

# Sujet de stage

## *Distance de transport et noyaux pour l'apprentissage en vision par ordinateur*

**Laboratoire :** LORIA (UMR 7503, Université de Lorraine), à Nancy.

**Équipe de recherche :** projet INRIA MAGRIT ([magrit.loria.fr](http://magrit.loria.fr)).

**Encadrant :** Frédéric SUR, maître de conférences à l'École des Mines de Nancy.

page web : [www.loria.fr/~sur/](http://www.loria.fr/~sur/)

courrier électronique : [sur@loria.fr](mailto:sur@loria.fr)

**Présentation du domaine.** La vision par ordinateur est un domaine de l'intelligence artificielle. Le but est d'extraire de l'information d'images ou de séquences vidéos afin de répondre automatiquement à des questions comme par exemple « un vélo est-il visible dans cette image particulière ? » ou « quelles sont les photographies de la tour Eiffel parmi cette collection d'images de Paris ? ». L'information est souvent extraite sous forme d'histogrammes représentant les statistiques de différents indicateurs. Comparer des histogrammes est un problème en soit et cela nécessite une distance adaptée, comme illustré sur la figure ci-après.

Une famille de distances adaptée pour la comparaison d'histogrammes est celle des distances de transport. En vision par ordinateur, il s'agit de la distance du terrassier (Earth Mover's distance) [RTG00] : intuitivement, il s'agit de minimiser le déplacement d'un terrassier qui déplacerait des quantités de matière des classes de l'histogramme 1 vers celle de l'histogramme 2. Différentes variations des distances de transport sont obtenues selon le support de l'histogramme (potentiellement multi-dimensionnel) et la métrique entre classes.

En apprentissage statistique, ce ne sont pas directement les distances qui sont utilisées, mais des *noyaux* basés sur ces distances. Ces noyaux représentent des produits scalaires dans des espaces de très grande dimension. En un certain sens, le but est de faciliter la classification des objets en les projetant dans un espace de grande dimension.

**Objectifs du stage.** Le but du stage est d'examiner sous quelles conditions les noyaux construits à partir des distances de transport sont valides pour l'apprentissage statistique (i.e. jouent bien le rôle d'un produit scalaire). Il s'agit d'un problème ouvert (cf [ZMLS07]). Le stage consistera d'abord en une étude bibliographique succincte sur les noyaux en apprentissage et sur les distances de transport [CG97]. Ensuite on étudiera quelques cas particuliers et néanmoins intéressants en pratique.

**Compétences souhaitées.** L'étudiant-e devra avoir de bonnes connaissances générales de niveau Licence en mathématiques (optimisation, espaces de Hilbert) et informatique (algorithmique) et un certain goût pour ces domaines. Aucune connaissance en vision par ordinateur ou en théorie de

hb

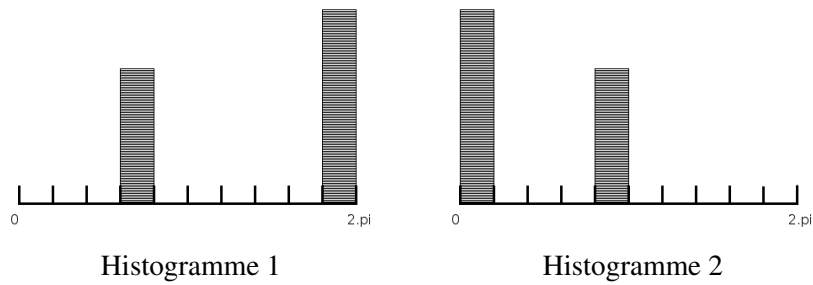


FIG. 1 – Comparaison d’histogrammes. Il s’agit ici d’histogrammes d’orientation d’un vecteur, mesurée sur  $[0, 2\pi]$  et quantifiée sur 10 classes (en abscisse). Les deux histogrammes sont proches dans le sens où une erreur de mesure a décalé les valeurs de l’histogramme 2 de  $2\pi/10$  par rapport à l’histogramme 1. Néanmoins, une distance entre histogrammes prenant uniquement en compte les différences d’effectifs entre classes (comme les distances classiques  $L^1$ ,  $L^2$ ,  $\chi^2$ ) sera grande, et donc non adaptée au problème.

l’apprentissage n’est requise. Au contraire, le stage peut être vu comme une première familiarisation avec ces domaines de recherche.

## Références

- [CG97] S. Cohen and L. Guibas. The Earth Mover’s distance : Lower bounds and invariance under translation. Technical Report CS-TR-97-1597, Stanford University, Stanford, CA, USA, 1997.
- [RTG00] Y. Rubner, C. Tomasi, and L.J. Guibas. The Earth Mover’s Distance as a metric for image retrieval. *International Journal of Computer Vision*, 40(2) :99–121, 2000.
- [ZMLS07] J. Zhang, M. Marszalek, S. Lazebnik, and C. Schmid. Local features and kernels for classification of texture and object categories : A comprehensive study. *International Journal of Computer Vision*, 73(2) :213–238, 2007.