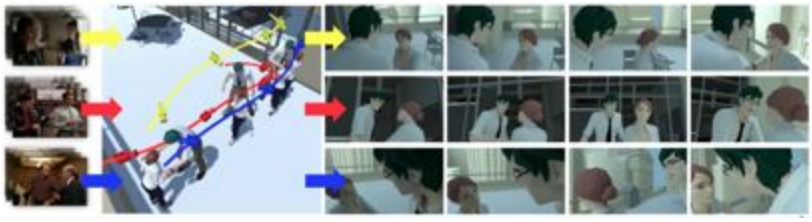
Master Thesis
Example-based suggestion system for virtual film prototyping



Advisors: Christophe Lino (Researcher at the GeoViC team of Ecole Polytechnique), Vicky Kalogeiton (Assistant Professor at the GeoViC team of Ecole Polytechnique), and Marc Christie (Assistant Professor in Mimetic team at the Cinematography group of IRISA/Inria Rennes)

Environment

This work will take place at the GeoViC team (Geometry & Visual Computing) of LIX, at Ecole Polytechnique. If sanitary conditions allow, the work will take place on-site; otherwise remotely.

Objective

Creating virtual films requires combining different shots taken from different cameras and viewpoints. This task is tedious and usually requires hours of manual watching clips and carefully selecting adequate combinations. To alleviate this, we are interested in developing techniques that enable fast editing of virtual films by exploiting real-life movies. Therefore, the primary goal of this thesis is to propose new deep-learning techniques that learn a mapping between an input filmed content and a directorial style (e.g. preferred viewpoints, motions, edits, or cutting rhythm). Given a new 3D animated content, the secondary goal is to learn how to generate a collection of interesting cinematographic suggestions. The successful candidate would need to work on deep-learning, virtual camera control and if needed filming conventions (with potential collaboration with the Pekin University and Beijing Film Academy).

How to apply

To apply, please contact Christophe Lino at Christophe.Lino@polytechnique.edu, Vicky Kalogeiton at vicky.kalogeiton@polytechnique.edu and Marc Christie at Marc.Christie@irisa.fr with "[Intenship Virtual Film Prototyping]" in the subject line and please provide a (1) CV, and (2) a short statement of research interest. If needed we may ask for two references. We particularly encourage applications from women, and from underrepresented groups in academia.

Motivation and context

Due to the growing need of compelling narrative experiences in virtual worlds, there is a pressing requirement to integrate dedicated content creation tools, especially ones that ease the reproduction of cinematographic conventions [1][2] on 3D animated contents.

Until recently, the virtual cinematography community had mostly focused on declarative approaches inspired by these conventions. Given user-declared constraints, they enable the automated computation of camera placements [5], camera motions [6] or film edits [7] (i.e. switching between cameras) through

optimization-based or geometric-solving-based techniques. Some interactive approaches have also enabled manipulating or interpolating viewpoints in image space [4], by using an intermediate camera parameter space (The Toric Space). With this space, one can also easily store collections of viewpoints or trajectories around two characters, and retarget them to new contents, yet in a naïve way. One drawback of these techniques is that they require a tedious process of manually specifying constraints or manually creating a collection of viewpoints or trajectories.

More recently, researchers have started investigating machine and deep learning techniques. For instance, some works learn camera behaviors (i.e. motions) to film two-characters animations [8], or film editing patterns between static cameras [9] from movie sequences, and to automatically re-apply such behaviors or patterns to new animated contents. Other works [10,11] use CNNs to classify shot types and analyse their usage depending on the cinematographic style each scene conveys.

However, current methods lack user-controllability. In this thesis, we would like to develop deep learning models that offer better cooperation between system and user (active learning, user-in-the-loop). For instance, the computer could provide suggestions to the user, who in turn would select one as a starting point. And, given the user's choices, or their previous films, the computer could also learn to adapt the suggestions to their own style.

Thesis organization

The work will first consist in reading and analyzing state of the art techniques in camera control and deep-learning for time series (RNN, LSTM), as well as controllability in machine learning through constraints or design of latent-driven control spaces.

After this first step, discussions with the supervisors will enable the identification of key features to extract, mappings to learn from, and dimensions of a potential control space. Implementation of prototypes will include i) feature extraction from real film sequences, ii) learning of temporal sequences of visual features, and iii) interactive control.

Required skills

- In pursuit of Masters degree in a relevant field
- Experience computer graphics and in particular 3D animation
- Proven experience in Python
- High level of innovation and motivation
- Hands-on experience with deep learning frameworks (PyTorch)
- Communication skills in English

References

[1] Grammar of the Shots. R. Thompson, C. Bowen. Taylor \& Francis, 2009.
[2] Grammar of the Edit. R. Thompson, C. Bowen. Taylor \& Francis, 2009.
[3] The Director's Lens. C. Lino, M. Christie, R. Ranon, W. Bares, 2011.
[4] Intuitive and Efficient Camera Control using the Toric Space. C. Lino and M. Chistie, In SIGGRAPH 2015.
[5] Improving the Efficiency of Viewpoint Composition. R. Ranon and T. Urli, In TVCG 2014.
[6] Camera-on-rails: Automated Computation of Constrained Camera Paths. Q. Galvane, M. Christie, C. Lino, R. Ronfard, In MIG 2015
[7] Continuity Editing for 3D Animations. Q. Galvane , R. Ronfard, C. Lino, M . Christie, AAAI 2015.
[8] Example-driven virtual cinematography by learning camera behaviors. H. Jiang, B. Wang, X. Wang, M. Christie, B. Chen. In TOG, 2020.

[9] A virtual director using hidden markov models. B. Merabti, M. Christie, K. Bouatouch. Computer Graphics Forum 2015.
[10] A Unified Framework for Shot Type Classification Based on Subject Centric Lens, A. Rao, J. Wang, L. Xu, X. Jiang, Q. Huang, B. Zhou, D. Lin, In ECCV 2020
[11] MovieNet: A Holistic Dataset for Movie Understanding, Q. Huang, Y. Xiong, A. Rao, J. Wang, D. Lin, In ECCV 2020