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

Doctoral Dissertation

Smoke Effect Generation and Design with Sketch Control using Diffusion Models

Hengyuan Chang

Supervisor Haoran Xie

Graduate School of Advanced Science and Technology
Japan Advanced Institute of Science and Technology
(Information Science)

March 2026

Abstract

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 a 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.