

Fingers' dance: une grammaire du geste pour la navigation et l'annotation de films de danse.

Vincent Puig, Yves Marie
Haussonne, Thibaut Cavalié,
Armen Khatchatourov
IRI (Institut de Recherche et
d'Innovation)
4 rue Aubry Le Boucher
75004 Paris
<http://www.iri.centrepompidou.fr>
prenom.nom@centrepompidou.fr

Dominique Sciamma, Ioana
Ocnarescu, Huieen Kim,
Annabelle Eugenia, Baptiste
Lanne, Vanessa Reiser
Strate Collège Designers
30, rue Troyon
92310 Sèvres
<http://www.stratecollege.fr>
ds@stratecollege.fr

Résumé : Cette présentation porte sur une première expérimentation d'IHM d'annotation par le geste de films de danse. En s'appuyant sur une grammaire de gestes élaborée par le réalisateur Thierry de Mey pour ses propres films et musiques, un repérage des principaux gestes chorégraphiques est opéré avec le logiciel d'annotation *Lignes de temps*. Pour chacun de ces gestes un tracé gestuel à deux mains sur une table multipoint est défini. Sur la table multipoint, on peut tracer les gestes précédemment définis et les séquences filmées correspondantes s'affichent dans Lignes de temps. Par exemple un geste rotatif produit avec le doigt sur la table graphique va lancer une requête vers toutes les séquences où des mouvements en rotation sont dansés. On peut également annoter une séquence en lui attachant un ou des gestes prédéfinis. Les enjeux propres à cette approche sont explicités.

Approche et positionnement.

Le champ des recherches sur l'analyse et l'annotation du geste est particulièrement actif. Deux grands axes sont l'analyse et la segmentation automatisées pour reconnaissance de geste, que cela soit à des fins d'annotation et d'analyse du mouvement [Chambers,2004] [Roh 2008], ou d'analyse langagière [Hélior 2005]. Le second axe est celui de l'annotation fine des gestes filmés, l'annotation se faisant par un expert au moyen du texte qui décrit les composantes élémentaires du geste, le plus souvent ici aussi pour l'analyse langagière.

Par ailleurs, de multiples travaux de recherche en IHM explorent le rôle du geste pour la commande de telle ou telle fonction, en explorant par exemple comme moyen de collaboration le geste fait à plusieurs [Morris, 2006]. Mais il faut aussi souligner une certaine maturité des technologies grand public : les interfaces gestuelles rentrent aujourd'hui dans une phase de standardisation dominée par les enjeux industriels conséquents au niveau des grammaires proposées (cf. la confrontation Apple vs. Synaptics&PC dans le cas du pad multi-touch). Des

initiatives commencent à voir le jour pour rassembler les « grammaires » existantes, voire les standardiser par W3C contre une telle appropriation par un seul producteur.

Notre approche se différencie de ces travaux et se propose d'explorer l'annotation du geste (ici du mouvement dansé) par le geste. La motivation qui guide cette démarche est double.

Du point de vue théorique, il s'agit de (re-) penser notre rapport psychologique, sensitif et critique aux œuvres cinématographiques et plus généralement au mouvement, en s'appuyant sur les possibilités d'appropriation et de perception active offertes par le numérique. Il s'agit ici de proposer un outil d'annotation et de regard critique sur le geste dansé. Dans le domaine du jugement critique porté sur des films, nos travaux antérieurs menés en collaboration avec des réalisateurs et des critiques confirment que les amateurs cinéphiles prolongent leur pratique du cinéma dans le monde numérique au moyen d'appareils critiques, qu'ils soient instrumentaux (presse critique, outils de navigation, d'annotation et de pratique créative) ou sociaux (équivalents dans les réseaux sociaux de ce que furent les cinéclubs). Il s'agit donc de développer les nouveaux outils critiques qui permettent, avec les moyens du numérique, ce que Gilbert Simondon appellerait le processus d'individuation psychique et collective [Puig 2008].

Du point de vue de l'IHM, la démarche s'inscrit dans le champ des interfaces qui font expressément appel à l'engagement sensori-moteur de l'utilisateur, c'est à dire des dispositifs sollicitant les boucles sensori-motrices ou la perception active.

Dans le contexte du cinéma les appareils critiques gestuels ont ainsi vocation à rejoindre les outils de création du film: il s'agit de permettre ainsi l'implication corporelle de l'utilisateur dans la navigation et

l'annotation du corpus, comme le font justement la caméra, la table de montage analogique, etc.

En accord avec cette approche, nous ne procédons pas par une analyse et une segmentation automatisées du geste, mais par sa segmentation manuelle réalisée par un expert du domaine, à l'aide du logiciel Ligne de Temps. De même, la reconnaissance du geste d'annotation est assurée par une couche logicielle standard (cf. infra). Notre effort porte plutôt sur l'articulation entre le geste enregistré et perçu d'un côté, et le geste d'annotation de l'autre. En d'autres mots, il s'agit d'explorer la grammatisation du geste.

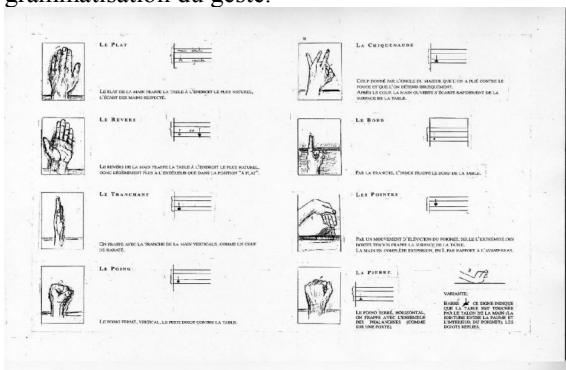


Fig1 : partition de *Musique de tables*

Grammaire chorégraphique et grammaire cinématographique

Le point de départ de cette recherche part du regard singulier que le réalisateur Thierry de Mey porte sur le geste, d'abord en tant que compositeur avec sa première partition gestuelle pour trois percussionnistes dans la pièce *Musique de tables* (1987).

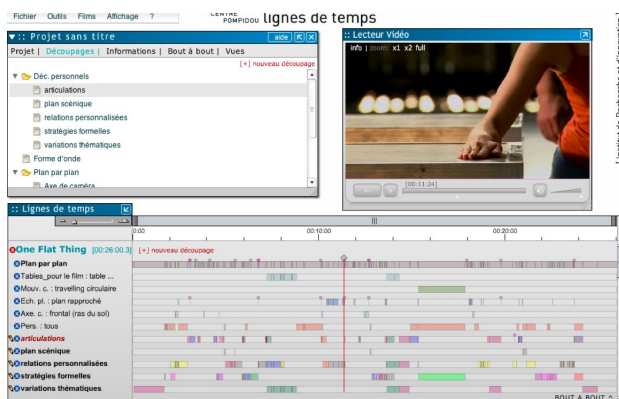


Fig2 : Analyse de *One flat thing* réalisée dans le logiciel Lignes de temps par Thierry de Mey

Il s'intéresse ensuite, en tant que réalisateur pour la danse, à l'articulation du langage cinématographique et des expressions chorégraphiques les plus contemporaines. C'est ainsi que la résidence de Thierry de Mey à l'IRI en 2007 a permis de procéder à une première analyse du film *One flat thing, reproduced*, une chorégraphie de William Forsythe en croisant la

perspective cinématographique (tailles de plan, mouvements de caméra) et chorégraphique (solos, groupes, gestes frappés, frottés, etc) à l'aide du logiciel d'annotation *Lignes de temps* développé depuis 2006 à l'IRI [Puig & al, 2007].

Fingers' dance : une interface pour l'annotation et la navigation par le geste

En septembre 2008, l'idée de concevoir une interface pour annoter et naviguer dans ce film s'est imposée dans le cadre d'une collaboration avec de jeunes designers du Strate Collège et grâce à un programme de résidence mis en place à l'IRI par Microsoft France.

L'outil choisi pour la commande gestuelle à des fins de navigation ou d'annotation est une table multipoint Microsoft Surface. Cet outil permet l'utilisation de gestes à deux mains et l'invention de formes collaboratives typiquement à deux, trois ou quatre personnes autour de la table. Le SDK fourni avec Microsoft Surface a été utilisé pour le développement de la commande gestuelle.

Après une phase de spécification avec Thierry de Mey, une recherche bibliographique a été entreprise sur les grammaires de geste actuellement développées, puis cinq maquettes ont été proposées pour valider le concept d'IHM le mieux adapté aux contraintes culturelles et technologiques. Une réduction du nombre de geste/signes a été ensuite opérée avec Thierry de Mey pour aboutir à deux grammaires de navigation/annotation limitées à quelques gestes de base. A l'issue de cette phase, un travail de design a été entrepris sur la base d'une interface multi-touch.

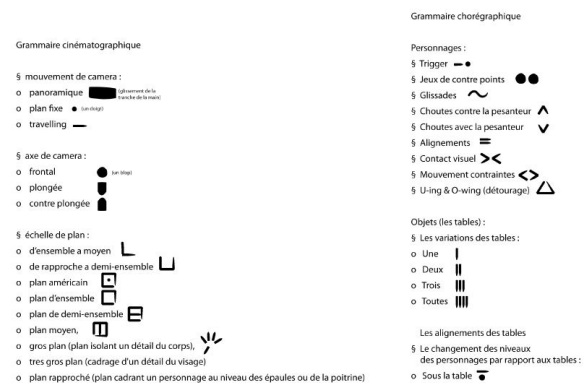


Fig3 : grammaire de gestes cinématographiques

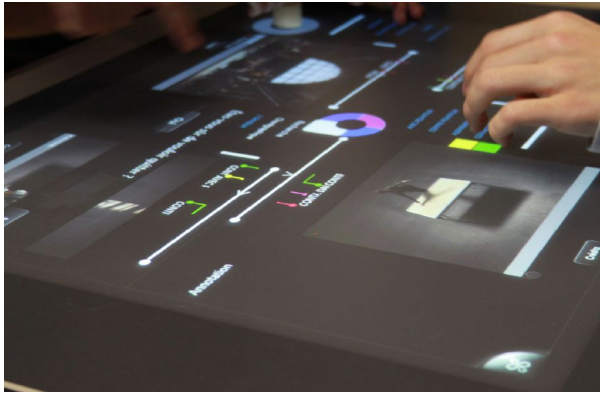


Fig4 : Interface utilisateur.

Mode d'annotation

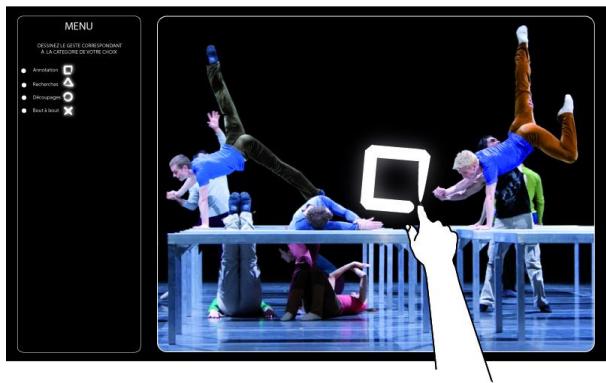


Fig5 : annotation d'une figure à quatre danseurs par tracé de geste sur l'écran multitouch

Sur la table multipoint, on peut tracer les gestes précédemment définis et les séquences filmées correspondantes s'affichent dans Lignes de temps. Par exemple un geste rotatif produit avec le doigt sur la table graphique va lancer une requête vers toutes les séquences où des mouvements en rotation sont dansés. On peut également annoter une séquence en lui attachant un ou des gestes prédéfinis. La table supporte aussi les objets tangibles pour la commande de certaines fonctions. Enfin, il est possible de partager la table en quatre zones de travail pour 1, 2 ou 4 personnes disposant chacune d'un écran de visualisation du film et d'outils pour la navigation et l'annotation à l'aide de gestes.

Plus généralement et à terme, l'objectif est bien de tenter de définir une grammaire de gestes suffisamment générique pour être utilisable sur un large ensemble de films de danse.

Enjeux de recherche et projets en préparation

Les enjeux de recherche que nous visons avec cette première expérimentation portent sur :

1. le geste comme mémoire du flux audiovisuel.

L'ensemble des travaux entrepris sur la grammaire chorégraphique montre le rôle enactif du geste et son pouvoir de mémorisation des repères spatiaux (position/postures des danseurs) alors que la grammaire cinématographique insiste sur les repères temporels. Ceci est particulièrement sensible dans l'analyse que fait Thierry de Mey des travellings de *One Flat Thing* (qui sont comme des « contre-points » des figures dansées) et des outils que nous avons par conséquent essayé de mettre entre les mains d'un large public d'amateurs. Mais évidemment le cinéma comme la danse jouent en permanence entre temps et espace et le geste se prête particulièrement bien à l'articulation entre les deux favorisant une mémorisation et une compréhension complémentaire de celle que l'on a déjà mis en place avec un outils d'annotation non-gestuel par mots-clés tel que Ligne de Temps.

2. la grammatisation du geste : le plus de généricité possible pour bénéficier à différentes applications. Procéder par réductionnisme sur un exemple singulier tel que *One flat thing* est une méthodologie qui n'est pas suffisante. On sait les enjeux de normalisation et de compétition industrielle que le pilotage d'un film par le geste pose déjà. Ici nous avons tenté d'opérer une généralisation pour les gestes de danse tout d'abord en testant la grammaire chorégraphique sur le film *Musique de tables* et ultérieurement sur un corpus de plus en plus large.

3. le geste collectif : productivité, surprise et bien être social.

L'interaction à plusieurs a commencé à être testée sous la forme de scénarios lors des ateliers proposés au public au festival VidéoDanse en novembre 2009, ainsi qu'au CNAM lors des Entretiens du Nouveau Monde Industriel et au cours des Carrefours des possibles de la FING (novembre 2009). D'or et déjà des effets de partage de tâches ou de partage de la mémoire du film sont observables et génèrent une interaction sociale productive qualitativement ou quantitativement.

Remerciements.

Nous remercions particulièrement Pierre-Louis Xech, Thomas Defaye et Thomas Serval de Microsoft France avec qui l'ensemble du projet depuis sa conception jusqu'aux phases de support technique a été mis en place. Remerciements également aux étudiants de l'EFREI et notamment Santiago Arias, Riley Leykni, Tarari Amine et Jonathan Pamphile qui ont participé au développement informatique.

Bibliographie

[Chambers 2004] Graeme S. Chambers, Svetha Venkatesh, Geoff A.W. West, and Hung H. Bui. Segmentation of Intentional Human Gestures for Sports Video Annotation. Proceedings of the 10th International Multimedia Modelling Conference (MMM'04)

- [Hélior 2005] Héloir, Gibet, Courty, Raynaud. Système d'annotation et de segmentation de gestes de communication capturés. TALN 2005
- [Morris 2006] Meredith Ringel Morris, Anqi Huang, Andreas Paepcke, Terry Winograd. Cooperative Gestures: Multi-User Gestural Interactions for Co-located Groupware, CHI 2006, April 22–28, 2006, Montréal, Québec, Canada.
- [Prié et Puig, 2008] Nouveaux modes de perception active de films annotés, Colloque Figures de l'Interactivité, Poitiers, novembre 2008 (à paraître).
- [Puig & al. 2007] Lignes de temps, une plateforme collaborative pour l'annotation de films et d'objets temporels, proceedings IHM, novembre 2007
- [Puig 2008] Amateurs du XXIème siècle, colloque Patrimoine et numérisation, novembre 2008, Culture et Recherche, novembre 2008
- [Roh 2008] Myung-Cheol Roh, Bill Christmas, Joseph Kittler and Seong-Whan Lee. Gesture spotting for low-resolution sports video annotation. Pattern Recognition. Volume 41, Issue 3, March 2008.