INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Project-Team RSIU

Image Understanding and Remote Sensing Applications

Liama - Beijing - Chine

THEME COG

Activity Report

2008

Table of contents

1. Team

Research Scientist

Véronique Prinet [Associate Professor, Head of RSIU]

PhD Student

Régis Behmo [PhD at Ecole Centrale de Paris,from Dec. 06; INRIA-Cordi]
Cyril Cassisa [PhD at Ecole Centrale de Lyon, from Sept. 06; MENESR]
Ting Peng [PhD at CASIA and Univ.Nice, from Sept. 05; MAE/Thales-alenia-space]
Fei Yuan [PhD at CASIA, from July 08]
LiangLiang He [Master at CASIA, from Sept. 06]
Xing Gong [Master at CASIA, from Sept. 08]
DongMin Ma [Master at CASIA, from Sept. 07]
Cui Zhang [Master at CASIA, from Sept. 08]

Visiting Scientist

Jean-Baptiste Bordes [ENST, from Oct. 2007 to March 2008]
Daniel Paulin [Intern, ECP, from June to Aug. 08]
Yves Piriou [Intern, Master2 at Sup'aero, from March to Aug. 08]

2. Overall Objectives

2.1. Overall Objectives

The main research activities of the RSIU project-team are Dynamic Scene Understanding an Object Recognition, with a specific focus on remote sensing related applications.

Image Understanding is the task which consists in *recognising* objects/persons in images, *localising* them, and labelling them with *semantic* attributes. In dynamic scenes, the additional information furnished by motion (of objects or persons) makes it possible the recognition of *actions*. Intrinsically, objects, persons or actions do not live in independent spaces but closely interact one with another: together, they form the *context* of a scene. The ultimous aim of computer vision is to mimic the role of the visual cortex, the occipital area (low level description) in relationship with the temporal lobe (memory, learning, recognsition) and parental lobe (characterisation, attention, spatial mapping, motion). Human vision is also closely related to speech and auditive functions, which together are at the basis of 'awareness' and 'consciousness'.

RSIU develops and validates new mathematical models and computational algorithms for the analysis of deformation, change, or motion observed from multi-temporal data set, particularly from satellite images and geographical data. Our research is motivated by the need to develop robust, automatic and fast information extraction approaches to better understand or survey the evolution of our environment.

We are mainly concerned with *object-category level recognition* from a *closed database*. A contrario to image search from world-wide-web, the range of objects/classes (buildings, roads, cars,...) we are seeking for and that can be found in image databases, is finite, well defined, and relatively small. We are not much interested in exact search of a given object instance, but rather in finding and localising all instances belonging to a generic class. In addition, in remote sensing images, individual objects present the specificity to have relatively well defined shapes (e.g. polygons for man-made structures), and the geometric/spatial relationship between nearby objects is relatively well constraint (because guided by natural laws or human planning), hence facilitating the task of modelling using prior constraints.

RSIU addresses the three main levels of an scene/video sequence understanding system. The lowest level consists in finding visual attributes which describe the input data and more importantly to produce a suitable representation, which embeds local and global characteristics, while being manipulable by the models. The intermediate level is to build appropriate physical or geometrical models to describe regions, shapes or motions. Tasks such as segmentation or motion estimation, are hence performed by minimisation of an energy functional. The highest and most challenging level aims at inferring semantic concepts from the data. It calls for efficient statistical or probabilistic learning techniques, in conjunction with robust data description and representation.
RSIU's main research axes are:

- **Visual features and Data representation**
  Based on the assumption that a scene can be represented by a set of key points and their associated features, we explore new ways to define multi-resolution local appearance, context dependencies, and spatial distribution. We pay particular attention to image sensor-independent and class-dependant features.
- **Energy minimisation modelling for image analysis**
  We are interested in the definition of energy functionals with respect to specific applications —segmentation, motion analysis. Our current research is aimed at defining robust regularisation constraints. Prior knowledge is integrated via the mean of exogenous data (specific prior) or is derived from physical or geometrical properties (generic constraint). Optimisation is achieved via variational, stochastic or graph-based approaches.
- **Scene understanding and recognition**
  Existing methods for visual object/scene inference are based either on the estimation of a conditional probability distribution that models the couple $(x = label, y = observation)$, or on the computation of a decision boundary that best separates positive $(y+)$ from negative $(y-)$ data samples in the feature space. We focus on probabilistic graphical models for their capacities to model local (or semi-local) dependencies and structured data. We also explore kernel based approaches.

2.2. Highlights

- Summer School on *Machine learning, Statistics, and Computer Vision*, co-organised with UCLA and Lotus Hill Institute for Computer Vision (Prof. Zhu SongChun), Hubei (China), July 2008.
- Participation to the *Pascal Visual Object Class Challenge* (classification competition, R. Behmo).
- PhD defense (Nov. 2008), jointly supervised research between RSIU and Ariana. Ting Peng received two best students awards in 2008.

3. Scientific Foundations

3.1. Scientific Foundations

3.1.1. Visual features and Data representation

Computer vision, as leaded by Marr in its early stage, used to consider the *pixel*, as the smallest unit of interest. Bottom-up approaches were then developed based on the principle that pixel-wise low-level characteristics (edges, corners, ...) should first be extracted from the image and used as support to build a high level representation of the scene. Pure bottom-up approaches failed due to their lack of robustness.

It is now widely accepted that an image can be robustly described and represented as a collection of 'atomic structures' (Julesz). These atomic structures are defined from statistical properties of patches (agglomerate of neighbouring pixels) extracted densely or sparsely from the image.

Image description and representation play an important role in the success or failure of a visual scene/object recognition system. Robust local appearance and shape descriptors are usually invariant to geometric or radiometric transformation. A commun approach is then to quantify the descriptors to generate a codebook. Efficient representation will then, either assume independance and interchangeability between local features (bag-of-words representation), or keep, to some degree, the relative spatial information by embeding the features in a graph.

3.1.2. Energy minimisation modeling for image analysis

Many computer vision tasks can be formulated as an energy minimisation problem (restoration, segmentation, stereo, motion estimation, ...) The energy functional is defined as the weighted sum of a *data term* and a *prior term*. The formulation of the data term (equivalent to a *likelihood*) is usually guided by the application ; it describes the observation. The prior term, per definition, characterises the knowledge we have a priori on the sought solution (e.g. constraint of smoothness): its role is to reduce the space of the solutions of the unknown random variable.

When the prior term is a convex function, the solution that minimises the total energy functional can be retrieved exactly from standard minimisation algorithms (variational approaches for continuous variable problems, graph-based or stochastic gradient for discrete problems). However, more and more computer vision problems call for non-convex functions (i.e. with multiple local minima), hence lead to NP-hard optimisation problems. To address this issue, tremendous efforts has been devoted these recent years to the development of optimal approximate solutions.

3.1.3. Scene understanding and recognition

Visual inference is one of the most challenging problems in computer vision and artificial intelligence. Representation, modelling and learning are the main common issues that need to be tackled in designing a system dedicated to visual recognition.

An ideal category-level recognition system would be able to cope with the following challenges : i) wide intra-class variations (large variety of shapes and appearances of objects belonging to a same class) and small inter-class variations (visual similarity of objects of different classes); ii) unconstraint acquisition conditions can lead to occlusion/truncation of the object, illumination variations in the image, etc; iii) the complexity of the model, the large dimension of the features space or the large size of the data set can lead to intractable computations or huge computational cost; iv) the number of categories of objects is estimated to be of $10.^4$; learning those classes individually requires tremendous effort.

Different learning and inference schemes exist for the purpose of recognition. On one hand, statistical classifiers (e.g. SVM, Adaboost, M3N) have demonstrated considerable successes in the recent years. Kernel based approaches in particular owe partly their popularity to their capacity to exploit high dimensional feature space. They are well adapted to image classification problems. Statistical classifier however often require a tidy experimental setting. On the other hand, graphical models offer the advantage to take into account the structural organisation (or spatial layout) of the data by modeling spatial dependancies between neigborood nodes of a graph. In probabilistic graphical models (e.g. MRF, CRF, DRF), the graph structure enables a spatialisation of the labeling. They are often prefered for detection/localisation or segmentation tasks.

4. Application Domains

4.1. Introduction

The applications are driven by needs expressed by governmental agencies or societal factors. We are currently dealing with the following application domains:

4.2. Digital Cartography Updating

Whereas the need for reliable and up-to-date maps is growing and the flow of remote sensing images is increasing, the bottleneck in the production of cartographic data lies in the manual processing applied to the data. Over the last twenty years a large number of works have been carried out aiming to automate map up-dating. But the utilisation of imagery as the sole source of information data has failed. “Updating” is the task to modify an obsolete digital map, based on the information given from external data —a high resolution optical satellite image. The challenge are: i) the profusion of detail, the presence of shadows and other artifacts, are all obstacles that cause the traditional IP and PR approaches to fail; ii) the “prior” knowledge given by the obsolete map is not entirely trustful and may induce the processing in error.

4.3. Multi-temporal Change Analysis

Change analysis from remote sensing data has been studied for more than thirty years by the Image Processing and Pattern Recognition communities. High resolution imagery raised new difficulties: occlusions, projective distortion, detail profusion or the presence of shadows, create “apparent changes” which do not correspond to real changes of the scene, and therefore make the interpretation difficult. Classical approaches based on pixel classification fail because they are not able to differentiate between “real change” and “apparent change”. To tackle this problem, we believe that it is necessary to better define and formulate the notion of change, to handle the problem via a representation of “object” or “structural pattern”, given some prior knowledge.

4.4. Content-based Image Indexation

The task of automatic image labelling and classification is becoming ever more crucial with the exponential increase of the amount of available images and their variety in terms of acquisition sensor and resolution. In particular, satellite image databases undergo a daily augmentation that is of the order of magnitude of the petabyte and it is not possible to assign human operators to the task of image labelling. Therefore, our ability to handle, manage and use large scale image databases will depend on our faculty to devise automatic methods for unsupervised content-based image indexation and retrieval. Image indexation/annotation is the operation that consists in extracting a signature, numerical or textual, that describes its semantic content in a precise and concise way, and therefore enables efficient search in a data base. Few works have been done on indexation from high resolution satellite images, even less on multi-sensor, multi-resolution database images. The challenge is to find a robust unique acquisition-independent representation of the objects/textures of interest.

5. New Results

5.1. Visual Features and Data Representation

Participants: Régis Behmo, Jean-Baptiste Bordes, Véronique Prinet.

5.1.1. Spectral graph and image indexing

Keywords: *Feature graph*, *graph commute times*, *image indexing*, *scene representation*, *spatial layout*, *spectral properties*.

This work is achieved in collaboration with Nikos Paragios [Ecole Centrale de Paris, Laboratoire MAS].

In order to understand visual content, it has to be organised in a structure that can be analysed, from which properties can be extracted. Repetitive and distinctive visual features have shown tremendous progress in the past few years. We build on this progress to structure visual content in feature graphs: this allows us to take into account the spatial layout of visual features as well as the relationships they share with their neighbours.

In high resolution satellite images, the spatial layout needs to be taken into account in order to represent the image content. This can be achieved by considering spectral properties of the feature graph constructed on extracted interest points [5]. To be more precise, we first construct the image feature graph of the image, and the commute time matrix of the collapsed feature graph constitutes our image representation. An appropriate machine learning algorithm is then used to perform image classification tasks. We have shown that this representation can provide a quantitative improvement over the orderless bag of words representation in the classification of Quickbird 0.6m images. Results confort us in the belief that graph structure can contribute greatly to the accuracy of image representation [4].

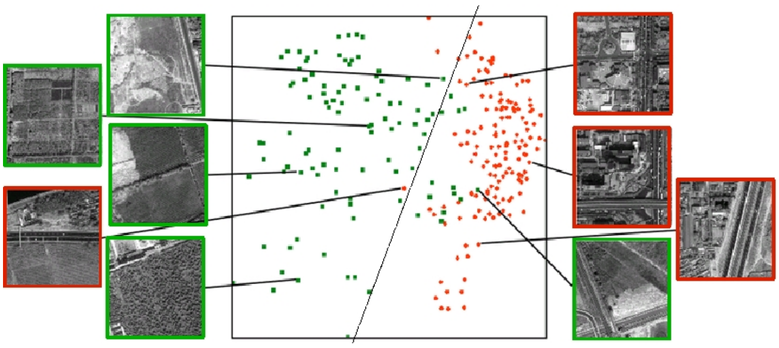


Figure 1. Computing the spectral properties of a graph enables to successfully discriminate, in a low dimensional space, urban class (red points/thumbnails) from rural class (green points/thumbnails).

5.1.2. Shape Descriptor for Image Classification from Mixture Distribution

Keywords: *Shape descriptor*, *bayesian model*, *mixture model*.

This work is achieved in collaboration with ENST.

For its simplicity and efficiency, the bag-of-words representation based on appearance features is widely used in image and text classification. Its drawback is that shape patterns of the image are neglected. We propose a novel image classification approach using a bag-of-words representation of textons while taking into account spatial information. The general idea from the pattern descriptor we introduce is to characterise the spatial distribution of keypoints within a given region by computing the Fourier transform of the function of their occurrences for a given radius. The descriptor is then described by the module of the first n coefficients of the FT, hence making it invariant to rotation transformation. We explain it as a transformation of a generalised correlogram taken at a keypoint.

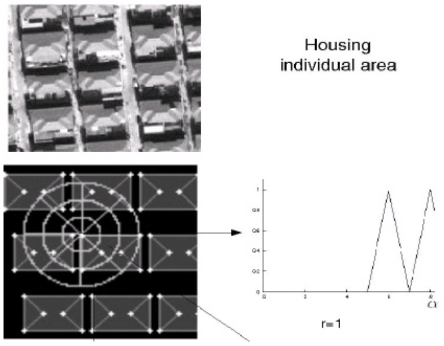
A generative probabilistic modelling of the distribution of textons is then proposed. The parameters of the mixture components are estimated using a EM algorithm. We show that the number of classes in a database can be found automatically and exactly by MDL. This modelling gives very good results for the task of weakly supervised classification in highly textured satellite images [7].

5.2. Energy Minimisation Modelling for Image Analysis

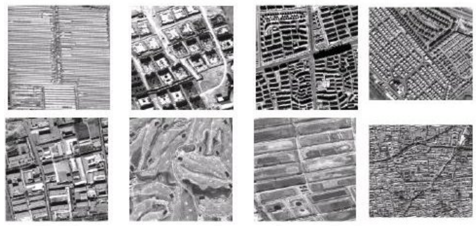
Participants: Cyril Cassisa, Daniel Paulin, Ting Peng, Véronique Prinet.

5.2.1. Shape Prior Modeling for Secondary Roads Extraction

This work is achieved in collaboration with Josiane Zerubia and Ian Jermyn [INRIA-Sophia, Ariana project-team].



(a)



(b)

Figure 2. (a) The shape texton characterises the spatial distribution of points of interest in a close neighbourhood. (b) Illustration of 8 visual classes.

We address the challenging issue of secondary roads detection in dense urban areas from a single QuickBird panchromatic image. The previously developed higher-order active contour (HOAC) prior energy [2], [3], though promising, appears insufficient because it constraints network branch width to be of the same order of magnitude as the branch maximum radius of curvature, thereby providing a poor model of networks with straight narrow branches.

To tackle the above mentioned problems, we proposed a non-linear non-local HOAC prior energy term [11] to increase the magnitude of interactions along the road, and obtained promising results at reduced resolutions. However, the computational cost of this non-linear term is high. Alternatively, we constructed a Euclidean invariant linear non-local HOAC prior energy term [10] as a replacement for the non-linear term. It encourages nearly aligned normal vectors inside the longer range along a nearly straight side to align; while it does not introduce the extra repulsion when the normal vectors of two points are exactly antiparallel. The interaction between pairs of points on the same side of a road has a longer range than that between pairs of points on opposite sides of a road. Based on a stability analysis of a bar with a desired width, we established constraints linking the parameters of the energy function. We explored the possible behaviours of the resulting prior energy as a function of the parameter settings, and showed that as well as separating the interactions between points on the same and opposite sides of a network branch, the new linear model permits the modelling of two widths simultaneously. The analysis also fixes some of the model parameters in terms of network width(s). Moreover, the linear non-local term is more efficient from a computational point of view, and can be applied to images at full resolution. For this reason, in the experiments on road network extraction from VHR satellite images (see Figure 3), the extraction accuracy has been generally improved, with the linear non-local term at full resolution.[1]

5.2.2. Minimisation techniques of HOAC energies

Keywords: *High-order energy*, *minimisation techniques*, *road extraction*.

We propose new ways to minimise the HOAC prior term of an energy function used for road recognition in high resolution satellite images. [1] is using a method based on Fourier transformation to calculate the energy derivative. This approach allows to compute the energy derivative in a very efficient and fast way, but it only works with energy functions which Fourier transform is easily computable. This limits the possible range of energy functions.

We propose instead a direct calculation of energy derivative, using the matrix calculation capabilities of Matlab, and the fact that the interaction functions have a limited range. We are using new energy functions and

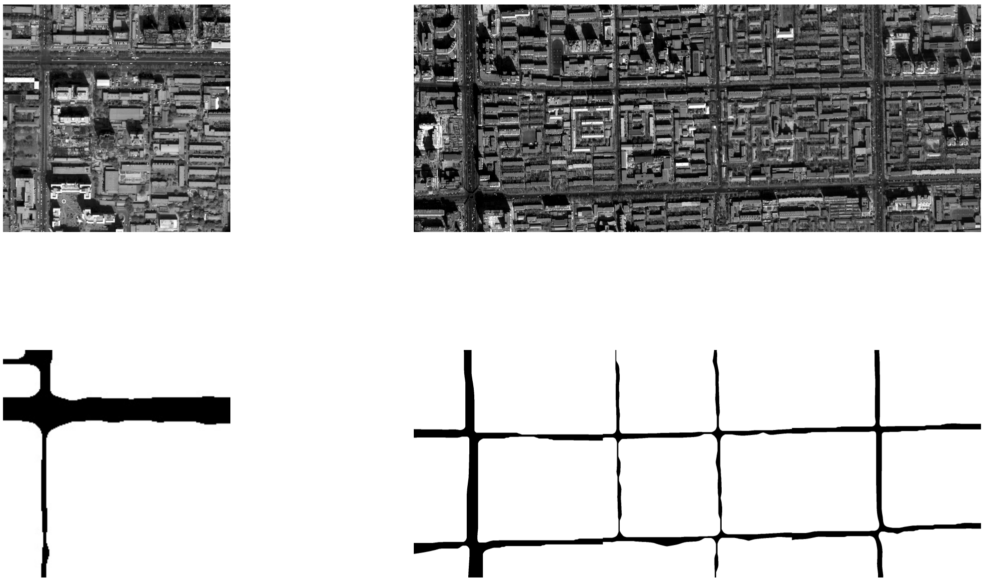


Figure 3. Main and secondary roads extraction using HOAC phase field model with the shape prior linear term. Top: Input QuickBird images (Beijing, 2003); Bottom: Extracted roads.

a more advanced convergence approach (instead of graduate descent with fixed time step), which minimises the energy in fewer model evaluations. In addition to these, we also propose an approximation of the integrals.

Thought this approach enables to use a larger range of energy functions, the computation load is yet very heavy [13].

5.2.3. Modelling turbulent fluid motion

Keywords: *Markov Random Fields*, *fluid motion analysis*, *turbulent flow*.

This work is achieved in collaboration with Liang Shao and Serge Simoens [Ecole Centrale de Lyon, Laboratoire de Mecanique des Fluides et d'Acoustique (LMFA)].
Fluid motion estimation from image sequences is a key issue in a many applications of fluid mechanics or atmospheric physics. In turbulent motion, classical approaches based on optical flow equation (OF) and illumination invariance [8] are insufficient. We propose instead to use the mass transport equation with diffusion —as commonly used in Large Eddy Simulation (LES). In addition to large scales observation variables —i.e., measured concentrations, pixel-wise intensity levels–, the model takes into account the effects of under-scale factors. A Smagorinsky sub-grid model enables to approximate the small scales effects by a statistical turbulent viscosity term. To retrieve the displacement field from fluid motion image sequence, we define a energy based on the sub-grid mass transport equation model as for the data term, and a Tikhonov regularisation. Discrete minimisation of the total energy is achieved by local search. Experimentations are done using fluid motion sequences acquired at LFMA.

5.3. Scene Understanding and Recognition

Participants: Régis Behmo, LiangLiang He, Yves Piriou, Véronique Prinet.

5.3.1. Robust change detection in dense urban areas via SVM classifier

Keywords: *Differential descriptor*, *change analysis*, *multi-temporal images*, *point-matching*.

This work is achieved in collaboration with I. Laptev [Project-team Vista, IRISA].

Change detection from satellite images is a key issue for many applications ranging from digital map updating to disaster rescue. Most classical methods fail because are unable to describe data with a sufficient degree of invariance wrt illumination changes (including shadowings) or geometrical changes, and are sensitive to 'visual noise' (e.g. vegetation effects).

To tackle these problems, we propose a general framework based on robust local features and statistical classifier. The rational behind this is that features computed at points of interest (i.e. differential Histogram of Oriented Gradients, dHOG) provide a description which is invariant to affine illumination change —and can be made invariant to rigid transformation. SVM classifiers are capable to cope with large intra-class/low inter-class variations and defacto are robust to local noise in images. Nevertheless, the shortcoming of local region features classification is to be highly sensitive to projective deformations that can appear between views of high altitude buildings. We thus introduce a point-matching procedure to find correct correspondence of points of interest between two images, before computing dHoG features. We model this task as a energy minimisation problem. The energy function includes appearance term, epipolar constraint term and continuity regularisation term; it is minimised by discrete optimisation. We show experimentally that our approach outperforms significantly the baseline method (Normalised Cross Correlation). When classifying images taken in dense urban areas, we demonstrate that the point-matching procedure is very performant [9].

5.3.2. Style Recognition from Digital Painting

Keywords: *Classification*, *digital art*, *local shape and global layout*, *style category*.

Our aim is to explore the potential of existing image descriptors and recognition techniques to digital paintings classification: to figure out the author of a painting, its style or its era. The issues we address are: i) whether the description and representation commonly used to characterise natural pictures are also suitable to describe artistic paintings. ii) what is a natural/intrinsic representation of styles/eras.

In this aim, we built a database of 5000 images of 70 *painters* from the XIIth to the XXth centuries, covering 17 *styles* in the History of Arts (such as Impressionism or Cubism) and 4 *eras* (from the Middle-Ages to the Modern-Art era). We created data manipulation tools and organise the database in a hierarchical structure, from painters to eras.

We studied the effect of the integration of local shape and global layout/structure descriptors, w.r.t a simple bag-of-words with appearance features. We show that, in a discriminative approach, both local and global shape/layout descriptors enhance significantly the *painter-level* correct classification rate; it reaches 0.65. We observe that classifying paintings in *style-level*, by learning from *painters* samples enable to retrieve similar results as classifying directly from *style* samples [14].

(a) (b)

Figure 4. (a) Illustration of eras-level recognition problem. (b) A style might be represented in a base of paintings.

5.3.3. Detection of Object Visual Parts

Keywords: *Classification*, *detection*, *feature graph*, *kernel*, *probabilistic graphical model*.

This work is achieved in collaboration with Nikos Paragios [Ecole Centrale de Paris, Laboratoire MAS].

The problem of detecting distinctive parts of an object class in an image is formulated as the extraction and the classification of distinctive sub-graphs of the image feature graph. The feature graph of an image connects interest points to their spatial neighbours with a certain level of connectivity. We introduce a directed probabilistic graphical model to compute the class marginal probability by non- loopy belief propagation. The optimal connectivity of the feature graph is learnt with respect to this model for each object class. The belief values of feature graph vertices are used to select candidate sub-graphs as visual object parts, thus avoiding several drawbacks caused by a sliding window approach [12].

Classifying these candidate sub-graphs is then presented as a problem of classifying small sets of features, which we call handfuls of features. A new distance measure between sets of features is introduced and used as a support vector machine kernel to produce detection results. Evaluation on the Graz-02 dataset show how this kernel, which does not require any feature quantisation step, outperforms state of the art bag of words-based classifiers. Finally, detection results are also produced on the PASCAL VOC2008 dataset.[6]

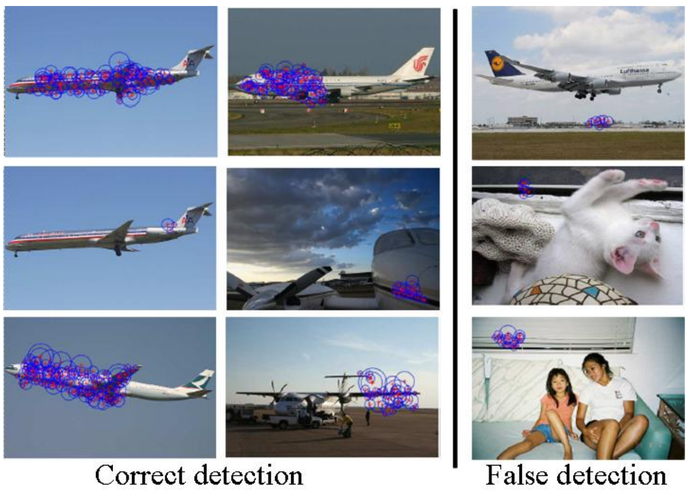


Figure 5. Correct and false detections of airplanes visual parts.

6. Contracts and Grants with Industry

6.1. Thales-Alenia-Space Grant

TAS, via a contract signed with Ariana project at INRIA, supports partially Mlle Peng Ting's stays in France (6 months a year). Ting Peng achieves her PhD in alternation between LIAMA-RSIU and INRIA-Ariana. Ting Peng's PhD research objective is to develop a new model based on high order active contours, for the extraction of road networks.

7. Other Grants and Activities

7.1. Grants and Awards

7.1.1. Awards

- Dai Ruwei Second Prize Scholarship, CASIA [Ting Peng]
- Excellent Student Awards, CASIA [Ting Peng]

7.1.2. Grants

7.1.2.1. French Embassy grant

Since Oct. 2005, Mlle Ting Peng receives a scholarship from the MAE via the French Embassy in China.

7.1.2.2. INRIA Cordi contract

Since Dec. 2006, Régis Behmo receives a scholarship from the INRIA-Cordi program. R. Behmo achieves his PhD in alternation between LIAMA-RSIU and ECP-MAS. His research subject is on object recognition and robust indexation.

7.1.2.3. MENESR scholarship

Since Sept. 2006, Cyril Cassisa receives a doctoral research allocation from the MENESR. C. Cassisa achieves his PhD in alternation between LIAMA-RSIU and ECL-LFMA. His research goal is particle motion estimation from image sequences .

7.2. International activities

7.2.1. PRA

Sino-French *Programe de recherches avancees* on “Analysis of deformation using Markov Random Fields : Application to Remote Sensing Images", in partnership with IRISA-Vista, University of Rennes 1 and INRIA-Ariana, from April 2006 to March 2008.

7.2.2. Asia ICT

Project "Survey of Catastrophes and Observation in Urban Territories" (SCOUT) founded by the ICT-Asia Regional program, 2006-2008. The project is leaded by the International Research Center MICA (CNRS UMI-2954), in partnership with INRIA, IGN, IRD, ITC (Cambodge) and Cargis (Vietnam).

8. Dissemination

8.1. Organisation of Conferences and Workshops

- *Summer School on Machine Learning, Statistics and Computer vision*, at Ezhou, Hubei, China, from June 30 to July 11 (with prof. Zhu SongChun, Departments of Statistics and CS, UCLA).
  Lecturers : Prof. Yuguo Chen (Department of Statistics, University of Illinois, Urbana Champaign), Prof. Ivan Laptev (Inria-Rennes, Vista), Prof. Bin Yu (Department of Statistics, University of California, Berkeley), Prof. Cordelia Schmid (Inria-Alpes, Lear). Prof. Bill Triggs (Laboratoire Jean Kuntzmann, Grenoble), Prof. Chengxiang Zhai (Department of CS, University of Illinois, Urbana Champaign), Prof. Ji Zhu (Department of Statistics, University of Michigan),
- *Seminar on Image processing and Optimization techniques*. November 18th 2008, CASIA (Beijing). Invited speakers: Prof. Patrick Louis Combettes (University Pierre and Marie Curie - Paris 6), Prof. Ian Jermyn (INRIA Sophia Antipolis), Prof. Nikos Paragios (Ecole Centrale de Paris & INRIA Saclay Ile-de-France), Prof. Wang Zengfu (University of Science and Technology of China),

8.2. Visiting Scientists

- Prof. Bill Triggs (CNRS researcher at Laboratoire Jean Kuntzmann, Grenoble, France), ’Scene Segmentation with Latent Topic Markov Field Models’, June 30.

- Prof. Cordelia Schmid (Project-Team Lear, INRIA Rhone-Alpes), 'Learning Human Actions', July 10.
- J.B. Bordes (PhD candidate at ENST) visited RSIU from Sept. 2007 to March 2008.

9. Bibliography

Year Publications

Doctoral Dissertations and Habilitation Theses

[1] T. PENG. *New higher-order active contour models, shape priors, and multiscale analysis: their application to road network extraction from very high resolution satellite images*, Ph. D. Thesis, Graduate University of Chinese Academy of Sciences (China) and University of Nice-Sophia Antipolis (France), November 2008.

Articles in International Peer-Reviewed Journal

[2] T. PENG, I. H. JERMYN, V. PRINET, J. ZERUBIA. *Incorporating generic and specific prior knowledge in a multi-scale phase field model for road extraction from VHR image*, in "IEEE Trans. Geoscience and Remote Sensing - Special Issue: Selected Topics in Applied Earth Observations and Remote Sensing", vol. 1 (2), 2008, p. 139-146.

Articles in National Peer-Reviewed Journal

[3] T. PENG, I. H. JERMYN, V. PRINET, J. ZERUBIA. *A robust framework based on phase field modeling for road network extraction from VHR satellite images*, in "Chinese Journal of Computers (in Chinese)", Under review, 2009.

International Peer-Reviewed Conference/Proceedings

[4] R. BEHMO, N. PARAGIOS, V. PRINET. *An Application of Graph Commute Times to Image Indexing*, in "IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Massachusetts, USA", July 2008.

[5] R. BEHMO, N. PARAGIOS, V. PRINET. *Graph Commute Times for Image Representation*, in "IEEE conference on Computer Vision and Pattern Recognition (CVPR), Alaska, USA", June 2008.

[6] R. BEHMO, N. PARAGIOS, V. PRINET. *Detection of Object Visual Parts in Image Feature Graphs*, in "Conference on Computer Vision and Pattern Recognition (CVPR)", Submitted, 2009.

[7] J.-B. BORDES, V. PRINET. *Mixture Distributions for Weakly Supervised Classification in Remote Sensing Images*, in "British Machine Vision Conference (BMVC), Leeds, UK", Sept. 2008.

[8] C. CASSISA, V. PRINET, L. SHAO, S. SIMOENS, C. LIU. *Optical flow robust estimation in a hybrid multi-resolution MRF Framework*, in "International Conference on Accoustic, Speech and Signal Processing (ICASSP), Nevada, USA", April 2008.

[9] L. L. HE, I. LAPTEV. *Robust Change Detection in Dense Urban Areas via SVM Classifier*, in "URBAN 2009", To appear, 2009.

[10] T. PENG, I. H. JERMYN, V. PRINET, J. ZERUBIA. *An extended phase field higher-order active contour model for networks and its application to road network extraction from VHR satellite image*, in "Proc. European Conference on Computer Vision (ECCV), France", October 2008.

[11] T. PENG, I. H. JERMYN, V. PRINET, J. ZERUBIA. *Extraction of main and secondary roads in VHR images using a higher-order phase field model*, in "Proc. XXI ISPRS Congress, Commission III, Part A, China", July 2008.

Other Publications

[12] M. EVERINGHAM, L. VAN GOOL, C. K. I. WILLIAMS, J. WINN, A. ZISSERMAN. *The PASCAL Visual Object Classes Challenge 2008 (VOC2008) Results*, 2008, http://www.pascal-network.org/challenges/VOC/voc2008/workshop/index.html.

[13] D. PAULIN. *Optimisation techniques in HOAC models*, 3rd year of Engineer School, Ecole centrale Paris, Sept. 2008.

[14] Y. PIRIOU. *Creation of a database of digital art image dedicated to image indexing*, Master2 report, Sup'aero, Sept. 2008.