PYGEEK

Vol. 1 · 2026

Proceedings of the Conference on PyGeek

2026 International Conference on PyGeek (PyGeek 2026)

ISSN (Online): 3140-5312

Published by Marketian

PyGeek

Proceedings of the Conference on PyGeek

2026 International Conference on PyGeek (PyGeek 2026)

Volume 1 · 2026

Published by Marketian

 ISSN (Online): 3140-5312

Automatic Generation of Emotion Retrospective Content Based on Generative AI

Junsun Chang[1], Janghwan Kim[2], Chaeyun Seo[3], R. Young Chul Kim[4]

[1]Graduate Student, Department of Software and Communications Engineering, Hongik University, South Korea

[2]Ph.D Candidate, Department of Software and Communications Engineering, Hongik University, South Korea

[3]Ph.D, Department of Software and Communications Engineering, Hongik University, South Korea

[4]Professor, Department of Software and Communications Engineering, Hongik University, USA

Corresponding author: R. Young Chul Kim (bob@hongik.ac.kr)

Acknowledgement: This research was supported by the Ministry of Science and ICT(MSIT), Korea, under the Sejong SW Convergence Cluster 2.0 Project supervised by the National IT Industry Promotion Agency (NIPA) (Grant No. PJTS08012604100700010003002 00).

Abstract

With advances in generative AI, the technology for automatically generating images and videos is rapidly developing. However, existing technology for creating videos from static images requires users to write video prompts directly, which requires specialized knowledge and has limitations in producing retrospective content that reflects personal memories and emotions. To solve this, we propose a generative AI-based framework that automatically creates emotional retrospective videos using two photos (static images) taken at the start and end. Our method analyzes the input photo to extract scene, object, and mood information using a large language model (LLM), then automatically generates video prompts that include actions, camera movements, environment, lighting, and emotional elements. The generated prompts are fed into a generative model that converts images into videos, creating emotional retrospective videos. Additionally, the entire generation process is automated to reduce the user's burden of writing prompts and simplify content creation. This study presents a method for the automatic creation of photo-based emotional retrospective content and suggests potential applications of digital retrospective services that utilize generative AI.

.

Keywords: Machine Learning, Dual Learning, Natural Language Processing

1. Introduction

Image-to-Video technology, enabled by recent advances in generative AI, is increasingly applied to digital reminiscence services alongside content creation. However, existing services remain limited to static content such as photos, failing to convey individuals' movements or the atmosphere of the time, and typically require users to write prompts—a barrier for senior users unfamiliar with generative AI [1].

To address these limitations, we propose an AI-based interface that automatically generates emotional reminiscence videos from a single photograph. The system analyzes the input photo, automatically generates video prompts using a large language model (LLM), and produces short reminiscence videos depicting human movement and environmental atmosphere via an Image-to-Video model. By eliminating the need for manual prompt writing, the proposed system transforms static photos into vivid videos, enhancing user immersion and demonstrating the potential of generative AI for emotional reminiscence content.

ISSN (Online): 3140-5312

Chapter 2 discusses DreamVideo as a related study, and Chapter 3 discusses the proposed Automatic Generation of Emotion Retrospective content. Finally, Chapter 4 discusses the conclusion and future research.

2. Related Works

DreamVideo [2] is an Image-to-Video (I2V) model that generates videos from a single input image conditioned on a text prompt. Since conventional I2V models rely solely on the semantic information of the input image, visual attributes such as facial features, clothing, and background often change inconsistently across frames. To address this, DreamVideo introduces an Image Retention Branch, which extracts CNN-based features from the input image and passes them to the Video Diffusion model to preserve fine-grained visual details, and a Double-condition Classifier-Free Guidance (CFG) scheme, which jointly conditions generation on both the input image and the text prompt—enabling diverse actions (e.g., walking, running, dancing) from the same image. Experimental results show that DreamVideo achieves better preservation of input image features and more natural video generation than existing I2V models.

However, DreamVideo primarily focuses on preserving visual consistency and does not account for semantic information such as a character's personality or role. It has limitations in generating long videos or in expressing complex storylines. Thus, in domains requiring character consistency and narrative continuity—such as educational storytelling—additional methods are needed to complement this approach.

3. Automatic Generation of Emotion Retrospective Content

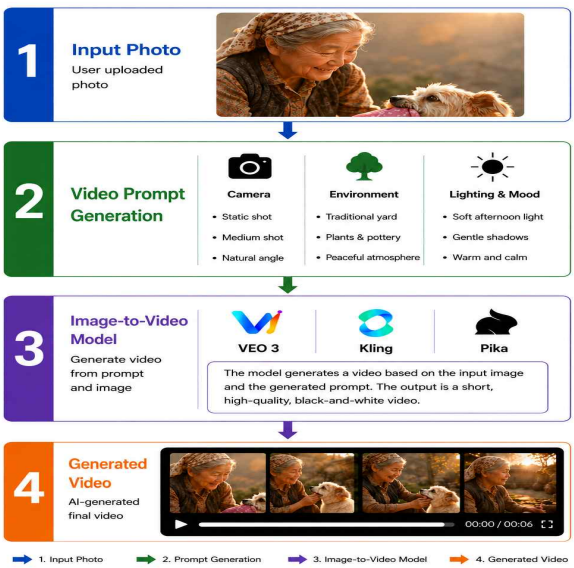

Fig. 1. Flowchart of data processing and analysisd

As shown in Figure 1, the proposed system consists of four stages. In Step 1 (Input Photo), the user uploads a single photograph, requiring no additional text input and making the system accessible even to users unfamiliar with generative AI.

In Step 2 (Video Prompt Generation), the LLM analyzes the image to extract Action, Camera, Environment, Lighting, and Mood/Style elements, converting them into video-generation prompts. The Camera element defines shot composition (e.g., static shot, medium shot), Environment defines the background context (e.g., traditional yard, peaceful atmosphere), and Lighting/Mood sets the overall tone (e.g., soft afternoon light, warm and calm). A Negative Prompt is also generated to minimize distortions and improve video quality.

In Step 3 (Image-to-Video Model), the generated prompts and the original image are jointly fed into the Image-to-Video model, which preserves the visual characteristics of the person and background while capturing natural movements such as walking or interaction.

In Step 4 (Generated Video), a short reminiscence video capturing both human movement and the atmosphere of the time is delivered to the user, conveying a sense of vividness that static photographs alone cannot provide.

Overall, this pipeline removes the burden of prompt writing while automatically generating emotional reminiscence videos from a single photograph.

4. Conclusion

We propose an AI-based interface that automatically generates emotional reminiscence videos from a single photograph. The system analyzes the input image, uses an LLM to generate video and negative prompts automatically, and applies an Image-to-Video model to produce short videos depicting human movement and the atmosphere of the time. By eliminating the need for manual prompt writing, the proposed system lowers the barrier for users unfamiliar with generative AI while enhancing immersion in digital reminiscence content.

However, the system does not account for semantic information such as a character's personality or narrative context and remains limited in generating long videos or complex storylines. Future work will focus on incorporating character-consistent generation techniques, extending the system to multi-photo, story-based video

 ISSN (Online): 3140-5312

generation, and validating its effectiveness through usability studies with senior users.

Conflict of Interest

The authors declare that there are no potential conflicts of interest related to this paper.

Acknowledgement

This research was supported by the Ministry of Science and ICT(MSIT), Korea, under the Sejong SW Convergence Cluster 2.0 Project supervised by the National IT Industry Promotion Agency (NIPA) (Grant No. PJTS08012604100700010003002 00).

References

[1] Z. Zhang, L. Xu, X. Wang, X. Zhang, & M. Fan, (2025). "Understanding and co-designing photo-based reminiscence with older adults.", *in Proceedings of the ACM on Human-Computer Interaction,* 9(2), 1-30.

[2] C. Wang, J. Gu, P. Hu, Y. Guo, X. Dong, H. Xu, & X. Liang (2025, April). "Dreamvideo: High-fidelity image-to-video generation with image retention and text guidance." *In Proceedings of the 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 1-5). IEEE.

Author Information

JunSun Chang
now: Graduate Student
Hongik University, South Korea
Email: cjsun0331@hanmail.net
Research Interest: Natural Language Processing, Machine Learning, Prompt Engineering

Janghwan Kim
now: Ph.D Candidate, Department of Software and Communications Engineering, Hongik University, South Korea
Email: lentoconstante@hongik.ac.kr
Research Interest: AI software validation, reinforcement learning-based software validation, software visualization, and requirements engineering.

Chaeyun Seo
2014: Ph.D, Department of Computer Engineering, Hongik University
Email: chaeyun@hongik.ac.kr
Research Interest: Metaverse, Metaverse Contents, Design Thinking, CBD, Software Engineering, BPM, AI, Business Process Modeling.

R. Young Chul Kim
2000: Ph.D. Degree,
Software Engineering from the Department of Computer Science, Illinois Institute of Technology, USA
Email: bob@hongik.ac.kr
Research Interest: Test Maturity Model, Model-Based Testing, Metamodeling, Software Process Model, Software Visualization, AI software validation, reinforcement learning-based software validation.