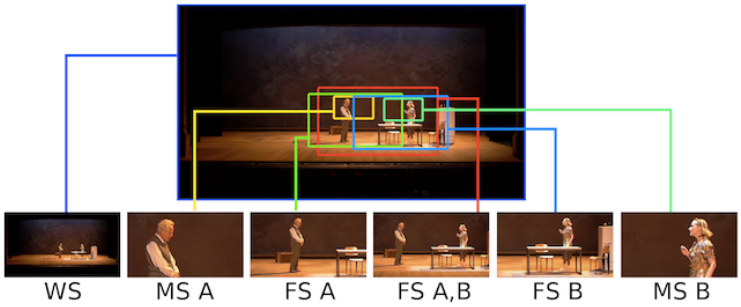
PhD Proposal : Computational Video Editing for Stage Performances

Supervisors : Remi Ronfard, remi.ronfard@inria.fr, LJK/IMAGINE, UGA, INRIA, https://team.inria.fr/imagine.

Benjamin Lecouteux, benjamin.lecouteux@univ-grenoble-alpes.fr, LIG/GETALP, UGA.



Context : This PhdD thesis is proposed as part of an ongoing collaboration between computer scientists and performings arts researchers at Univ. Grenoble Alpes and INRIA to use video in teaching and researching the performing arts. In a previous project, the IMAGINE team at INRIA has developped methods for automatic generation of cinematic rushes from ultra high definition video recordings of stage performances [1]. Here, we would like to propose techniques for making documentary movies from the generated rushes, based on an analysis of the script of the performance and a formalization of the rules of film editing. Ideally, the proposed techniques should be completely non-invasive (not requiring sensors on actors or on stage) and intuitive enough to be used by performing arts students, professors and researchers, without any expertise in video production.

Description: The goal of the PhD thesis will be to propose novel interaction techniques to students, professors and researchers in the performing arts for making movies from stage performances recorded on stage. On the one hand, we will propose novel algorithms for editing cinematographic rushes together into movie clips automatically, based on computational models of film editing « idioms » and machine analysis of the actors speech and motion. On the other hand, we will propose novel user interfaces for easily choosing between available idioms as in [2] and creating new idioms for the specific purpose of teaching and researching mise en scene and acting techniques.

More specifically, we will draw on previous work for automatically creating well-edited movies from a 3D animation [3]. The proposed framework uses dynamic programming to minimize, under a semi-Markovian hypothesis, the errors made along three criteria: the quality of the shots (with respect to the unfolding actions), the respect of continuity editing rules, and the respect of a well-founded model of rhythm (cutting pace). The same model is applicable to live-action video because it is based on measurable qualities in camera screen coordinates, which can be obtained with audio and video analysis. The main difference to be worked out during the PhD is that those measurements may be unreliable in the case of live-action video, due to tracking and audio or video analysis errors. We will propose a probabilistic version of our algorithm, and tools to solve them using Viterbi decoding. Further, we will propose novel methods for motion and speech analysis to support our models. The recent work of Leake et al. [2] will be used as a baseline for experimental validation.

During his/her thesis, the PhD student will create an extensive database of stage performance recordings, as part of a collaboration with the performing arts department at Univ. Grenoble Alpes and associated theatre companies. The raw recordings and the generated movies will be used as supporting material for teaching mise-en-scène and acting techniques, and for researching multiple aspects of expressive human motion, verbal and non-verbal communication, and dramaturgic techniques, as part of the new cross-disciplinary research project « Alps Performance Lab ».

References:

[1] Vineet Gandhi, Rémi Ronfard, Michael Gleicher. Multi-Clip Video Editing from a Single Viewpoint. CVMP 2014 - European Conference on Visual Media Production, Nov 2014.
[2] Mackenzie Leake, Abe Davis, Anh Truong, and Maneesh Agrawala. Computational video editing for dialogue-driven scenes. ACM Trans. Graph. 36, 4, July 2017.
[3] Quentin Galvane, Rémi Ronfard, Christophe Lino, and Marc Christie. Continuity Editing for 3D Animation. In AAAI Conference on Artificial Intelligence, AAAI Press, Austin, Texas, United States, January 2015.