

영상 정보 기반의 채널 로고 인식 기법

장원동 이철우 김창수

고려대학교 전기전자전파공학부

{wdjang, chulwoo, cskim}@mcl.korea.ac.kr

Channel Logo Recognition based on Visual Information

Won-Dong Jang Chulwoo Lee Chang-Su Kim

School of Electrical Engineering, Korea University

요약

최근 다양한 모바일 디바이스의 개발 및 대중화로 인해, 사용자가 콘텐츠에 노출되는 시간이 급증하고 있다. 이와 함께, 사용자가 필요한 정보를 선택적으로 제공하기 위한 사용자 맞춤형 서비스의 개발도 많은 관심을 받고 있다. 이를 위해서는 콘텐츠가 내포하는 정보를 자동적으로 추출하여 인식하는 기술이 필수적이며, 채널 정보는 지능적 상황 인지를 위해 방송 영상이 가지는 중요한 정보 중 하나이다.

이에 본 논문은 영상 정보 기반의 채널 로고 인식 기법을 제안한다. 본 기법은 채널 로고가 포함된 동영상에서 색상 정보와 그라디언트를 추출하여 동영상 내의 각 좌표가 로고에 해당할 가능성을 의미하는 Logosity 개념을 정의한다. Logosity를 기반으로 채널 로고 후보 영역을 탐지하고, 각 로고 후보에 대해 채널 종류를 판단하는 분류를 수행한다. 실험을 통하여 제안하는 알고리즘이 다양한 채널의 영상에 대해 뛰어난 인식 및 분류 성능을 나타냄을 확인한다.

1. 서론

최근 멀티미디어 통신 기술의 발달과 모바일 디바이스의 대중화로 인해 사용자가 콘텐츠에 노출되는 시간이 급증하고 있다. 사용자들은 필요한 정보를 능동적으로 판단하지만 방대한 양의 콘텐츠로 인해, 유용한 정보를 취사선택하는 데에 어려움을 느끼고 있다. 따라서 사용자가 필요한 정보를 콘텐츠에 기반하여 제공하는 지능형 상황인지 기술에 대한 중요성이 대두되고 있다.

채널 정보는 장르와 더불어 콘텐츠를 분류하는데 중요한 정보 중 하나이다. 가정에서 TV를 시청할 경우, 송출되는 전파를 통해 채널 정보를 쉽게 얻을 수 있지만, 모바일 디바이스 상에서 동영상 내의 채널 정보를 인식하려면, 시각적인 정보에만 의존하여야 한다. 뿐만 아니라 동영상에는 채널 정보 외의 시각적인 정보들이 다수 포함되어 있기 때문에, 시각적인 정보만을 바탕으로 채널 정보를 인식하기는 쉽지 않다. 따라서 본 논문에서는 영상 정보 기반의 채널 로고 인식 기법을 제안한다.

본 논문의 구성은 다음과 같다. 2장에서는 제안하는 채널 로고 인식 기법에 대해서 기술한다. 3장에서는 제안하는 알고리즘의 성능에 대해 검토한다. 그리고 4장에서 본 논문의 결론을 맺는다.

2. 제안하는 채널 로고 인식 기법

본 논문에서 제안하는 채널 로고 인식 기법의 구조는 그림 1과 같

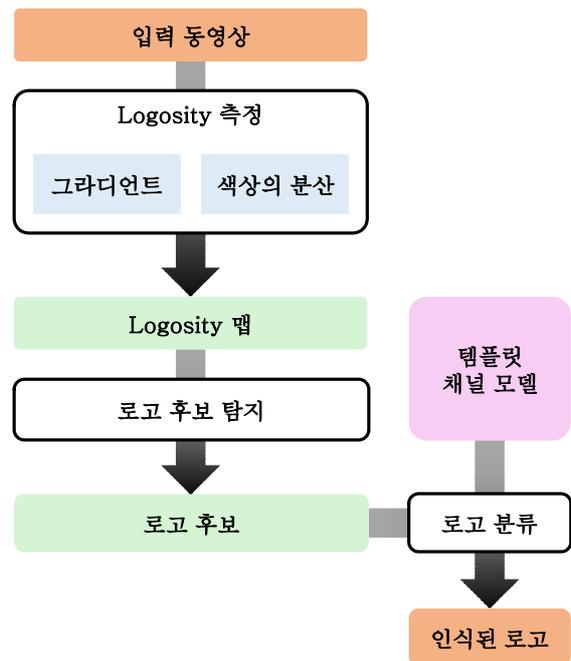


그림 1. 제안하는 채널 로고 인식 기법 구조

다. 제안하는 구조는 크게 채널 로고 탐지와 채널 로고 분류로 나뉜다. 채널 로고 탐지에서 채널 로고가 될 만한 후보 영역들을 추출하고, 채널 로고 분류에서 채널 로고 후보가 속하는 채널을 결정한다.

제안하는 기법은 영상을 R, G, B 채널로 분리하여 독립적으로 처리하고, 이하 내용은 R 채널을 기준으로 기술한다. 채널로고 주변에는

이 논문은 2013년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원(No.2009-0083495) 및 2013년도 삼성전자의 지원을 받아 수행된 연구임.



그림 2. 채널 로고 후보 탐지 결과

그라디언트가 항상 일정 크기 이상으로 분포하는 특징이 있으므로, 매 프레임 t 마다 합하여 아래와 같이 누적 그라디언트를 계산한다.

$$G_R(x, y) = \sum_{t=1}^T \sqrt{g_R^h(x, y, t) + g_R^v(x, y, t)} \quad (1)$$

여기서 (x, y) 는 영상에서의 2차원 좌표를 나타내고, $g_R^h(x, y, t)$ 와 $g_R^v(x, y, t)$ 는 각각 수평, 수직 성분의 그라디언트를 나타낸다. 그라디언트 추출에는 Sobel 연산자를 사용하였다. 채널 로고 탐지를 위한 또 다른 요소로 화소값의 유지도를 측정하였다. 채널 로고는 다른 위치에 비해 화소값이 일정하게 유지되는 특성을 가진다. 화소값 유지도는 시간 축에 대한 화소값의 분산을 기반으로 아래 식과 같이 구할 수 있다.

$$V_R(x, y) = 1 / \left(\frac{1}{T} \sum_{t=1}^T (I_R(x, y, t) - \bar{I}_R(x, y))^2 \right) \quad (2)$$

여기서 $I_R(x, y, t)$ 는 R채널의 화소값, $\bar{I}_R(x, y)$ 는 (x, y) 위치에서의 시간 축에 대한 평균 화소값을 나타낸다. 그라디언트와 화소값 유지도를 기반으로 Logosity라는 개념을 아래와 같이 정의한다.

$$L(x, y) = \max \left(\begin{matrix} V_R(x, y) G_R(x, y), \\ V_G(x, y) G_G(x, y), V_B(x, y) G_B(x, y) \end{matrix} \right) \quad (3)$$

Logosity는 동영상의 각 위치가 채널 로고를 구성할 가능성을 뜻한다.

Logosity 값을 바탕으로 큰 logosity 주변에 또 다른 큰 logosity가 존재하면 동일한 그룹으로 포함시켜주는 그룹화를 수행하여 채널 로고 후보를 그림 2와 같이 생성한다. 채널 로고 분류를 위해 생성된 채널 로고 후보에 대해 특징을 추출한다. 특징으로는 logosity와 지역 이진 패턴 [1]을 사용한다. 지역 이진 패턴은 주변 화소와 화소값 비교를 수행하여 1 또는 0으로 표현하고, 이를 패턴화하여 특징으로 사용한다. Logosity에 대한 벡터는 \mathbf{v}_l , 지역 이진 패턴에 대한 특징 벡터는 \mathbf{v}_p 로 나타낸다.

제안하는 알고리즘에서는 드문 분포(sparse) 표현 기반의 분류 [2]를 수행한다. 분류에 앞서 템플릿 채널 로고를 기반으로 분류의 기반 모델을 생성해야 한다. 총 N 개의 템플릿 로고에 대해 특징 벡터를 측정, 아래와 같은 행렬로 나타낸다.

$$A = \begin{bmatrix} \left[\begin{matrix} \mathbf{v}_l^{n_1} \\ \mathbf{v}_p^{n_1} \end{matrix} \right] & \left[\begin{matrix} \mathbf{v}_l^{n_2} \\ \mathbf{v}_p^{n_2} \end{matrix} \right] & \dots & \left[\begin{matrix} \mathbf{v}_l^{n_N} \\ \mathbf{v}_p^{n_N} \end{matrix} \right] \end{bmatrix} \quad (4)$$

실험 동영상에 포함된 로고가 1번 템플릿 로고에 해당하는 로고라면, 실험 동영상에서 측정된 특징 벡터 $\mathbf{v} = [\mathbf{v}_l \ \mathbf{v}_p]^T$ 는 $\left[\begin{matrix} \mathbf{v}_l^{n_1} \\ \mathbf{v}_p^{n_1} \end{matrix} \right]^T$ 와 유사하다. 따라서 $\mathbf{Ac} = \mathbf{v}$ 라는 시스템에서, 계수 벡터 \mathbf{c} 를 구하면, $[1 \ 0 \dots 0]$ 에 가까운 형태로 나오게 된다. 의사역행렬을 통해 아래와 같이 계수 벡터 \mathbf{c} 를 구할 수 있다.

$$\mathbf{c} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{v} \quad (5)$$

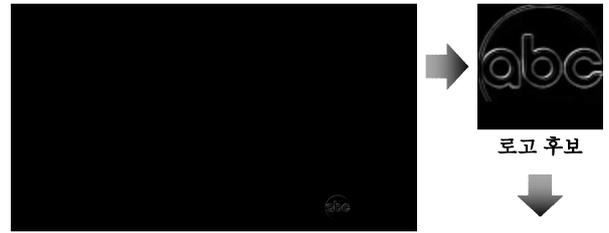


그림 3. 채널 인식 결과

가장 가까운 템플릿 로고를 아래의 식으로 찾아서 최종적으로 채널 로고의 종류 n^* 를 인식한다.

$$n^* = \operatorname{argmin}_n \| \mathbf{c} - \mathbf{1}_n \| \quad (6)$$

여기서 $\mathbf{1}_n$ 은 n 번째 요소만 1이고, 나머지는 0인 벡터를 나타낸다.

그림 3은 채널 인식 결과를 도시화한 것이다. Logosity 값을 기반으로 화소를 그룹화하여 로고 후보를 탐지하는데 성공하였다. 또한 분류 단계에서 템플릿 로고 중에 해당하는 로고로 정확히 분류한 것을 확인할 수 있다.

3. 실험결과

채널 종류 개수	테스트 개수	성공 개수	인식률
57	2110	2055	97.4%

표 1. 다양한 채널에 대한 인식 성공률

채널 로고 인식 실험 결과 표 1에 도시하였다. 동영상의 길이는 1분으로 제한하여, 사용자가 채널을 돌렸을 때의 상황을 제한 조건으로 가정하였다. 실험결과, 그 57개 채널 종류 2110개 동영상에 대해 2055개의 동영상에서 채널을 인식하는데 성공했고, 가장 낮은 인식률을 보인 채널에 대한 인식률은 85.2%였다. 따라서 제안하는 알고리즘이 안정적으로 채널을 인식하는 것을 확인할 수 있다.

4. 결론

본 논문은 영상 정보 기반의 채널 로고 탐지 기법을 제안하였다. 제안하는 기법은 실험결과를 통하여 다양한 채널 종류에 대해서 높은 인식률을 보이는 것을 확인하였다. 제안하는 기법이 적용된 기기에서 사용자에게 유용한 채널 정보를 제공하는 것을 기대해 볼 수 있다.

5. 참고문헌

[1] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, no.7 pp. 971-987, July 2002.
 [2] J. Wright, A. Yang, A. Ganesh, S. Sastry, and Y. Ma, "Robust Face Recognition via Sparse Representation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 201-227, Feb. 2009.