

生成 AI を用いたシナリオデータに基づく動画生成手法の提案

第 1 グループ 河地 駿太郎

1. はじめに

近年、動画配信サービスが盛んになり、動画コンテンツの重要性が増している。その中にはネットコミュニティにおける共有知である表現を用いることで必要な素材の数を減らし、容易に制作が可能なコンテンツが存在する。

このような表現はミームと呼ばれる。ミームとは、R・ドーキンスにより提唱された文化における情報の伝達単位であり、ここではインターネット上で話題となったコンテンツを指す。ミームには、多くの人が共通で認識するネタやコラージュ画像などが含まれている。

本研究では、シナリオデータから必要な情報を抽出し、公知のミームを用いた動画を生成する手法を提案する。シナリオデータからの動画の生成が容易になれば、動画の生成に関して時間の節約や作成のコスト削減、大量のコンテンツの作成が可能になる。

2. 要素技術

2.1. Generative Pre-trained Transformer (GPT)

GPT [1] は、OpenAI が開発した言語モデルである。自然言語や生成タスクにおいて、高速で効率的な処理を実現し、複雑な文脈を理解する能力を有する。

本研究では、情報の抽出やミーム素材の決定において、gpt-4 というモデルを使用し、temperature は 0.5 としている。

2.2. Structured Query Language (SQL)

SQL は、リレーショナルデータベースのデータを操作するための言語である。データの検索、追加、更新、削除などをするための標準的な構文を提供する。本研究では、MySQL を使用した。

2.3. Stable Diffusion

Stable Diffusion は Stability AI が提供する画像生成 AI モデルで、テキスト入力から高品質な画像を生成する。

2.4. Adobe After Effects (AE)

Adobe After Effects は、Adobe が提供する高度なデジタル映像編集・合成ソフトウェアである。主にモーショングラフィックス、ビジュアルエフェクトを動画に加えることに利用され、動画、テレビ番組、ウェブコンテンツ、広告などで活用されている。操作には、Adobe 製品向けの JavaScript ベースのスクリプト言語である ExtendScript を使用した。

3. 提案手法

本研究では、シナリオデータを解析し、動画を自動生成する手法を提案する。

- GPT を使用して、シナリオデータから情報を抽出
- ミーム素材¹に関するデータを入力した MySQL のデータベースを使用し、適切なファイルパスを入手
- Stable Diffusion を使用して、背景画像を生成
- ファイルパス、テキスト情報をもとに ExtendScript を実行

以下に、使用したシナリオデータを示す。
「学生にとって学校での勉強はとても重要である。部屋で長時間勉強していると、疲れてしまうこともあるだろう。そんな時は、少し休憩すると気分がとても楽になる。」

¹https://sasalabo.net/2024/02/27/猫ミームとは/#google_vignette

表 1: シナリオから抽出した情報

| | 時間 | 場所 | 登場人物の状態 | テキスト情報 |
|-----|------|----|---------|---------|
| 文 1 | | 学校 | 勉強中 | 勉強の重要性 |
| 文 2 | 長時間 | 部屋 | 疲れてしまう | 勉強中 |
| 文 3 | そんな時 | | 休憩中 | 気分が楽になる |

表 2: ミーム素材の決定

| | 登場人物の状態 | 特徴名 | ミーム名 |
|-----|---------|-------|---------|
| 文 1 | 勉強中 | タイピング | PC 猫 |
| 文 2 | 疲れてしまう | 落ち込む | しょんぼり猫 |
| 文 3 | 休憩中 | 爆睡 | いびきをかく猫 |

3.1. Step 1: シナリオからの情報の抽出

シナリオデータを文単位に分割し、各文に対して GPT を使用して時間、場所、登場人物の状態、テキスト情報を抽出する。

3.2. Step 2: ミーム素材の決定

ミーム素材をまとめた MySQL のデータベースから、抽出した登場人物の状態に適する素材を決定する。

3.3. Step 3: 背景画像の生成

抽出した時間、場所の情報をもとに GPT で背景画像を説明するプロンプトを作成し、Stable Diffusion を使用して背景画像を生成する。

3.4. Step 4: 動画の生成

Step 2, Step 3 で決定した背景画像とミーム素材、Step 1 で抽出した情報をもとに ExtendScript を作成し、実行する。

4. 実験

シナリオデータを使用して、実際に 1 本の動画を生成した。ここでは、Step 1, Step 2 について評価する。

表 1 に、Step 1 で抽出した情報を示す。空欄は特定できなかったことを表す。

表 2 に、Step 2 で決定したミーム素材に関する情報を示す。勉強中に対して PC 猫、疲れてしまうに対してしょんぼり猫、のように登場人物の状態とミームの雰囲気一致している。

5. まとめと今後の課題

本研究では、シナリオの解析、ミーム素材の決定、背景画像の生成を実施し、これを使用した動画の生成を自動化するスクリプトを作成し、実行した。

今後の課題として、ミーム素材の適切なラベル付けや動画生成 AI の利用についても検討したい。

参考文献

- [1] OpenAI. Models. <https://platform.openai.com/docs/models/gpt-4-turbo-and-gpt-4>. Accessed: 2024-12-1.