

Une machine à raconter des histoires ? Analyse et modélisation du processus de ré-éditorialisation de vidéos

Xavier BOST

Encadrement :

Georges Linarès (dir.) : Professeur, informatique.

Damien Malinas : Maître de Conférences, sciences de l'information et de la communication.

Serigne Gueye : Maître de Conférences, informatique.

Thèse soutenue le 23 novembre 2016

- Objet et résultats

La thèse porte sur le résumé automatique de vidéos, avec une attention particulière au résumé automatique de séries télévisées par composition d'extraits de la vidéo source.

Par opposition aux séries TV classiques, les séries modernes les plus populaires tendent à développer des intrigues continues d'un épisode à l'autre. Néanmoins, cette continuité narrative entre frontalement en conflit avec les conditions usuelles du visionnage moderne : les modes de visionnage issus de l'essor d'internet (streaming, téléchargement) offrent au spectateur l'opportunité de se libérer du rythme de diffusion officiel ; sitôt disponibles, les nouvelles saisons des séries TV sont généralement visionnées sur une période de temps relativement brève : il s'écoule donc inévitablement plusieurs mois entre le visionnage de deux saisons consécutives, et le public a généralement besoin de se ré-immérer, tant cognitivement qu'affectivement, dans l'univers de la série avant d'en découvrir la suite.

Une telle situation donne au résumé automatique de films de fiction une opportunité applicative inédite, en lui procurant des cas d'usage à la fois réalistes et bien définis. La prise en compte du spectateur et de ses attentes dans la génération automatique des résumés de séries TV conduit donc naturellement au dialogue entre l'informatique et la sociologie de la réception des œuvres par leurs publics, et confère à cette thèse l'aspect interdisciplinaire propre aux thèses financées par la Fédération de Recherche *Agorantic*.

Les principales contributions de ce travail sont les suivantes :

- Modélisation originale de l'intrigue par le réseau dynamique des personnages en interaction : l'objectif de couverture sémantique du contenu de l'intrigue par le résumé est ainsi atteint par la construction d'une représentation intermédiaire aisément calculable d'un point de vue informatique.
- Utilisation de descripteurs de saillance émotionnelle issus de la grammaire cinématographique pour obtenir l'effet de ré-immersion affective visé.
- Évaluation à une large échelle et dans des conditions réalistes des résumés générés, par diffusion d'un questionnaire auprès de 187 personnes, avec deux objectifs : évaluation de résumés automatiques de *Game of Thrones* quelques semaines avant la diffusion de la nouvelle saison ; explicitation des pratiques des publics en matière de séries TV.

- Apports (et limites) du travail interdisciplinaire

Difficile à définir indépendamment d'un scénario d'utilisation, le résumé de vidéos exige de se centrer sur le spectateur et ses besoins. De ce point de vue, la réflexion sur les usages est première, et la méthodologie expérimentale (enquête par questionnaire) utilisée par la sociologie de la réception des œuvres par leurs publics s'est révélée cruciale : une fois identifiés, les usages du résumé vidéo de séries TV ont orienté le travail de modélisation informatique du flux source (choix de descripteurs), et inspiré un protocole d'évaluation originale, en situation réelle, des résumés automatiquement générés.

La difficulté a peut-être alors été de faire en sorte que soient représentées dans le questionnaire les attentes scientifiques des deux disciplines sans trop en alourdir la passation, mais un compromis a pu être aisément trouvé après discussion avec mes divers co-encadrants.

Peut-être une autre difficulté est-elle de prendre assez tôt la décision de ne pas vouloir apporter de contributions théoriques dans toutes les disciplines représentées, mais seulement dans la discipline dominante : c'est alors plus dans sa méthodologie propre que l'autre discipline peut apporter sa pleine contribution.

– Intérêt par rapport à l'axe identitaire "Culture, Patrimoines, Sociétés numériques"

Le succès des séries TV, œuvres culturelles par excellence, est probablement lié à l'essor d'internet et au renouvellement des pratiques de consommation qu'il implique : la multiplication des canaux de diffusion crée un appel d'air considérable pour la création filmique. De ce point de vue, notre travail sur le résumé de séries TV s'inscrit pleinement dans l'axe identitaire d'Agorantic : la diffusion numérique des séries TV induit de nouvelles pratiques de visionnage propre à motiver la génération automatique des résumés de séries TV.

Situation actuelle et la mission de mon travail

Je suis Docteur Ingénieur Logiciel chez Orkis SA, une startup basée à Aix-en-Provence.

J'y suis responsable du pôle innovation, et travaille sur l'extraction d'informations de contenus multimédias.