

MoPhrase: 大規模言語モデルを用いた自然言語指示によるモーションパス編集システム

千葉 裕斗^{1,a)} 渡邊 恵太¹

概要: モーショングラフィックス制作では、手描き入力動作をリアルタイムにモーションパスとして記録するワークフローが用いられている。「ここをもう少し膨らませる」「徐々に動きを激しくする」といった編集意図を反映するには、ベジェ曲線の制御点やイージング曲線の直接操作が必要となる。そこで本研究では、自然言語によってモーションパスを局所的に編集するモーショングラフィックスエディタ MoPhrase を提案する。MoPhrase は、ユーザーが選択した区間に対し、自然言語で入力された意図に基づいて、大規模言語モデルを介し編集案を提示するシステムである。これにより、従来の直接操作に加えて、セマンティックな指示を活用したアニメーション編集を目指す。

1. はじめに

モーショングラフィックスは、図表、文字、図形といった視覚要素に動きを付与することで、情報を効果的に提示する映像表現手法である。モーショングラフィックスの制作において、創作者のアイデアを直感的に反映するアプローチとして、手描き入力をリアルタイムにアニメーションとして記録するワークフローが用いられている。例えば Adobe After Effects^{*1} や Procreate Dreams^{*2} は、ユーザーのスケッチを時系列の軌跡として取り込み、モーションパスの形状や速度変化を直接生成する機能を備えている。

スケッチによる入力はユーザーの想定と実際の入力結果が乖離しやすく、事後の整形が重要となる。これまで、手描き軌跡の平滑化手法 [1] や、幾何学的な制約に基づくストロークの整形手法 [2], [3] が提案されてきた。これらの手法は、手ブレの除去や規則的な形状への補正には有効である。一方、「ここをもう少し膨らませる」「徐々に動きを激しくする」といったセマンティックな編集を直接反映することは困難である。そのため、創作者がこれらの意図を反映するには、ベジェ曲線の制御点やイージング曲線を直接操作する必要がある。

そこで本研究では、自然言語によってモーションパスの形状および速度変化を局所的に編集するモーショングラフィックスエディタ「MoPhrase」を提案する (図 1)。MoPhrase は、ユーザーが選択した編集区間に対し、自然言



図 1 MoPhrase の編集例。手描き入力によるモーションパス (上) に対し、自然言語で「徐々に減衰させて」と指示することで、選択区間の形状を LLM が補正する (下)。

語で入力された意図に基づいて、大規模言語モデル (LLM) を介した編集案を提示するシステムである。本システムにより、従来の直接操作に加えてセマンティックな指示を活用し、ユーザーの意図を効率的かつ直感的にアニメーションへ反映させることを目指す。

¹ 明治大学 総合数理学部 先端メディアサイエンス学科

^{a)} ev230560@meiji.ac.jp

^{*1} <https://www.adobe.com/products/aftereffects.html>

^{*2} <https://procreate.com/dreams>

2. 関連研究

2.1 ストロークの整形手法

マウスやタブレットを用いたフリーハンド入力、高度な描画スキルを持たないユーザにとって制御が難しく、入力結果が曖昧なものとなりやすい。Schneider は、入力された座標点列に対してベジェ曲線を再帰的にフィッティングするアルゴリズムを提案し、この手法は軌跡の平滑化に広く用いられている [1]。Igarashi らは、入力ストロークに対して接続・平行・直交などの幾何学的な拘束を適用し、整形候補を提示するプロトタイプシステム「Pegasus」を提案した [2]。また Fiser らは、Pegasus の概念を拡張し、一般的なベジェ曲線をサポートした Adobe Illustrator^{*3} プラグイン「Shipshape」を提案した [3]。

これらの手法は、幾何学的な正解が存在する図形描画や、手ブレ等のノイズ除去には有効である。しかし、これらはルールベースの処理であり、対応できる操作は平滑化や単純化、幾何学的拘束の適用に限られる。「ここをもう少し膨らませる」「より鋭くする」といったセマンティックな編集を行うには、依然としてユーザが制御点を直接操作する必要がある。

2.2 自然言語によるアニメーション制作支援

創作支援の分野では、アニメーション制作を効率化するため、直接操作と自然言語指示を統合したインターフェースが提案されている。Tseng らは、SVG 形式のベクター画像に対して自然言語で指示を与え、CSS アニメーションコードを生成するツール「Keyframer」を提案した [4]。Bourgault らは、直接操作と自然言語による指示を統合した手描きアニメーションシステム「AniMate」を提案した [5]。Rosenberg らは、スケッチと発話を組み合わせるインタラクティブなプロトタイプを制御するシステム「DrawTalking」を提案した [6]。

これらの研究は、アニメーションの生成やオブジェクト間の振る舞いの定義に焦点を当てており、ユーザが描画したモーションパス自体の形状や速度変化に対するセマンティックな編集は扱っていない。

2.3 本研究の位置づけ

既存研究における未解決の課題は、モーションパスにおけるセマンティックな編集をいかに直感的に実現するかという点にある。本研究の独自性は、LLM の解釈能力を活用し、自然言語による指示を幾何学的パラメータの操作へ変換することで、モーションパスの形状や速度変化の直感的な編集を実現する点にある。

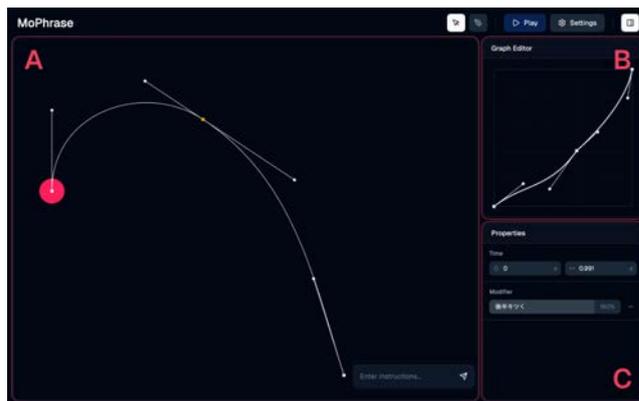


図 2 本システムのインターフェース。主に (A) キャンバスエディタ、(B) グラフエディタ、(C) プロパティエディタから構成される。

3. 提案システム

MoPhrase は、自然言語による指示に基づいてモーションパスを局所的に補正できるモーショングラフィックスエディタである。本システムは、モーションパスの形状を扱うキャンバスエディタと、イージング曲線を扱うグラフエディタから構成される (図 2)。

3.1 システム構成

本システムは、Web ベースのプロトタイプとして実装した。開発言語には TypeScript を用い、ビルドには Vite を採用した。キャンバス上の描画とインタラクションには p5.js を用い、UI は HTML と Tailwind CSS により構築した。補正案の生成には OpenAI GPT-5.2^{*4} を用い、Structured Outputs により JSON 形式の候補を生成した。

3.2 エディタ構成と基本操作

キャンバスエディタは、オブジェクトが移動する軌跡を定義するモーションパスを描画・編集するためのインターフェースである。グラフエディタは、オブジェクトの時間的な変化を定義するイージング曲線を編集するためのインターフェースであり、横軸を時間、縦軸を進捗率としたカーブによってモーションの緩急を補正する。プロパティエディタは、アニメーションの開始時間や長さなどの基本的なパラメータに加え、自然言語による編集を管理するインターフェースである。

ユーザはまず、キャンバスエディタ上でドラッグ操作を行い、オブジェクトの動きをスケッチする。システムは入力された点群をフィッティングし、ベジェ曲線によるモーションパスの概形を生成する [1]。同時に、描画にかかった時間と距離から速度変化を解析し、イージング曲線の初期形状としてグラフエディタに反映する。

*3 <https://www.adobe.com/products/illustrator.html>

*4 <https://platform.openai.com/docs/models/gpt-5.2>

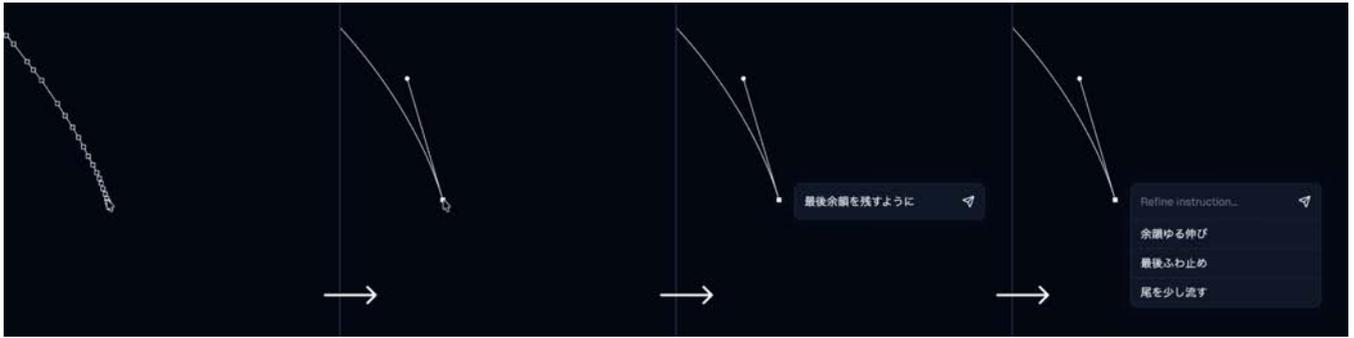


図 3 自然言語による編集フロー．ユーザは補正区間を選択してポップアップに編集の意図を入力すると、システムは LLM を介して複数の補正案を提示する。



図 4 補正案の選択および再調整のインタラクション．(左) 生成された補正案をスライダー操作により強度を調整しながら選択する．(右) 適用した補正はモディファイアとして保持され、後から再調整できる。

3.3 自然言語に基づく補正案の生成と編集

生成されたモーションパスやイージングに対して、ユーザは自然言語を用いた補正を行える。ユーザがキャンバスエディタ上のパスから補正したい区間を選択すると、システムはその近傍にポップアップを表示する。ユーザがポップアップ内のテキストボックスに「だんだん動きを激しく」「もう少し波打つ感じに」といった編集の意図を入力すると、システムは LLM を介して 3 通りの異なる補正案を生成し、ポップアップ内にリスト形式で提示する (図 3)。

提示された各候補案は、リスト項目そのものが適用度を操作するスライダーとして機能する。ユーザは領域上を水平にドラッグすることで、直感的に適用の強度を調整できる。このインタラクションにより、ユーザは LLM による提案と元の形状から、自身の意図に合致する補正案を探索的に決定できる。ユーザが適用した補正案は、モディファイアとしてプロパティエディタ内に表示される。これにより、ユーザは後から特定のモディファイアの適用度を再調整したり、補正を無効化したり、削除したりできる (図 4)。

3.4 補正案の生成アルゴリズム

本システムは、ユーザが選択した区間に対し、LLM を

用いて複数の補正案を生成する。LLM への入力データは、選択区間に含まれる幾何情報 (アンカーポイントおよびコントロールポイントの座標列) と、ユーザの自然言語による指示の 2 点である。

システムは、座標データを LLM に入力する前に次の前処理を行う。まず、選択された編集区間内の座標を 0 から 1 の範囲に正規化する。次に、各コントロールポイントの座標を、属するアンカーポイントを原点とする極座標に変換する。ベジェ曲線では、アンカーポイントに対するコントロールポイントの位置によって曲線の形状が定まる。極座標表現では、その距離が曲線の膨らみの程度を、角度が曲線の向きを表す。これにより、LLM がベジェ曲線の構造を解釈しやすくすることを目指した。

前処理後の座標データと自然言語による指示は JSON 形式で LLM に入力し、補正後の座標データも同形式で受け取る。LLM は異なる補正案を複数出力し、エディタ上でプレビュー可能な形で提示する。単一の出力ではユーザの意図との乖離が生じうるため、多様なパラメータを持つ候補を複数生成する設計とした。

4. 議論

4.1 LLM による空間把握の限界

LLM は本来テキスト処理を目的としており、座標データのような幾何情報の処理を得意としない可能性がある。本システムは、座標の正規化や極座標変換といった前処理によってこの課題の緩和を試みているが、複雑な形状の理解には限界が残る。

4.2 応答遅延とインタラクションの即時性

既存のアニメーション制作ツールにおける直接操作は、ユーザのアクションに対して即時的なフィードバックを返す。対して、本システムのように LLM による生成を用いたアプローチでは、生成時に数秒から十数秒の遅延が生じる。この遅延は、ユーザの作業への没入や集中を阻害する可能性がある。

4.3 評価と今後の展望

本稿は、プロトタイプの実装とコンセプトの提案を主たる目的としたため、定量的な評価やユーザスタディを実施していない。本システムの有用性を検証するためには、モーショングラフィックスデザイナーを対象としたタスク評価やユーザスタディが必要である。特に、提案手法によってユーザの意図通りの補正が実現できているか、また既存の映像制作のワークフローに円滑に組み込めるかを検証し、適用範囲と限界を明らかにする必要がある。

5. 結論

本稿では、自然言語による指示に基づきモーシオンパスの形状や速度変化を局所的に補正できるモーショングラフィックスエディタ「MoPhrase」を提案した。本システムは、ユーザが指定した区間に対して LLM が複数の補正案を生成し、プレビュー上での比較を通じて修正できる編集ワークフローを提供する。これにより、従来の直接操作では困難であったセマンティックな編集を実現した。今後は、定量的な評価およびユーザスタディを実施し、実際の制作フローにおける本システムの有用性を明らかにする。

参考文献

- [1] Philip J. Schneider. *An algorithm for automatically fitting digitized curves*, p. 612–626. Academic Press Professional, Inc., USA, 1990.
- [2] Takeo Igarashi, Satoshi Matsuoka, Sachiko Kawachiya, and Hidehiko Tanaka. Interactive beautification: a technique for rapid geometric design. In *Proceedings of the 10th Annual ACM Symposium on User Interface Software and Technology*, UIST '97, p. 105–114, New York, NY, USA, 1997. Association for Computing Machinery.
- [3] J. Fišer, P. Asente, and D. Sýkora. Shipshape: a drawing beautification assistant. In *Proceedings of the Workshop on Sketch-Based Interfaces and Modeling*, SBIM '15, p. 49–57, Goslar, DEU, 2015. Eurographics Association.
- [4] Tiffany Tseng, Ruijia Cheng, and Jeffrey Nichols. Keyframer: Empowering animation design using large language models, 2025.
- [5] Samuelle Bourgault, Li-Yi Wei, Jennifer Jacobs, and Rubaiat Habib Kazi. Narrative motion blocks: Combining direct manipulation and natural language interactions for animation creation. In *Proceedings of the 2025 ACM Designing Interactive Systems Conference*, DIS '25, p. 1366–1386, New York, NY, USA, 2025. Association for Computing Machinery.
- [6] Karl Toby Rosenberg, Rubaiat Habib Kazi, Li-Yi Wei, Haijun Xia, and Ken Perlin. Drawtalking: Building interactive worlds by sketching and speaking. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology*, UIST '24, New York, NY, USA, 2024. Association for Computing Machinery.