RLVisualizer: 강화학습의 문제의 학습궤적을 시각화하는 응용

RLVisualizer:
An application for Visualizing Trajectories of Reinforcement Learning Problem

정 태 충, Le Pham Tuyen
경희대학교 컴퓨터공학과 교수 및 박사과정

Chung, TaeChoong (Professor),
Le Pham Tuyen(PhD candidate)
KyungHee University, Computer Engineering Dept.

요약

딥러닝을 비롯한 전반적인 인공지능에 대한 관심이 뜨겁다. 특정 영역에 영향력을 주었던 과거와 다르게 인공지능의 영향력은 인류문명 전체에 변화를 주고 있다. 예술 분야도 영향을 받고 있는데, 그 중 한 분야는 과학적 실험의 자료를 어떻게 시각화 하느냐의 문제를 풀다가 나오기도 한다. 자료를 시각화하는 것은 실험과정 및 결과를 과학자 및 독자들에게 쉽게 전달하기위한 것이다. 그런데, 그 시각화된 영상 중에는 미적인 아름다움이 있는 경우가 있다. 본 연구자는 강화학습의 정책이 어떻게 개선되고 있는지 보기위해 강화학습의 과정을 시각화 해서 검증하는 시도를 했다. 그 과정에서 만든 자료가 미술적인 관점에서도 아름다움이 있는 작품을 만들 수 있다는 확신이 들어서 강화학습용 디지탈예술 도구를 만들어 작품을 생성해 보았다.

1. 서론

강화학습(Reinforcement learning: RL) [1]은 문제 영역의 지식이 많이 있지 않아도, 어떤 상황에서 어떻게 행동해야 할지를 알려주는, 정책(policy: 폴리시)을 학습할 수 있게 해주는 강력한 방법이다. RL 은 로봇, 게임 및 제어 등의 분야에 적용하기 좋은 방법으로, 여러 분야에서 많은 성과를 거두었다.

학습하는 과정에서 폴리시는 점진적으로 개선이 되며, 결국 최적의 폴리시에 도달하게 된다. 개선되는 것을 보여주려면 알고리즘의 성과 (즉, 보상)을 보여 주어야 한다. 그 외에도 그 보상 값이 어떻게 나왔는지 사람의 시각으로 확인할 수 있다면 더 좋을 것이다. 예를 들어, 로봇에이전트가 자전거를 배운다면, 얼마나 중심을 잡고 오래 타는지 그 시간의 길이가 늘어나는 자료도 의미가 있지만, 자전거의 궤적이 어떻게 개선되고 있는지 눈으로 볼 수 있다면, 알고리즘을 좀 더 잘 개선할 수 있을 것이다. 본 논문에서는 배-운전[2]의 궤적을 시각화해서 알고리즘을 개선한 연구를 보여주려 한다.

자료의 시각화에 대한 많은 응용들이 있다. [2] 맨눈으로는 볼 수 없는 과학적인 자료로부터 아름다운 시각영상을 만들어 낸 좋은 결과들도 많다. [3]. 그러나, 아직까지 강화학습문제의 궤적들을 시각화하는 논문은 볼 수 없었다.

즉, 본 논문의 기여점은 강화학습문제들의 궤적을 시각화하는 도구를 최초로 개발하는 것이다. 이 도구를 이용하여 강화학습 알고리즘의 개선 뿐만 아니라 디지탈 예술 작품을 만들어 내는 도구로 활용될 수 있다.

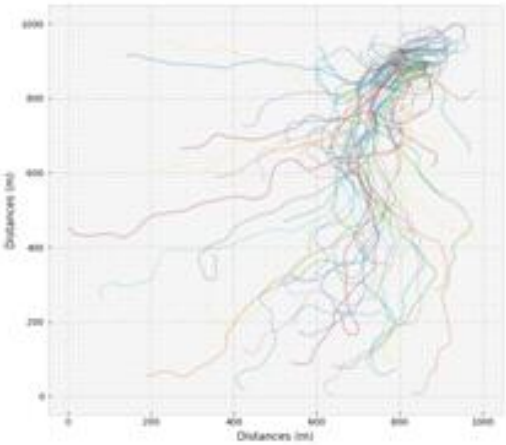
2. RLVisualizer

▶▶ 그림 1. RLVisualizer 화면

주요 기능은 도메인 문제의 궤적을 시각화하는 것이다. 학습과정에서 축적된 궤적 데이타를 사용자가 다양한 목적으로 궤적을 선택/편집/속성값수정/채색 등을 대화식으로 할 수 있게 만든 것이다. 예를 들어, 10의 배수 번째 에피소드의 궤적들만 골라서 보라색으로 색갈을 바꾸어 그릴 수 있는 것이다. 처리를 한 다음 svg, png, pdf 와 같은 다양한 포맷으로 저장할 수 있다.

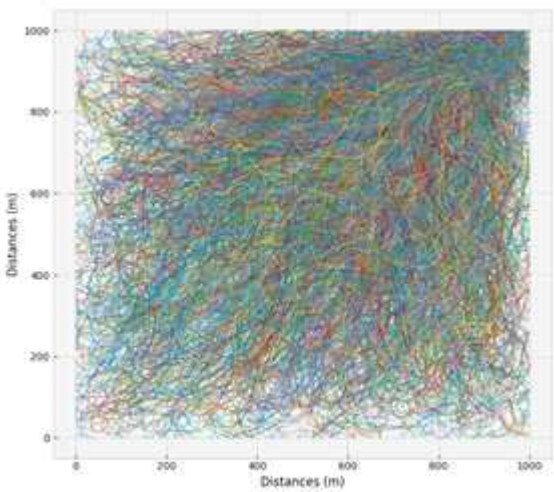
3. 실험

임의의 지점에서 출발한 배는 오른쪽 상단 (900, 900)의 목적지를 도달하도록 학습하는 강화학습 알고리즘을 수행하고 그 궤적을 저장한다. 그림 2는 배-운전 학습과정 동안의 후기단계 궤적들 모음이고 그림 3는 학습과정의 전체 궤적들을 표현하고 있다.



▶▶ 그림 2. 학습과정 마지막 궤적들 50개

4. 궤적들의 예술적 가치



▶▶ 그림 3. 배-운전 학습과정 전체 궤적들

미술의 큰 분야가 추상화이며, 김환기 작가의 그림 4는 엄청난 고가로 평가 받고 있다.[4] 사람에 따라서는 그림 3이 그림 4보다 더 아름다움을 줄 수 있다고 본다. 왜냐하면 그림 4가 무엇을 표현하는지 모르지만, 그림 3에는 문명을 바꾸고 있는 강화학습을 표현하고 있기 때문이다.

▶▶ 그림 4. 경매 54억 낙찰된 김환기 '무제'

5. 결론

이 논문은 강화학습 과정을 시각화하는 도구를 만들고 활용해 학습이 잘 되고 있는지 검증하는 도구로 사용될 뿐만 아니라 예술 작품으로 활용할 수 있는 길을 열어보이는 시도를 내용으로 한다.

사사

본 연구는 과학재단 기초연구과제 (NRF-2014R1A1A2057735)와 경희대학교 교내연구과제 (KHU-20160601)의 지원으로 수행되었습니다.

■ 참 고 문 헌 ■

[1] Sutton, Richard S., and Andrew G. Barto. Reinforcement learning: An introduction. Vol. 1. No. 1. Cambridge: MIT press, 1998.
[2] Data visualization tools. http://www.creativebloq.com/design-tools/data-visualization-712402
[3] https://www.nsf.gov/news/special_reports/scivis/winners_2016.jsp
[4] http://www.hankookilbo.com/v/80e3caac1b26456aa4300b4205e37310