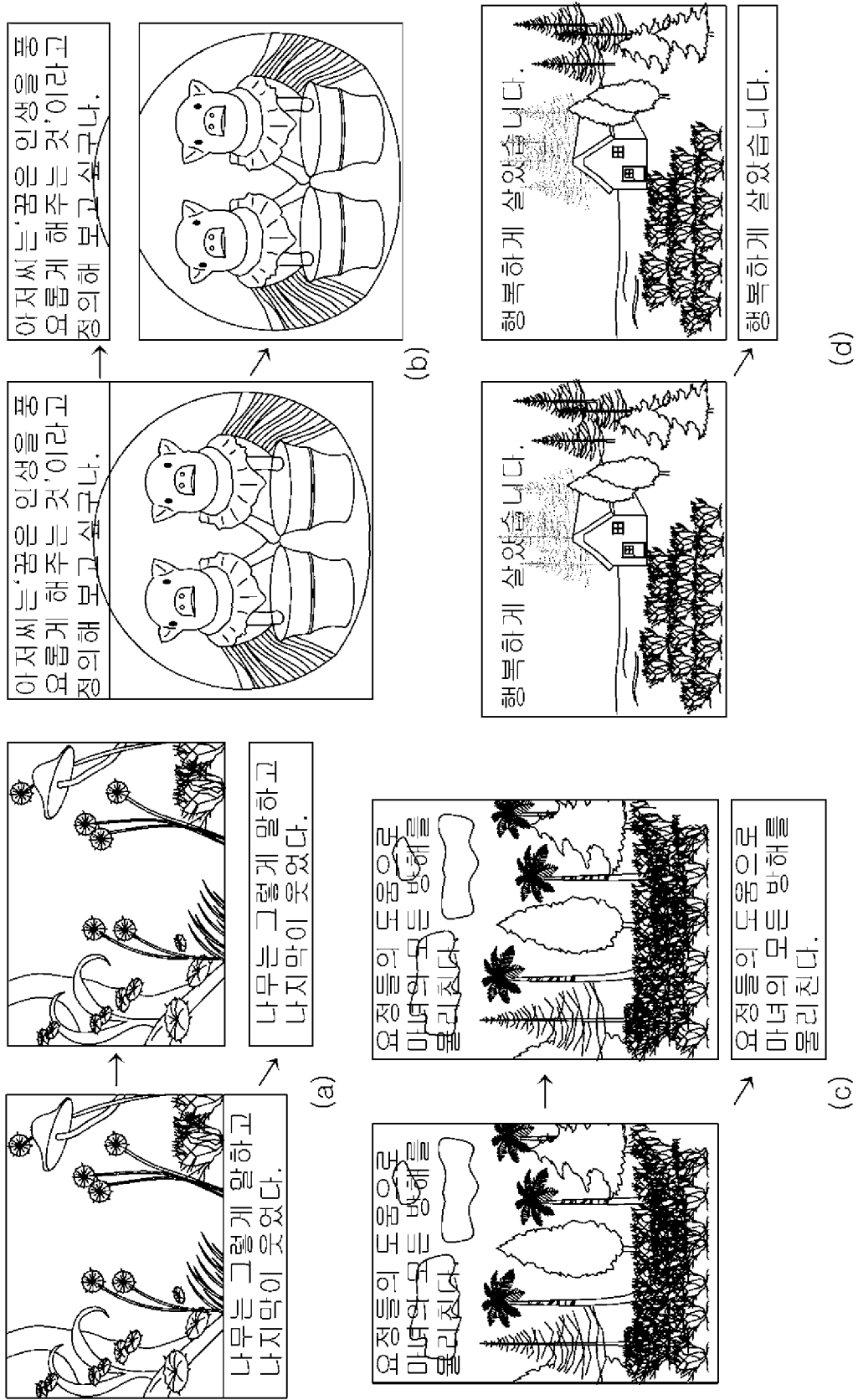
- 각기 다른 기준선을 기반으로 그림 영역과 문자 영역을 분리하는 것을 특징으로 하는 전자책 콘텐츠를 생성하기 위한 장치.
- [청구항 6] 제2 항에 있어서,
상기 영역 분리부는,
상기 그림 영역과 상기 문자 영역이 전부 겹쳐져 있는 경우에는 문자 영역을 분리한 후 전체 영역을 그림 영역으로 취하는 것을 특징으로 하는 전자책 콘텐츠를 생성하기 위한 장치.
- [청구항 7] 제1 항에 있어서,
상기 문자파일 생성부는,
상기 문자 영역으로부터 줄 단위로 문자 영역을 추출하고, 그 추출된 문자 영역 각각을 OCR(Optical Character Recognition) 기법을 통해 문자를 인식하여 하나의 문자 파일로 생성하는 것을 특징으로 하는 전자책 콘텐츠를 생성하기 위한 장치.
- [청구항 8] 제1 항에 있어서,
상기 콘텐츠 생성부는,
상기 전자책 콘텐츠를 화면의 가로세로 비율 및 크기에 따라 자동적으로 화면에 맞추어 보여주는 자동공간조정이 가능한 특성을 갖는 ePub(electronic Publication) 포맷으로 생성하는 것을 특징으로 하는 전자책 콘텐츠를 생성하기 위한 장치.
- [청구항 9] (a) 그림책의 페이지 이미지를 입력받으면, 입력된 상기 페이지 이미지를 그림 영역과 문자 영역으로 분리하는 단계;
(b) 분리된 상기 그림 영역을 그림 파일로 생성하는 단계;
(c) 분리된 상기 문자 영역을 문자 인식을 통해 문자 파일로 생성하는 단계; 및
(d) 생성된 상기 그림 파일과 상기 문자 파일을 결합하여 하나의 전자책 콘텐츠를 생성하는 단계;
를 포함하고, 상기 (d) 단계는 상기 그림 파일의 형태에 따라 상기 문자 파일에 포함된 문자를 배치하여 상기 전자책 콘텐츠를 생성하는 전자책 콘텐츠를 생성하기 위한 방법.
- [청구항 10] 제9 항에 있어서,
상기 (a) 단계는,
상기 그림책의 페이지 이미지가 컬러이면, 컬러의 페이지 이미지를 그레이의 페이지 이미지로 변환하고, 변환된 상기 그레이의 페이지 이미지에 비등방성 확산필터를 적용하여 노이즈를 제거하며,
노이즈가 제거된 상기 페이지 이미지의 대조도를 증가시킨 후에 이진화를 통해 얻어 이진 데이터에 대해 행투영 데이터와 열투영 데이터를 획득하며,

- 획득한 상기 행투영 데이터와 상기 열투영 데이터를 분석하여 그림 영역과 그림 영역의 나눌 수 있는 기준선을 구하고, 구한 기준선을 기반으로 상기 그림 영역과 상기 문자 영역을 분리하는 것을 특징으로 하는 전자책 콘텐츠를 생성하기 위한 방법.
- [청구항 11] 제10 항에 있어서,
상기 이진 데이터는,
데이터가 존재하는 영역에서 1로 설정되고, 데이터가 존재하지 않는 영역에서 0으로 설정되는 것을 특징으로 하는 전자책 콘텐츠를 생성하기 위한 방법.
- [청구항 12] 제11 항에 있어서,
상기 행투영 데이터는 가로축으로 이진 데이터가 1로 설정된 영역의 개수를 나타내고, 상기 열투영 데이터는 세로축으로 이진 데이터가 1로 설정된 영역의 개수를 나타내는 것을 특징으로 하는 전자책 콘텐츠를 생성하기 위한 방법.
- [청구항 13] 제10 항에 있어서,
상기 (a) 단계는,
상기 그림 영역과 상기 문자 영역이 일부 겹쳐져 있는 경우에는 각기 다른 기준선을 기반으로 그림 영역과 문자 영역을 분리하는 것을 특징으로 하는 전자책 콘텐츠를 생성하기 위한 방법.
- [청구항 14] 제10 항에 있어서,
상기 (a) 단계는,
상기 그림 영역과 상기 문자 영역이 전부 겹쳐져 있는 경우에는 문자 영역을 분리한 후 전체 영역을 그림 영역으로 취하는 것을 특징으로 하는 전자책 콘텐츠를 생성하기 위한 방법.
- [청구항 15] 제9 항에 있어서,
상기 (c) 단계는,
상기 문자 영역으로부터 줄 단위로 문자 영역을 추출하고, 그 추출된 문자 영역 각각을 OCR(Optical Character Recognition) 기법을 통해 문자를 인식하여 하나의 문자 파일로 생성하는 것을 특징으로 하는 전자책 콘텐츠를 생성하기 위한 방법.
- [청구항 16] 제9 항에 있어서,
상기 (d) 단계는,
상기 전자책 콘텐츠를 화면의 가로세로 비율 및 크기에 따라 자동적으로 화면에 맞추어 보여주는 자동공간조정이 가능한 특성을 갖는 ePub(electronic Publication) 포맷으로 생성하는 것을 특징으로 하는 전자책 콘텐츠를 생성하기 위한 방법.

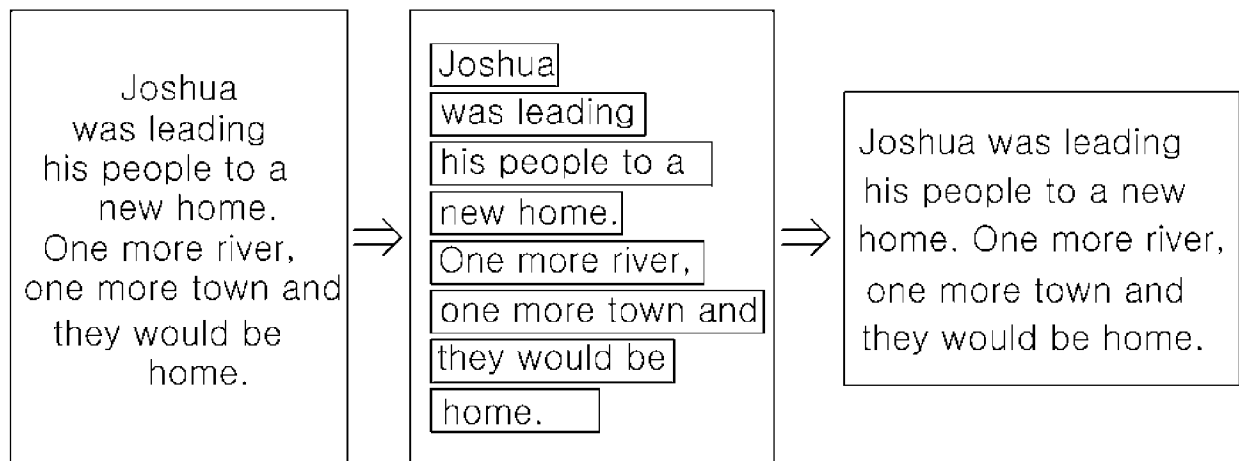
[Fig. 1]



[Fig. 2]



[Fig. 3]



[Fig. 4]

Simon Baker as Patrick Jane? th
protagonist and the titular Men
psychic until Red John murdered
and daughter in retaliation for



(a)



It has been suggested that this
section be merged into Mentalis
This article is about the perfo
other uses of the word, see men
(disambiguation). Theatrical po

(b)

The show's real-time nature
is emphasized by an on-screen
digital clock which appears
before and after commercial
breaks. Smaller, silent,
clock displays also appear at
other points in the narrative
between commercial breaks. The
time shown is the in-universe



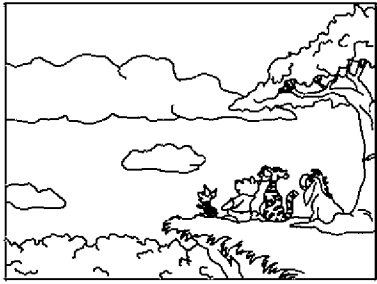
(c)



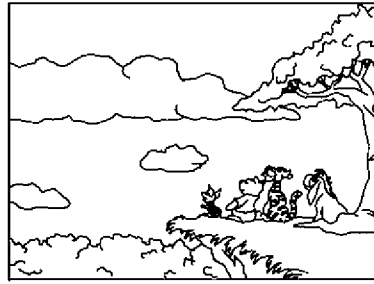
The show has won numerous awards,
including Best Drama Series at the
2003 Golden Globe Awards and
Outstanding Drama Series at the
2006 Primetime Emmy Awards. Kiefer
Sutherland has been awarded Best
Actor? Television Series Drama at
the 2001 Golden Globe Awards, and
Outstanding Lead Actor? Television

(d)

[Fig. 5]

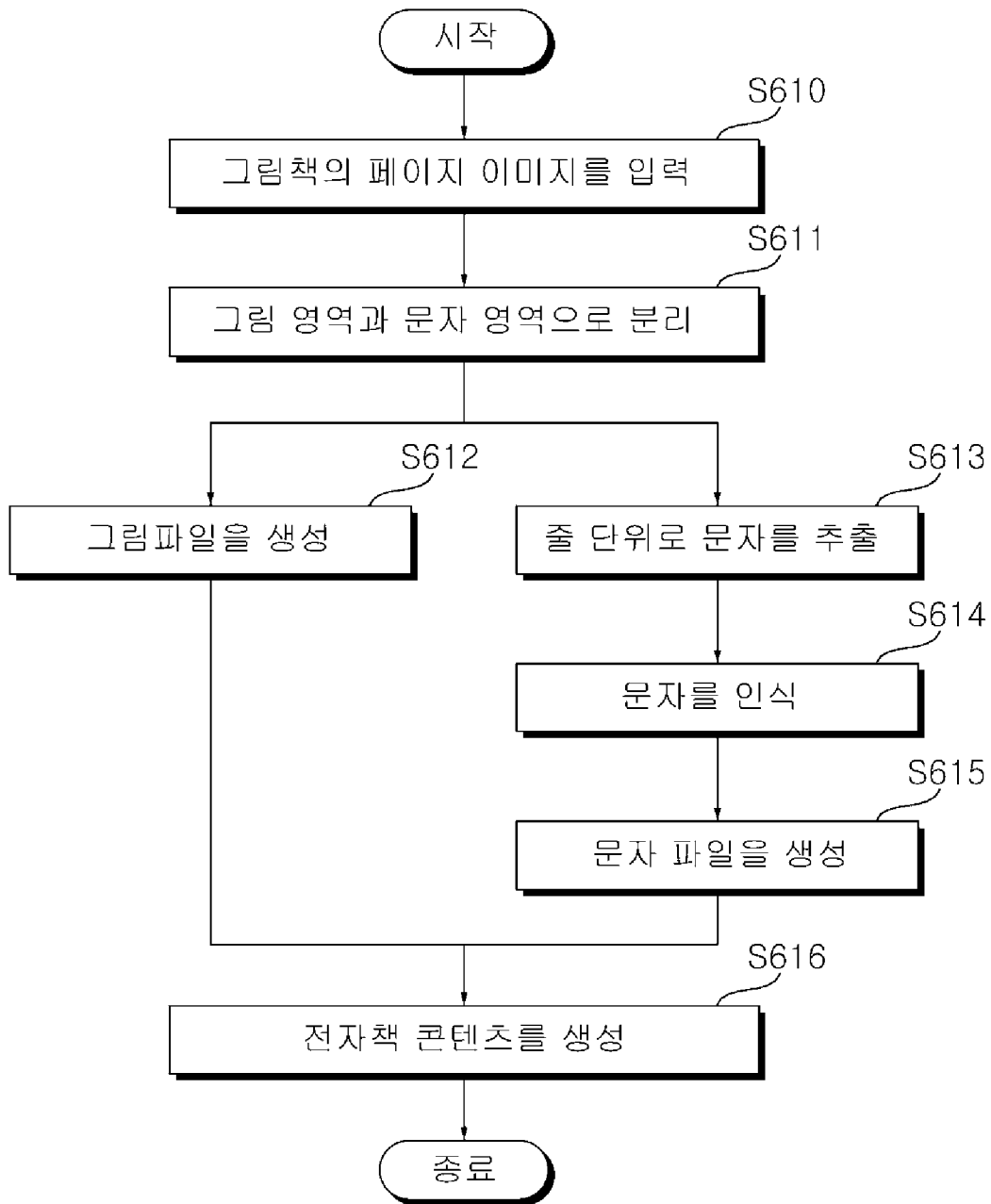


Joshua was leading his people to a new home. One more river, one more town and they would be home.



Joshua was leading his people to a new home. One more river, one more town and they would be home.

[Fig. 6]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2011/002722

A. CLASSIFICATION OF SUBJECT MATTER

G06Q 50/00(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06Q 50/00; H04N 1/387; H04N 7/26

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
Korean Utility models and applications for Utility models: IPC as above
Japanese Utility models and applications for Utility models: IPC as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

eKOMPASS (KIPO internal) & Keywords: e-book, picture book, e-book, image, image, picture book, character, text, OCR, ratio, size

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	KR 10-2006-0088175 A (JANG, WAN HO et al.) 04 August 2006 See abstract; claims 7-9; figures 1-4.	1-16
Y	KR 10-2010-0065694 A (SEOJEON MEDIA TECH INC.) 17 June 2010 See abstract; claims 1,3-5,7-9; figures 1,3-6.	1-16
A	KR 10-2008-0069887 A (PURDUE RESEARCH FOUNDATION et al.) 29 July 2008 See abstract; claims 1,7; figure 2.	1-16
A	JP 2008-519329 A (SLOMKOWSKI STEVEN, WHITESEL MICHAEL D.) 05 June 2008 See abstract; claims 14-22; figures 1-5.	1-16
A	KR 10-0980748 B1 (RICOH CO., LTD.) 07 September 2010 See abstract; claims 1-4; figures 1a-2a, 2e-4.	1-16

 Further documents are listed in the continuation of Box C. See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"I" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

17 APRIL 2012 (17.04.2012)

Date of mailing of the international search report

18 APRIL 2012 (18.04.2012)

Name and mailing address of the ISA/KR

Korean Intellectual Property Office
Government Complex-Daejeon, 139 Seonsa-ro, Daejeon 302-701,
Republic of Korea

Facsimile No. 82-42-472-7140

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/KR2011/002722

Patent document cited in search report	Publication date	Patent family member	Publication date
KR 10-2006-0088175 A	2006.08.04	NONE	
KR 10-2010-0065694 A	2010.06.17	NONE	
KR 10-2008-0069887 A	2008.07.29	CN 101231698 A EP 1950692 A1 EP 1950692 B1 US 2008-0175476 A1 US 7899247 B2	2008.07.30 2008.07.30 2011.12.28 2008.07.24 2011.03.01
JP 2008-519329 A	2008.06.05	NONE	
KR 10-0980748 B1	2010.09.07	NONE	

A. 발명이 속하는 기술분류(국제특허분류(IPC))

G06Q 50/00(2006.01)i

B. 조사된 분야

조사된 최소문헌(국제특허분류를 기재)
G06Q 50/00; H04N 1/387; H04N 7/26

조사된 기술분야에 속하는 최소문헌 이외의 문헌
한국등록실용신안공보 및 한국공개실용신안공보: 조사된 최소문헌란에 기재된 IPC
일본등록실용신안공보 및 일본공개실용신안공보: 조사된 최소문헌란에 기재된 IPC

국제조사에 이용된 전산 데이터베이스(데이터베이스의 명칭 및 검색어(해당하는 경우))
eKOMPASS(특허청 내부 검색시스템) & 키워드: 전자책, 그림책, E-BOOK, IMAGE, 이미지, 그림, 문자, 텍스트, OCR, 비율, 크기

C. 관련 문헌

카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
Y	KR 10-2006-0088175 A (장완호 외 1명) 2006.08.04 요약; 청구항 7-9; 도 1-4 참조.	1-16
Y	KR 10-2010-0065694 A ((주)서전미디어텍) 2010.06.17 요약; 청구항 1,3-5,7-9; 도 1, 3-6 참조.	1-16
A	KR 10-2008-0069887 A (퍼듀 리서치 파운데이션 외 1명) 2008.07.29 요약; 청구항 1,7; 도 2 참조.	1-16
A	JP 2008-519329 A (SLOMKOWSKI STEVEN, WHITESEL MICHAEL D.) 2008.06.05 요약; 청구항 14-22; 도 1-5 참조.	1-16
A	KR 10-0980748 B1 (가부시기가이샤 리코) 2010.09.07 요약; 청구항 1-4; 도 1a-2a,2e-4 참조.	1-16

추가 문헌이 C(계속)에 기재되어 있습니다. 대응특허에 관한 별지를 참조하십시오.

* 인용된 문헌의 특별 카테고리:
 “A” 특별히 관련이 없는 것으로 보이는 일반적인 기술수준을 정의한 문헌 “T” 국제출원일 또는 우선일 후에 공개된 문헌으로, 출원과 상충하지 않으며 발명의 기초가 되는 원리나 이론을 이해하기 위해 인용된 문헌
 “E” 국제출원일보다 빠른 출원일 또는 우선일을 가지나 국제출원일 이후에 공개된 선출원 또는 특허 문헌 “X” 특별한 관련이 있는 문헌. 해당 문헌 하나만으로 청구된 발명의 신규성 또는 진보성이 없는 것으로 본다.
 “L” 우선권 주장에 의문을 제기하는 문헌 또는 다른 인용문헌의 공개일 또는 다른 특별한 이유(이유를 명시)를 밝히기 위하여 인용된 문헌 “Y” 특별한 관련이 있는 문헌. 해당 문헌이 하나 이상의 다른 문헌과 조합하는 경우로 그 조합이 당업자에게 자명한 경우 청구된 발명은 진보성이 없는 것으로 본다.
 “O” 구두 개시, 사용, 전시 또는 기타 수단을 언급하고 있는 문헌 “&” 동일한 대응특허문헌에 속하는 문헌
 “P” 우선일 이후에 공개되었으나 국제출원일 이전에 공개된 문헌

국제조사의 실제 완료일 2012년 04월 17일 (17.04.2012)	국제조사보고서 발송일 2012년 04월 18일 (18.04.2012)
--------------------------------------------	--------------------------------------------------

ISA/KR의 명칭 및 우편주소 대한민국 특허청 (302-701) 대전광역시 서구 청사로 189, 정부대전청사 팩스 번호 82-42-472-7140	심사관 박미정 전화번호 82-42-481-8379
------------------------------------------------------------------------------------------------	-----------------------------------



국제조사보고서에서 인용된 특허문헌	공개일	대응특허문헌	공개일
KR 10-2006-0088175 A	2006.08.04	없음	
KR 10-2010-0065694 A	2010.06.17	없음	
KR 10-2008-0069887 A	2008.07.29	CN 101231698 A EP 1950692 A1 EP 1950692 B1 US 2008-0175476 A1 US 7899247 B2	2008.07.30 2008.07.30 2011.12.28 2008.07.24 2011.03.01
JP 2008-519329 A	2008.06.05	없음	
KR 10-0980748 B1	2010.09.07	없음	