



# Automatic Generation of 3D Scene Animation Based on Dynamic Knowledge Graphs and Contextual Encoding

Wenfeng Song<sup>1</sup> · Xinyu Zhang<sup>1</sup> · Yuting Guo<sup>2</sup> · Shuai Li<sup>2,3</sup> · Aimin Hao<sup>2,4</sup> · Hong Qin<sup>5</sup>

Received: 3 August 2022 / Accepted: 14 June 2023 / Published online: 1 July 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023, corrected publication 2023

## Abstract

Although novel 3D animation techniques could be boosted by a large variety of deep learning methods, flexible automatic 3D applications (involving animated figures such as humans and low-life animals) are still rarely studied in 3D computer vision. This is due to lacking of arbitrary 3D data acquisition environment, especially those involving human populated scenes. Given a single image, the 3D animation aided by contextual inference is still plagued by limited reconstruction clues without prior knowledge pertinent to the identified figures/objects and/or their possible relationship w.r.t. the environment. To alleviate such difficulty in time-varying 3D animation, we devise a dynamic scene creation framework via a dynamic knowledge graph (DKG). The DKG encodes both temporal and spatial contextual clues to enable and facilitate human interactions with the affordance environment. Furthermore, we construct the DKG-driven variational auto-encoder (DVAE) upon animation kinematics knowledge conveyed by meta-motion sequences, which are disentangled from videos of prior scenes. It is then possible to utilize the DKG to induce the animations in certain scenes, thus, we could automatically and physically generate plausible 3D animations that afford vivid interactions among humans, low-and life animals in the environment. The extensive experimental results and comprehensive evaluations confirm our DKGs' representation and modeling power towards new animation production in 3D graphics and vision applications.

**Keywords** 3D scene animation · Dynamic knowledge graphs · Contextual encoding · Automatic animation generation

---

Communicated by Wanli Ouyang.

✉ Shuai Li  
lishuai@buaa.edu.cn

✉ Hong Qin  
qin@cs.stonybrook.edu

Wenfeng Song  
songwenfenga@gmail.com

Xinyu Zhang  
zhangxinyu1@bistu.edu.cn

Yuting Guo  
guoyuting@buaa.edu.cn

Aimin Hao  
ham@buaa.edu.cn

<sup>1</sup> Computer Science, Beijing Information Science and Technology University, Beijing 100081, People's Republic of China

<sup>2</sup> State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing 100191, People's Republic of China

## 1 Introduction

The creation of human animations is essential in various applications, such as augmented and virtual reality, robotics, and human-computer interaction. In these applications, the movements of humans have interactive relations with objects in the environment. However, most existing works (Kocabas et al., 2020; Yuan & Kitani, 2020; Kolotouros et al., 2019) in this field overlook the prior knowledge of the world related to possible interactions among humans and the environment. Therefore, the goal of this paper is to explore the possibility of automatic animation generation from a single image. Generating 3D scene animations is a relatively unexplored area, to

<sup>3</sup> Zhongguancun Laboratory, Beijing, People's Republic of China

<sup>4</sup> Research Unit of Virtual Body and Virtual Surgery Technologies, Chinese Academy of Medical Sciences, 2019RU004, Beijing, People's Republic of China

<sup>5</sup> Department of Computer Science, Stony Brook University Stony Brook, New York, NY 11794-2424, USA