

Visual Rhythm과 Pattern Matching을

이용한 EPG 시간 정보 갱신

윤형화, 김정림, 설상훈

고려대학교 전자공학과

Updating Time Table of EPG using Visual Rhythm and pattern matching

Hyong-hwa Yoon Jung-Rim Kim, Sanghoon Sull

School of Electrical Engineering, Korea Univ

{frog, mrbone, sull}@mpeg.korea.ac.kr

요 약

최근 방송의 디지털화에 따라 수많은 채널의 수용이 가능하게 되어, 현재와 같이 신문의 방송편성표를 보고 사용자가 원하는 프로그램의 시작 시간을 알아내는 것은 힘들어 졌다. 디지털 방송에서는 신문 방송 편성표와 같은 기능을 하는 방송 프로그램에 대한 정보를 보여주는 EPG(Electronic Program Guide)를 제공하고 있다.

현재 서비스 되고 있는 EPG를 보면, 프로그램의 시작 시간을 GUI로 사용자가 쉽게 알아 볼 수 있어 하고 있어 신문 편성표나 인터넷으로 제공되는 각 방송국의 편성표와 달리 TV 화면에서 프로그램 정보를 검색할 수 있어 편리해졌다. 그러나 이 EPG에서 제공되는 시간은 매우 부정확하기 때문에 시청자가 불편을 느낄 수 있다.

본 논문에서는 방송의 시작 시간을 정확히 알아내어 EPG를 정확히 갱신하는데 목적이 있다. 이를 위해 방송의 시작부분을 미리 저장한 후, 시작 부분을 Visual Rhythm을 이용해 shot들을 추출하여 데이터베이스를 구성하고, 이와 같은 방식으로 방송 스트림의 shot들도 추출하여 데이터베이스와의 Pattern Matching을 통해 보다 정확한 방송시간으로 EPG를 갱신하고자 한다.

I. 서 론

최근 디지털 비디오 기술은 정보통신의 발달에 힘입어 빠른 속도로 진보하고 있다. 현재 TV는 아날로그 방식과 디지털 방식을 병행해서 송출하고 있다. 국내의 디지털 방송은 시험 방송을 거쳐 본 방송을 시작하였다. 2005년까지 건국에 단계적으로 확대 실시할 예정이며, 2010년까지 아날로그/디지털 동시 방송을 할 예정이다. 최근 이러한 흐름에 발맞추어 디지털 비디오 처리 및 디지털 방송에 관한 많은 연구들이 진행 중이다.

본 논문에서는 비주얼 리듬을 이용해 검출된 shot을 이용해 각 방송 프로그램의 시작 부분 비디오와 실시간 전송되는 방송 비디오와의 matching 기법을 통해 정확한 방

송 프로그램의 시작 시간을 알아내어 EPG의 시간 정보를 갱신하는 방법을 제안한다.

본 논문의 구성은 다음과 같다. 제 II장에서는 EPG의 개념 및 비주얼 리듬을 이용한 shot 검출 기법과 데이터베이스화에 대해 알아보고 제 III장에서 본 논문에서 제안하는 방송 프로그램의 시작 부분을 찾아내는 pattern matching 기법을 소개하고 그 결과를 제 IV장에서 기술한다. 마지막으로 제 V장에서 최종적인 결론 및 향후 연구과제에 대해 기술한다.

II. EPG 와 비주얼 리듬

1. EPG

Electronic Program Guide(전자 프로그램 가이드)의 약자이다. EPG는 방송의 여유채널을 이용하여 방송 프로그램에 대한 시간표를 제공하는 기능이다. 방송 채널의 증가에 따라 시청자가 원하는 채널을 빠르게 검색하고 쉽게 방송을 시청할 수 있도록 해주는 것이 EPG의 주기능이다. 그림 1은 EPG를 통한 채널 검색 및 시청자에게 각 방송의 정보의 내용을 전달해주는 장면이다.

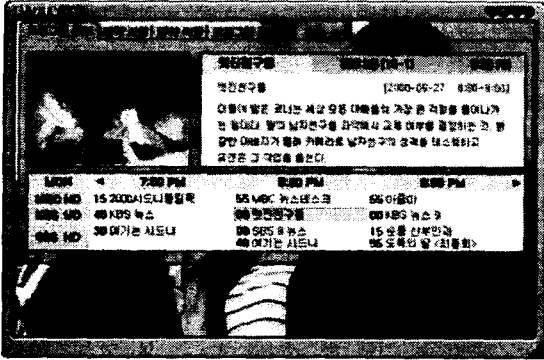


그림 1 EPG의 예

EPG의 역할로는 먼저 포털채널기능을 들 수 있다. 시청자가 원하는 채널을 인터넷의 포털 사이트와 같은 방식으로 쉽게 접근하도록 도와준다. 디지털 TV는 수백 개의 채널을 동시 전송할 수 있기 때문에 시청자의 기호에 따라 맞춤형 TV프로그램 편성표를 구성할 수 있다. 또 디지털 저장 매체에 저장하는 경우 빠르고 쉽게 EPG상에 있는 프로그램 방송 시간 정보를 이용해 간단한 조작으로 프로그램의 예약녹화 등을 할 수 있는 기능이 제공된다.

시청자는 EPG의 부정확한 시간 정보 때문에 시청자는 방송이 시작할 때까지 기다리거나, 예전 시간 보다 먼저 방송이 시작 해버려 불편을 느끼게 된다. 이러한 이유등으로 EPG시간 정보의 정확성은 꼭 필요하다.

2. 비주얼 리듬(Visual Rhythm)

원본 비디오 V 에서의 임의의 프레임(frame) t 의 (x, y) 의 픽셀 값을 $f_V(x, y, t)$ 라고 하면, 비주얼 리듬 VR은 다음과 같이 정의 된다.

$$VR = \{ f_{VR}(z, t) \} = \{ f_V(x(z), y(z), t) \} \quad (1)$$

여기서 $x(z)$ 와 $y(z)$ 는 독립변수 z 에 대한 1차원 함수이다. (z 와 t 는 각각 비주얼 리듬에서 세로축과 가로축이 된다)

간단히 말하면, 비디오의 프레임을 대각선 방향으로 샘플링 하여 각 프레임의 시간 순서에 따라 표시한 것이 비주얼 리듬이다. 다음 그림 2는 비주얼 리듬의 샘플링 방식을 그림으로 나타낸 것이다.

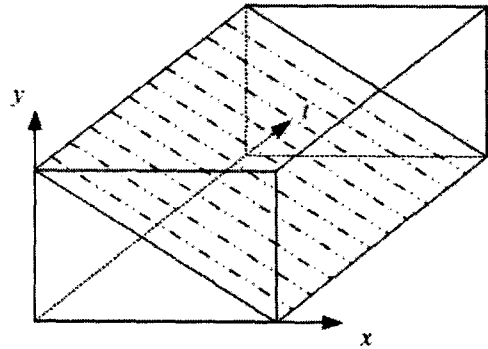


그림 2 비주얼 리듬의 대각선 샘플링 방식

비주얼 리듬은 비디오의 각 프레임으로부터 샘플링 된 픽셀들의 sequence형태로 구성되어 있기 때문에 이웃한 두 프레임 간의 불연속성을 이용한 shot 검출에 있어서 탁월한 성능을 발휘한다. 그림 3은 뮤직비디오의 비주얼 리듬결과이다. 그림을 통해 불연속적인 부분을 확인할 수 있다. 연속적인 각각의 부분을 shot(장면)이라 한다.

본 논문에서는

$$d(z, t) = |f_{VR}(z, t) - f_{VR}(z, t-1)| \quad (2)$$

인접 두 프레임간의 비주얼 리듬의 절대값을 이용한 [1]에서 제안된 방법으로 shot을 검출하였다.



그림 3 불연속적 비주얼 리듬의 예

비주얼 리듬을 통한 shot 검출 방법은 인코딩 당시의 bit-rate, fps(frame per seconds)에 따라 다르게 shot을 검출해 낼 수 있다. 즉 비주얼 리듬이 언제나 같은 shot을 검출해 낼 수 없다는 것이다. 그러나 이러한 경우는 인접한 frame간의 불연속도가 작을 때 존재하게 된다.

shot수가 작은 경우에 있어서 연속적인 패턴 검출이 어렵기 때문에 shot이 5shot 이상 되는 시작부분 비디오만 데이터베이스에 저장하였다. 미리 저장된 각각의 방송 시작 비디오는 비주얼 리듬을 이용한 shot 검출 후, 각 shot의 비주얼 리듬에서 색상 히스토그램을 이용한 특징 벡터를 추출한다. 특징 벡터는 [2]에서 제안된 LBG(The Linde-Buzo-Gray)알고리즘을 이용하여 2D로 구성 및 clustering을 하였다.

III. 방송 시작 부분 검색

이 장에서는 방송 시작 부분을 검색하기 위한 방송 시작 비디오의 비주얼 리듬과 방송 비디오의 비주얼 리듬과의 색상 히스토그램 비교를 통한 pattern matching 기법

에 대해 기술한다.

방송 비디오가 실시간으로 전달되고, 그 방송 비디오를 비주얼 리듬을 이용해 shot 검출이 완료되면, 각 shot의 비주얼 리듬에 대한 색상 히스토그램을 만든다. 그 후, 방송 비디오 shot의 히스토그램과 데이터베이스에 저장된 shot들의 히스토그램과의 거리(distance, dissimilarity) 비교를 통해 가장 가까운 데이터베이스의 원소만 유사 shot vector에 저장한다.

그림 4와 같이, 방송 비디오의 shot들의 집합 S와 방송 시작 부분의 shot들의 집합을 각각 T1, T2, T3...라하고, 각 방송 시작 부분의 shot에 두 번째 숫자를 표시해 준다. 예를 들면 T1의 첫 번째 검출 shot은 T11이라 할 수 있고, T2의 세 번째 shot은 T23라 한다. shot S1과 T31은 거리가 가깝은 것을 알 수 있다. S3의 경우 데이터베이스에 저장된 방송 시작 부분의 히스토그램과 거리가 가까운 shot은 없다.

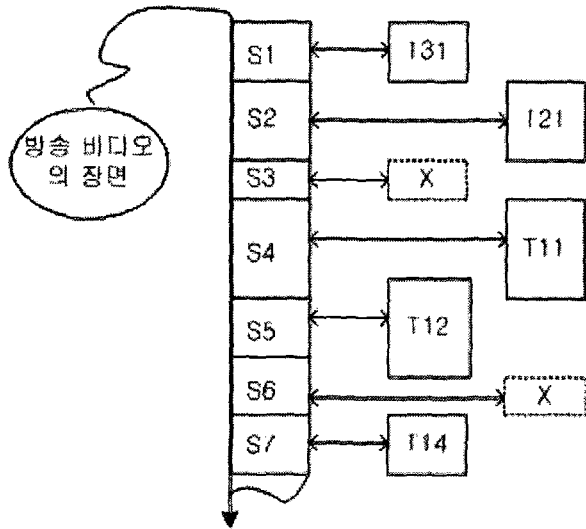


그림 4. 방송 비디오 shot들의 matching pattern

S4의 경우 T11과 거리가 가깝고, 그 다음 shot S5의 경우도 T12와 거리가 가깝다. S7의 경우에도 T14와의 거리가 가깝다.

이미 언급한 바와 같이 비주얼 리듬을 이용한 shot 추출시 fps나 encoding bit-rate에 따라서 다른 shot이 추출될 가능성이 있다. 이러한 경우 S4부터 S10까지 T1의 모든 shot이 일대일 대응이 되지 않는다. 이런 경우도 감안하여 다음과 같은 Pattern Matching 기법을 제안한다.

유사 shot 벡터에 연속적으로 3 shot 이상의 거리가 가까운 shot이 없는 경우에 있어 데이터베이스와 거리가 가까운 처음 shot부터 끝 shot까지의 수를 N_T 라고 하고, 데이터베이스에서 방송 비디오와의 거리가 가까운 shot의 수를 N_S 라고 할때,

$$\varepsilon_T = \frac{N_T}{N_S} \quad (4)$$

로 정의한다.

이때 $\varepsilon_T \geq 0.5$ 일 때, 데이터베이스와의 matching이 일어나는 첫 번째 방송 비디오의 shot을 시작부분으로 검출해 낸다.

IV. 실험결과

본 장에서는 위에서 제안한 알고리즘을 사용하여 방송 비디오와 방송 시작 부분 비디오의 히스토그램 비교를 통해 얻어진 유사 shot 벡터와 이를 이용한 방송 시작부분 검색 결과를 기술한다.

1. 비주얼 리듬을 이용한 shot 검출

본 실험을 위한 비디오 자료는 마이크로소프트사의 ASF codec을 이용하였고, 320 * 240 해상도에서 30 fps로 인코딩 하였다. 프로그램 시작 부분의 비디오는 인코딩한 방송의 부분을 잘라내어 구성하였다. 실험용으로 2002년 10월 9일부터 18일까지의 10일간의 MBC의 방송을 7시부터 각각 약 5시간씩 녹화하여 사용하였다.

표1은 18일 방송에서 잘라낸 방송시작 부분 비디오를 데이터베이스화해서 18일 방송 비디오와의 matching 결과를 유사 shot 벡터로 나타낸 것의 일부이다.

표 1. 18일 방송의 유사 shot 벡터

shot 번호	DB의 index	비고
:	:	:
1868	T11	T1시작
1869	T12	
1870	T13	
:	:	:
2884	T41	T4시작
2885	T42	
계속된	일치를 보였다.	
2893	T49	

이와 같은 경우에서 shot번호 1868과 shot번호 1869에서 시작하는 프로그램은 $\varepsilon_T = 1$ 로 프로그램의 시작 부분이 됨을 알 수 있다.

shot 번호를 통해 정확한 방송 시간을 알아 낼 수 있었고, 이를 통해 수정된 방송 시간은 다음과 같다.

표 2 수정된 18일 방송 시간 정보

예정 시작 시간	실제 시작 시간	방송 구분
8:20	8:16:37	드라마
8:55	8:55:06	뉴스
9:50	9:49:22	스포츠 뉴스

표 3은 18일 방송에서 잘라낸 방송시작 부분 비디오를 데이터베이스화해서 11일 방송 비디오와의 matching 결과를 나타낸 것이다.

표 3 수정된 11일 방송 시간 정보

예정 시작 시간	실제 시작 시간	방송 구분
8:20	8:17:08	드라마
8:55	8:55:41	뉴스
9:50	9:48:18	스포츠 뉴스

V. 결론 및 향후과제

본 논문은 visual rhythm과 pattern matching 기법을 통한 방송 프로그램의 시작 부분을 실시간으로 검출해 내는 방법을 제안하였다. 본 연구를 통해 EPG의 시간 정보를 정확하게 갱신해줄 수 있다.

EPG 갱신을 통한 Video Editing 기법은 본 연구를 통해 다루어지지 않았다. Video Editing을 위한 알고리즘의 수정 및 구체적인 구현은 향후 과제로 남긴다.

참고문헌

- [1] M. Chung, H. Kim, and S. M. Song, A scene boundary detection method, Proc. IEEE International Conference on Image Processing, vol. 3 (2000) 933-936.
- [2] Ruey-Feng Chang, Wen-Tsuen Chen, Jia-Shung Wang, A fast finite-state algorithm for vector quantizer design, Signal Processing, IEEE Transactions on , Volume: 40 Issue: 1 , Jan. 1992
- [3] Yeung, M.M, Bede Liu, Efficient matching and clustering of video shots, Image Processing, 1995. Proceedings., International Conference on , Volume: 1 , 1995 Page(s): 338 -341 vol.1