

모바일 단말기를 위한 만화 영상 자동 변환

한은정^o, 전성국, 정기철
 송실대학교 정보과학대학 미디어학과 HCI Lab.
 {hanej^o,k612051,kcjung}@ssu.ar.kr

Automatic Conversion of Large Comic Pictures on Mobile Devices

EunJung Han^o, Sungkuk Chun, Keechul Jung
 HCI Lab., School of Media, College of Information Science, Soongsil University

요 약

모바일 기술의 발전으로 인쇄 매체가 아닌 온라인 매체로 만화 콘텐츠를 제공받는 수요층이 늘어나고 있고, 모바일 단말기의 작은 화면 위에 기존 오프라인 만화를 브라우징하기 위한 새로운 기술에 대한 연구가 많이 진행되고 있다. 본 논문에서는 큰 만화 콘텐츠를 영상의 의미 정보(배경이 아닌 모든 부분: 사람, 중요한 오브젝트, 문자)의 손실을 최소화하여, 기존의 오프라인 만화 콘텐츠를 모바일 단말기에 맞게 자동 변환 시스템 (Automatic Comics Conversion System: ACCS)을 제안한다. ACCS는 기존 만화 콘텐츠 영상의 각 페이지를 프레임 단위로 자르고 잘라진 프레임을 분할, 확대, 축소 등을 통해 모바일 단말기 화면 크기에 맞게 자동으로 변환한다.

1. 서 론

최근 IT기술의 빠른 발전으로 모바일 환경에서 다양한 정보서비스가 유비쿼터스 환경에 맞게 진화하고 있으며, 현재의 서비스에서 더욱더 발전된 모습으로 변화되고 있다. 그 중 만화 콘텐츠는 하나의 문화 장르로서 자리 매김하고 있으며 다양한 매체와 폭넓은 수요층을 형성하고 있다. 최근 들어서는 인터넷과 모바일 기술의 발달로 인쇄 매체가 아닌 실시간 온라인 매체를 통해 소비자에게 만화 콘텐츠의 공급이 시작되었고, 이에 필요한 영상 인식, 영상 처리, 데이터 전송, 과금 시스템, 사용자 및 만화 콘텐츠 DB 시스템 등의 요소를 필요로 한다. 그러나 현재, 모바일 단말기에 제공된 만화 콘텐츠는 대부분 기존 오프라인 만화 콘텐츠를 스캔받아 만화 콘텐츠의 임의 영역을 지정해 모바일 단말기 화면 크기에 맞게 수작업으로 편집/수정하여 제공해 주고 있다. 이러한 수작업을 하기 위해서는 많은 비용과 시간이 소요되며, 또 그만큼 학습자들의 요구에 부합할 것인가 하는 위험 부담이 크다

이를 해결하기 위해 기존 오프라인 콘텐츠를 모바일 단말기에 맞게 자동으로 변환해주는 시스템에 관한 연구가 많이 진행되고 있다. Chen[1] 등은 웹 페이지를 무선 단말기 위에 보여주기 위한 방법으로 큰 영상을 작은 영상으로 변환하는 썸네일 (thumbnail)방식을 사용하였고, Lizhen[2] 등은 웹 문서들을 자동으로 변환하는 시스템으로 온라인에서 웹 콘텐츠 마이닝을 통해 콘텐츠 탐색을 향상시켜 주는 기법을 제공하였으며, 김원철[3] 등은 방송사 웹 페이지 내에서 실시간으로 사용자가 선호하는 방송사 웹 페이지의 섹션을 추출하고, 모바일 환경에 적합하도록 각 섹션의 순서를 재구성하여 모바일 단말기에 제공해주는 기법을 사용하였다. 강신상[4] 등은 모바일 단말기에서 영상 및 폰트를 효율적으로

처리할 수 있게 비트맵으로 처리하는 기법을 사용 하였다. Zhang[5] 등은 큰 영상들을 모바일 단말기에 자동으로 변환하는 연구로써 영상 정보의 손실을 최소화하며, 특정 영역 부분만 순차적으로 보여주는 기법을 연구하였다. 또한, 국내 모바일 만화 콘텐츠 제공하는 업체 중 스튜디오에서는 기존 오프라인 만화 콘텐츠를 특정 영역에 애니메이션을 적용하여 모바일로 서비스를 제공한다[6]. 이와 같이 모바일 단말기에 제공되는 콘텐츠의 질적 향상과 자동 변환에 관한 연구가 많이 진행되고 있다.

본 논문에서는 다양한 형태를 지닌 만화 콘텐츠를 모바일 단말기에 맞게 자동 변환 시스템 (Automatic Comics Conversion System: ACCS)을 제안한다. ACCS는 인쇄 만화 콘텐츠의 스캔한 한 페이지의 입력받은 영상을 페이지 단위로 분석하여, 모바일 단말기 화면에 맞게 자동 변환한다. 우리는 영상의 주요 정보 (배경이 아닌 모든 부분: 사람, 중요한 오브젝트, 문자)의 손실을 최소화하여 만화 영상을 의미론적 구조 (Semantics Structure)로 추출하여 화면에 맞게 잘라서 제공한다(그림 1).

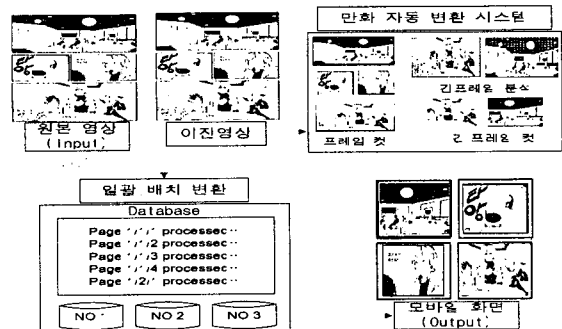


그림 1. 만화 ACCS

2. ACCS 구성

우리는 기존 오프라인 콘텐츠의 만화 내용이나 영상의 의미 정보의 손실을 최소화하여 모바일 단말기에 만화를 제공한다. 만화 페이지 분석은 기존 만화 콘텐츠의 영상을 각각의 의미론적 구조(semantic structure)로 각 픽셀의 질감을 분할(texture segment)한다. 의미론적 구조의 계층으로 표현되어 있는 전체 오브젝트에서 각각의 계층 단위를 “프레임(frame)”으로 지칭한다. 또한, 만화의 다양한 형태 중에서 프레임의 형태가 고정되어 있는 영상을 자동 변환한다.

2.1. 페이지의 프레임 자르기

우리는 모바일 단말기 화면에 맞게 영상 한 페이지를 각 프레임 단위로 자른다. 기존 오프라인 만화 영상을 Recursive X-Y cut 알고리즘을 이용하여 프레임을 추출하며, 이는 탑-다운(Top-down) 방식으로 픽셀을 읽어 들여 영상을 사각형으로 반복적으로 블록화해 여러 조각으로 분할하는 기술로써, 각 과정에서 픽셀로 잡힌 윤곽들이 계산된다. 그림 2(a)는 이진화한 영상이고, 그림 2(b)는 분석한 페이지를 프레임 단위로 자른 영상이다.

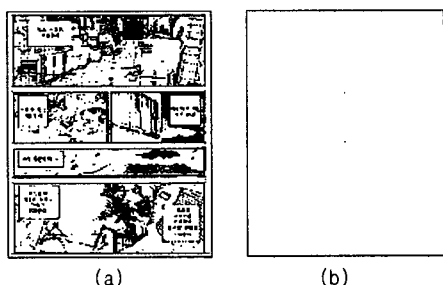


그림 2. 영상 분석: (a) 이진 영상, (b) 프레임 단위로 자른 영상

2.2. 의미론적 구조를 고려한 긴 프레임 자르기

모바일 단말기의 화면은 가로 길이보다 세로의 길이가 긴 반면에, 만화 영상은 가로의 길이가 긴 영상보다 세로의 길이가 긴 영상이 더 많다. 즉 모바일 단말기 화면에 영상이 보여질 때 과도한 횡적 축소가 일어난다. 이 논문에서는 이를 방지하기 위하여 프레임이 다음과 같은 조건을 만족할 경우 프레임을 자르는 과정을 수행한다.

$$\frac{I_W}{I_H} \geq \frac{M_W}{M_H} \times 1.3 \quad (1)$$

I_W 는 만화 영상의 가로, I_H 는 만화 영상의 세로, M_W 는 모바일 단말기 화면의 가로, M_H 는 모바일 단말기 화면의 세로 길이를 의미한다. 즉, 만화 영상의 세로 길이를 모바일 단말기 화면의 세로 길이에 맞추어 영상을 확대 또는 축소 시에 변환된 영상의 가로 길이가 화면의 가로 길이보다 130% 이상일 경우 프레임을 자르는 과정을 수행한다. 130% 이하의 영상의 경우 모바일 단말기 화면에 제공하였을 시에 영상의 내용을 깨뜨리지 않는 정도의 축소가 발생한다. 그리고 130% 이상의 영상의 경우 모바일 단말기 화면에 적용하기 위한 과도한 축소가 발생한다. 이를 극복하기 위해 영상의 의미 정보가 아닌 외곽 부분을 모바일 단말기 화면에 맞게 제공하기 위하여 영상의 정보를 분리한다[7].

2.2.1. k-means 알고리즘

본 논문에서의 만화 영상은 주로 흑백 영상이며 만화 영상의 각 세로줄의 정보량은 각 줄의 서로 다른 질감(texture) 수와 비례한다고 가정한다. 그림 3에서는 영상에서 같은 질감을 갖는 블록들을 K-means 알고리즘을 사용하여 군집화 한다. n 은 각 블록의 질감(v)의 개수이며, k 는 군집의 개수이다. 우리는 v 중 무작위로 k 개를 선택하여 군집의 초기 mean값으로 설정한다. 각 v 들을 군집들의 초기 mean값과의 거리를 계산하여 가장 작은 군집에 속하게 한다. 그림 4의 긴 프레임 만화 영상과 같이 각 군집에 속해진 v 값들로 군집의 평균값을 새롭게 정해준다. 새로 정한 평균값과의 유클리디언 거리(Euclidean distance)를 측정하여 가장 가까운 군집으로 다시 저장한다. 이는 평균값의 변화가 없어 질 때까지 작업을 수행한다.

```

begin initialize  $n, k$ 
divide the image to  $n$  components
initialize textures of  $n$  components using pixels values
randomly choose  $k$  means among  $n$  components
do classify each  $n$  component to the cluster which has the nearest mean from the component
  recalculate mean of cluster
until mean equals to previous mean of cluster
end
    
```

그림 3. k-means 알고리즘

그림 4(a)와 같이 긴 프레임 만화 영상 I 를 $M \times N$ 개의 블록으로 나누고, 각 블록의 크기는 (i, j) 로 표현한다. 그림 4의 블록의 질감 v 는 블록의 중심을 기준으로 가로, 세로, 대각선에 위에 존재하는 $2 \times (i + j) - 4$ 개의 픽셀이다[8]. 각 픽셀들의 그레이(Gray)값을 구하여 블록의 질감(v)으로 결정한다. 그림 4(b)에서는 영상의 군집을 통해 각각의 다른 정보의 질감 정보로 분석한 영상이다.

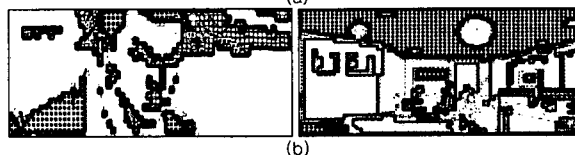
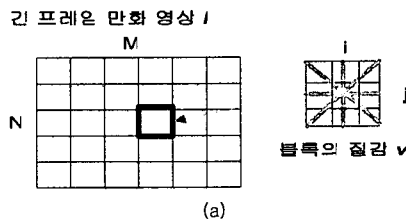


그림 4. 질감 분석: (a) 블록의 질감 정보 분석, (b) 같은 정보의 픽셀로 분석한 영상

K-means 알고리즘에서 군집의 수가 많을수록 군집간의 질감 차이는 작아진다. 즉, 비슷한 질감을 가지는 군집이 다른 군집으로 나뉘질 수 있다. 본 논문에서는 비슷한 질감을 가지는 다른 군집을 하나의 군집으로 합치기 위하여 재 군집화(Reclustering)를 수행한다.

2.2.2. 만화 영상의 자동 변환

K-means 알고리즘과 재 군집화 과정을 통해 파악된 만화 영상의 의미론적 구조를 바탕으로 모바일 단말기에 적절한 만화 콘텐츠를 생성한다. 그림 4와 같이 같은 정보의 질감을 가진 컴포넌트들이 모여 군집화된다. 따라서 우리는 다른 질감을 가지는 군집들이 많이 모여있는 영역을 보존하고 그 외의 영역을 삭제하는 방법을 사용한다.

그림 5에서는 긴 프레임용 모바일 단말기에 적용하기 위해 영상 상의 컴포넌트로 분할된 세로줄(Y)을 가로줄(X)로 분석하여, 유사한 질감을 가진 블록들끼리 모여 몇 개의 군집으로 이루어졌는지 분석한다. t는 세로줄에 존재하는 재 군집한 서로 다른 군집의 총 개수이며, 세로줄의 t를 각각 줄마다 비교하여, t의 개수가 많을 때 복잡도가 높다고 인식한다. 영상 외곽에서부터 영상이 작은 군집의 개수가 나온 곳(S(시작점)과 E(끝점)부분만 분할하여 모바일 화면에 제공한다.

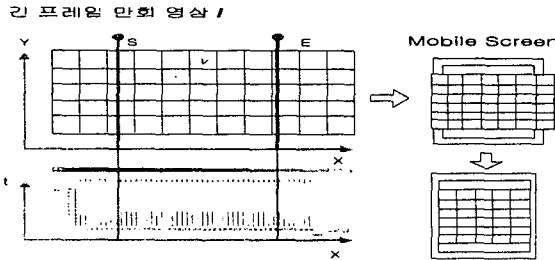


그림 5. 긴 프레임 영상 모바일 단말기에 자동 변환

3. 실험 및 결과

본 시스템에 사용한 PDA는 Pocket-PC 2003기반의 카메라가 장착된 POZ x301을 사용하였다. POZ x301 모델은 400MHz 인텔 프로세서를 사용하며, 62MB SDRAM /160MB Flash ROM을 사용한다. 그림 6은 긴 프레임 만화 영상을 모바일 단말기 화면에 자동 변환한 결과 화면이다. 그림 6(a)는 원본 영상이며, 그림 6(b)는 분석한 영상에서 배경이 아닌 모든 부분만 자동 분할하여 모바일 단말기 화면상에 제공된 영상이다.

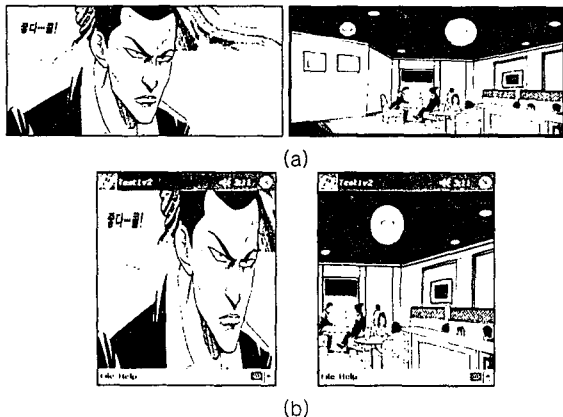


그림 6. 결과 화면: (a) 원본 영상, (b) ACCS를 적용한 모바일 단말기 화면상에 제공된 영상

그림 6과 같은 실험을 50장을 해본 결과 90%는 모바일 화면 상에서 자동 변환 되었으나, 성공하지 못한 10%로 중에는 그림 7(a)와 같이 영상이 매우 복잡한 의미론적 구조를 가질 경우는 분할이 되지 않았다. 또한, 그림 7(b)와 같이 문자가 있는 말풍선 부분이 외곽에 있을 경우는 문자가 잘리는 문제점을 가지고 있다.



그림 7. 문제 영상 구조: (a) 복잡한 구조, (b) 외곽에 문자 있는 구조

4. 결론

기존의 만화 콘텐츠는 인쇄 매체의 디지털화에 필요한 인력과 비용, 시간의 비효율성과 함께 소비자가 쉽고 편리하게 모바일 단말기에서 만화를 볼 수 있게 하는 도구가 부족하다는 단점이 있다. 이를 해결하기 위해 본 논문에서는 기존 만화 콘텐츠를 모바일 단말기에서 자동으로 변환하는 시스템을 제공한다.

향후 연구 과제로는 만화 콘텐츠 영상의 연결성분 (connected component)을 이용하여 고정되지 않는 프레임 단위를 추출할 수 있게 적용할 것이며, 문자 부분이 잘리는 단점은 문자 영역의 추출을 통해 문자의 이동/변환 방법을 적용시킴으로써 이러한 문제점을 해결할 것이다. 이를 확장하여, 만화 콘텐츠 자동 변환 시스템의 지속적인 개발로 보다 효과적인 콘텐츠와 저렴한 비용의 만화 콘텐츠 영상의 공급에 관한 연구를 진행할 예정이다.

참고문헌

- [1] Y.Chen, W.Y.Ma, H.J.Zhang, "Detecting Web Page Structure for Adaptive Viewing on Small Form Factor Devices," Proceedings of the International WWW Conference, Budapest, Hungary, ACM 1-58113-680-3/03/0005, pp.225-233, 2003.
- [2] Baldonado Lizhen Lin, Junjie Chen and Hantao Song. "The research of Web mining." Proceedings of the 4th World Congress on Intelligent Control and Automation, Vol.3, pp.2333-23337, 2002.
- [3] 김원철, 이수철, 황인준, 변광준, "모바일 환경에서 사용자 검색 성향을 반영한 웹 방송 정보 재구성 기법." 정보처리학회논문지 D, 제 11-D 권, 제 5 호, pp.1149-1158. 2004.10.
- [4] 강신상, 옥경달, 이상범, "모바일 단말기 상에서의 효율적인 영상 및 폰트 처리." 한국정보과학회지, Vol.31, No.01, pp.0685-0687, 2004.
- [5] Hao Liu, Xing Xie, Wei-Ying Ma, Hong-Jiang Zhang, "Automatic Browsing of Large Pictures on Mobile Devices." Proceedings of the MM, ACM 1-58113-722-2/03/0011, pp.148-153, 2003.
- [6] <http://cstudio.co.kr>.
- [7] V. Guralnik, G. Karypis, "Workshop on Data Mining in Bioinformatics," pp.73-80, 2001.
- [8] Anil K. Jain, Kalle Karu, "Learning Texture Discrimination Masks," IEEE Transactions on pattern analysis and machine intelligence, Vol.18, No.2, pp.195-205, 1996.