

Prédiction de séries temporelles et applications à l'analyse de séquences vidéo

Rémi Auguste, Ahmed El Ghini, Marius Bilasco, Chabane Djeraba

LIFL/CNRS-UMR 8022 - Université Lille 1

{remi.auguste, ahmed.elghini, marius.bilasco, chabane.djeraba} @lifl.fr

L'utilisation de méthodes de classification avec différentes mesures de similarité a été fortement étudiée pour l'analyse de séquences vidéo (Veeraraghavan et al. (2005), Chen et Zhang (2008) et Simovici et Djeraba (2008)). Ces mesures de similarité appliquées aux séquences d'images sont fondamentales pour la mise en place de systèmes de recherche et d'exploration de vidéos par le contenu visuel. L'approche stochastique basée sur des modèles spatio-temporels relève d'une importance cruciale car elle tient compte des propriétés à la fois spatiales et temporelles de la séquence d'images. Dans cet article, nous étudions le rapprochement entre des séquences vidéo à l'aide de distances entre modèles paramétriques.

Nous utilisons le modèle paramétrique $\text{ARMA}(p, q)$ pour modéliser l'évolution des pixels dans le temps, qui peut être considérée comme une série temporelle. Dans la littérature, plusieurs méthodes, plus ou moins complexes, permettent de mesurer la similitude entre deux modèles. Nous proposons d'utiliser la distance euclidienne mentionnée dans l'article de Sekita et al. (1992), généralisée aux modèles $\text{ARMA}(p, q)$. Nous comparons cette distance avec une distance plus complexe : la distance de Martin (Martin (2000)) utilisée dans l'article de Chen et Zhang (2008). La distance de Martin est moins précise mais plus robuste que la distance euclidienne pour comparer deux modèles. Elle offre en plus une certaine souplesse, ce qui s'avère utile lorsque les vidéos modélisées sont sujettes au bruit ou à des conditions d'éclairage différentes. En revanche, le coût de calcul de la distance euclidienne est moindre ($O(n)$) comparé à celui de la distance de Martin ($O(n^4)$), cette différence peut être particulièrement utile dans le cadre d'une application temps réel. Notre idée est de combiner ces deux distances de façon à obtenir une mesure de similarité qui concilie à la fois vitesse, simplicité et précision. Nous calculons tout d'abord la distance euclidienne, si cette mesure montre que les deux modèles sont semblables alors une mesure plus précise est calculée grâce à la distance de Martin. Fixer les seuils pour considérer deux modèles semblables ou non reste un inconvénient majeur quelque soit la distance utilisée. Une piste possible est de se baser sur la moyenne des paramètres des modèles pour déterminer ces seuils.

Nous allons maintenant nous intéresser à la façon d'exploiter les propositions précédentes dans la conception d'une application de reconnaissance d'objet ou de personne. Voici les différentes étapes que nous proposons pour construire une telle application :

Segmentation de l'image La modélisation complète de l'image n'offre pas de bonnes performances ni de données exploitables, la solution souvent employée est de découper l'image en sous images (segments) qui sont modélisées séparément. Une des difficultés rencontrée est de dimensionner correctement ces sous-images pour obtenir un bon rapport précision/performance.