

条件付き変分オートエンコーダを用いたビデオゲームのレベル生成

吉貞 心[†] 飯間 等[‡]京都工芸繊維大学 情報工学専攻[†] 同 情報工学・人間科学系[‡]

1. はじめに

近年のビデオゲーム開発は大規模化しており、ゲームコンテンツの一つであるゲームレベルを制作するためのコストも増大している。そこで、機械学習を活用することにより開発効率化を図ることが期待されている。従来研究では、敵対的生成ネットワーク (GAN) を用いて多様なゲームレベルを生成したり [1], GPT を用いてゲームレベルをテキストとして処理し、開発者の要求に近いゲームレベルを生成したり [2] するなど、コーディングベースではなくシンボル・テキストベースで生成を行うモデルが提案されている。

本論文では生成 AI の 1 手法である条件付き変分オートエンコーダ (CVAE) [3] を活用し、少数のゲームレベルとゲーム仕様の設定により、開発者の要求通りのゲームレベルを生成する手法を提案する。提案手法を用いてゲームレベルを生成する実験を行った結果、ある程度要求を満たしたゲームレベルを生成することに成功し、提案手法が有効であることが確認された。

2. 提案手法

2.1 概要

ゲームレベル生成の学習には CVAE を活用する。CVAE はニューラルネットワークの一種で、画像などの情報とその画像に対するラベルを圧縮して符号化したものを潜在変数とし、その潜在変数とラベル情報から元の画像に類似したものを生成するモデルである。提案手法ではゲームレベルを CVAE の入出力とする。

初期の学習データとなるゲームレベルは、手作業で作る必要がある性質上大量に用意することは困難である。そこで、提案手法による学習途中で生成したゲームレベルも学習対象として加え、学習を進めていく程学習用データが増加していくようにする [1]。

生成するゲームレベルは、対象としているビデオゲームの仕様を満たす必要がある。例えば、スタート位置からゴールを目指すゲームであれば、ゴールまでの経路が必ず存在する必要がある。

Generation of Video Game Levels Using a Conditional Variational Autoencoder

[†]Makoto Yoshisada, Information Science, Kyoto Institute of Technology

[‡]Hitoshi Iima, Information and Human Sciences, Kyoto Institute of Technology

る。しかしながら、CVAE で生成するゲームレベルが仕様をすべて満たすとは限らない。そこで、生成したゲームレベルに対して、仕様をすべて満たすように変換処理を行う。

ゲームレベルの生成では、開発者の要求通りのものを生成できると有用である。このために、CVAE のラベルを活用する。学習データセット内のゲームレベルや生成したゲームレベルに対して、開発者が要求した項目を満たしているかどうかを示すラベルを付与する。CVAE を用いてラベル毎のゲームレベルを生成できるように学習させることを通して、要求項目を満たすゲームレベルを生成できるようにする。

2.2 ゲームレベル生成のための学習法

以降では例として、3 節に示す実験に採用したビデオゲーム『Zelda』を用いて説明する。Zelda は図 1 の左に示すように方眼状にオブジェクトが配置されたゲームレベルにおいて脱出を目指すゲームで、どこかにある鍵を入手して扉に到達すればゲームクリアとなる。ところが、図 1 左のような画像をそのまま CVAE の入出力に用いると、オブジェクトの形状や色などの情報量が多く、正確に生成することが困難である。そこで、図 1 右に示すように、ゲームレベルの各オブジェクトを 0, 1 の 2 値データに変換することで、情報量を削減する。

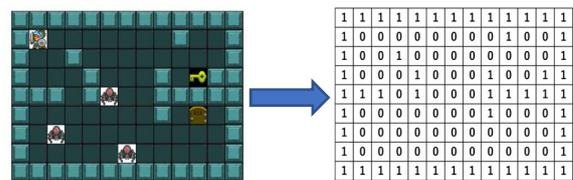


図 1. Zelda のゲームレベル(左)と、壁オブジェクトに対する変換後データ(右)

本論文で開発者が要求する項目を表 1 に示す。各要求は 0~1 の実数値で与え、値が大きいほど数や距離が大きいことを示す。これらの要求を満たすゲームレベルを生成するために、初期のゲームレベルや学習過程で生成されるゲームレベルを評価し、ラベル付けする。ラベルは、ゲームレベルに対してオブジェクトの個数や経路探索などの手法を用いて、要求を満たしてれば 1 を、要求に満たなければ 0 をラベリングする。ただし、要求した値に正確に一致することは困

難であるので、要求値に近ければ要求を満たしていると判定する。このラベルが4項目すべてにおいて1となるゲームレベルが、開発者の要求を満たすゲームレベルとなる。また、このラベルは学習段階でも CVAE のエンコーダへの入力として用いる。

表 1. 開発者が要求できる項目

要求項目	関連するオブジェクト
壁の数	壁
敵の数	敵
アバターと鍵との距離	アバター, 鍵, 床
アバターと扉との距離	アバター, 扉, 床

学習データの構成について説明する。学習においては要求項目に関連するオブジェクトを用いて、要求項目毎に CVAE を用意して学習させる。関連するオブジェクトを表 1 に示す。例えば、壁の数であればゲームレベルの壁オブジェクトの位置のみ 1 とし、それ以外の位置を 0 としたものを学習データとする。

要求毎に CVAE を用意して学習し、デコーダから出力を得る。出力はゲームレベルと同じサイズで 0~1 の実数値行列である。この行列からゲームレベルを生成するために、要求毎に閾値を設け、実数値との大小によって関連するオブジェクトを生成するかどうかを決定する。

この閾値 $Thres$ の適切な値は未知であり、またその適切な値は学習データ数に応じて変化していく可能性があることから、学習と並行して次式で更新させる。

$$Thres = Thres + m \cdot update$$

$$update = \gamma \cdot update$$

ただし、 $\gamma (< 1)$ は学習率、 m は要求に対する成功の如何を示す 3 値 (成功:0, 要求より小さい:-1, 要求より大きい:1), $update$ は更新幅である。エポック数が少ないうちは更新幅を大きくし、学習が進むにつれて更新幅を小さくすることにより、最適な閾値に収束させようとしている。

3. 実験

提案手法においてエポック数 1200, 初期データ数 5 とし、3 種類の要求を与えてゲームレベルを生成する実験を行った。その結果を表 2 に示す。

表 2. 各要求に対するゲームレベル生成結果

要求項目	要求値 : 要求達成率(%)		
	(a) 要求1	(b) 要求2	(c) 要求3
壁の数	0.5 : 59.522	0.6 : 47.164	0.45 : 57.368
敵の数	0.2 : 55.046	0.1 : 89.222	0.4 : 52.602
アバターと鍵との距離	0.5 : 14.492	0.4 : 8.671	0.4 : 15.095
アバターと扉との距離	0.4 : 9.162	0.3 : 12.075	0.6 : 8.853
生成総数	5272	1234	5015
要求達成総数 : 達成率(%)	28 : 0.531	4 : 0.324	32 : 0.638

いずれの要求に対しても、壁の数と敵の数に関しては比較的高い割合で満たすことができている。一方で、距離の項目については性能が悪く、この影響で要求をすべて満たすゲームレベルの生成数が少なくなっている。

また、実際に生成されたゲームレベルの一部を図 2 に示す。いずれの要求においても、開発者の要望に近いゲームレベルを生成することができている。しかし、アバターが侵入できない床オブジェクトが生成されているなど不十分な部分もあり、より最適な生成をするための工夫が必要である。



(a) 要求1 (b) 要求2 (c) 要求3

図 2. 生成したゲームレベル

4. おわりに

CVAE を用いて、少数のゲームレベルから学習を始め、開発者の要求通りのゲームレベルを生成する手法を提案した。実験で提案手法の有効性を確認できた。一方で、要求をすべて満たすゲームレベルの生成割合は小さく、要求が4項目と限定的であるなど、改良の余地があることが確認された。さらには Zelda 以外のビデオゲームに対する手法や有用性の検証を今後検討していく必要がある。

なお、本研究は JSPS 科研費 JP22K04172 の助成を受けたものである。

参考文献

- [1] R. R. Torrado, A. Khalifa, M. C. Green, N. Justesen, S. Risi, and J. Togelius, "Bootstrapping Conditional GANs for Video Game Level Generation," Proceedings of 2020 IEEE Conference on Games, pp. 41-48, 2020.
- [2] S. Sudhakaran, M. Gonzalez-Duque, C. Glanois, M. Freiberger, E. Najarro, and S. Risi, "MarioGPT: Open-Ended Text2Level Generation through Large Language Models," Proceedings of the 37th Annual Conference on Neural Information Processing Systems, 2023.
- [3] D. P. Kingma, S. Mohamed, D. J. Rezende, and M. Welling, "Semi-supervised learning with deep generative models," Advances in Neural Information Processing Systems, Vol. 27, pp. 3581-3589, 2014.