

개인화된 비디오 요약 방송 서비스

*진성호, *조준호, *배태면, *노용만, **강경옥, **홍진우
*한국정보통신대학교 공학부 멀티미디어 그룹
**한국전자통신연구원, 디지털방송연구단
wh966@icu.ac.kr

Broadcasting Service for Personalized Video Skim

*Sung Ho Jin, *Jun Ho Cho, *Tea Meon Bea, *Yong Man Ro, **Kyeongok Kang, and
**Jinwoo Hong

*Multimedia Group, Information and Communications University (ICU)
**Electronics and Telecommunications Research Institute (ETRI)

요약

본 논문은 시청자 개개인의 선호도 정보(user preference)를 이용하여, 콘텐츠 제공자(content provider)로부터 개인화된 비디오 요약을 제공받을 수 있는 맞춤형 방송 서비스를 제안한다. 제안하는 방송 서비스는 다채널의 디지털 방송 환경에서 시청자들에게는 채널 선택의 편의성과, 콘텐츠 제공자들에게는 시청자들이 자신들의 콘텐츠에 대한 소비를 촉진시키는 기능을 제공한다. 따라서, 본 논문에서는 개인화된 비디오 요약 서비스를 제공하기 위해, TV-Anytime에 기반한 사용자 선호도 정보를 이용하는 시스템 스킴(scheme)을 제안한다. 제안하는 방송 서비스의 유효성을 테스트하기 위해 영화 장르의 비디오에 대해서 이벤트(event)들을 세그먼트(segment)하고, 해당 선호도 정보에서 추론된 선호 이벤트 정보에 따라 비디오 요약(video skim)을 생성하고, 시청자 단말에 제공하는 실험을 수행한다.

1. 서론

디지털 방송은 다채널을 제공할 뿐 아니라, 아날로그 방송보다 진보된 방송 서비스를 제공한다. 현재 개발되고 있는 디지털 방송 서비스들은 콘텐츠들의 정보를 제공하는 데이터 방송(data broadcasting)과, 리턴 채널(return channel)을 이용하여, 시청자들의 요구를 적극적으로 반영하는 양방향 방송 서비스(Bidirectional broadcasting)들로 구성된다. 이와 더불어 중요하게 대두되고 있는 방송 서비스로는 맞춤형 방송 서비스들이 제안되고 있다[1-3]. 맞춤형 방송 서비스는 TV 단말에서 시청하고 있는 시청자에게 가장 최적화된 방송 서비스를 제공하는 방송 서비스를 의미한다. 이를 위해, 시청자 개개인의 선호도 정보(user preference)가 핵심 정보로 이용된다.

일반적으로 서비스 제공자(service provider)나 콘텐츠 제공자(content provider)들은 다채널의 디지털 방송환경에서 자신들의 콘텐츠들의 소비를 촉진시키기 위해서, 시청자들에게 해당 콘텐츠에 대한 비디오 요약(video skim)을 제공한다. 이런 비디오 요약들은 각각의 시청자들의 기호나 취향에 맞게 제작된 것이 아니라 콘텐츠 제공자에 의해 일률적으로 만들어진다. 그러나, 하나의 콘텐츠에 대해서

시청자들이 선호하는 장면이나 이벤트들은 각각 서로 다르다. 예를 들어, 영화 콘텐츠에 대해서 액션장면을 선호하는 시청자, 로맨틱한 장면을 선호하는 시청자와 같이 여러 시청자 군으로 나눌 수 있다. 따라서, 시청자가 흥미있는 방송 콘텐츠 선택을 위해, 원하지 않는 장면들이 포함되거나, 원하는 장면이 포함되지 않은 방송 콘텐츠들의 비디오 요약들을 모두 시청하는 것은 시청자에게 시간의 낭비를 초래한다.

그러므로, 시청자들의 콘텐츠 선택의 효율성과 콘텐츠 제공자의 콘텐츠 소비 촉진 목적을 위해 본 논문에서 개인화된 비디오 요약 서비스를 제안하고자 한다. 이 서비스는 맞춤형 방송 서비스들 중 하나로써, 디지털 방송 환경에서 시청자들에게 맞춰진 비디오 요약을 제공하는 방송 서비스이다. 제안하는 방법에서의 맞춤형 서비스는 콘텐츠 제공자에서 해당 콘텐츠에 대해서 미리 세그먼트(segment)된 이벤트 정보들을 이용하여, 각각의 시청자들의 선호도 정보에 따라 서로 다른 비디오 요약들을 생성하고 제공한다.

본 논문은 제안하는 방송 서비스를 위한 시스템 스킴(scheme)과 사용자 선호도 및 비디오 생성 틀에 대해서 설명과, 제안하는 시스템을 통한 실험 및 결론으로 구성된다.

2. 개인화된 비디오 요약 서비스를 위한 시스템 구조

본 논문에서는 제안하는 개인화된 비디오 요약 서비스를 제공할 수 있는 시스템 스킴(scheme)은 그림 1에서 보여준다.

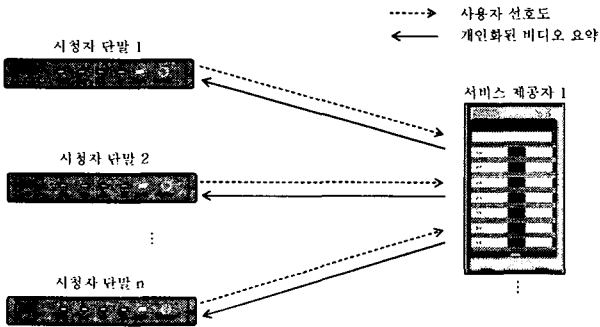


그림 1. 개인화된 비디오 요약 방송 서비스를 위한 시스템 구성도

시청자 단말은 시청자가 방송중인 콘텐츠를 시청할 수 있고 원하는 방송 콘텐츠 및 부가 정보를 저장할 수 있는 저장 공간을 가진 디지털 TV 수신 환경을 의미하며, 서비스 제공자는 시청자에게 방송 콘텐츠나 부가 정보를 제공하는 시스템을 의미한다. 각각의 시청자 단말에서 서로 다른 취향을 가진 시청자들의 선호도 정보를 서비스 제공자에 전달하면, 서비스 제공자는 이에 대응되는 응답으로 개인화된 비디오 요약을 전송한다.

그림 2는 시청자 요구에 의해서 시청자 단말과 서비스 제공자 사이에서 개인화된 비디오 요약 방송 서비스가 제공되는 순서도를 보여준다.

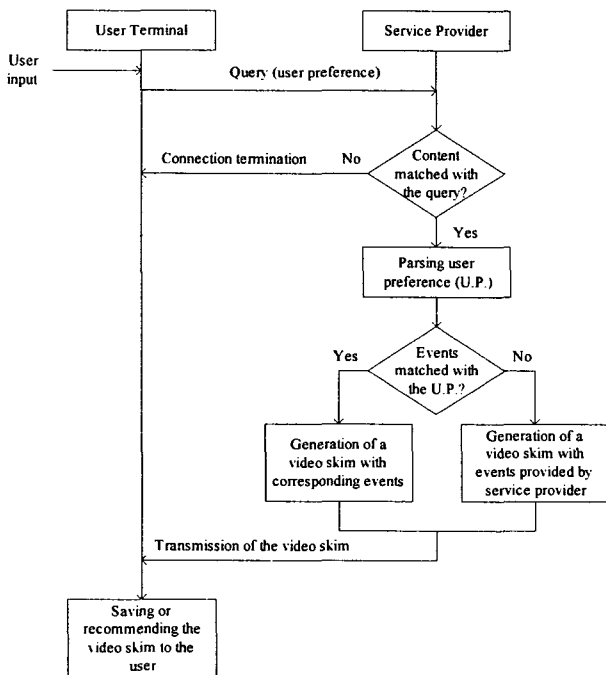


그림 2. 개인화된 비디오 요약 방송 서비스의 순서도

먼저, 각각의 시청자 단말에서 시청자의 요구에 의해서, 사용자 선호도를 서비스 제공자나 콘텐츠 제공자로 질의를 한다. 이때, 질의는 사용자 선호도 정보를 포함한다. 서비스 제공자는 먼저, 해당되는 콘텐츠가 자신의 콘텐츠들의 목록과 비교하여, 해당 콘텐츠가 없는 경우 연결을 종료하고, 해당 콘텐츠를 보유하고 있는 경우, 사용자 선호도를 파싱(parsing)한다. 파싱된 정보로부터 획득된 선호 이벤트 정보를 가지고, 해당 콘텐츠에 대해, 미리 파싱된 이벤트 정보와 비교를 한다. 사용자의 선호 이벤트가 미리 파싱된 이벤트 목록에 속하면, 해당 이벤트를 가지고, 비디오 요약을 생성한다. 이벤트 목록에 해당 이벤트가 없는 경우, 서비스 제공자에 의해서 제공되는 비디오 요약을 생성하게 된다. 생성된 비디오 요약을 시청자 단말에 전송이 되게 되고, 전송된 비디오 요약은 사용자 터미널측에서 저장을 하거나, 혹은 직접적으로 추천할 수 있게 된다.

3. 사용자 선호도

사용자 선호도란 시청자가 시청자 단말에서 행하는 행동 패턴과 행동 패턴에 따라 예측되는 시청자의 방송 콘텐츠에 대한 시청자의 성향을 의미한다[3,4].

본 논문에서 적용하는 사용자 선호는 TV-Anytime에 기반하며, 콘텐츠 레벨의 선호 장면 정보 전달을 위하여, TV-Anytime 선호도 정보를 확장한다.

그림 3에서 보여주는 것과 같이 TV-Anytime의 사용자 선호도 정보 중 ClassificationPreferences DS는 사용자가 선호하는 장르 또는 언어와 같이 콘텐츠에 대한 여러가지 classification들에 관한 선호도를 나타낸다[4]. 개인화된 비디오 요약 방송 서비스를 위해서 필요한 이벤트 정보는 ClassificationPreferences DS의 자식 엘리먼트(child element)들에 'Events'로 삽입되어 확장된다.

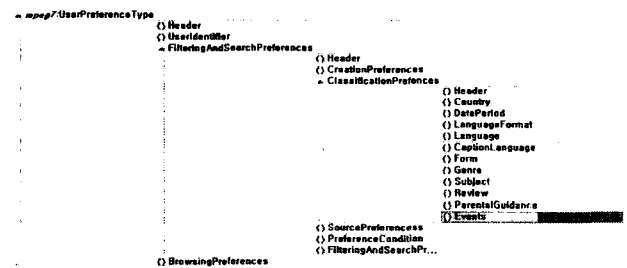


그림 3. 확장된 사용자 선호도 정보의 enhanced grid view

'Events'의 내용은 콘텐츠의 각 장르에 따라 달라진다. 예를 들어 영화 장르의 경우, "Action", "Romantic", "Comic" 등과 같은 이벤트로 나눌 수 있다. 이벤트의 분류는 선호도 정보의 업데이트를 고려하여, TV-Anytime의 메타데이터들의 규격 중 Segmentation DS에서 규정하는 이벤트 타입을 따른다.

여기서, 규정된 사용자 선호도를 시청자 단말에서 업데이트하는 방법을 고려할 수 있다. 시청자 단말에서 시청자 A에 대한 이벤트에 대한 정보를 업데이트 하는 방법은 다음과 같다. 장르에 의해서 구분되는 각 콘텐츠에 대한 이벤트 선호도는 시청자에 의한 직접 입력 방식과, 시청자의 액션 히스토리 정보를 이용하여 업데이트가 된다. 직접 입력 방식은 시청자가 자신이 원하는 각 장르에 대한 선호 이벤트를 기록하고, 이 정보를 중심 정보로 이용하게 된다.

또한, 시청자가 반복 시청 구간에 대한 빈도수를 이용하여, 이벤트에 관한 선호도 정보를 업데이트한다. 다음과 같은 수식으로 업데이트된다. 각 콘텐츠의 장르를 G라고 하고, 장르에 대한 이벤트들을 E_i라고 하면,

$$P\{E_i\} = W_0 * I + W_1 * N_w + W_2 * N_m, \quad (1)$$

여기서 P{E_i}는 이벤트 E_i의 선호도 값을 의미한다. W₀, W₁, W₂는 W₁>W₂>W₀를 만족하는 가중치 값들이 된다. I는 시청자가 직접 입력한 선호도 지수를 나타내며, N_w는 지난 일주일내에 해당 이벤트에 대해서 'Play' 및 'Replay'를 한 횟수를 나타낸다. 또한 N_m는 지난 한달 동안 해당 이벤트에 대해서 'Play' 및 'Replay'를 한 횟수를 나타낸다.

'Play'와 'Replay'의 횟수에 대한 정보는 TV-Anytime 사용자 선호도 정보 중 UserActionHistory DS를 참조하여 획득하게 된다. 이 DS가 가지고 있는 UserActionList DS에 사용자가 행한 액션의 목록 및 횟수가 저장되기 때문이다.

4. 비디오 요약

서비스 제공자에서 전송된 사용자 선호도를 이용하여, 각각 해당되는 비디오 요약을 생성하기 위해서, 앞서 언급한 것처럼 콘텐츠에 대해서 미리 이벤트 별로 세그먼트되어야 한다. 현재까지 많은 연구들을 통해서 알 수 있듯이 [4,5], 모든 장르별 및 콘텐츠별 이벤트 검출은 자동으로 하기가 불가능하다. 본 연구에서는 서비스 제공자측에서 각 콘텐츠들의 이벤트를 세그먼트하는 방식은 반자동 혹은 수동에 의해서 생성되는 것으로 가정하였다.

본 연구에서는 반 자동으로 콘텐츠의 이벤트를 검출하는 방법으로 다음과 같은 방법을 이용하였다. 비디오 콘텐츠에 대해서 샷(shot)을 검출하고, 이웃하는 샷을 통합하여 하나의 이벤트를 생성하는 방법을 이용하였다.

샷을 검출하는 알고리즘은 다음과 같다. 샷 검출 알고리즘은 MPEG-7 비주얼 서술자들 중 HT(Homogeneous Texture), SC(Scalable Color) 그리고 EH(Edge Histogram)을 사용하여 생성된다. 식 (2)은 샷 변이(shot transition)를 결정하는 방법을 보여준다.

$$Shot_Transition_i = \begin{cases} True, & \text{if } Frame_{i_diff} > Th_1 \text{ and} \\ & |Frame_{i-1_diff} - Frame_{i_diff}| > Th_2 \\ False, & \text{Otherwise} \end{cases} \quad (2)$$

여기서, Frame_{i_diff}은 (i-10)번째 프레임과 i번째 프레임 사이의 프레임 차이이다. 그리고, Th₁과 Th₂는 경험적인 임계값들이다. 제안하는 샷 검출 알고리즘은, 현재 비디오 프레임과 10번째 앞의 프레임을 계산하도록 10 프레임의 간격을 갖는 마스크를 사용하여 프레임간의 차이를 계산한다. Shot_Transition_i이 True값을 가지면, i번째 프레임에서 샷 변이를 검출된다. Frame_{i_diff}은 다음의 식들을 통해 계산된다.

$$Frame_{i_diff} = Diff_SC_i + Diff_EH_i + Diff_HT_i, \quad (3)$$

$$Diff_SC_i = (SC_frame_{i-10} - SC_frame_i)^2 / Max_Diff_SC, \quad (4)$$

$$Diff_EH_i = (EH_frame_{i-10} - EH_frame_i)^2 / Max_Diff_EH, \quad (5)$$

$$Diff_HT_i = (HT_frame_{i-10} - HT_frame_i)^2 / Max_Diff_HT. \quad (6)$$

여기서 Diff_SC_i, Diff_EH_i과 Diff_HT_i은 (i-10)번째 프레임과 i프레임에서 각각 SC, EH와 HT의 정규화된 차이를 나타낸다. 그리고, SC_{frame_i}, EH_{frame_i}과 HT_{frame_i}은 i번째 프레임에서 각 서술자들의 값을 말한다. Max_Diff_SC, Max_Diff_EH과 Max_Diff_HT은 Diff_SC_i, Diff_EH_i과 Diff_HT_i의 최대값이다.

상기 샷 검출 알고리즘을 통해 획득된 샷들 중 의미있는 샷들을 통합하여 각각의 이벤트들을 생성하게 된다. 그림 5는 비디오 요약 생성 틀을 이용하여 샷을 이벤트로 병합하는 모습을 보여준다.

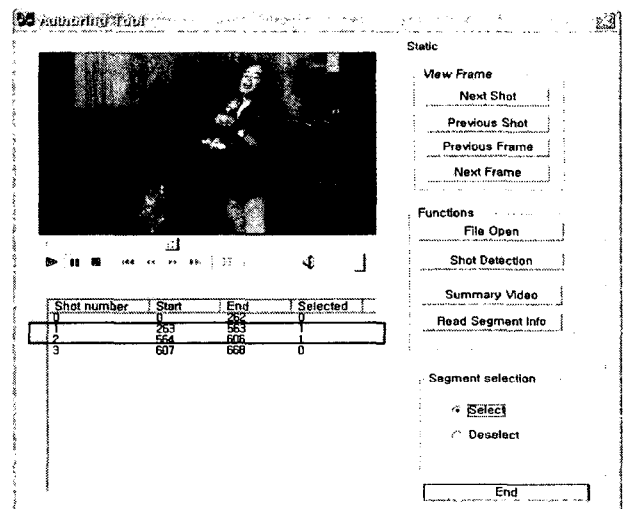


그림 4. 샷 검출 알고리즘에 의해 검출된 샷들을 이벤트로 병합하는 비디오 요약 틀

병합된 이벤트들은 서비스 제공자의 해당

컨텐츠들에 대한 이벤트들로 저장되고, 3장에서 언급되었던 ‘Events’ 내용과의 매칭에 따라서 비디오 요약의 생성에 적용된다.

5. 실험 및 결론

제안하는 개인화된 비디오 요약 방송 서비스의 유효성을 테스트하기 위해 영화 장르의 비디오에 대해서 이벤트(event)들을 세그먼트(segment)하고, 해당 선호도 정보에서 추론된 선호 이벤트 정보에 따라 비디오 요약(video skim)을 생성하고, 시청자 단말에 제공하는 실험을 하였다.

네트워크를 통해서 사용자 3명에 동일한 컨텐츠(영화)대한 사용자 선호도들을 입력받은 서비스 제공자는 그림 5에서 보여주는 것과 같은 시퀀스들로 구성된 비디오 요약을 각각의 시청자 단말에게 전송하였다. 각 사용자들은 선호 이벤트들이 액션 이벤트, 자동차 및 스포츠 이벤트, 그리고 로맨틱 이벤트들로 구성되어 있다. 아래 그림은 영화 ‘Cat woman’에 대한 개인화된 비디오 요약 시퀀스를 보여준다.

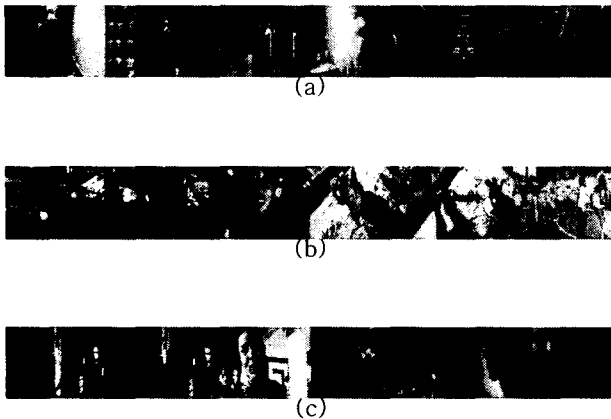


그림 5. 사용자 선호도에 따라 보여주는 서로 다른 비디오 요약 예, (a) 액션 이벤트로 구성된 요약, (b)자동차 추격 이벤트 및 스포츠 이벤트로 구성된 요약, (c) 로맨틱 이벤트로 구성된 요약.

상기 실험을 통해서, 제안하는 맞춤형 방송 서비스 시나리오를 통해서 시청자 개개인에게 차별화된 비디오 요약을 제공할 수 있음을 확인하였다. 따라서, 본 맞춤형 비디오 요약 서비스는 컨텐츠에 대한 흥미를 보다 유발 시킬수 있으며, 시청자들의 만족도도 증가와 함께 컨텐츠 소비를 유발할 것이며, 향후 에이전트를 이용한 방송 시스템과 연동하여, 시청자의 지시나 행위 없이 시청자에게 맞추어진 방송 컨텐츠의 비디오 요약본을 서비스 제공자가 생성하고 시청자 단말에서 시청자에게 추천이 가능해 질것으로 예상된다. 본 논문에서 제안하는 연구의 아이디어를 확장하여 실제의 디지털 방송 서비스에 충분히 적용 가능할 것으로 예상된다.

6. 참고문헌

- [1] 이석필, “맞춤형 방송 서비스와 PVR 동향”, 전자공학회지 제 31권 5호, 2004.
- [2] 강대갑, “지상파 디지털 방송 및 데이터 방송 기술”, 전자공학회지 제 29권 7호, 2002.
- [3] 강경욱, 김진웅, “메타데이터를 이용한 지능형 방송 기술”, 전자공학회지 제 29권 7호, 2002.
- [4] SP003v12 PartA, “Metadata Specification,” *the TV-Anytime Forum*, 2002.
- [5] Ying Li et al., “Video Content Analysis using Multimodal Information: For Movie Content Extraction, Indexing and Representation,” *Kluwer Academic Publishers*, 2003.
- [6] Remco C. Velkamp, et al., “State-of-the-Art in Content-Based Image and Video Retrieval.” *Kluwer Academic Publishers*, 2001.
- [7] B. S. Manjunath, et al., “Introduction to MPEG-7: Multimedia Content Description Interface”, John Wiley & Sons, LTD, 2002.
- [8] Sung Ho Jin, Tae Meon Bae, Jin Ho Choo, Yong Man Ro and Kyeongok Kang, “A MPEG-7 and TV-Anytime based Multi-Agents Platform for Interactive Broadcasting System”, SPIE2003, Vol. pp., 2003.

Acknowledgements

본 연구는 한국전자통신연구원(ETRI)의 지원받은 수행중인 SmarTV과제의 연구 결과 중 일부분임.