

Eye-Tracker를 이용한 사용자 집중도 분석

김은솔[○], 김지섭, 장병탁

{eskim, jkim, btzhang} @ bi.snu.ac.kr

서울대학교 컴퓨터공학부 바이오지능연구실

Eye-Tracking Analysis of Attention in TV Drama Viewers

Eun-Sol Kim[○], Jiseob Kim, Byoung-Tak Zhang

Biointelligence Lab, School of Computer Science and Engineering, Seoul National University

요 약

이미지 처리와 관련한 기계 학습 분야의 주제는 크게 이미지 인식과 생성으로 나눌 수 있다. 그 중 이미지 생성문제는 기계 학습 분야의 오래된 문제로 예를 들면 '비행기'라는 입력을 받고 기계가 비행기 그림을 만들어 내는 것이다. 이 문제는 기계를 학습시켜 해결하기 매우 어려운 문제인 반면 사람에게서는 매우 자연스러운 문제이다. 이에 우리는 초기 연구 결과로서 이미지 생성 문제에 적용하기 위한 사람 실험을 제시한다. Eye-Tracker는 사람의 안구 움직임을 통해 사람의 집중도와 피로도를 측정할 수 있는 장치이다. 우리는 Eye-Tracker를 이용하여 드라마를 시청할 때 사용자의 안구 움직임을 기록하고 이를 이용하여 사용자의 드라마 시청 패턴을 분석하였다. 이 결과를 바탕으로 사람의 안구 운동 데이터를 기계로 이용한 이미지 재생성 문제에 적용할 수 있는 모델을 제시한다.

1. 서 론

사람은 언어나 음악 또는 냄새와 같은 자극을 외부에서 받으면, 그에 대한 이미지를 머릿속에서 그려볼 수 있다. '비행기'라는 단어를 보면 순식간에 여러 종류의 비행기를 머릿속에서 그릴 수 있고, 신나는 음악을 들을 때에는 정확한 물체는 아니지만 자신의 감정을 나타내는 어렴풋한 이미지를 만들어 낼 수 있다. 이처럼 이미지를 생성하는 일은 사람에게 매우 자연스럽지만 현재의 기술로 기계가 이미지를 재생성해 내기란 어려운 일이다. 그림을 보고 어떤 그림인지 맞추거나 그림들을 일정한 분류에 맞추어 나누는 문제는 지금의 기계 학습 알고리즘으로 충분히 해결할 수 있지만, 단어를 보고 그에 해당하는 그림을 그려보는 일은 매우 어려운 일이다. 이러한 차이는 기존의 기계 학습 알고리즘이 컴퓨터를 기반으로 문제에 접근하여 사람이 정보를 처리하는 방법과 매우 다르게 고안되었기 때문이다.

사람의 정보처리 원리와 기계의 정보처리 원리가 근본적으로 다르기 때문에, 사람의 정보 처리 방법을 연구하여 기계 학습에 적용하려는 시도가 많이 이루어지고 있

다. 사람의 사고를 모델링하는 방법은 접근 방법에 따라 다음과 같이 크게 3개로 나눌 수 있다[1].

- 1) fMRI, MEG, EEG 등을 이용한 뇌의 신호 분석
- 2) Eye-Tracker, Motion Tracker 를 이용한 embodiment cognition 분석
- 3) 언어, 사고와 같은 상위 레벨의 인지 기능 분석

우리가 제시하는 방법은 위의 접근 방법에서 2번 단계로, Eye-Tracker를 이용하여 사용자 개인의 특성을 프로파일링 하고 이를 기계 학습 알고리즘에 적용하는 방법을 제시한다.

Eye-Tracker는 사람의 안구 운동(eye gaze)을 기록하기 위한 장비로 전통적으로 난독증이나 정신 질환의 감정 및 치료에 많이 쓰이는 장비이다. 최근에는 Eye-Tracker가 다양한 분야에 쓰이는데 심리학 연구에도 많이 사용되고 있으며 연구용도 뿐만 아니라 광고에서 소비자가 어느 요소에 흥미를 느끼는지 알아보는 것과 같은 응용 목적으로도 많이 사용되고 있다. 우리는 사용자가 드라마를 볼 때의 집중도를 파악하기 위해 Eye-Tracker를 사용하였다. 이 데이터를 통해 사용자 개인의 집중도를 프로파일링하고 이미지 재생성 문제에 적용할 수 있는 알고리즘을 제시한다.



[그림 1] 거실 환경으로 꾸며진 가상 현실 실험 세트에서 Eye-Tracker를 착용하여 드라마를 시청하는 실험 모습

2. 실험 설계

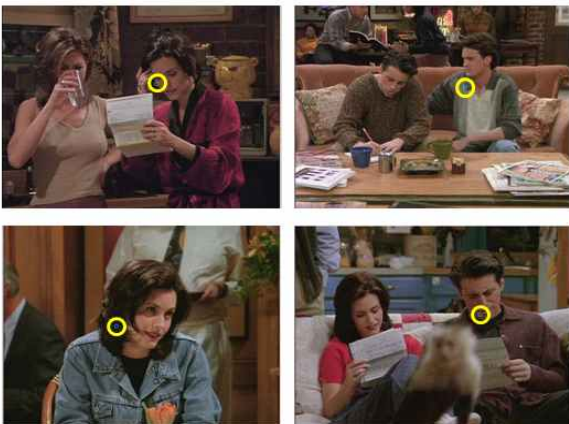
우리는 드라마를 시청하는 동안 피험자의 안구 움직임을 Eye-Tracker로 기록하여 피험자의 집중도를 프로파일링 할 수 있는 실험을 제시한다. 피험자는 [그림 1]과 같이 거실 환경으로 꾸며진 가상 현실 세트에서 미국 유명 TV 드라마 Friends를 (시즌1 20화) 약 20분 동안 시청한다. 동시에 안경형태의 Eye-Tracker를 착용하여 피험자의 안구 움직임을 측정한다.

Eye-Tracker 장비는 Arrington research 사의 BS007을 이용하였고 60 fps로 양쪽 안구의 움직임을 기록하였다.

3. 사용자 집중도 프로파일링

3.1 Eye-Tracker를 이용한 안구 운동 분석

Eye-Tracker를 이용하여 얻은 안구 운동은 1초에 60번 기록 되는데, 이를 60 fps의 동영상에 표시해보면 [그림 2]과



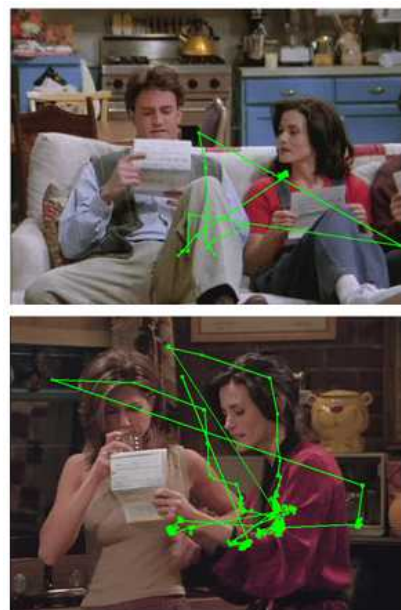
[그림 2] 동영상에서 안구의 초점

같이 하나의 프레임마다 하나의 점으로 표시할 수 있다. 전체 드라마에서 안구의 움직임을 이와 같이 분석하면 시간에 따른 사용자의 관심도를 파악할 수 있는데, 이에 대해 3.2절에서 논의하였다.

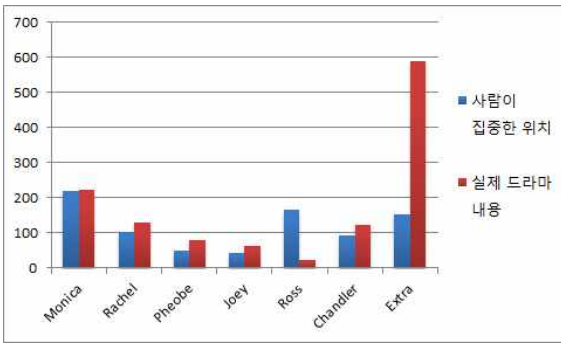
또, 정지된 이미지 상에서 피험자의 집중도를 분석할 수 있다. 동영상에서 이미지를 빠른 속도로 캡처하면, 연속한 이미지가 비슷한데, 비슷한 이미지를 묶어 그 그룹을 대표하는 하나의 이미지를 뽑았다. 즉 이미지 하나는 어떤 시간 구간을 대표하고, 그 시간에 기록된 안구 움직임을 이미지 하나에서 분석하는 것이다. 이렇게 분석하여 [그림 3]과 같이 하나의 이미지에서 움직이는 피험자의 시선의 경로를 확인할 수 있다. 이것으로 3.3 절에서는 사용자의 집중도 지도를 그렸는데 드라마를 시청할 때 사용자 집중도의 spatial한 특징을 파악할 수 있다.

3.2 안구 운동에 나타난 사용자의 드라마 시청 패턴

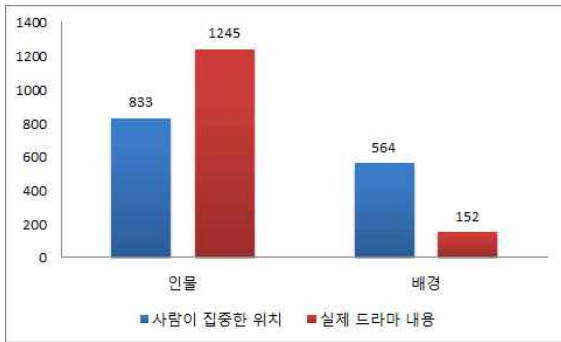
우선, 시간에 따른 안구 움직임을 분석하여 피험자의 드라마 시청 패턴을 분석하였다. 동영상을 캡처하여 이미지 세트로 만든 뒤, 피험자의 시선이 고정된 영역을 인물과 배경을 중심으로 나누어본 결과를 [그림 4]에 표시하였다. 비교를 위하여 얼굴 인식과 음성 인식을 이용하여 비디오 내용을 자동으로 프로파일링한 결과를 같이 제시하였는데, 사람은 인물 뿐 아니라 핵심이 되는 배경이나 사물에 초점을 많이 두는 것으로 나타났다. 사람은 인물 못지않게 대화의 주제가 되는 사물이나 배경에도 많은 관심을 보이는 것을 확인할 수 있었다.



[그림 3] 정지된 화면에서 안구의 이동 경로



[그림 4] 실제 동영상에서 사람이 집중한 내용을 등장인물 별로 분석한 그래프



[그림 5] 실제 이미지에 등장한 인물과 배경의 빈도수와 사람이 이미지에서 집중한 인물, 배경의 빈도수 비교

또한 인물을 주인공 6명 및 엑스트라로 나누어 사람의 집중도 분포를 [그림 5]에 표시하였는데, 대체로 엑스트라보다 주인공에 많은 집중을 보이는 것을 확인할 수 있다. 이러한 사실은 전체의 드라마 흐름 안에서 피험자의 집중 패턴을 프로파일링 하는 데에 사용될 수 있다.

3.3 사용자 집중도 지도

피험자의 집중도는 특정 부분에 시선이 머무른 시간과 시선이 고정된 횟수를 통해 특정할 수 있다[2]. 두 가지 요인 중에 어느 것이 피험자의 집중도를 정확히 반영할 수 있는지에 대해서는 아직까지 명확히 알려지지 않았는데 우리는 이미지 영역에 시선이 고정된 횟수를 통해 사용자 집중도 분포를 표현하였다.

하나의 이미지에 해당하는 안구 움직임 기록이 N개 있을 때, 각각의 안구 고정위치에 대해 가우시안 분포를 그리고 N개에 대한 GMM(Gaussian Mixture Model)을 만들었다[3]. 이 집중도 분포를 해당 이미지에서 확인하면 [그림 6]과 같다. 이러한 사용자 집중도 지도는 광고계 혹은 난독증 치료와 관련한 연구에서 주로 사용 되던 것인데, 우리는 [그림 6]와 같이 사용자 집중도 지도를 그려서 피험자의 집중에



[그림 6] 사용자 집중도 지도



[그림 7] 자막에 편중된 사용자 집중도

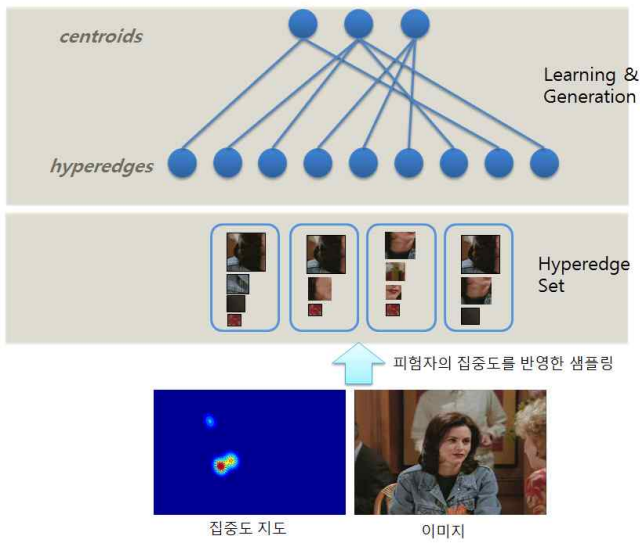
관련한 spatial한 특성을 프로파일링 하였다. 우리는 4장에서 이를 이용하여 이미지를 생성하는 알고리즘에 대해 논의하고자 한다.

한편, 이미지마다 집중도 지도를 분석해 보면 재미있는 결과를 얻을 수 있는데, [그림 7]은 한국인이 미국 드라마를 볼 때의 성향을 보여준다. 피험자가 드라마를 볼 때에 자막이 같이 제시되었는데, [그림 7]을 보면 사용자의 집중도가 자막에 편중됨을 확인할 수 있다.

4. 사용자 안구 운동을 적용한 이미지 생성 알고리즘

위에서 논의한 사용자 집중도 지도를 이용하여 이미지를 생성하는 기계 학습 알고리즘의 프레임워크를 [그림 8]에 나타내었다. 제시하는 알고리즘은 population coding에 기반한 Hypernetwork 모델[4]로 Hypernetwork는 hyperedge들의 집합으로 표현되고 각각의 hyperedge들은 가중치를 가진다. 그리고 Hypernetwork는 진화연산을 이용하여 hyperedge들의 가중치와 그것들의 최적의 조합을 학습한다.

이미지를 다루는 대부분의 알고리즘은 이미지의 모든 영역에 똑같은 가중치를 주고 학습을 시작하는데[5], [그림 8]은 사용자의 집중도를 반영하여 가중치를 초기화하는 Hypernetwork 모델을 보여준다. 즉, Eye-Tracker를 이용하여 파악한 사용자 집중도에 비례하게 hyperedge들의 초기 가중치를 부여하게 된다.



[그림 8] 집중도를 반영하여 이미지를 샘플링하는 Hypernetwork 모델의 프레임 워크

5. 논의

우리는 이미지 재생성을 위한 기계 학습 알고리즘 연구에 대한 초기 연구 결과로 Eye-Tracker를 이용하여 사용자 집중도를 프로파일링하고 알고리즘에 적용하는 방법을 제시하였다. Eye-Tracker를 이용하여 사용자의 드라마 시청 패턴을 밝히고 집중도 지도를 통하여 이미지에서 spatial한 집중 패턴을 알아보았다. 이것을 population coding 기법에 적용하여 이미지를 재생성하는 프레임워크를 제시하였는데, 제시된 방법으로 사용자가 많이 집중했던 내용 중심의 이미지 재생성이 가능할 것이라고 기대한다.

감사의 글

이 논문은 교육과학기술부의 재원으로 국가연구재단의 지원을 받아 수행된 연구(2010-0018950, 뇌정보처리 기반 사용자 의도 변화 모델링 및 예측 기술 개발), (0421-20110032, 지능형 추천 서비스를 위한 인지 기반 기계학습 및 추론기술, Videome)이며, 지식경제부 산업원천 기술개발사업(10035348, 모바일 플랫폼 기반 계획 및 학습 인지 모델 프레임워크 기술 개발) 및 교육과학기술부의 BK21-IT사업에 의해 일부 지원되었음.

참고문헌

[1] D.H.Ballard, M.M.Hayhoe, P.K.Pook, R.P.Rao, "Deictic codes for the embodiment of cognition", *Behavioral and*

Brain Sciences, 20, 723~767, 1997

[2] C.Shagass, R.A.Roemer, M.Amadeo, "Eye-Tracking Performance and Engagement of Attention", *Arch Gen Psychiatry*, 33(1), 121-125, 1976

[3] K.Holmqvist, M.Nystroem, R.Dewhurst, H.Jarodzka, J.Weijer, *Eye Tracking*, Oxford, 2011

[4] B.T.Zhang, "Hypernetworks:A molecular evolutionary architecture for cognitive learning and memory", *IEEE Computational Intelligence Magazine*, 3(3), 49-63, 2008

[5] H.Lee, R.Grosse, R.Ranganath, A.Y.Ng, "Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations", *Proceedings of the 26th Annual International Conference on Machine Learning*, 609-616, 2009