

Haar-like 특징을 이용한 유해영상 분류

박민수*, 김용민*, 박찬우*, 박기태**, 문영식*¹

*한양대학교 컴퓨터공학과

**한양대학교 BK21

e-mail: ysmoon@cse.hanyang.ac.kr

An adult image classification using Haar-like feature

Min Su Park*, Yong Min Kim*, Chan Woo Park*, Ki Tae Park**, Young Shik Moon*

*Dept. of Computer Science and Engineering, Hanyang University

**Institute of BK21, Hanyang University

요 약

인터넷 매체가 급증함에 따라 많은 이들에게 쉽게 노출 되어 유포되고 있는 유해 영상을 검출하기 위해 다양한 분류 방법에 대한 연구들이 이루어지고 있다. 본 논문에서 유해 영상 내의 피부색 영역에서의 Haar-like 특징을 추출하여 유해 영상을 분류하는 방법을 제안한다. 이를 위해, 첫 번째 단계에는 샘플 영상에 대하여 기존에 제안된 피부색 검출 방법을 적용하고, 두 번째 단계에는 검출된 피부색 영역 내의 Haar-like 특징을 추출한다. 각 샘플 영상에서 추출한 특징들은 SVM(Support Vector Machine)을 이용하여 각각 2000 장의 유해, 무해 영상을 학습한다. 학습된 모델은 유해 및 무해 영상이 혼합되어 있는 영상 집합들을 분류하는데 사용한다.

1. 서론

인터넷의 사용이 보편화 되면서 무분별한 정보 제공으로 인해 미성년자들이 유해정보를 쉽게 접할 수 있게 되었고, 이에 따라 미성년자 성범죄가 증가하고 있어 유해 정보 접근을 규제하고 차단하는 방법이 필요하게 되었다. 기존의 유해 정보 차단 방법들은 유해 정보를 가진 IP 주소와 키워드가 저장된 데이터베이스를 이용하여 유해 정보 접근을 차단하는 방법들이 있다. 하지만 이러한 방법들은 기존의 수집된 데이터베이스만을 이용하기 때문에 새로운 유해 사이트의 IP 주소와 키워드를 지속적으로 업데이트해야 하는 어려움이 있다. 그래서 최근에는 유해 사이트의 주요 콘텐츠인 유해 영상의 특징을 이용하는 방법에 대한 연구들이 이루어지고 있다.

2. RGB 컬러 공간에서의 피부색 검출 방법

유해 영상에서 많은 영역을 차지하는 피부색 검출을 위하여 본 논문에서는 식(1)에서 제안한 Kovac의 RGB 컬러 공간에서의 방법을 사용하여 RGB 컬러 각 채널에 대한 임계 값을 설정하여 피부색을 검출한다 [1].

$$\begin{aligned} R > 95, G > 40, B > 20 \text{ AND} \\ \text{MAX}(R, G, B) - \text{MIN}(R, G, B) > 15 \text{ AND} \\ |R - G| > 15, R > G, R > B \end{aligned} \quad (1)$$

3. Haar-like 특징을 이용한 특징추출

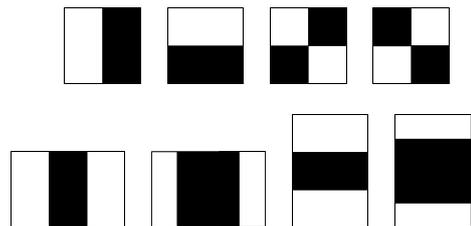
영상을 유해 혹은 무해로 분류하는 방법으로 피부색 영역의 비율과 모양 그리고 위치 특성을 이용하는 연구들이 있다[2-4]

제안하는 방법은 샘플 영상을 32 x 32 로 정규화한 후 RGB 컬러 공간에서의 임계 값을 통해 얻은 피부색 영역에서 사람의 형태특징 정보를 추출하기 위해 Viola 가 제안한 그림 1 의 Haar-like 특징을 이용한다 [5].

다양한 유해 영상들로부터 추출한 Haar-like 특징들에 대해서 SVM(Support Vector Machine)을 이용하여 유해 영상과 무해 영상 각각 2000 장에 대해 학습한다.

4. 실험결과

그림 2 은 유해 영상과 무해 영상의 피부색 검출 결과이며, 학습된 모델을 사용하여 인터넷을 통해 획득한 유해 영상과 무해 영상 각각 1000 장씩 분류 실험에 사용하였다.



(그림 1) Haar-like 특징

¹ 교신저자(Corresponding author)



(a) 입력 영상 (b) 검출 결과

(그림 2) 유해 영상의 피부색 검출 영역과 무해 영상의 피부색 검출 영역의 비교

제안하는 방법으로 실험한 결과 각 영상에 대하여 유해 영상 98%, 무해 영상 83%의 분류 율을 얻을 수 있었다.

5. 결론

본 논문에서는 RGB 컬러 공간에서 피부색 영역을 검출하고 검출된 영역에서 Haar-like 특징을 이용하여 유해 영상내의 특징을 얻는 분류 방법을 기술하였다.

다양한 환경에서의 유해 영상 특징들을 학습하기 위해서는 상당한 양의 샘플 영상이 요구되는 단점이 있지만, 샘플의 특징들이 많아질수록 더욱 효과적인 분류 결과를 보였다. 향후 과제로서 유해 영상들이 가지는 고유의 특성을 정확히 판별 할 수 있는 연구가 필요하다.

ACKNOWLEDGMENT

본 연구는 지식경제부 및 정보통신산업진흥원의 대학 IT 연구센터 지원사업의 연구결과로 수행되었음 (NIPA-2011-C1090-1131-0004)

참고문헌

- [1] J. Kovac, P. Peer, and F. Solina, "2D versus 3D colour space face detection," Proc. of EURASIP Conf. Focused on Video/Image Processing and Multimedia Communications, pp. 449-454, July 2003.
- [2] J. S. Lee, Y. M. Kuo, and P. C. Chung, "Naked image detection based on adaptive and extensible skin color model," *Pattern Recognition*, Vol. 40, No. 8, pp. 2261-2270, 2007.
- [3] J. L. Shih, C. H. Lee, and C. S. Yang, "An adult image identification system employing image retrieval technique," *Pattern Recognition Letters*, Vol. 28, No. 16, pp. 2367-2374, 2007.
- [4] L. Duan, G. Cui, W. Gao, and H. Zhang, "Adult image detection method base-on skin color model and support vector machine," Proc. of Asian Conf. on Computer Vision, pp. 797-800, January 2002.
- [5] P. Viola and M. J. Jones, "Robust real-time face detection," *International Journal of Computer Vision*, Vol. 57, No. 2, pp. 137-154, 2004.