

Eye-Tracker를 이용한 사건 순서 기억의 특징 분석
Analysis of Event Order Memory Using Eye-Tracker

지도교수 : 장병탁

이 논문을 공학학사 학위 논문으로 제출함.

2011년 12월 22일

서울대학교 공과대학
컴퓨터공학부
안홍렬

2012년 2월

목차

1. 서론

2. 시선과 영상에 대한 연구

2.1. 시지각의 특징

2.2. 영상을 볼 때의 시선의 특성

3. 실험 설계

3.1. Multimodal Memory Game(MMG)

3.2. Eye-Tracker

3.3. 실험 방법

4. 실험 결과 및 분석

4.1. MMG 게임에서의 eye-movement

4.2. 이미지와 텍스트 비교 분석

5. 결론

참고문헌

Eye-tracker를 이용한 사건 순서 기억의 특징 분석

안 홍 렬

기억의 문제와 통시적 문제는 기계학습 분야의 새로운 도전과제이다. 우리는 이들 문제에 대한 기초연구로써 사람의 사건 순서 기억을 평가하는 멀티모달 메모리 게임(Multimodal Memory Game)을 설계하고 동영상 시청과 MMG 게임 과정에서 피험자의 시선 데이터를 수집하였다. 시각에 대한 선행 연구에서 사람의 시선은 관심 영역에 고정되며 이 관심 영역은 사람 사이에 비슷함을 보여준다. 우리는 실험 결과를 통해 MMG 게임 중에서 앞선 연구와 같이 피험자의 관심 영역이 특정 영역에 집중되고 이 영역은 피험자 간에 일치함을 확인할 수 있었다. 그것에 더하여 우리는 사건 순서 기억과 이미지와 텍스트의 상관관계를 탐구하고자 했지만 분명한 결론을 얻지는 못하였다.

1. 서 론

컴퓨터로 순차적 프로그램을 실행시키면 매우 빠르고 정확하게 결과물을 산출한다. 그러나 많은 문제들은 순차적 프로그램으로 해법을 구현하기 어려운 한계가 있다. 사람의 경우를 보면 순차적 프로그램을 처리하는 데는 처리 속도가 컴퓨터에 훨씬 못 미치지만 패턴인식 같은 컴퓨터로 해결하기 어려운 문제에 대해서는 매우 즉각적이고 정확한 판단을 내린다. 이런 식으로 기계가 잘 풀 수 있는 문제와 사람이 잘 풀 수 있는 문제는 그 성격이 다르다는 것이 일반적인 통설이다.

최근에 기계학습 분야에서 집중적으로 연구되고 있는 과제는 통시적 문제에 대한 풀이이다. 기계학습 분야의 초창기에 순차적 프로그램 방법론이 패턴 매칭과 같은 문제에 적용되지 않았듯 현재의 공시적 문제를 해결하는 기계학습 알고리즘은 통시적 문제 해결에 적절하지 않다. 통시적 문제의 해결에 있어서 기존의 알고리즘이 아닌 새로운 알고리즘의 필요성이 대두된다. 우리는 통시적 문제를 기계로 풀 수 있는 알고리즘 연구를 위한 기초 연구로 사람이 통시적 문제를 어떻게 푸는지 알아보려는 목적으로 본 연구를 진행하였다.

우리는 광범위한 통시적 문제 영역에서 단순하고 분명한 사건 순서 기억에 대한 연구로 연구 대상을 축소하고 사람의 사건 순서 기억의 특징을 연구하였다. 우리는 사람의 시각에 초점을 맞추어 사건 기억에 관련하여 사람에게 체

특된 사건 순서 기억의 방법을 연구하고자 하였다. 사람의 시각은 여러 연구에서 논의된 바대로 의식과 뇌의 통제를 단선적으로 받는 것이 아니라 그것이 스스로 유용한 정보를 얻어 낼 수 있도록 동작함이 밝혀져 있다. 이는 사람의 시각 연구가 사람의 정보처리 능력의 원리를 밝힐 수 있는 단초를 제공할 수 있다는 것을 암시한다.

우리 연구의 목적은 크게 두 가지이다. 하나는 사람이 사건 순서를 기억하는 과정에 어떤 시각적 특징이 존재하는가를 알아보는 것이다. 우리는 이를 위해 사람의 기억을 연구하는 플랫폼인 MMG를 이용하여 사건 기억을 시험할 수 있는 실험을 구축하고 실험 간에 Eye-tracker를 통해 피험자의 시선 데이터를 수집하여 분석하였다. 우리 연구의 다른 목적은 사람이 사건 기억 과정에서 이미지와 텍스트의 상관관계가 있는지 알아보는 것이다. 우리는 이를 위해 이 관계를 판단할 수 있도록 MMG 게임 세션을 설계하였다.

2. 시선과 영상에 대한 연구

2.1. 시지각의 특징

사람의 시지각 활동은 안구 운동에 의한 고정과 탐색 작업의 반복으로 이루어진다. 고정은 시지각 과정에서 대상물에 약 100~300ms의 짧은 시간 동안 시선이 정지되는 것이다. 이 고정의 짧은 순간 망막의 중앙 1.5~2.0도 사이에 명확한 상이 맺히게 되고 이 상에 대해 시지각 정보의 처리가 이루어진다. 복잡한 사물에 대해서 시선은 작은 간격을 움직이면서 몇 개의 고정점을 형성하여 복잡한 사물을 부분적으로 인지하게 되는데 이 과정이 탐색이다. 시각은 고정과 탐색 외에도 단속적(saccadic)운동을 하는데 이것은 시선이 급속하게 움직일 때 일어나는 안구의 미세한 떨림을 의미한다. 사람이 단속적 움직임을 할 때 정보처리는 이루어지지 않는다.

사람의 시선 움직임은 시선 고정의 이동으로 설명되며 시선의 고정은 두 가지 원인에 의해 일어난다. 그 하나는 '하향식 주의(Top-down attention)'로 의식적으로 원하는 장소에 시선을 집중 시키는 것을 말한다. 다른 하나는 '상향식 주의(Bottom-up attention)'로 주위 자극에 의해 시선이 집중되는 것을 말한다. 하향식 주의를 우리의 의식이 시선을 결정함을 보여주지만 하향식 주의를 눈이 의지의 통제를 넘어서 정보습득을 위해 능동적으로 참여하고 있음을 보여준다.

이 같은 내용으로 Arnheim(1966)은 사람의 시지각이 수동적으로 정보를 수용할 뿐 아니라 능동적으로 인지하는 특성이 있으며 목적성(Purposive)과 선택성(Selectivity)을 가진다는 것을 밝혀내었다. 계속하여 Arnheim(1969)는 시인지 체계를 안구운동, 시각적 사고, 인지적 사고의 세 단계로 나누어 시지각의 계층

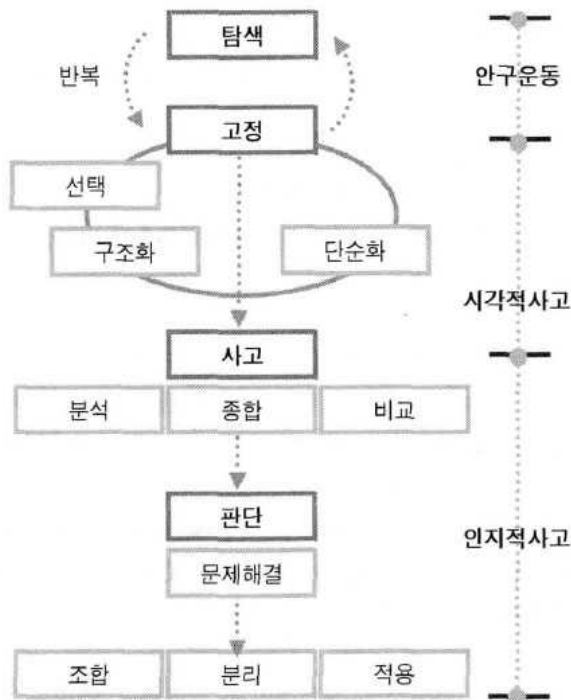


그림 1. Arnheim의 시지각 계층 구조

적 구조를 제시하였다.<그림 1>¹⁾ Arnheim의 시인지 체계에서 전체 단계는 사람의 문제 해결을 목적으로 하여 긴밀히 동작하며 특히 이때 안구 운동과 인지적 사고 사이에서 시각적 사고는 구조화와 단순화를 반복하며 거의 자동적으로 인지적 사고에 필요한 정보를 수집하도록 안구 운동에 관여한다.

2.2. 영상을 볼 때의 시선의 특성

상향식 주의를 유발하는 자극으로 '크기', '색', '위치', '특이함', '움직임'이 있다. 여기서 크기는 자극하는 물체의 크기를 말하고 색은 주변과 대조적으로 두드러지는 색을 말하고 위치는 물체가 중심에 놓였는지 가에 놓였는지를

말하고 특이함은 지각하는 사람에게 익숙하지 않거나 주변의 물체들과 대조적이어서 불편함을 유발하는 것을 말한다. 그리고 움직임이란 형태가 변하거나 위치가 변하는 것을 말한다. 이런 것들은 사람의 시선의 주위를 고정시키는 원인이 된다.

위의 시각 연구의 결과로써 Koch(1997)는 일반적으로 사람이 영상을 볼 때 그 영상에서 몇몇 영역에만 주의를 기울인다는 것을 발견했다. 그의 연구 결과는 피험자들에게 영상을 보는 시간이 무제한으로 주어진다 해도 영상의 모든 부분을 세밀하게 살피지 않고 몇몇 부분에만 관심을 기울이는 실험 결과를 보여준다. 또, Goldstein(2006)은 영화를 시청하는 다수의 시선 정보를 획득 분석하여 대부분의 사람들이 영화를 시청할 때 같은 곳을 본다는 것을 보였다. 이밖에 Dorr(2010)는 실외 장면을 촬영한 고화질 영상을 시청하는 동안 시선정보를 분석하여 약 80%가 같은 영역을 보고 있음을 보였다. 우리는 이 같은 연구 결과를 바탕으로 사람의 사건 기억과 관련하여 어떤 시지각적 특징이 있는지 알아보려고 한다.

1) 최민영, "사용자의 시각적 인지 특성에 기초한 시선추적 분석 방법에 대한 연구: 제품 디자인을 위한 분석 프레임 개발을 중심으로," 디자인학연구, Vol. 16, No. 4, pp.197-206, 11. 2003.에서 재인용

3. 실험 설계

3.1. Multimodal Memory Game(MMG)

멀티모달 메모리 게임(MMG)은 인간의 회상 기억을 예측하고 이해하기 위해 설계된 학습게임 플랫폼이다. MMG는 피험자는 비디오 영상을 시청하고 영상에 대한 기억을 테스트하는 게임을 수행하게 된다. 이때 MMG 설계자는 연구 목적에 따라 게임의 방식을 다르게 설계함으로써 인간 기억의 다양한 특징을 탐구할 수 있다. 장병탁(2010)은 MMG를 제안하면서 장면(image)을 보여주고 대응하는 대사(text)를 알아맞히는 I2T(image to text) 게임과 대사를 보여주고 대응하는 장면을 알아맞히는 T2I(text to image) 게임을 수행하였다. 그리고 그 결과로 피험자가 수행한 세션의 수가 증가할수록 정답률이 향상됨을 확인하였다. 또 I2T 게임의 정답률이 T2I의 정답률보다 빠르게 증가함을 확인하였고 이를 통해 대사에서 장면을 추론하는 것이 장면에서 대사를 추론하는 것보다 상대적으로 쉽다고 주장하였다. 우리는 MMG 플랫폼에서 사건의 시간 순서를 선택하도록 게임을 변형하여 사람의 사건 순서 기억의 특징을 알아보고자 한다.

3.2. Eye-Tracker

Eye-Tracker는 사람의 시선이 눈동자의 방향에 동기화된다는 것에 착안하여 눈동자의 이동을 통해 시선의 위치를 추적·기록하는 장비이다. Eye-Tracker의 작동 원리는 다음과 같다. 피험자의 몸과 머리를 고정하여 시선의 이동이 오직 눈동자의 이동으로 이루어지도록 환경을 준비한다. 그리고 피험자가 모니터 화면의 모서리 같은 특정 지점을 보도록 지시한 후 그 때의 눈동자의 위치를 카메라로 찍어 얻어낸다. 이 같은 반복을 통해 시선-눈동자 위치를 동기화 하고 실험 중에 피험자의 눈동자 위치를 카메라로 찍어 피험자의 시선의 위치 데이터를 얻어낸다. 우리는 이 Eye-Tracker의 기능을 이용하여 사람의 시선 데이터를 수집하고 사람의 사건 순서 기억의 특징에 대해 시각 정보에 중점을 두어 탐구를 진행할 것이다.

3.3. 실험 방법

우리는 2가지의 관심사에 대해 연구를 진행하였으며 그 첫 번째는 사람이 MMG 게임을 수행할 때 관심 영역이 일치하는지 알아보는 것이었다. 이는 기존의 시선과 인지에 관한 연구의 사람들이 공통된 관심영역에 집중한다는 실험적 결과가 MMG 게임을 할 때 적용되는지 알아보기 위함이었다. 이를 통해 사람이 사건 순서 기억을 할 때 특정한 인식 패턴을 가지고 있으며 그것이 공통적인지 알아보고자 한 것이다. 우리의 두 번째 관심사는 사람이 사건 순서

를 기억할 때 이미지 정보와 문자 정보 중에서 어느 정보에 치중하는 지 알아보는 것이었다. 이는 사건 순서를 기억할 때 무엇을 집중하는가를 밝힘으로써 사람에게 사건 순서를 기억될 때 그 기억되는 것이 어떤 형태인지에 대한 단초를 얻을 수 있을 것이라는 기대에 의한 것이다.

우리는 3명의 피험자를 대상으로 영상을 시청하고 MMG를 푸는 동안 시선 데이터를 수집하였다. 우리가 보여준 영상은 NBC에서 1994년부터 2004년까지 방영된 'Friends'의 첫 번째 시즌 24화 영상이다. 이 영상은 재생시간이 24분 정도며 우리는 이미지 정보와 문자 정보의 대비를 극대화하기 위해서 피험자에게 영상과 함께 자막을 보여주었다. 그리고 우리는 사건 순서 기억을 연구할 수 있도록 MMG 게임을 수정하였다. 수정된 MMG는 총 3세션으로 구성되며 각 세션은 20문제로 이루어진다. 첫 번째 세션에서 피험자는 주어진 이미지를 보고 2개의 보기 중에서 가장 가까운 미래에 해당하는 이미지를 선택하게 된다.<그림 2(a)> 이때 피험자는 모르겠음을 선택할 수 있다. 두 번째 세션에서 피험자는 대사에 대해 첫 번째 세션과 같은 방법으로 문제를 풀게 된다.<그림 2(b)> 세 번째 세션에서 피험자에게 이미지와 대사 쌍이 주어지고 피험자는 첫 번째 세션과 같이 가장 가까운 미래에 해당되는 이미지와 대사 쌍을 고르게 된다. 그런데 이때 문제로 주어지는 이미지와 텍스트 쌍은 그 시간적 위치가 동일하지 않다. 이것은 우리가 세 번째 실험에서 사람이 사건 순서를 기억 할 때 이미지와 문자 정보 중 어떤 정보를 더 치중하는지 알아보고자 의도적으로 문제를 설계한 것이다. 이 시간적 위치가 동일하지 않은 문제에 대해 첫 번째 보기는 문제의 이미지의 다음 사건에 해당하는 보기이며 두 번째 보기는 문제의 대사의 다음 사건에 해당하는 보기이다.<그림 2(c)>

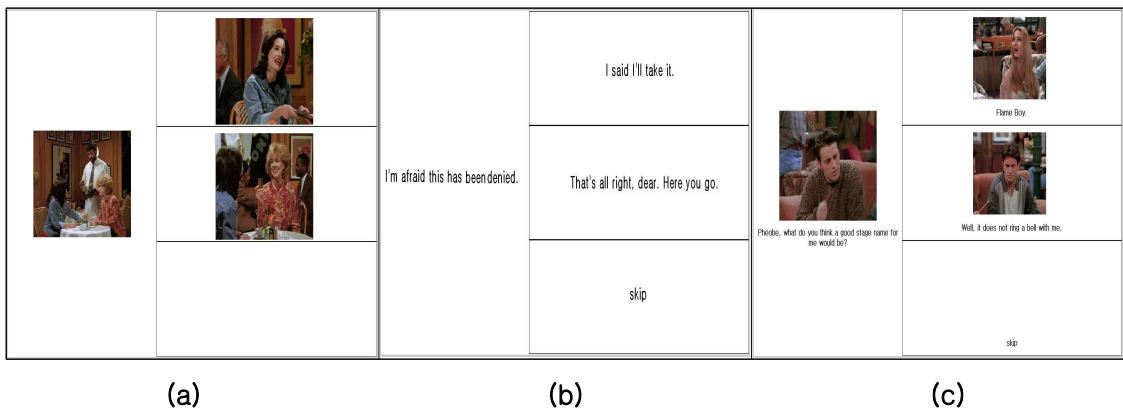


그림 2. 사건 순서 기억을 평가하기 위해 설계된 MMG 게임(a, b, c는 각 1, 2, 3 세션)

4. 실험 결과 및 분석

4.1. MMG 게임에서의 eye-movement

우리는 MMG의 첫 번째 세션의 20문제에 대해서 MMG 문제를 풀 때의 피험자의 시점을 분석하였다. 그 결과 대부분의 문제에서 피험자의 관심영역이 일치하며 피험자들은 그 관심영역을 제외하고 다른 영역에 거의 시선을 돌리지 않음을 확인할 수 있었다.<그림 3> 또 피험자들은 평균적으로 5번의 장면 간 시선이동을 하였다. 즉, 피험자들은 실험 중 평균적으로 한 장면을 두 번 보았음을 확인하였다. 그러나 피험자 간의 장면 이동 순서는 특정한 공통점을 보이지 않았다.<그림 4 ~ 그림 6>



그림 3. MMG 게임 세션 1에서의 피험자 시선 좌표(빨간색, 초록색, 파란색은 각 피험자)



그림 4. 피험자 A의 시선의 이동(노란색 별은 처음과 마지막 시선)

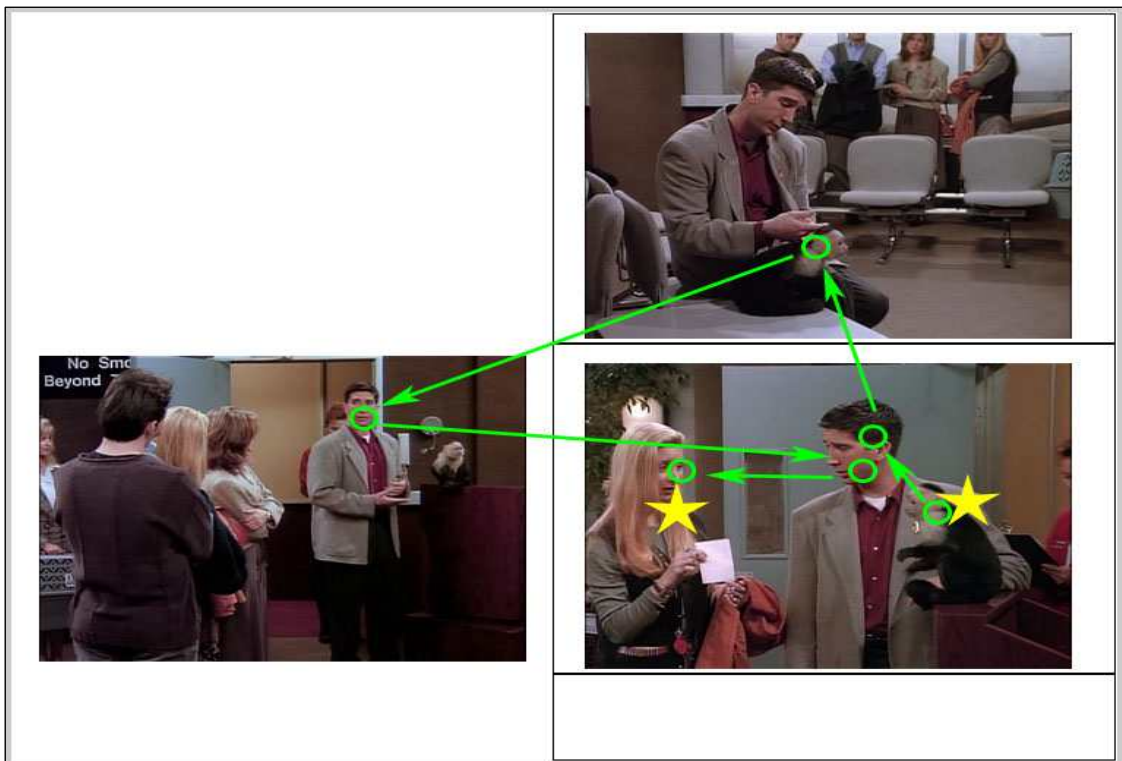


그림 5. 피험자 B의 시선의 이동(노란색 별은 처음과 마지막 시선)

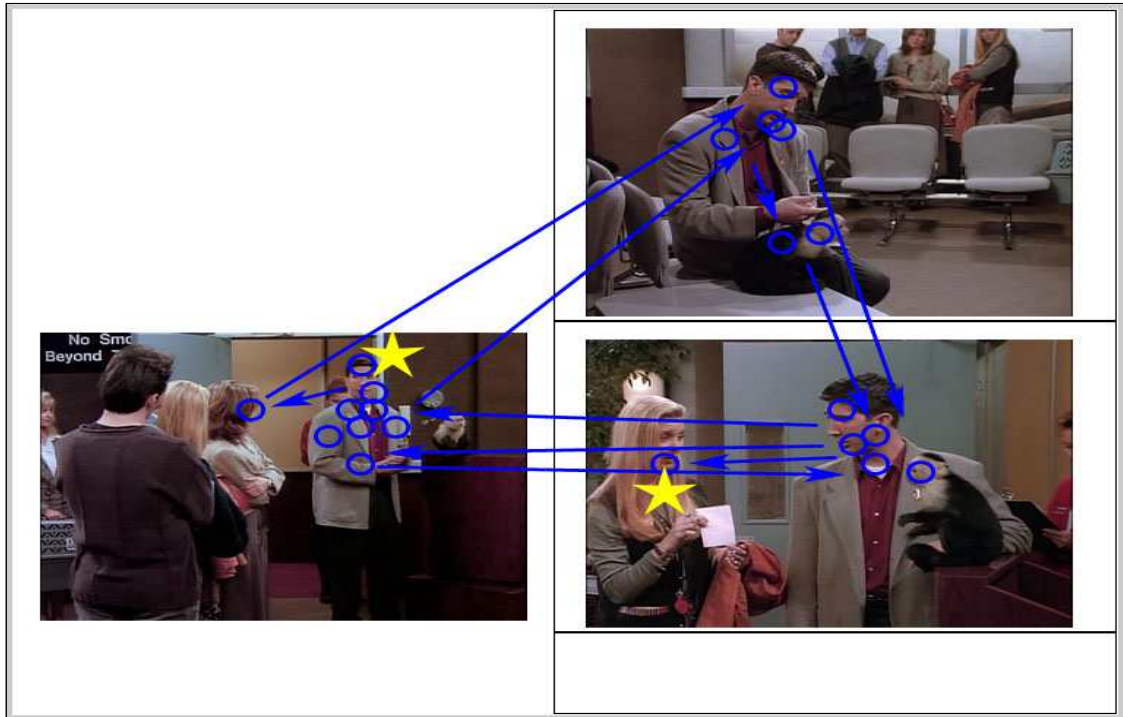


그림 6. 피험자 C의 시선의 이동(노란색 별은 처음과 마지막 시선)

4.2. 이미지와 텍스트 비교 분석

우리는 각 세션에서 피험자가 MMG 문제를 푸는 데 걸린 시간을 측정하였다. 그 결과 세션 1과 2에서 정답의 응답시간이 가장 짧았으며 미세하게 정답보다 오답의 응답시간이 길었고 모름의 응답시간은 둘과 비교해서 차이가 있음을 확인할 수 있었다.<표 1> 이는 문제를 보았을 때 사건 기억의 단서에 대해 곧장 발견하는 경우 즉각 정답을 고른 것으로 보이며 이렇게 곧장 단서를 발견하지 못했을 경우에는 잘못 판단하거나 모른다고 답하게 될 가능성이 많음을 보여준다. 이를 통할 때 사람이 사건 기억의 단서를 발견하는 과정은 최초의 단서 발견 여부에 관계가 있음을 짐작할 수 있다.

세션	정답	오답	모름
세션1	6.331	7.093	9.408
세션2	10.396	11.679	19.772
세션3	12.293 (이미지선택)	9.362 (텍스트선택)	10.197

표 1. MMG 세 섹션에 대한 응답시간

우리는 또한 영상을 시청하는 동안에 피험자가 얼마나 자막을 보았는지 분석하여 피험자 별로 각 세션의 선택을 분석하였다.<표 2> 그러나 실험 분석 결과 기대한 결과를 얻지는 못했다. 동영상 시청에서 장면을 많이 보는 사람은 영상에 치중함으로 세션 1에 좋은 성적을 보이고 세션 3에서 이미지를 기준으로 답을 고를 것이라 예상했는데 결과를 보면 자막 대비 영상 시청률과 각 세션의 답 선택 간에 상관관계는 보이지 않는다. 이를 통해 자막 시청의 여부와 게임 수행 여부의 관계에 어떤 단정을 하기는 힘들다. 너무 표본이 적고 표본 간에 외적 환경이 다른 것이 문제가 되기 때문이다.

영상시청률% (전체 시청 대비)	세션1 (정답률)	세션2 (정답률)	세션3 (이미지 선택률)
A(62)	0.95	0.75	0.80
B(81)	0.55	0.60	0.10
C(69)	0.60	0.45	0.30

표 2. 영상 시청률 대비 정답 및 이미지 선택률

5. 결론

우리는 통시적 문제에 대한 기계학습 알고리즘에 대한 기초 연구로써 사람의 사건 순서 기억 특징을 알아보았다. 특히 Eye-tracker를 사용하여 시선 데이터를 얻어 사람의 시각 인지 특성을 분석하였다. 실험 결과 우리는 사람이 사건 순서 기억을 시험하도록 설계된 MMG 게임 과정에서 피험자들이 비슷한 관심영역에 집중하는 것을 확인하였다. 그리고 그 외에 다른 영역에 대해서는 거의 시선이 고정되는 일이 없음을 확인하였다. 그러나 피험자들 간 시선의 이동 순서에 대해서는 공통점이 발견되지 않았다. 그러나 우리 연구는 실험 대상자가 매우 적은 한계를 지니고 있다. 또 사건 순서 기억에서 이미지와 텍스트의 역할과 관계를 알아보고자 한 실험 설계는 분명한 결론을 얻지 못하였다. 따라서 후속 연구를 위한 제언으로 논문을 마무리하고자 한다. 다음 연구에서는 충분한 수의 실험 대상자를 모집하고 이미지와 텍스트의 차이를 분명하게 드러나는 기준으로 집단을 분류하여 실험한다면 의미 있는 결과를 얻을 수 있을 것이라고 생각된다. 또한 관심 영역의 대상이 무엇인지 분석한다면 관심 영역이 사건 기억에 어떤 역할을 하는 지에 대한 고찰까지 나아갈 수 있을 것이라 생각된다.

참고문헌

- 임찬, 최규호, 김완석, 윤재선, 민병철, 백로라, "아이트래킹을 통한 효과적인 스토리텔링", *디자인학연구*, Vol. 23, No. 4, pp.243-252, 2010.
- 최경주, 박민철, "인간의 주의시각에 기반한 시각정보 선택 방법", *한국멀티미디어학회지*, Vol 14, No. 3, pp. 378-391, 2011.
- 최경주, 이일병, "인간의 상향식 시각적 주의 특성에 바탕을 둔 현저한 영역 탐지", *정보과학회논문집*, Vol. 31, No. 2, pp.189-202, 2004.
- 최민영, "사용자의 시각적 인지 특성에 기초한 시선추적 분석 방법에 대한 연구: 제품 디자인을 위한 분석 프레임 개발을 중심으로", *디자인학연구*, Vol. 16, No. 4, pp.197-206, 2003.
- Alexander C. Schütz, Doris I. Braun, Karl R. Gegenfurtner, "Eye movements and perception: a selective review", *Journal of Vision*, vol. 11, no. 2011.
- David E. Irwin, Gregory J. Zelinsky, "Eye movements and scene perception: Memory for things observed", *Attention, Perception, & Psychophysics*, Vol. 64, No. 6, pp. 882-895, 2011.
- E. Niebur and C. Koch. "Computational Architectures for Attention", *The Attentive Brain*, MIT Press, 1997.
- Rudolf Arnheim, "*Toward a Psychology of Art*", California Press, 1966.
- Rudolf Arnheim, "*Visual Thinking*", California Press, 1969.
- Robert B. Goldstein, Russell L. Woods, Eli Peli, "Where people look when watching movies: Do all viewers look at the same place?", *Computers in Biology and Medicine*, Vol. 37, Issue 7, Pages 957-964, 2007.
- U. Fareed and B.-T. Zhang, "MMG: A learning game platform for understanding and predicting human recall memory", *The 11th International Workshop on Knowledge Management and Acquisition for Smart Systems and Services (PKAW 2010), Lecture Notes in Artificial Intelligence*, 6232:300-309, 2010.