

Traitement et modélisation pour l'image (TMI)

Introduction

Marie-Odile Berger

<http://members.loria.fr/moberger>

September 23, 2024

Présentation de TMI

Un cours plus spécifiquement dédié à l'image avec beaucoup de techniques applicables à la reconnaissance des formes en général

- ▶ Techniques de modélisation utilisées en reconnaissance de formes (MRF, 24h). Intervenants: Marie-Odile Berger et Erwan Kerrien

objectifs

Etudier différents outils géométriques, statistiques et numériques qui forment la base des méthodes de reconnaissance des formes et interviennent dans la construction effective de modèles à partir d'une collection de données ou de mesures

- ▶ Traitement et d'analyse d'images (TAI, 24h) . Intervenants: Isabelle Debled et Phuc Ngo

objectifs

Etude des méthodes de segmentation d'images et des modèles d'objets associés aux images numériques : leurs représentations (structures discrètes) et leurs exploitation (filtrage).

Qu'est ce que la reconnaissance des formes(RF)?

Objectif: Doter les machines des capacités de l'homme à reconnaître et analyser des **formes** : caractères, objets, sons, signes, signaux temporels...

Quelques définitions historiques:

- ▶ "Pattern recognition, a field concerned with machine recognition of meaningful regularities in noisy and complex environments" (Duda et Hart 1973, extrait de leur livre "Pattern recognition" [2])
- ▶ "The field of pattern recognition is concerned with the automatic discovery of regularities in data through the use of computer algorithms and with the use of these regularities to take actions **such as** classifying the data into different categories." (Pattern Recognition and machine learning, C. Bishop 2007 [1])

Un petit coup d'oeil sur le contenu

- COVER
- Preface
- Mathematical notation
- Contents
- ▶ 1. Introduction
- ▶ 2. Probability Distributions
- ▶ 3. Linear Models for Regression
- ▶ 4. Linear Models for Classification
- ▶ 5. Neural Networks
- ▶ 6. Kernel Methods
- ▶ 7. Sparse Kernel Machines
- ▶ 8. Graphical Models
- ▶ 9. Mixture Models and EM
- ▶ 10. Approximate Inference
- ▶ 11. Sampling Methods
- ▶ 12. Continuous Latent Variables
- ▶ 13. Sequential Data
- ▶ 14. Combining Models
- Appendix A. Data Sets
- Appendix B. Probability Distributions
- Appendix C. Properties of Matrices
- Appendix D. Calculus of Variations
- Appendix E. Lagrange Multipliers
- References
- Index



PATTERN RECOGNITION AND MACHINE LEARNING CHRISTOPHER M. BISHOP

Un petit coup d'oeil sur le contenu

- COVER
- Preface
- Mathematical notation
- Contents
- ▶ 1. Introduction
- ▶ 2. Probability Distributions
- ▶ 3. Linear Models for Regression
- ▶ 4. Linear Models for Classification
- ▶ 5. Neural Networks
- ▶ 6. Kernel Methods
- ▶ 7. Sparse Kernel Machines
- ▶ 8. Graphical Models
- ▶ 9. Mixture Models and EM
- ▶ 10. Approximate Inference
- ▶ 11. Sampling Methods
- ▶ 12. Continuous Latent Variables
- ▶ 13. Sequential Data
- ▶ **14. Combining Models**
- Appendix A. Data Sets
- Appendix B. Probability Distributions
- Appendix C. Properties of Matrices
- Appendix D. Calculus of Variations
- Appendix E. Lagrange Multipliers
- References
- Index



PATTERN RECOGNITION AND MACHINE LEARNING CHRISTOPHER M. BISHOP

Un domaine à l'interface de
l'informatique, des
mathématiques et de
l'apprentissage !

En quoi la RF est elle si complexe?

L'humain sait reconnaître

- ▶ des visages
- ▶ des sons
- ▶ des formes

et ceci **indépendamment**

- ▶ du point de vue sous lesquels on les observe
- ▶ des conditions d'observation
- ▶ de leur variabilité

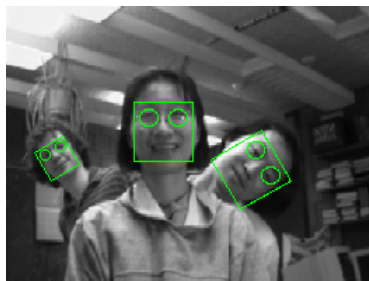
Obtenir ces invariances est un problème majeur de la RF!

Le domaine est en forte progression depuis l'émergence de l'apprentissage et des réseaux convolutionnels.

- ▶ c'est par exemple
 - ▶ Identifier la présence d'objets dans une image
 - ▶ localiser l'objet dans l'image
 - ▶ le segmenter finement
- ▶ mais aussi...
 - ▶ identifier des structures (souvent paramétriques) dans des données (images, nuages de points, mesures industrielles...)
 - ▶ estimer l'état d'un modèle à partir d'observations directes ou indirectes
 - ▶ construire des modèles de formes
 - ▶ valider l'appartenance d'une donnée à un modèle
 - ▶ ...

Quelques exemples déjà anciens en images

La reconnaissance de chiffres ou de caractères (classification), la localisation de visages, de piétons:



Exemple plus récent de classification: Segmentation sémantique

but: attribuer à chaque pixel une classe d'appartenance (sol, arbre, voiture,...)

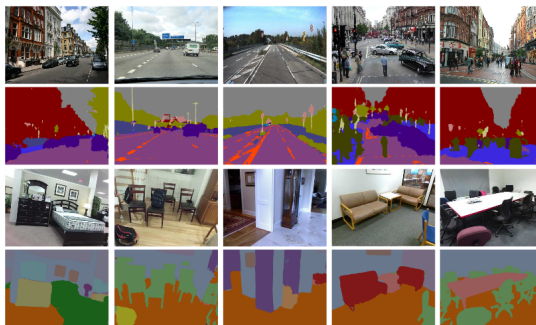
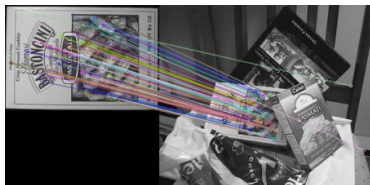


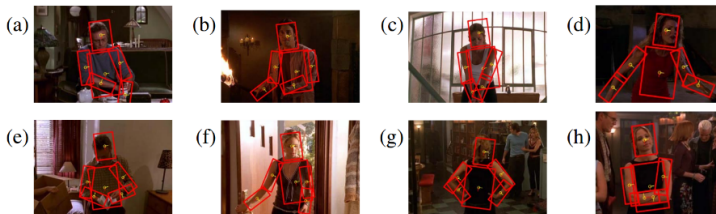
Figure: Segnet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation [Badrinarayanan 2016]

La RF ce n'est pas que de la classification...

exemples: mise en correspondance entre images, reconstruction multi-vues



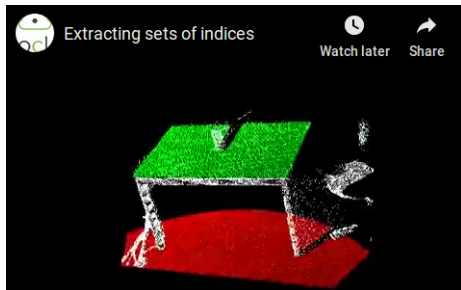
reconstruction de poses humaines



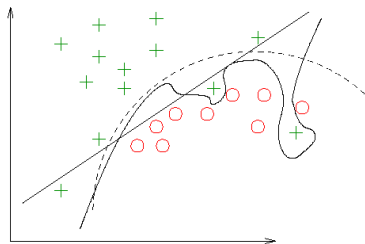
C'est aussi la mise en oeuvre d'algorithmes robustes

Constat: les données contiennent souvent des mesures **aberrantes**

Objectif: construire des algorithmes qui ne sont pas perturbés par la présence de ces données



Extraction robuste de plans dans un nuage de points



Eviter le sur-apprentissage

Joint reconstruction and semantic segmentation [Hane CVPR 2013]

Où comment la résolution conjointe de deux problèmes améliore la reconstruction
→ régularisation spécifique à chaque classe

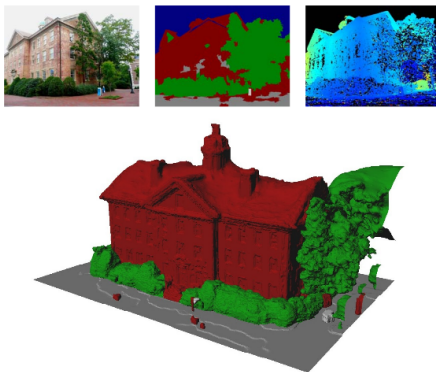


Figure 1: Top: Example of input image, standard image classification result, depthmap. Bottom: Our proposed joint optimization combines class segmentation and geometry resulting in an accurately labeled 3D reconstruction

Quelles méthodes pour la RF?

Au début de la RF, il y avait débat entre **Cognitivistes** et **Comportementalistes**:

- ▶ le cognitiviste s'inspire de nos connaissances sur la perception humaine pour concevoir des systèmes d'inspiration cybernétique.
- ▶ le comportementaliste ne cherche pas à analyser le concept mais essaie de collecter un maximum de données différentes pour en extraire des régularités et des moyens de classification

avec le développement des techniques de machine learning, c'est l'approche travaillant directement sur les données qui a pris l'ascendant.

Exemple de l'évolution de la reconnaissance de visages

- ▶ on est passé d'approches visant à extraire yeux, bouches, lèvres (représentation explicite)...
- ▶ à des approches où on apprend une représentation linéaire d'après une base de données (voir le cours sur l'ACP et le TP sur les eigenfaces [4])...
- ▶ puis à des approches avec des CNNs (DeepFace [3] et [poster](#) associé) où on donne des millions d'exemples et le système construit une représentation interne des visages (exemple base de données d'Oxford, dont la dernière <https://www.robots.ox.ac.uk/vgg/data/voxceleb/>)
- ▶ **Précautions d'usage:** Comment savoir si les données recueillies sont bien représentatives du phénomène? Consulter l'article 2018 de [Sciences et Avenir](#): Intelligence artificielle : la reconnaissance faciale est-elle misogyne et raciste ?

Voir aussi [Une brève histoire de la reconnaissance de visages](#) avec les avancées technologiques et usages récents

Problèmes typiques de la RF

Une série de problèmes typiques de la RF que nous aborderons via les **méthodes conventionnelles** ou les **réseaux convolutionnels**.

- ▶ **Reconnaître** et **Comparer des formes**: réduction de la dimensionalité, mesure de similarité, invariance au contexte...
- ▶ **Modéliser/représenter les formes**: par des méthodes explicites ou des réseaux de neurones
- ▶ **Robustesse**: les résultats des algorithmes doivent (devraient) être insensibles à la présence de données aberrantes (régularisation, estimation robuste)
- ▶ **Estimation**: calculer les paramètres d'un modèle à partir de données (problème d'optimisation car la plupart des problèmes modélisés par le minimum d'une fonction de coût)

Contenu du cours sur la partie modélisation et reconnaissance des formes (erwan/moi)

- ▶ Représenter et comparer les formes: des descripteurs construits à la main aux descripteurs CNN (MOB)
- ▶ Représentation de l'incertain (EK)
- ▶ Modélisation: le cas linéaire (MOB)
 - ▶ Régression linéaire: moindres carrés
 - ▶ l'analyse en composante principales (ACP)
- ▶ Modéliser et regresser avec des réseaux de neurones
 - ▶ MLP, réseaux siamois, réseaux adversoriaux (GAN)
 - ▶ Comment fonctionne un détecteur d'objets?
 - ▶ Quelques idées sur les transformers
- ▶ Estimation et optimisation
 - ▶ Les méthodes classiques d'estimation: moindres carrés, max vraisemblance, ...
 - ▶ Méthodes classiques d'optimisation (descente de gradient) , Algo EM
 - ▶ Estimation robuste

- [1] Christopher M. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. 1st ed. Springer, 2007. ISBN: 0387310738. URL: <http://www.amazon.com/Pattern-Recognition-Learning-Information-Statistics/dp/0387310738%3FSubscriptionId%3D13CT5CVB80YFWJEPWS02%26tag%3Dws%26linkCode%3Dxm2%26camp%3D2025%26creative%3D165953%26creativeASIN%3D0387310738>.
- [2] R. O. Duda and P. E. Hart. *Pattern Classification and Scene Analysis*. Wiley-InterScience, 1973.
- [3] Y. Taigman et al. "DeepFace: Closing the Gap to Human-Level Performance in Face Verification". In: *2014 IEEE Conference on Computer Vision and Pattern Recognition*. June 2014, pp. 1701–1708. DOI: [10.1109/CVPR.2014.220](https://doi.org/10.1109/CVPR.2014.220).
- [4] M. Turk and A. Pentland. "Eigenfaces for Recognition". In: *Journal of Cognitive Neuroscience Volume 3, Number 1 3.1* (1991), pp. 71–86.