

Structuration automatique des flux télévisuels par apprentissage non supervisé des répétitions

Rakia Jaziri^{1,2}, Mustapha Lebbah¹
Younés Bennani¹, Jean-Hugues Chenot²

¹Université Paris 13, LIPN-UMR 7030 - CNRS
99, av. J-B Clément, F-93430 Villetaneuse
{younes.bennani, mustapha.lebbah, rakia.jaziri}@lipn.univ-paris13.fr

²Institut National de l'Audiovisuel
4, av. de l'Europe 94 366 Bry-sur-Marne
{rjaziri, Jhchenot}@ina.fr

1 Introduction

La classification et le typage de flux importants de télévision posent de nombreux problèmes pratiques. Dans le cas des centres d'archivage, l'INA a développé à cet effet, des solutions techniques destinées à détecter des contenus audiovisuels, qui permettent de surveiller en temps réel 10 chaînes de télévision et d'y identifier des segments de programmes d'une durée de quelques secondes similaires à des extraits d'une base de références de plusieurs milliers d'heures. Les résultats de détection de répétitions sont très difficilement interprétables et la prise en compte des aspects temporels (généalogie des contenus) et structurels (motifs de répétitions) n'est pas évidente. Les solutions existantes sont avant tout manuelles ou semi-automatiques (Bennett 1999). Le principal objectif de ce travail est de s'appuyer sur la masse des données récoltées pour en extraire autant d'informations que possible et d'utiliser les outils d'apprentissage pour que l'extraction des motifs puisse renvoyer à une pertinence interprétative.

2 Classification et visualisation des données

Avant l'application des méthodes de classification, un processus de prétraitement des données est appliqué. Cette étape permet de générer des vecteurs. L'objectif est de pouvoir déterminer des types de répétitions qui possèdent une pertinence éditoriale pour la compréhension des flux vidéo. Les modèles des cartes auto-organisatrices sont utilisés pour la visualisation et la classification. La classification de Kohonen est suivie d'une classification automatique des référents de manière à les regrouper en clusters. À l'issue de cette étape, il convient de déterminer quelle signification doit être accordée aux clusters résultants (film, pub, feuilleton...).

Nous avons opté, dans un premier temps, pour une validation visuelle par un expert à partir d'un échantillonnage de séquences audiovisuelles correspondant à chaque cluster. Puis, nous avons aussi validé les résultats en comparant sur des échantillons les informations obtenues par le guide de la télévision disponible sur ce lien:

Structuration automatique des flux télévisuels par apprentissage

<http://inatheque.ina.fr/Ina/ws/dltv/dlweb/general/MultiSearchNames9>. La figure 1.(a) montre les variables qui caractérisent le mieux chaque cluster. En vérifiant à chaque fois le programme auquel appartient le segment, on constate que les clusters sont constitués de données homogènes (figure 1.(b)).

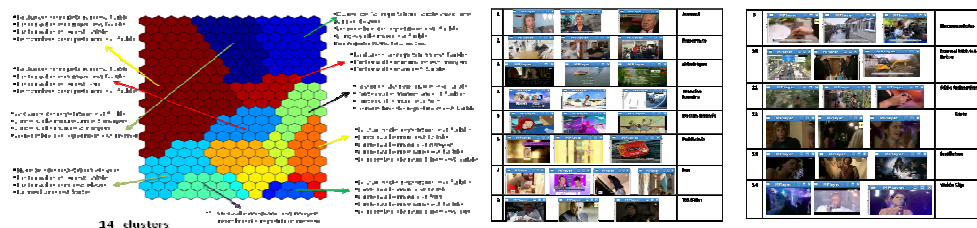


FIG. 1 – (a) La caractérisation des clusters, (b) Croisement des clusters avec les séquences correspondantes

3 Conclusion

L'objectif de notre travail est de montrer la faisabilité d'une solution de classification entièrement automatique de flux de télévision. Bien qu'il reste encore beaucoup de travail à faire, les résultats obtenus sont très encourageants. Différentes voies sont envisagées en ce qui concerne l'amélioration de cette approche, dont l'usage des données de répétitions sur une durée plus longue (des feuilletons, des émissions complètes).

Références

- Bennett, K. P. U. Fayyad, and D. Geiger, "Density-based indexing for approximate nearest-neighbor queries," in Proc. of Conf. on Knowledge Discovery in Data, 1999, pp. 233–243.
- T. Kohonen (1997), Self-Organizing Maps, Springer Series in Information Sciences, Vol. 30, 2nd edition.
- Xiaomeng Wu, Ichiro Ide, Shin'ichi Satoh: PageRank with Text Similarity and Video Near-Duplicate Constraints for News Story Re-ranking. MMM 2010: 533-544

Summary

The accumulation of large volumes of TV flows requires the implementation of exploration tools such as clustering and visualization methods to understand the mechanisms involved in the processes generating the data. The objective of our work is to provide a method for identifying television programs and detecting cross-program links from repetition data extracted from several television streams. We first describe our representation model of the flow of television. In particular, we insist on tools to extract a large set of variables. Then, we show that the use of clustering techniques allow the unsupervised detection several classes of TV contents.