

Réduction de dimension pour l'analyse de données vidéo

Nicolas Verbeke*, Nicole Vincent*

*Laboratoire CRIP5-SIP, Université René Descartes Paris V,
45 rue des Saints-Pères, 75270 Paris Cedex 06, France
{nicolas.verbeke,nicole.vincent}@math-info.univ-paris5.fr,
<http://www.sip-crip5.org>

Résumé. Les données vidéo ont la particularité d'être très volumineuses alors qu'elles contiennent peu d'information sémantique. Pour les analyser, il faut réduire la quantité d'information dans l'espace de recherche. Les données vidéo sont souvent considérées comme l'ensemble des pixels d'une succession d'images analysées séquentiellement. Dans cet article, nous proposons d'utiliser une analyse en composantes principales (ACP) pour réduire la dimensionnalité des informations sans perdre la nature tridimensionnelle des données initiales. Nous commençons par considérer des sous-séquences, dont le nombre de trames est le nombre de dimensions dans l'espace de représentation. Nous appliquons une ACP pour obtenir un espace de faible dimension où les points similaires sémantiquement sont proches. La sous-séquence est ensuite divisée en blocs tridimensionnels dont on projette l'ellipsoïde d'inertie dans le premier plan factoriel. Nous déduisons enfin le mouvement présent dans les blocs à partir des ellipses ainsi obtenues. Nous présenterons les résultats obtenus pour un problème de vidéosurveillance.

1 Introduction

De nos jours, le stockage de grands volumes de données est devenu possible et abordable. Ainsi, à des problématiques aussi diverses que l'analyse statistique de la fréquentation d'un lieu, la sécurisation de l'accès à des bâtiments, la surveillance de malades épileptiques dans des hôpitaux, ou encore la facturation des véhicules aux péages des autoroutes, les industriels proposent de plus en plus de solutions techniques basées sur l'acquisition numérique de séquences vidéo. Ces données vidéo sont tridimensionnelles (deux dimensions spatiales, et une dimension temporelle). Il s'agit donc d'un volume 2D+T tel que représenté sur la Figure 1(b)¹. Quelle que soit l'application, la première tâche d'un système d'analyse de séquences vidéo est toujours la détection de mouvement, et si possible, la détection (segmentation) des objets mobiles. La difficulté de cette tâche est très variable selon les conditions d'acquisition, la précision et la rapidité du traitement escomptées. Une liste relativement exhaustive des difficultés liées à l'acquisition et au contenu de la scène peut être trouvée dans Toyama et al. (1999). Dans cet article, nous ne nous intéresserons qu'au cas d'une acquisition par caméra fixe.

¹La séquence vidéo utilisée pour créer la Figure 1 a été fournie par le laboratoire CVLAB de l'École Polytechnique Fédérale de Lausanne (Suisse).