

Title	拡散モデルを用いたスケッチによる煙エフェクトの生成と制御
Author(s)	常, 恒遠
Citation	
Issue Date	2026-03
Type	Thesis or Dissertation
Text version	ETD
URL	https://hdl.handle.net/10119/20592
Rights	
Description	Supervisor: 謝 浩然, 先端科学技術研究科, 博士

氏名	CHANG Hengyuan		
学位の種類	博士 (情報科学)		
学位記番号	博情第 579 号		
学位授与年月日	令和 8 年 3 月 25 日		
論文題目	Smoke Effect Generation and Design with Sketch Control using Diffusion Models		
論文審査委員	謝 浩然	北陸先端科学技術大学院大学	准教授
	長谷川 忍	同	教授
	岡田 将吾	同	教授
	井之上 直也	同	准教授
	藤澤 誠	筑波大学	准教授

論文の内容の要旨

2D smoke effects are ubiquitous visual elements in graphic design, illustration-oriented creation and animation. In practice, conventional 2D smoke design is generated using procedural physical simulations with various parameters, which requires substantial expert knowledge of fluid dynamics. Non-expert users are required to transfer the creative intentions into modification of simulation parameters, which causes a semantic gap between design intent and the resulting smoke effect. In contrast, hand-drawn sketches provide an intuitive and abstract representation of visual intent. For smoke effects, a few strokes can convey both salient structures and coarse motion cues, making abstract intent more concrete.

Fluid synthesis has been reformulated as a data-driven generation task due to advancements in conditional generative models. The generative models learn an end-to-end mapping from sketches to target outputs by leveraging paired training data. Still, previous sketch-guided flow field generation methods adopt a one-stage conditional generative adversarial network flow field generation framework using sketch as direct input. The training process may be unstable due to the adversarial generation process. One-stage mapping also offers insufficient control constraints during the generation process. Inferring dense flow fields from sparse sketches is ill-posed, one-stage generation process amplifies the uncertainty inherent in mapping sketches to flow fields. Two-stage flow field generation study invokes intermediate representation extracted from flow field to provide explicit constraint during the generation process. The prior studies rely on regional constraints including mask or filter, which lacks explicit consistency constraints on the motion information contained from flow field.

To address the aforementioned challenges, this dissertation connects the sketch and the smoke effect using diffusion model, and constrains the generation process by extracting underlying physical information from the flow field. The dissertation

structure is organized as follows:

The first stage introduces sketch-guided flow field generation using diffusion models. The velocity field used for training are obtained via smoke simulation with simplified streamlines extracted as corresponding sketch representations. Diffusion models circumvent adversarial training between generator and discriminator by adopting a denoising framework for data distribution estimation, ensuring a stable training process. The proposed framework enables users to directly synthesize velocity fields from hand-drawn sketches.

The second stage introduces two-stage sketch-guided smoke illustration generation using Lagrangian coherent structure (LCS). The process of generating a velocity field from a sketch is decomposed into two simplified subprocesses. The first process generates LCS from input sketches. The second process generates velocity field under the guidance of LCS. The LCS serves as an intermediate representation between sketch and velocity field. The LCS enhances regional control ability of the diffusion model. The generated velocity field can be used as a guiding force, which can be directly integrated into existing fluid simulation tools for smoke effect design.

The third stage introduces two-stage sketch-guided smoke illustration generation using stream function. The stream function carries the rotational information of the flow field and the geometry of the fluid. The stream function is used as an intermediate representation between the sketch and the velocity field, incorporating motion priors into velocity field generation. The generated velocity field is used as guidance force for smoke effect generation. The artistic stylization of smoke is achieved for 2D smoke illustration synthesis through a stylized pre-trained model.

The final stage introduces two-stage sketch-guided smoke video generation using stream function. In the first process, the sketch is used as input condition for generating the smoke video first frame and stream function using diffusion model. In the second process, the generated first frame, visualized stream function and text prompts are combined to control the smoke video generation through video diffusion model. A motion control module is incorporated to leverage the stream function as motion guidance for smoke video generation. In addition, a pre-trained stylization model is optionally provided for stylized smoke video synthesis.

This dissertation provides a sketch-guided approach to create smoke effects, offers a new perspective on balancing user creative intent with maintaining the consistency of generated physical contents. First, this dissertation utilizes diffusion model instead of generative adversarial network as training framework, improves the generation stability by circumventing the adversarial process. Second, this dissertation introduces two-stage generation strategy to decrease the velocity field generation ambiguity from sketch inputs. Third, this dissertation enhances the coherence of

generated velocity fields by incorporating explicit motion guidance, moving beyond the reliance on regional constraints. Finally, this dissertation synthesizes 2D smoke animations via video diffusion model conditioned on motion guidance derived from sketch inputs, thereby obviating the requirements of simulation tools.

Keywords: Fluid Simulation, Diffusion Model, Sketch-Guided Generation, Lagrangian Coherent Structure, Stream Function, Video Generation.

論文審査の結果の要旨

従来の煙や流体のエフェクト CG 制作においては、設計意図と煙エフェクトの間の複雑なマッピングが大きな障壁となっており、非専門家にとっては参入が困難であった。本論文は、生成 AI モデルと手描きスケッチを用いて伝統的な流体シミュレーションに対して、スケッチベースのデータ駆動煙エフェクト生成システムを提案・構築したものである。本論文は、この課題に対して拡散モデルに基づく一連の手法を段階的に開発し、その有効性を体系的に実証した。

第一段階：スケッチから速度場への直接生成においては、ユーザが描いたスケッチの形状および流れの方向と整合性を持つ速度場の直接生成を実現した。本研究では、潜在拡散モデルを用いてスケッチと速度場間のマッピングを学習する試みを初めて行った。従来の敵対生成ネットワーク GAN ベースの速度場生成手法との比較実験を通じて、拡散モデルが優れた安定性を発揮することを実証した。

第二段階：スケッチから速度場への複雑なマッピングを二段階の生成に分解することで、速度場生成における曖昧性の低減を図った。具体的には、Finite-Time Lyapunov Exponent から導出される Lagrangian Coherent Structures(LCS)領域を中間表現として採用し、スケッチと速度場を接続するための媒介として機能させた。一段階モデルとの比較実験により、二段階モデルが制御条件を明示的な制約として活用することで、より精密な速度場を生成できることが確認された。

第三段階：速度場生成の一貫性をさらに改善するために流れ関数 (Stream Function) を導入した。LCS 領域が速度場を制約するための二値マスクとしての役割にとどまるのに対し、流れ関数は速度場の回転ダイナミクス特徴および構造的情報の双方を包含するため、より豊かな物理的制約を生成プロセスに付与することが可能となる。これにより、生成される速度場の物理的整合性が大幅に向上することが示された。

第四段階：煙エフェクトの動画生成においては、入力スケッチを誘導として用いた動画生成のパイプラインを構築した。本研究では、まずスケッチから映像の初期フレームを生成し、同時に流れ関数を導出してモーション制御モジュールに統合することで、運動制約として機能させた。さらに、初期フレームにスタイル転移を適用することで、アニメスタイルの映像生成を実現した。先行研究との定量的比較分析により、提案フレームワークの有効性が実証された。

以上、本論文は、拡散モデルを基盤として、物理的にもっともらしい煙エフェクトと意味的なスケッチ入力を効果的に結合することに成功した。スケッチベースの流体生成分野における今後の発展に寄与するものであり、学術的に貢献するところが大きい。よって博士 (情報科学) の学位論文として十分価値あるものと認めた。