

Advanced Documents 저작 도구

홍광진, 정기철
 숭실대학교, 정보과학대학, 미디어학과
 e-mail : {hongmsz, kcjung}@ssu.ac.kr

Advanced Documents Authoring Tool

Kwangjin Hong, Keechul Jung
 School of Media, College of Information Science, Soongsil University

요 약

프로젝션 디스플레이 시스템(PDS)을 사용한 Advanced Paper Document (APD)는 오프라인 문서와 온라인 문서의 장점을 모두 가지고 있다. PDS는 가상의 물체를 실제 세계에 공간적으로 연결하여 줌으로써 사용자에게 더욱 풍부한 정보를 제공할 수 있다. 본 논문은 단순하게 오프라인 문서에 대한 온라인 정보를 보여주는 APD의 기능을 확장한, 사용자가 직접 오프라인 문서에 온라인 정보들을 삽입, 삭제, 수정할 수 있는 Advanced Documents Authoring Tool (ADAT)을 제안한다. ADAT는 PDS 위에 놓여진 실제의 오프라인 문서와 가상의 온라인 정보를 직관적으로 연결해준다. 사용자가 전자펜을 통하여 문서를 선택하면, 작업 공간 위쪽에 설치된 카메라를 통하여 문서의 영상을 입력받고, 해당 문서의 기존에 저장되어있는 정보를 검색한 후, 프로젝터를 통하여 문서 위에 투영해줌으로써, 사용자에게 온라인 정보를 제공한다.

1. 서론

컴퓨터의 발전과 보급으로 인해, 다양한 정보를 담을 수 있는 온라인의 디지털 문서가 오프라인의 종이 문서의 역할을 대신하고 있지만 높은 가독성과 휴대의 편리성 때문에 여전히 사람들은 오프라인 문서를 선호한다. 따라서 최근 오프라인의 문서에 대한 온라인 정보의 삽입, 삭제, 수정이 가능한 저작도구 시스템에 대한 연구가 활발히 이루어지고 있다. 특히 실제의 물체에 가상의 정보를 연결하기 용이한 AR을 이용한 연구가 다양하게 이루어지고 있다. 기존의 연구는 표 1과 같이 오프라인 문서를 검색하는 방법에 따라 특정한 마커를 이용하는 방법과 문서 영상을 이용하는 방법으로 나눌 수 있고, 또한 문서 내의 좌표를 검색하는 방법에 따라 전자펜 등의 입력 장치를 이용하는 방법과 다른 장치없이 영상 처리 알고리즘을 이용하여 위치를 검색하는 방법으로 나눌 수 있다. 특정한 마커를 이용하여 문서를 검색하는 연구에는 Paper++ [2], Listen Reader [3], EnhancedDesk [5], AugmentedDesk [6] 등이 있고, 문서 영상을 이용하여 문서를 검색하는 연구에는 Advanced Paper Document (APD) [1], DigitalDesk [4], ScreenCrayons [7], Paper Augmented Digital Documents (PADD) [8] 등이 있다.

표 1. AR 저작도구에 관한 기존 연구

Retrieval method	Position detecting method	Author [Paper]	Year
External Marks	Using input devices	Paul Luff [2]	2004
		Maribeth Back [6]	2001
	Using camera	Hideki Koike [8]	2000
		Yoichi Sato [9]	2002
Image processing	Using input devices	François Guimbretière [11]	2003
		Dan R. Olsen Jr. [10]	2004
	Using camera	Kwangjin Hong [1]	2004
		Pierre Wellner [7]	1991

홍광진과 정기철의 APD, Pierre Wellner의 DigitalDesk, Hideki Koike 등의 EnhancedDesk, Yoichi Sato의 AugmentedDesk는 오프라인 문서 위에 온라인의 정보를 첨가하기 위해서 프로젝션 디스플레이 시스템 (PDS)을 사용하고, 사용자의 손을 이용하여 온라인의 정보를 제어함으로써, 사용자에게 자연스럽게 직관적인 환경을 제공해준다. 그러나, 이들 시스템은 다양한 동작을 인식하기 어려워 제한된 명령만 수행할 수 있다는 단점을 가지고 있다. Maribeth Back의 Listen Reader는 책 기능 향상을 위해 소리를 사용한다. Listen Reader는 Radio Frequency Identification (RFID)를 사용하여 책을 인식하기 때문에 쉽게 오프라인 문서와 온라인 정보를 연결할 수 있지만, RFID가 없는 책

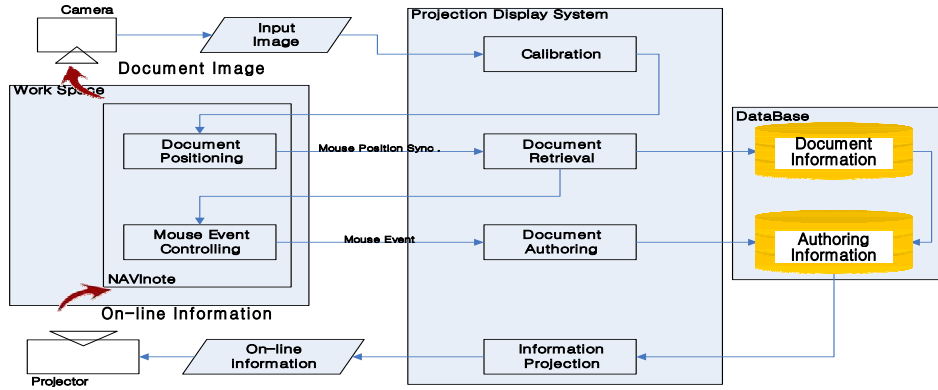


그림 1. ADAT 시스템 구성도.

은 인식하지 못하고, 책에 새로운 정보를 첨부할 수 없다는 단점을 가지고 있다. Moira C. Norrie 등의 Paper++와 Guimbretière François의 PADD는 오프라인 문서와 온라인 문서를 동기화하기 위해서 전자펜을 이용한다. 그러나, 이들 시스템은 특정한 패턴이 미리 출력된 용지를 사용해야 하고, 오프라인 문서에 답을 수 있는 정보만 입력할 수 있다는 단점을 가지고 있다.

본 논문은 사용자가 직접 오프라인 문서에 온라인 정보들을 삽입, 삭제, 수정할 수 있는 Advanced Document Authoring Tool (ADAT)을 제안한다. ADAT는 사용자가 책상 위의 오프라인 문서를 선택하면, 해당 문서와 연결된 온라인 정보 프로젝션을 위해 ADAT 시스템과 동기화한다. 동기화된 문서는 자동적으로 검색 작업이 이루어지고, 검색된 문서에 대한 기존의 온라인 정보가 문서 위에 프로젝션되어 온라인 정보에 대한 삽입, 수정, 삭제 작업이 가능하다. 오프라인 문서 위에서 이루어진 모든 작업은 실시간으로 확인 가능하고, DB에 저장되어 이후에 문서에 대한 내용을 확인하고 다른 내용을 삽입, 삭제하는 등 재사용이 가능하다.

2. ADAT

그림 1은 ADAT의 간략한 내부 구성도이다. 제안된 시스템을 이용하여 오프라인 문서를 편집하기 위해서는 오프라인 문서와 온라인 문서의 동기화, 동기화된 문서의 기존 저장 정보 확인을 위한 검색, 해당 문서에서 이루어지는 각종 작업에 대한 저장된 온라인 정보의 효율적 관리가 필요하다.

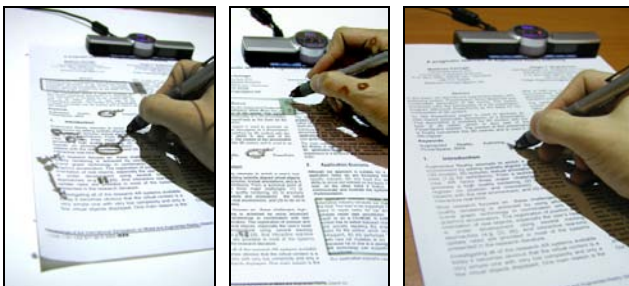


그림 2. 전자펜을 이용한 온라인 정보 입력.

ADAT를 사용함으로써, 사용자는 그림 2와 같이 전자

펜을 이용하여 오프라인 문서를 편집하듯이 직관적 이면서도 자연스럽게 온라인 정보를 사용할 수 있는 환경을 제공받게 된다.

2-1. 전처리

본 시스템은 오프라인 문서에 온라인 정보를 추가하기 위해 PDS를 사용한다. 그리고 오프라인 문서와 일치하는 기존에 저장되어있는 문서를 검색하기 위해 카메라를 사용한다. 일반적으로 카메라로 입력 받은 영상은 여러 가지 왜곡을 가지고 있어서 실제의 장면과 많은 차이를 보인다. 카메라 영상 내의 왜곡 중에서 가장 주된 왜곡 두 가지를 기하 왜곡과 색상 왜곡이라고 가정한다[10]. 우리는 문서 영역 검출을 위해 입력 영상의 이진화 연산을 사용하기 때문에 입력 영상의 색상 정보는 문서 검색에 아무런 영향을 미치지 않는다. 따라서 본 논문에서는 색상 왜곡을 고려하지 않고, 기하 왜곡만을 해결한다. 기하 왜곡의 해결을 위해, Ashdown, Robinson[11]의 좌표축 변환을 위한 투영 변환을 사용한다. 기하 왜곡을 투영 변환 관계에 있는 두 픽셀의 위치로 모델링 해보면, 식(1)과 같이 표현된다.

$$(x, y, w) = \mathbf{H} \begin{pmatrix} u \\ v \\ z \end{pmatrix}, \quad \mathbf{H} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \quad (1)$$

$$\begin{bmatrix} u_0 & v_0 & 1 & 0 & 0 & 0 & -u_0x_0 & -v_0x_0 & -x_0 \\ u_1 & v_1 & 1 & 0 & 0 & 0 & -u_1x_1 & -v_1x_1 & -x_1 \\ u_2 & v_2 & 1 & 0 & 0 & 0 & -u_2x_2 & -v_2x_2 & -x_2 \\ u_3 & v_3 & 1 & 0 & 0 & 0 & -u_3x_3 & -v_3x_3 & -x_3 \\ 0 & 0 & 0 & u_0 & v_0 & 1 & -u_0y_0 & -v_0y_0 & -y_0 \\ 0 & 0 & 0 & u_1 & v_1 & 1 & -u_1y_1 & -v_1y_1 & -y_1 \\ 0 & 0 & 0 & u_2 & v_2 & 1 & -u_2y_2 & -v_2y_2 & -y_2 \\ 0 & 0 & 0 & u_3 & v_3 & 1 & -u_3y_3 & -v_3y_3 & -y_3 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \\ i \end{bmatrix} = 0 \quad (2)$$

9개의 상수로 이루어진 투영변환 행렬은 4개 이상의 프레임 버퍼 영상과 카메라 영상의 대응하는 픽셀들에 의해서 계산된다. 식 (2)에서 행렬을 \mathbf{A} 라고 하고 \mathbf{A} 는 4개 이상의 대응하는 픽셀들에 의해서 만들어진다. 투영 행렬 \mathbf{H} 의 상수 a 에서 i 까지는 $\mathbf{A}^T \mathbf{A}$ 의 최소 고유 값과 고유 벡터에 의해 결정된다 [13]. 이

렇게 구해진 **H**행렬을 이용하여 카메라 영상 내의 좌표 (u, v) 는 프레임 버퍼 내의 좌표 (x, y) 로 일치된다.

2-2. 문서 동기화

오프라인 문서 상의 사용자가 원하는 위치에 온라인 정보를 추가하기 위해서는 책상 위에 놓인 오프라인 문서와 책상 위에 프로젝션 된 화면이 동기화가 되어야 한다. 오프라인 문서와 프로젝션 된 화면을 동기화 시키기 위해서 우리는 전자펜을 사용한다. 전자펜은 오프라인 문서 내에서의 펜이 위치한 절대 좌표를 검출할 수 있다. 먼저 우리는 책상 위에 놓여진 문서가 기울어 지지 않고, 문서에 온라인 정보의 삽입, 삭제, 수정하는 작업을 마칠 때까지 문서의 위치 변화가 없다고 가정을 한다. 이러한 가정 아래, 오프라인 문서를 ADAT 시스템에 동기화 시키기 위해 오프라인 문서가 책상 위에 놓여지면, 문서의 왼쪽 상단과 오른쪽 하단 모서리를 전자펜을 이용하여 선택을 한다. 문서의 위치가 지정되면, 선택된 두 모서리의 좌표를 이용하여 오프라인 문서 둘레에 온라인 정보를 삽입, 삭제, 수정할 수 있는 영역이 표시된다.

2-3. 문서 검색

문서 영역이 선택되면 ADAT는 선택된 문서와 일치하는 저장된 문서를 검색한다. 본 논문은 카메라를 통해 입력되는 저해상도의 이진 영상을 사용하여 문서를 검색한다. 입력 영상은 기하 보정을 통해 왜곡을 보정한 카메라 영상의 이진영상을 사용하고, 입력 영상과의 비교에 사용할 영상은 스캐너를 사용하여 저장시킨 문서의 이진영상을 사용한다. 저해상도의 이진 영상을 사용한 문서 검색의 경우 영상 내에서 글자를 인식한다거나 단어 별로 문서의 검색에 앞서서, X-Y recursive cut 알고리즘[12]을 사용하여, 스캔 영상과 카메라 영상에서 문자영역을 분할한다.

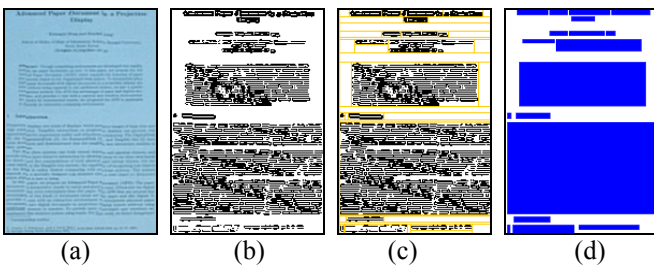


그림 3. X-Y recursive cut 알고리즘을 이용한 문서분할 과정: (a)카메라 영상, (b)이진화 영상, (c)문서 분할, (d)결과 영상.

분할된 각각의 컴포넌트들을 비교함으로써, 사용자가 선택한 문서와 가장 유사한 기존에 저장된 문서를 검색한다. 문서의 유사도를 구하기 위해 입력 받은 영상과 저장되어있는 영상의 서로 대응하는 픽셀을 비교하는 방법을 사용한다. 먼저 두 영상을 비교할 때 발생할 수 있는 오차를 줄이기 위해, 문서의 단락을 둘러싸고 있는 여백을 최소화하고, 비교하는 두 영상의 크기를 일치시켜, 서로 대응하는 픽셀을 비교한다. 입력 받은 영상의 m 번째 픽셀을 I^m 이라 하고, 기존에 저장되어있는 영상의 m 번째 픽셀을 S^m 이

라 하면, 두 영상의 유사도 $D(I, S)$ 는 식 (3)을 이용하여 구할 수 있다.

$$K^m = \begin{cases} 1 & \text{if } I^m = S^m \\ 0 & \text{otherwise} \end{cases}, D(I, S) = \sum_m K^m \quad (3)$$

본 논문에서는 문서를 검색하는데 걸리는 시간을 줄이기 위해 비교하는 이미지의 크기를 19×26 픽셀과 190×260 픽셀의 두 가지를 사용한다. 따라서, 위의 식 (3)에서 비교하는 픽셀의 개수를 나타내는 m 의 범위는 각각 $1 \leq m \leq 494$ 와 $1 \leq m \leq 49400$ 으로 정해진다. 이러한 방법을 사용하여 유사도가 가장 높은 문서를 선택하고, 선택된 문서의 이전에 입력된 온라인 정보들을 화면에 보여준다. 기존에 저장된 정보들이 없을 경우 화면에는 오프라인 문서만 보여진다.

2-4. 문서 정보 관리

우리가 온라인 정보를 추가하기 위해 사용하는 전자펜은 문서 내에서 절대 좌표를 검출할 수 있고 이를 이용하여 원하는 위치에 온라인 정보를 연결할 수 있다. 온라인 정보는 그림, 동영상, 음악 등의 링크된 파일과 사용자가 직접 입력하는 메모로 나눌 수 있다. ADAT는 온라인 파일을 링크시킴으로써 사용자는 오프라인 문서를 통해 얻는 것보다 더 다양하고 많은 정보를 얻을 수 있도록 해준다. 또한 메모의 경우 사용자가 직접 펜으로 문서에 쓰거나, 가상의 키보드를 통해 입력이 가능하며, 문서에 대한 수정 사항이나 추가 사항 등을 기입함으로써 이후에 해당 문서를 수정할 때 참고할 수 있도록 해준다.

이러한 온라인 정보들은 ADAT 내의 데이터베이스에 저장된다. 데이터베이스 내에 저장된 정보들은 그림 4과 같은 관계를 가지는 Contents, Memos, Files 의 3개 테이블로 구성되어있다.

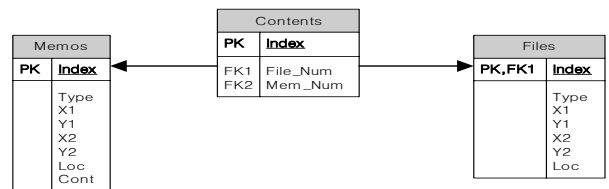


그림 4. 데이터베이스 구성.

Contents 테이블은 문서의 번호 인덱스와 문서에 대한 메모와 파일 테이블에 대한 인덱스로 구성되어있다. Memos 테이블과 Files 테이블은 메모와 파일의 종류를 나타내는 Type과 메모가 연결된 문서 상의 위치(X1, Y1, X2, Y2), 그리고 파일의 위치(Loc)로 구성되고, Memos 테이블의 경우, 가상 키보드로 입력한 메모를 관리하는 Cont가 추가적으로 들어있다.

3. 실험 및 결과

실험에 사용된 시스템은 프로젝션 된 화면과 종이 문서를 입력 받기 위한 캠코더, 오프라인 문서에 온라인 정보를 연결시켜 주기 위한 전자펜, 프로젝터, 책상, 그리고 입력 받은 영상들을 이용하여 원하는 정보를 얻어 내기 위한 일반 PC로 이루어져있다. 영상처리시스템은 Intel® Pentium4 2.66Hz CPU와, 1Gbyte

RAM, ATI® Radeon 9600으로 구성 되어있고, 프로젝션 화면은 1280×1024 픽셀의 32비트 트루 컬러 영상을 사용한다. 카메라 영상은 Sony® DCR-VX2000를 사용하여 720×480 픽셀의 영상을 입력 받는다. 카메라 영상 내에서 문서의 크기는 190×260 픽셀이고, DB에 저장된 문서의 크기는 평균 1119×1576 픽셀이며, ADAT 시스템에서는 총 100개의 스캔받은 문서 영상을 사용하여 문서를 검색한다. 각종 온라인 정보를 입력하는데 사용하는 전자펜은 NAVIis® NAVInote를 사용하고, 미리 저장된 문서는 HP® Scanjet 5P 스캐너를 사용하여 스캔 받는다.

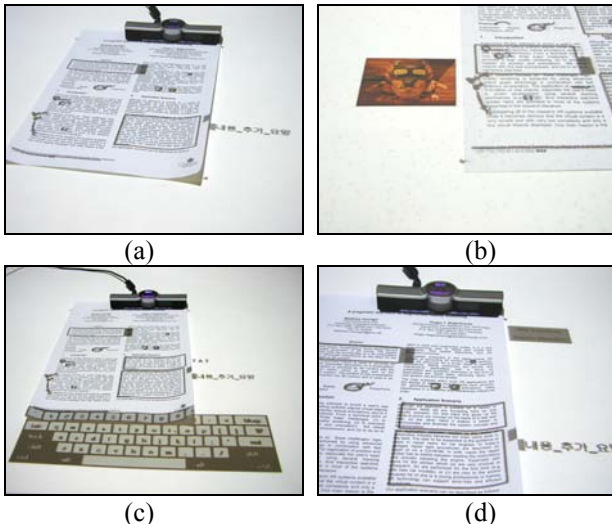


그림 6. ADAT 실행 화면: (a)오프라인 문서 위에 온라인 정보 투영, (b)문서에 연결된 그림파일 실행, (c)메모입력을 위한 가상키보드 실행, (d)영역 선택하여 메뉴상자 보여주기.

그림 6은 ADAT를 실행한 화면을 보여준다. 사용자가 시스템 내에 오프라인 문서를 올리고 문서의 위치를 지정하면, ADAT는 그림 6의 (a)와 같이 문서 위에 온라인 정보를 보여준다. 그림 6의 (b)는 문서 내에 연결된 파일을 실행하는 것을 보여준다. 문서 내에 파일이 연결되면 연결 영역 좌측에 해당 파일의 실행과 파일의 연결 해제를 위한 버튼이 표시된다. 그림 6의 (c)는 가상 키보드가 실행된 화면을 보여준다. 문서의 일정 영역을 선택하면, 그림 6의 (d)와 같이 문서 좌측에 삽입할 파일을 찾거나 메모 입력을 위한 가상 키보드를 실행할 수 있는 메뉴 상자를 보여준다.

4. 결론

본 논문에서 우리는 전자펜과 PDS를 이용하여 오프라인 문서에 온라인 정보를 삽입, 삭제 수정하는 것이 가능한 ADAT를 제안하였다. ADAT는 사용자가 오프라인 문서에 직접 펜을 이용하여 온라인 정보를 추가할 수 있기 때문에 사용자에게 직관적이고 자연스러운 환경을 제공한다. 따라서, ADAT를 사무 환경이나 교육 환경에 적용했을 때, 시간 절약이나 작업 능률 향상, 보다 높은 학습 효과 등의 이득을 얻을 수 있을 것으로 기대한다. 우리가 제안한 ADAT는 저해상도의 이진 영상을 사용하여 문서를 검색하기 때

문에, 문서 검색의 정확도가 고해상도의 이미지를 사용하는 경우에 비해 낮다는 단점이 있다.

우리는 이후에 오프라인 문서와 온라인 문서의 내용을 동기화하여 실시간으로 문서를 수정하고 그 결과를 프린트하여 확인할 수 있는 확장된 ADAT에 관한 연구를 할 것이다. 또한 문서의 위치와 기울어진 정도에 제약을 받지않고 문서에 온라인 정보를 프로젝션할 수 있는 방법에 관한 연구와 문서 검색의 정확도를 높일 수 있는 방법에 관한 연구를 지속적으로 하여 사용자에게 보다 자연스러운 환경을 제공할 것이다.

참고문헌

- [1] K. Hong, K. Jung, "Advanced Paper Document in Projection Display," 5th Pacific-Rim Conference on Multimedia (2004) pp. 81-87
- [2] L. Paul, H. Christian, M. C. Norrie, B. Signer, H. Peter, "Only Touching the Surface: Creating Affinities between Digital Content and Paper," the 2004 Conference on Computer Supported Cooperative Work (2004) pp. 523-532
- [3] M. Back, J. Cohen, R. Gold, S. Harrison, S. Minneman, "Listen Reader: an Electronically Augmented Paper-based Book," the SIGCHI Conference on Human Factors in Computing Systems (2001) pp. 23-29
- [4] P. Wellner, "The DigitalDesk Calculator: Tangible Manipulation on a Desktop Display," the 4th annual ACM Symposium on User Interface Software and Technology (1991) pp. 27-33
- [5] H. Koike, Y. Sato, Y. Kobayashi, H. Tobita, M. Kobayashi, "Interactive Textbook and Interactive Venn Diagram: Natural and Intuitive Interfaces on Augmented Desk System," the SIGCHI Conference on Human Factors in Computing Systems (2000) pp. 121-128
- [6] Y. Sato, Y. Kobayashi, H. Koike, "Fast Tracking of Hands and Fingertips in Infrared Images for Augmented Desk Interface," the 4th IEEE International Conference on Automatic Face and Gesture Recognition (2000) pp. 462
- [7] D. Olsen, T. Taufer, J. A. Fails, "ScreenCrayons: Annotating Anything," the 17th annual ACM Symposium on User Interface Software and Technology (2004) pp. 165-174
- [8] G. François, "Paper Augmented Digital Documents," the 16th annual ACM Symposium on User Interface Software and Technology (2003) pp. 51-60
- [9] H. Kang, S. Kim, C. Lee, K. Jung, M. H. Park, "Foreground Object Detection in Projection Display," the Journal of the Institute of Electronics Engineers of Korea (2004) pp. 27-37
- [10] K. Hong, K. Jung, "Foreground Object Detection in Projection Display using Color Calibration and Stereo Information," the 31st KISS Spring Conference (2004) pp. 784-786
- [11] M. Ashdown, P. Robinson, "The Escritoire: A Personal Projected Display," the Journal of International Conferences in Central Europe on Computer Graphics, Visualization and Computer Vision (2003) pp. 33-40
- [12] K. Jung, J. Han, "Hybrid Approach to Efficient Text Extraction in Complex Color Images," the Pattern Recognition Letters (2004) pp. 679-699
- [13] R. Hartley, A. Zisserman, "Multiple View Geometry in Computer Vision," Cambridge University Press, (2001)