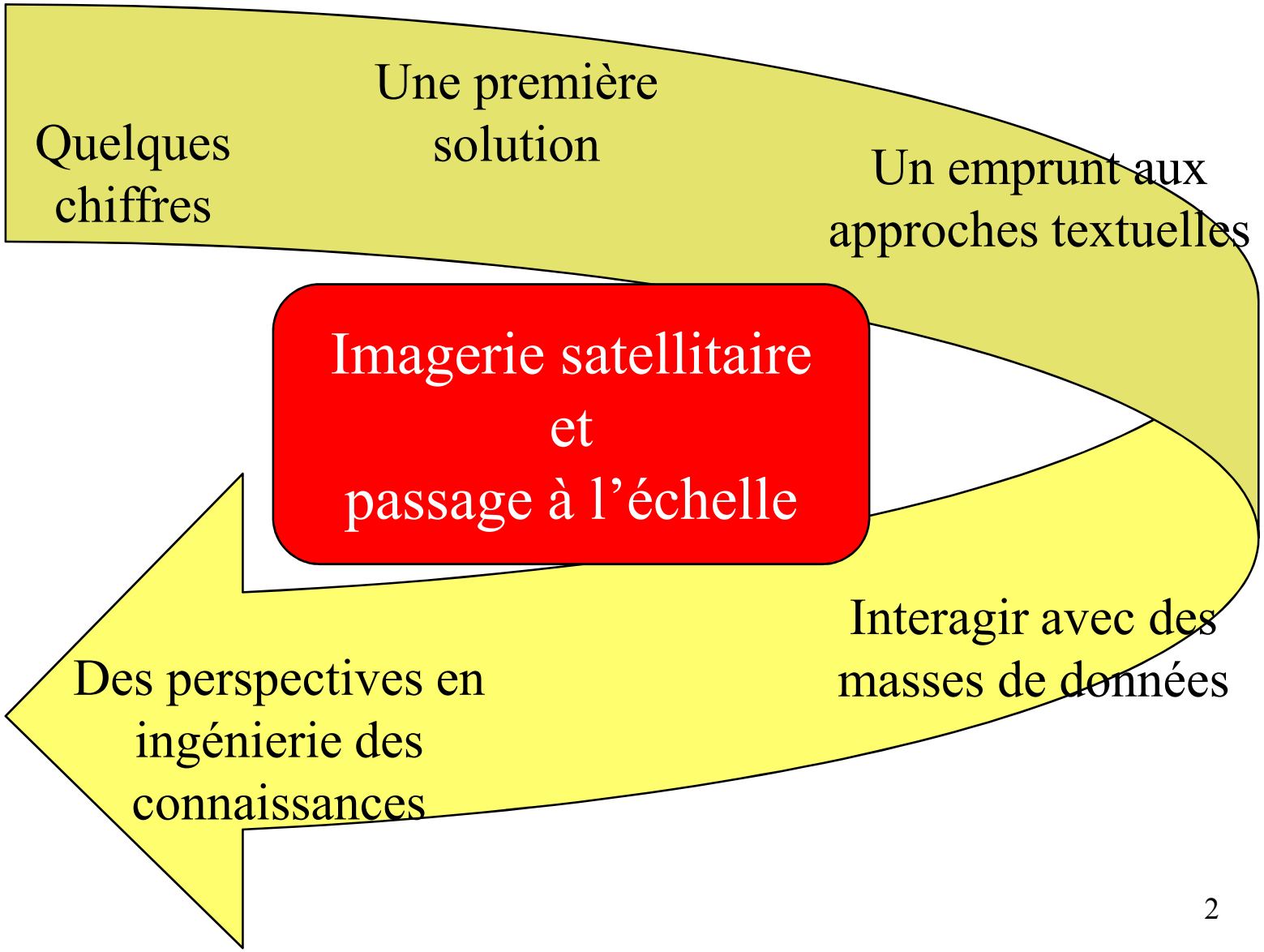
Competence Centre on Information Extraction and Image Understanding for Earth Observation

Imagerie satellitaire et masse de données : le passage à l'échelle est-il possible ?

Marine Campedel
9 juin 2009

Centre de Competence (COC)
CNES/DLR/Télécom ParisTech

PLAN


Quelques chiffres
Une première solution
Un emprunt aux approches textuelles
Imagerie satellitaire et passage à l'échelle
Interagir avec des masses de données
Des perspectives en ingénierie des connaissances

Quelques chiffres

- **Taille d'une seule image**
  - **SPOT5 panchromatique 5m/pixel 12000x12000 pixels 60kmx60km**
  - **1 image = 144 Mo + metadata 2,4 Mo**
  - **Besoin : outils de visualisation/manipulation adaptée**
- **Nombre d'images**
  - **Pléiades (2010) : 450 images par jour à 50cm/pixel**
  - **eCORCE (Blue Planet) : en 2012, toute la planète à 1m, toutes les semaines**
  - **Besoin : catalogues, procédures d'indexations, applications ?**
- **Richesse sémantique des images**
  - **Faible exploitation (moins de 3% selon le CNES)**
  - **Besoin : outils adaptés aux professionnels (photo-interprètes)**

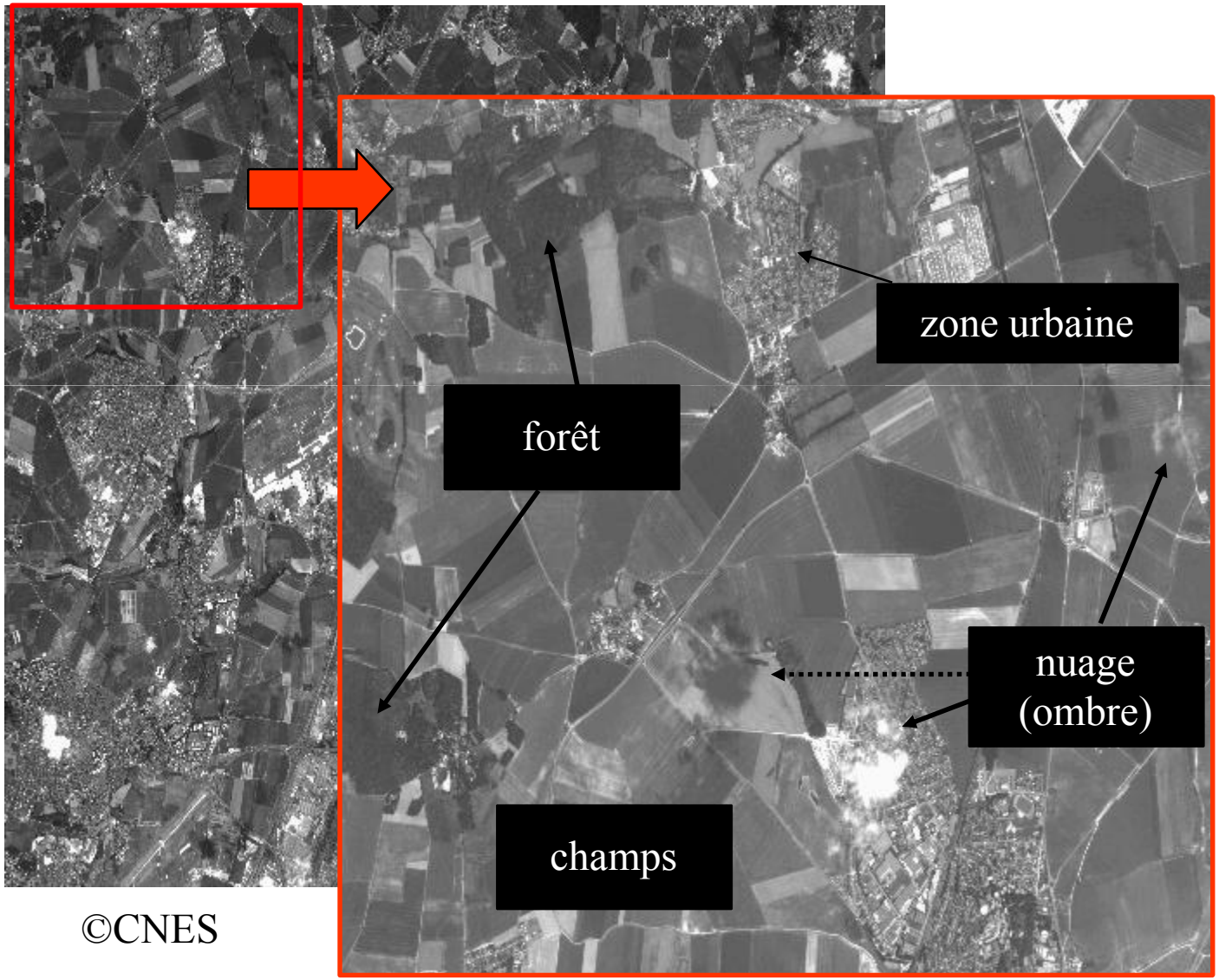
Exemple

Image SPOT 5
Aéroport
CDG1 et CDG2



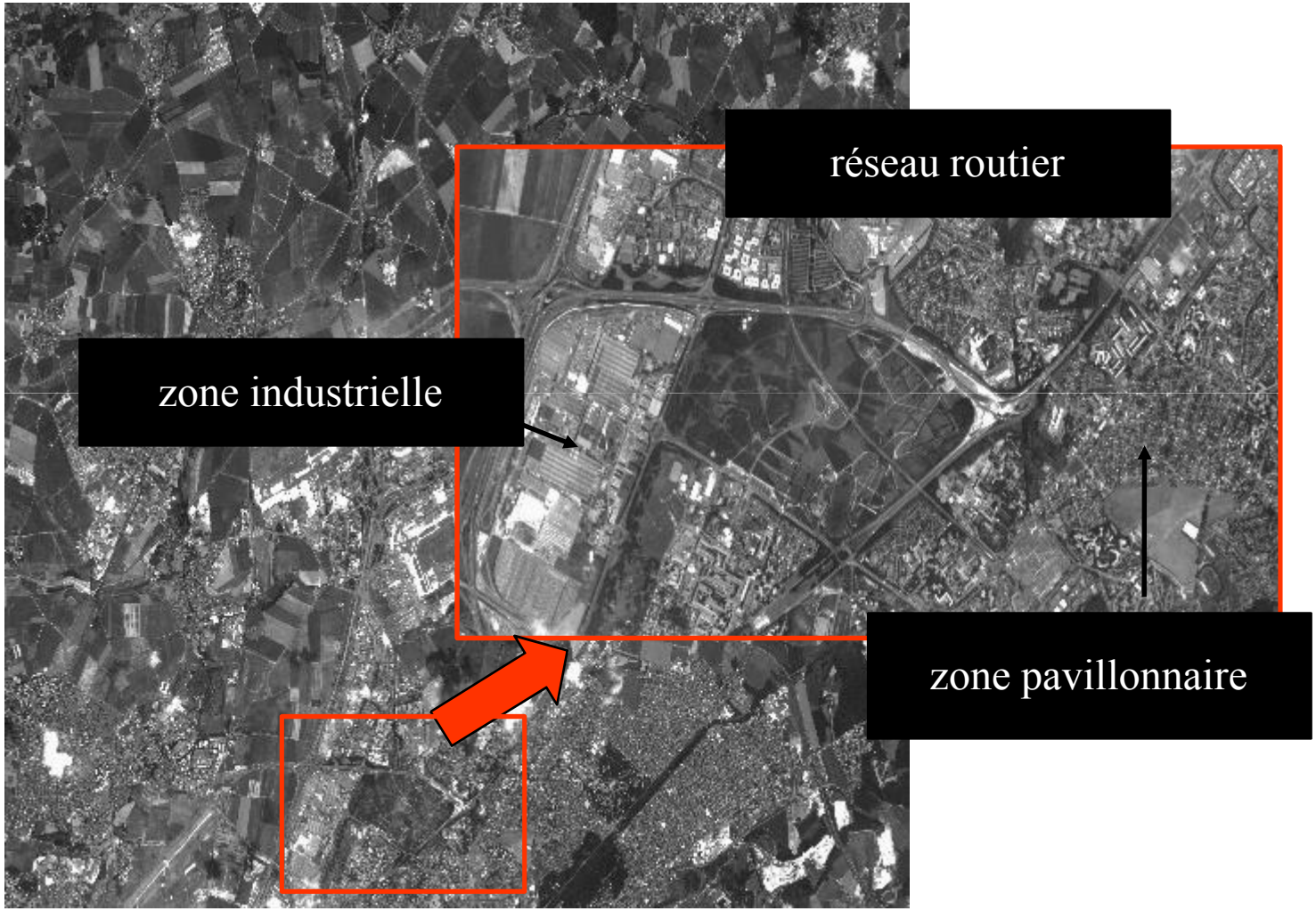
©CNES

Exemple



©CNES

Exemple



©CNES

Quickbird


www.satimagingcorp.com
Copyright (c) DigitalGlobe - All rights reserved

Competence Centre on Information Extraction and Image Understanding for Earth Observation

COC

Indexation

Information
Texturale/Géométrique/Spectrale,
Extraction et sélection d'attributs,
Attributs locaux ou distribués,
Codage des attributs, …

Organisation de l'information

Classification supervisée ou non
(fouille),
Détection d'objets,
Taxonomie/Ontologie,
Analyse sémantique, …

Échelle et résolution

Notion d'échelle caractéristique,
Prédiction des caractérisations
des objets pour une résolution nouvelle,
…

Interaction avec l'utilisateur

Apprentissage des requêtes par
l'exemple,
Boucle de pertinence,
Apprentissage/désapprentissage actif,
Modélisation de l'utilisateur,…

Séquences temporelles

Compression et indexation à l'aide
des mêmes attributs,
Segmentation,
Détection d'événements, …

Masse de données (ingénierie)

Gestion des images et informations
extraites,
Choix d'un SGBD, PLATO
Etude du passage des algorithmes de
traitement à une grande quantité de
données, grilles de calculs, …

Approche « classique »

KIM : knowledge Image Mining (DLR, M. Datcu)

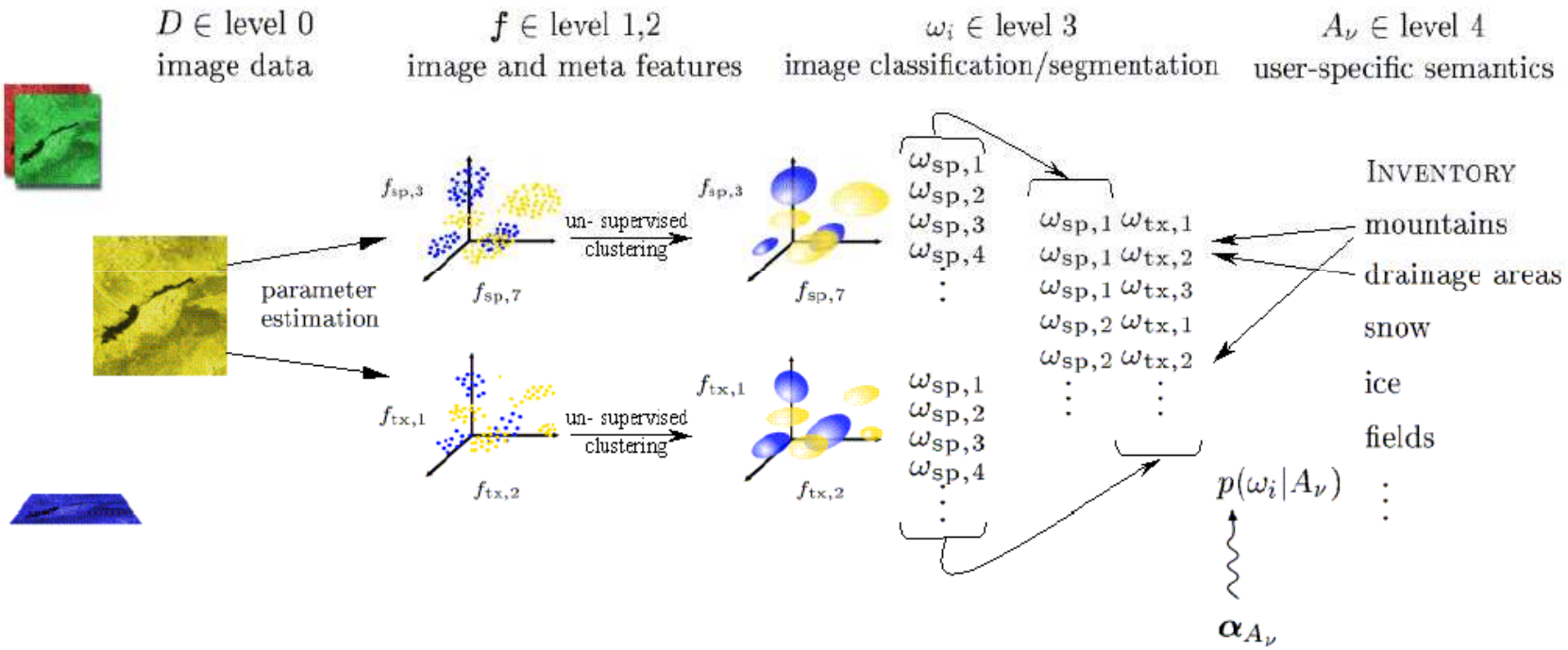


Figure 2. The hierarchy of information used to enable information extraction and knowledge representation adapted to the user needs.

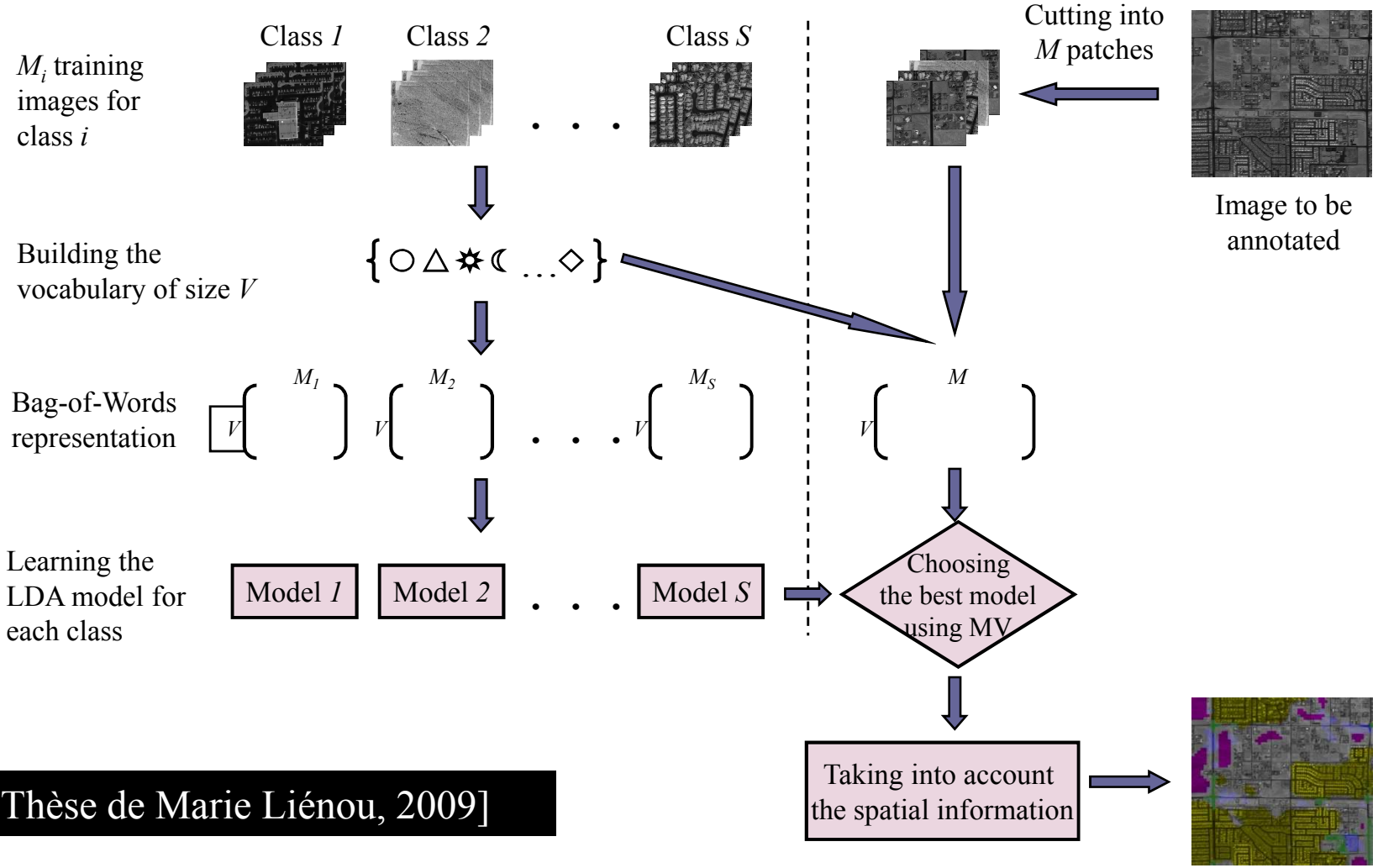
Approche « classique »

- Découpage de l'image
  - Arbitraire (fenêtre glissante)

    SPOT5 : Si imagette de taille 64x64 et pas de 32, génère environ 140 000 positions à considérer dans une image
  - Segmentation
- Extraction d'attributs locaux et sélection d'attributs (supervisée ou non)
  - Texture, couleur (spectral), géométrie (forme)
  - Sélection : compromis redondance / « pertinence »
  - Manipulation de graphes = coûteuse
- Apprentissage sur une base étiquetée (SVM, Bayes)
- Exploitation du classificateur sur une scène d'étude

Les mots visuels + LDA

une approche empruntée à l'indexation textuelle



[Thèse de Marie Liénou, 2009]

Interaction avec les masses de données

- Masses
  - **Les informations brutes (images, métadonnées associées)**
  - **Les informations extraites**

- l'usager n'est pas expert en traitement des images

Solution « informatique »

PLATO

Solution par « apprentissage actif »

PLATO

- Solution « informatique »
  - **Plateforme de fouille de données multimédia**
  - **Base de données + interface web**

- Objectifs
  - **Centraliser les données (images, metadata, outils, données générées, …), sécuriser**
  - **Donner à voir, naviguer**
  - **Expérimenter**
  - **Démontrer, valoriser**

Boucle de pertinence

- Solution par « apprentissage »
  - **Fossé sémantique**
  - **Stratégie proche du travail de l'interprète**

- Principe itératif
  - **annotation manuelle de quelques exemples (régions de l'image étudiée)**
  - **apprentissage binaire +/-**
  - **présentation des résultats**

Ingénierie des connaissances

- Une nouvelle approche
  - **Exploiter les connaissances de la scène imagée**
- Approche mixte
  - **« bottom-up » = approche « sac de mots »**
  - **« top-down » = ontologie et raisonnement logique**
- Projet ANR DAFOE
  - www.dafoe4app.fr
- Application à l'aide à l'annotation (interprétation) des images de télédétection

Conclusion

- **URGENCE !**

- **COC : fin juin 2010 (renouvellement possible de 2010 à 2015)**
- **Beaucoup de problématiques de recherche et d'ingénierie**
- **Beaucoup d'applications d'envergure :**
  - **Risques, cartographie après catastrophe naturelle (EXITER, SAFER)**