

Traitement de l'information et vision artificielle

(cours aussi appelé « **Mathématiques pour l'image** »)
ENPC – département IMI – VITR1

Responsable : Renaud MARLET (renaud.marlet@enpc.fr)

[03/10/2013]

Attention, le cours évolue par rapport à l'année 2012-2013

Objectif

Le but de ce cours est de donner des outils mathématiques et informatiques permettant d'aborder avec discernement un certain nombre de problèmes qui apparaissent dans le **traitement de l'information et du signal multivarié** (c.-à-d. multidimensionnel), notamment dans ses applications en **traitement d'image** et en **vision par ordinateur**. L'ambition est de transmettre des **fondements mathématiques** et de les mettre en œuvre sur des **applications concrètes** en guise de motivation.

2D, 3D, vidéo..., l'image est partout : comprenez-la, maîtrisez-la.

Contenu

Ce cours est divisé en deux grandes parties. Il étudie tout d'abord la problématique du **traitement du signal multivarié** : représentation, débruitage, compression, analyse, défloutage, etc. Il aborde ensuite des questions d'optimisation discrète qui ont vu leur essor notamment grâce à leur succès sur des problèmes de **vision artificielle** : segmentation, détection et reconnaissance d'objet, etc.

De fait, le cours est principalement illustré par des problèmes de traitement d'image et de vision artificielle. Les concepts et techniques présentés sont néanmoins les concepts fondamentaux pour le traitement et l'analyse automatique de la plupart des signaux multivariés, qu'il s'agisse d'images médicales (imagerie par résonance magnétique, imagerie cérébrale (MEG, EEG, IRMf), scanner, échographie, imagerie microscopique), de données audio et vidéo, ou encore d'images satellitaires, hyperspectrales, 3D, radar, sonar, thermiques, sismiques, etc.

Lien avec d'autres formations

Les questions abordées dans ce cours seront l'occasion d'illustrer ou de traiter sous un nouvel angle des notions enseignées par ailleurs aux Ponts, qui concernent notamment l'optimisation continue, l'optimisation discrète, le traitement du signal, les statistiques, l'apprentissage automatique (statistique), et les processus aléatoires.

Ce cours constitue une excellente **préparation** pour les élèves qui suivront en dernière année le **Master Vision et Apprentissage (MVA)**. Pour les X-FCI qui suivent MVA la même année, il constitue également un très bon complément.

Prérequis

Le cours est accessible avec les connaissances mathématiques acquises en première année ou précédemment. La plupart des notions de base utiles pour le cours sont redéfinies dans le cours même.

Les TP sont construits de manière à ne requérir qu'un niveau de programmation modeste, pas supérieur à celui des élèves de fin de première année des Ponts. Ils sont pour la plupart faits en Matlab (logiciel de calcul numérique dont le langage est introduit ou rappelé dans le cours), ce qui réduit aussi la charge de programmation. Les TP laissent cependant suffisamment de latitude aux élèves intéressés par l'informatique pour pousser leurs développements.

Plan du cours

Le cours est constitué de 12 séances, plus une séance d'évaluation. Les séances sont décrites ci-dessous. L'enseignant pour la séance (voir la section « Enseignants ») est indiqué par ses initiales entre crochets. Voir aussi la section « Langue du cours ».

A) Signaux et information (cours en français)

1. [PM] Les bases du traitement d'image : pixels, changement de contraste, convolution, débruitage simple, détection de bords...
2. [GO] Théorie de Fourier. Les images et de nombreuses autres formes de signaux sont obtenues par échantillonnage discret d'un signal continu. La théorie de Fourier permet d'expliquer les distorsions induites dans le signal par cet échantillonnage, comme le phénomène de Gibbs et l'aliasing, ou de les éviter comme le théorème d'échantillonnage de Shannon-Nyquist.
3. [GO] Réduction de dimension. Les méthodes de réduction de dimension sont fondamentales pour pouvoir encoder, représenter, débruiter et analyser les signaux multivariés de façon adaptative. Nous verrons l'analyse en composantes principales (ACP), aussi appelée transformée de Karhunen-Loève, qui est appliquée très largement en traitement de données ou traitement des signaux.
4. [GO] Théorie de l'information. La théorie de l'information de Shannon est une théorie riche et élégante qui se base sur une mesure de la quantité d'information contenue dans un signal. Elle offre à la fois un fondement à la théorie de la compression de signaux et à la théorie de la transmission de signaux à travers un canal bruité. Mais ses concepts sont tout aussi fondamentaux pour les domaines de l'apprentissage automatique et des statistiques qui ont précisément pour but d'extraire de l'information des données. Ce cours d'introduction sera plutôt orienté vers la problématique de la compression.
5. [GO] Théorie des ondelettes 1. Pour de nombreux types de signaux, il est crucial d'obtenir une représentation compacte qui soit à la fois locale en temps et/ou en espace d'une part et locale en fréquence d'autre part, ce qui est a priori difficile du fait du principe d'incertitude d'Heisenberg. La théorie des ondelettes permet de construire des représentations des signaux qui répondent à ces desiderata avec des applications importantes en compression et débruitage.
6. [GO] Théorie des ondelettes 2.

B) Techniques d'optimisation discrète pour la vision par ordinateur (cours en anglais)

7. [NK] Discrete optimization in computer vision: introduction/basic concepts, Markov Random Fields (MRFs).
8. [NK] Dynamic programming & belief propagation, message-passing-based optimization algorithms.
9. [NK] Max-flow, min-cut, and submodularity.
10. [NK] Graph-cut based discrete optimization methods.
11. [NK] Segmentation & clustering.
12. [NK] Structured-output learning & max-margin parameter estimation for MRFs

C) Validation

13. Examen sur table.

Organisation

Le cours est constitué de 12 séances, plus une séance d'évaluation. Chaque séance de cours dure 3 heures. Certaines séances comportent une première moitié **théorique** (cours magistral de 1h30), suivie d'un **travail pratique** (TP de 1h30). D'autres séances sont entièrement théoriques (3h), mais peuvent comporter malgré tout une partie pratique sous forme de « devoir à la maison ».

Les travaux pratiques et devoirs à la maison sont calibrés pour ancrer les connaissances présentées en séance sans accroître démesurément la charge de travail pour le cours.

Travaux pratiques

Les TP se font pour la plupart en **Matlab** (logiciel de calcul numérique). Les enseignants fournissent pour les TP un canevas ou des éléments de bibliothèque de programmes, plus ou moins détaillés, permettant de se focaliser sur le point de cours traité.

Les TP ont vocation à être réalisés de manière individuelle mais peuvent toutefois être faits en binôme, en particulier si le nombre de postes de travail en salle machine est insuffisant.

Devoirs à la maison

Certains cours s'accompagnent d'un devoir à la maison (DM). Certains de ces DM sont à rendre, généralement dans la semaine qui suit le cours, et sont **notés** (voir « Validation du cours »). Ces DM doivent être réalisés de manière **individuelle**.

Un DM peut être constitué de tout ou partie d'un TP qui, s'il n'est pas fini lors de la séance, est à terminer à la maison. En ce cas, si le TP a été commencé en binôme, il doit être terminé de manière individuelle.

Validation

La validation du cours fait appel pour une large part à une forme de **contrôle continu**. Les notes des devoirs à la maison, y compris les fins de TP, comptent pour 60% de la note finale. Les DM et TP à rendre sont donc indispensables pour avoir la moyenne. Il est fortement recommandé de les rendre tous. Ils sont « accessibles » et permettent souvent de rattraper une mauvaise performance à l'examen.

La 13^e et dernière séance consiste en un **examen de 3h sur papier** (sans machine) qui compte pour 40% de la note finale. Bien que cet devoir sur table ne nécessite pas de programmer un ordinateur, il peut comporter des parties plus informatiques qui demandent l'écriture d'un algorithme. Les documents sont autorisés. Ils pourront être consultés sur ordinateur portable personnel pour éviter d'être imprimés, mais le wifi devra être désactivé : l'accès à Internet (y compris mail, chat, etc.) sera interdit.

Il n'y a pas de projet explicitement associé à ce cours, mais les élèves intéressés par cette thématique ont la possibilité d'y consacrer leur projet IMI. Ils ont intérêt pour cela à en discuter avec les enseignants du cours, qui pourront les conseiller.

Travail individuel

La discussion entre élèves autour d'un TP ou d'un DM n'est pas interdite, au contraire, mais le rendu doit cependant montrer par sa **formulation personnelle** que le problème est maîtrisé par l'élève, et qu'il y a eu travail individuel.

Il est rappelé que **le plagiat est très facile à identifier** sur le rendu d'un programme et que tout plagiat est **sanctionné**. Dans le cas où un TP est commencé à deux au cours d'une séance, il est donc important (1) de le terminer seul à la maison, (2) de rédiger seul le compte rendu, et (3) d'indiquer en le rendant avec qui le TP a été commencé.

Enseignants

Les enseignants pour ce cours sont :

- [PM] Pascal MONASSE (chercheur au LIGM-IMAGINE),
- [GO] Guillaume OBOZINSKI (chercheur au LIGM-IMAGINE),
- [NK] Nikos KOMODAKIS (chercheur au LIGM-IMAGINE).

Le responsable du cours est Renaud MARLET (chercheur au LIGM et directeur du projet IMAGINE).

Langues du cours

Dans ce domaine relativement jeune dont certaines parties touchent encore des problèmes de recherche, une très large part de la bibliographie est en anglais et les termes anglo-saxons prédominent. Durant leur cours, les enseignants donnent les clés des deux terminologies, anglaise et française (lorsqu'elle existe).

Les séances enseignées par Pascal Monasse [PM] et Guillaume Obozinski [GO] sont en **français**. Celles de Nikos Komodakis [NK] sont en **anglais**.

Les devoirs (DM et TP) ainsi que l'examen peuvent indifféremment être rendus en français ou en anglais.

Support de cours

Il n'y a pas de photocopié pour ce cours, mais des notes de cours et des transparents sont disponibles en ligne. Des livres sont également disponibles à la bibliothèque.

Les livres de cours et éléments de bibliographie principaux sont les suivants :

- Mallat, Stéphane. *A Wavelet Tour of Signal Processing*, 3rd ed., Academic Press, dec. 2008.
- Szeliski, Richard. *Computer Vision: Algorithms and Applications*, Springer, 2011.