



GDR 720 ISIS du CNRS

Information, Signal, Image, viSion

Demande de renouvellement

2019 - 2023

Préambule. Ce document intitulé « Demande de renouvellement » décrit les perspectives scientifiques et d'animation du GdR ISIS pour le mandat 2019-2023, l'organisation proposée afin de mettre en œuvre ce projet, ainsi que l'équipe de direction. Il fait suite à un document intitulé « Rapport d'activité », présentant le bilan pour la période 2014-2018.

Table des matières

| | |
|---|-----------|
| Préambule. | 3 |
| 1 Présentation générale | 1 |
| 1.1 Positionnement scientifique du GdR ISIS | 1 |
| 1.2 Évolutions | 1 |
| 2 Proposition pour le GdR ISIS 2019-2023 | 3 |
| 2.1 Les objectifs | 3 |
| 2.2 Fonctionnement interne | 3 |
| 2.2.1 Outils d'animation scientifique | 4 |
| 2.2.2 Le comité de direction | 5 |
| 2.3 Fonctionnement externe | 7 |
| 2.4 Club des partenaires | 7 |
| 2.5 Réseau des doctorants | 8 |
| 2.6 Inter-GdR, international | 9 |
| 3 Présentation par thème | 11 |
| 3.1 Thème A : Méthodes et modèles en TdSI | 11 |
| 3.1.1 Présentation générale du Thème A | 11 |
| 3.1.2 Perspectives du Thème A | 12 |
| 3.1.3 Actions proposées | 18 |
| 3.2 Thème B : Image et Vision | 22 |
| 3.2.1 Présentation générale du Thème B | 22 |
| 3.2.2 Perspectives du Thème B | 23 |
| 3.2.3 Actions proposées | 25 |
| 3.3 Thème C : AAA pour le TdSI | 31 |
| 3.3.1 Présentation générale du Thème C | 31 |
| 3.3.2 Perspectives du Thème C | 31 |
| 3.3.3 Actions proposées | 32 |
| 3.4 Thème D : Télécommunications | 34 |
| 3.4.1 Présentation générale du Thème D | 34 |
| 3.4.2 Perspectives du Thème D | 35 |
| 3.4.3 Actions proposées | 39 |
| 3.5 Thème T : Apprentissage | 42 |
| 3.5.1 Présentation générale du Thème T | 42 |
| 3.5.2 Perspectives du Thème T | 43 |
| 3.5.3 Actions proposées | 46 |

1 Présentation générale

1.1 Positionnement scientifique du GdR ISIS

Le traitement du signal au sens large est une discipline scientifique dont l'objet principal est le développement de théories, méthodes et technologies pour l'acquisition de données, leur analyse et l'extraction des informations utiles qu'elles véhiculent, en vue de leur interprétation et d'une exploitation finale. Au service de tous les grands secteurs d'activité (santé, environnement, énergie, transports, télécommunications, etc.), la discipline a su se renouveler pour mieux s'adapter aux évolutions et révolutions successives dans ces domaines, grâce au fil conducteur qu'elle tisse du support initial de l'information jusqu'à l'application. Central et singulier dans les sciences de l'information, le traitement du signal est donc par nature une discipline d'interface¹.

De l'acquisition des signaux à leur exploitation finale, la séquence des opérations à mettre en œuvre est longue et variée. Le GdR ISIS s'intéresse aux méthodes de traitement de l'Information, du Signal, de l'Image et de la viSion, comprenant la mise en forme des signaux, l'analyse et la modélisation, la compression et le codage, le stockage et la transmission. Le contenu sous-jacent à cette nomenclature pérenne est évidemment réinventé au rythme des innovations dans les sciences et technologies de l'information. Le GdR ISIS a pour vocation de se saisir des changements de paradigme de la discipline afin de maintenir son niveau d'excellence dans l'animation et la formation de sa communauté. Les 5 thèmes qui structureront ses activités lors du prochain contrat sont :

- Thème A : Méthodes et modèles en traitement du signal et de l'image
- Thème B : Image et Vision
- Thème C : Adéquation Algorithme Architecture
- Thème D : Télécommunications : compression, protection, transmission
- Thème Transverse : Apprentissage pour l'analyse du signal et des images

La description scientifique détaillée de chacun de ces thèmes, et la vision de leurs responsables scientifiques sur les évolutions récentes et à venir de la discipline dans chacun de ces onglets, figurent au chapitre 3 accompagnés de leur projet d'animation.

1.2 Évolutions

On a assisté au cours des cinq dernières années à un engouement majeur et persistant pour le traitement des mégadonnées, en écho à une mutation radicale de nos sociétés de l'information, associée à des usages nouveaux, et soutenue par une variété de capteurs et d'infrastructures en réseaux qui produisent, diffusent et stockent des données continuellement. La disponibilité de cette data-masse, en conjonction avec le développement de moyens de calcul d'une puissance sans précédents, favorisés par l'émergence d'un écosystème économique réunissant les plus importants acteurs du secteur des technologies de l'information et de la communication, ont sans aucun doute contribué à la remontée en puissance de l'apprentissage automatique. Fort de ses premiers succès en reconnaissance de la parole, puis en vision par ordinateur avant de gagner de nombreux autres domaines, celui-ci constitue aujourd'hui une force motrice dans les récentes évolutions de

1. P. Flandrin, Le traitement du signal, au cœur de la science et de notre vie quotidienne, 2017.

notre discipline. Il charrie également nombre d'incertitudes et difficultés de mise en œuvre qui constituent autant d'opportunités pour notre communauté.

Le GdR ISIS a établi sa prochaine feuille de route en conservant un socle de connaissances fondamentales qui ont garanti son succès et sa pérennité, tout en cherchant à identifier les verrous et questionnements soulevés par les percées les plus récentes de la discipline qu'il conviendra tôt ou tard de lever. Les principaux sujets nouvellement abordés au cours du prochain contrat sont :

- Thème A : Méthodes et modèles en traitement du signal et de l'image
 - Données massives et grandes dimensions
 - Signal et géométrie : information, variétés et graphes
 - Modèles structurés et inversion
 - Signal, image et apprentissage
- Thème B : Image et Vision
 - Capteurs émergents, co-conception
 - Vision 3D, géométrie et apprentissage
 - Visage, geste et comportement
 - Interaction avec l'humain : qualité et perception
 - Analyse de données massives en imagerie du vivant
 - Apprentissage profond pour l'indexation multimédia
- Thème C : Adéquation Algorithme Architecture
 - Architectures dédiées vs. processeurs multi-cœurs et many-cœurs
 - Méthodes et outils pour le choix de nouvelles plateformes matérielles
 - Capteurs et caméras
- Thème D : Télécommunications : compression, protection, transmission
 - Nouvelles tendances en compression des signaux visuels
 - Sécurité des données multimédia
 - Apprentissage pour la couche physique
 - Transfert et récupération d'énergie
 - Théorie des graphes et décision distribuée dans les réseaux télécoms
 - Architectures Cloud RAN et Fog RAN, virtualisation et distribution

Une liste exhaustive, accompagnée d'une description détaillée, figure au chapitre 3.

Compte tenu du nombre sans cesse croissant de chantiers ouverts qui sont dédiés à l'apprentissage et du très fort intérêt exprimé par le club des partenaires, le GdR ISIS a souhaité lui consacrer un thème spécifique : le Thème T. Les intérêts que le GdR ISIS porte à l'apprentissage sont de deux ordres :

- promouvoir l'utilisation des techniques d'apprentissage les plus récentes dans des tâches de traitement du signal qui y avaient peu recours jusqu'ici ;
- développer et analyser de nouvelles classes de méthodes pour l'apprentissage.

Le Thème T est par nature transverse au sein de l'organisation du GdR, compte tenu de ses multiples interactions possibles avec les quatre autres Thèmes, comme en atteste la liste ci-dessus. En prise permanente avec le domaine concerné, les responsables du Thème T seront impliqués dans l'organisation des journées relatives à l'apprentissage automatique émanant des autres thèmes. Ils seront également force de proposition en organisant des journées propres au Thème T sur des sujets tels que :

- Thème Transverse : Apprentissage pour l'analyse du signal et des images
 - Formulation d'apprentissage et architecture
 - Passage à l'échelle et transfert
 - Compréhension formelle des réseaux profonds
 - Traitement de données multimodales

Le Thème T aura par ailleurs des interactions fortes avec le GDR IA (« Aspects Formels et Algorithmiques de l'Intelligence Artificielle »). Une partie des activités sera co-organisée par les deux structures, impliquant plusieurs actions du Thème T. Des discussions sont en cours à ce sujet.

2 Proposition pour le GdR ISIS 2019-2023

2.1 Les objectifs

Les objectifs du GdR ISIS du CNRS seront pour le mandat 2019-2023 les mêmes que ceux présentés dans le rapport de bilan :

- faire de l'**animation scientifique** pour et par la communauté française en Information, Signal, Image, viSion ;
- **féderer** les recherches académiques de tous les laboratoires, quels que soient leur taille, leurs moyens et leur visibilité, en lien avec les grands organismes et au contact des partenaires industriels ;
- mener des actions de **formation** sur les thématiques Information, Signal, Image et viSion, en particulier en direction des doctorants mais pas uniquement ;
- réaliser des études de **prospective** pour le CNRS ;
- **représenter** et porter les thématiques Information, Signal, Image, viSion en France au sein du CNRS, de l'ANR ou des alliances (Allistène en particulier) avec une vision internationale.

Le GdR ISIS a la volonté de fournir un tremplin d'actions au niveau national, en particulier aux doctorants et aux jeunes chercheurs dans le domaine de l'information, du signal, des images et de la vision. Il a aussi la volonté de rassembler la communauté par un fonctionnement ouvert et consensuel que nous voulons préserver.

2.2 Structure et fonctionnement interne

Pour mettre en œuvre ces objectifs, la structure et le fonctionnement proposés pour le GdR ISIS sont sensiblement ceux du mandat précédent.

2.2.1 Outils d'animation scientifique

Les outils d'animation du GdR ISIS sont les suivants :

- La journée d'animation scientifique
L'élément d'activité principal est l'organisation de journées d'animation scientifique, avec une volonté de programmer conjointement des exposés tutoriels présentés par des chercheurs confirmés de renommée internationale, et des exposés plus pointus proposés par des doctorants. Il est systématiquement proposé aux acteurs industriels du Club des Partenaires de s'associer à chacune de ses journées.
- Les actions
Il s'agit d'une initiative structurante dans les activités du GdR ISIS. Une action est portée par un ou plusieurs animateurs qui s'engagent à assurer une animation scientifique sur une thématique ciblée et identifiée comme prioritaire. Ceci permet d'identifier des lignes d'actions fortes, programmées dans le temps. Les actions ont vocation à être renouvelées à l'occasion du bilan dressé à chaque assemblée générale.
- L'assemblée générale
Organisée lors des demandes de renouvellement du GdR ISIS ainsi qu'à mi-mandat, elle permet de réunir régulièrement l'ensemble de la communauté ISIS. L'objectif principal est d'élaborer ensemble les perspectives de chaque thème pour la période à venir, et de redéfinir éventuellement la structure d'organisation du GdR ISIS afin qu'elle puisse servir les axes prospectifs dégagés. Elle permet en particulier de préparer la demande de renouvellement de manière collégiale.
- Prix de thèse
Pour promouvoir la recherche, renforcer la visibilité et encourager les études doctorales dans le domaine du Signal, de l'Image et de la Vision, le GdR ISIS co-organise conjointement avec le Club EEA² et le GRETSI³ un Prix de Thèse distinguant des travaux particulièrement originaux et novateurs. Ce prix, doté de 1000 euros, est délivré annuellement. Les lauréats sont invités à présenter leurs travaux lors du colloque biennal GRETSI. L'organisation est décrite dans l'appel 2018⁴.
- École d'été
Le GdR ISIS organise, conjointement avec le GRETSI, l'école d'été de Peyresq en Traitement du Signal et des Images. Cette école annuelle créée en 2006 réunit 45 participants durant une semaine. La présence des doctorants est largement majoritaire, et celle des jeunes chercheurs encouragée. Des industriels participent aussi régulièrement à cette école. La programmation est assurée par un directeur scientifique, renouvelé chaque année. Pour l'édition 2018, le sujet est : Signaux, images et science des données⁵.
- Les projets de recherche exploratoire
Le GdR ISIS publie périodiquement des appels à projet, selon les ressources financières dont il dispose. Les aides accordées, d'un montant de 7000 euros environ pour une durée de 2 ans, sont destinées à soutenir des travaux de recherche à caractère très prospectif dans les domaines de l'information, du signal, de l'image et de

2. <http://www.clubeea.org/>

3. <http://www.gretsi.fr/>

4. <http://gretsi.fr/prix-de-these2018/>

5. voir <http://www.gretsi.fr/peyresq18/>

la vision. Seuls les projets amont, en rupture avec les travaux connus dans le champ scientifique du projet, peuvent être retenus. Les projets doivent associer deux laboratoires au minimum, dont au moins une unité mixte de recherche (UMR) du CNRS, membre du GdR ISIS. La sélection des projets est faite sur la qualité scientifique des propositions. Trois catégories de projets sont considérées : projet industriel associant un membre du Club des Partenaires, projet junior n'impliquant que des chercheurs de moins de 35 ans, projet interdisciplinaire associant des laboratoires de, et hors de, la communauté ISIS.

2.2.2 Le comité de direction

Le lieu de décision du GdR ISIS est son comité de direction. Il comprend le directeur et les deux directeurs adjoints, les responsables scientifiques des thèmes, les responsables du réseau des doctorants, le président du Club des Partenaires industriels, la responsable des relations inter-GdRs et internationales, le responsable du système d'information, la responsable administrative et de gestion. Le comité se réunit trois à quatre fois par an, décide des grandes orientations à donner, planifie les journées d'animation scientifique à venir, prépare l'Assemblée Générale, décide des actions pour les doctorants, propose le sujet de l'école d'été annuelle de Peyresq. Ses principales lignes d'action ainsi que sa composition sont présentées et discutées en Assemblée Générale. Le comité de direction est garant de la qualité scientifique des activités du GdR ISIS. Il est aussi garant de la mise en œuvre du projet et décide des priorités en fonction du budget. Il est à l'écoute des propositions d'animation émanant de la communauté.

Le comité de direction proposé pour le mandat 2019-2023, validé à l'Assemblée Générale qui s'est tenue du 14 au 16 novembre 2017 à Sète, est le suivant :

- Directeur :
Cédric Richard
PU61, Laboratoire Lagrange (UMR CNRS 7293), Université Nice – Sophia Antipolis
- Directeurs adjoints :
Yannick Berthoumieu
PU61, IMS (UMR CNRS 5218), INP de Bordeaux, Université de Bordeaux
Jérôme Idier
DR CNRS 07, LS2N (UMR CNRS 6004), Nantes
Sous la supervision de C. Richard, Y. Berthoumieu sera en charge des relations avec le Club des Partenaires représenté par son Président, tandis que J. Idier se consacrera aux relations avec les instances.
- Directeurs scientifiques adjoints, responsables des thèmes scientifiques :
 - Thème A : Méthodes et modèles en traitement du signal et de l'image
Rémy Boyer
MCF61, L2S (UMR CNRS 8506), CentraleSupélec, Gif-sur-Yvette
François Desbouvries
PU, SAMOVAR (UMR 5157), Telecom SudParis, Evry
Nicolas Le Bihan
CR CNRS 07, GIPSA-lab (UMR CNRS 5216), Grenoble

– Thème B : Image et vision

Cédric Demonceaux

PU61, Le2i, VIBOT (ERL CNRS 6000), Université Bourgogne Franche-Comté

François Goudail

PU63, Laboratoire Charles Fabry (UMR CNRS 8501), Institut d'Optique Graduate School

Edoardo Provenzi

PU26, IMB (UMR 5251), Université de Bordeaux

Ruan Su

PU61, LITIS, Normastic (FR CNRS 3638), Université de Rouen Normandie

– Thème C : Adéquation algorithme architecture pour le TdSI

Christophe Jégo

PU61, IMS (UMR CNRS 5218), INP de Bordeaux, Université de Bordeaux

Jean-François Nezan

PU61, IETR (UMR CNRS 6164), INSA de Rennes

– Thème D : Télécommunications – compression, protection, transmission

Chaker Larabi

MCF61, XLIM (UMR CNRS 7252), Université de Poitiers

William Puech

PU27, LIRMM (UMR CNRS 5506), Université de Montpellier

Ghaya Rekaya-Ben Othman

PU, Communications and Electronics Department, Telecom ParisTech

Maxime Guillaud

Huawei Technologies, Paris

– Thème T : Apprentissage pour l'analyse du signal et des images

Christian Wolf

MCF HDR, LIRIS (UMR CNRS 5205), INSA de Lyon

Nicolas Thome

PU, CNAM, Paris

● Président du Club des Partenaires :

Daniel Duclos

SAGEM DS, Argenteuil

● Réseau des doctorants :

Audrey Giremus

MCF61, IMS (UMR CNRS 5218), INP de Bordeaux, Université de Bordeaux

Barbara Nicolas

CR CNRS 07, Creatis (UMR CNRS 5220), Lyon

● Actions inter-GdRs, ouverture internationale :

Nelly Pustelnik

CR CNRS 07, Laboratoire de Physique (UMR CNRS 5672), ENS Lyon

● Responsable du système d'information :

Michel Jordan

Ingénieur de recherche, ETIS (UMR CNRS 8051), ENSEA

- Responsable administrative et gestion :

Alison Desachy

Pôle de soutien aux unités, DR01

Le comité de direction accueille 13 nouveaux membres sur les 24 qui le constituent, soit plus de la moitié de son effectif. Ce roulement garantit le renouvellement des idées tout en assurant la continuité dans la réalisation des missions.

2.3 Fonctionnement externe

La communication du GdR ISIS s'effectue principalement par son site web⁶ où sont portés les annonces de réunions du GdR, les comptes rendus, et tout autre activité de la communauté. La liste de diffusion est utilisée pour annoncer les journées d'animation du GdR ISIS et les annonces des membres du club des partenaires.

Afin d'assurer une participation large de la communauté scientifique aux réunions d'animation, le GdR ISIS finance, par laboratoire inscrit ou membre de droit au GdR ISIS, le déplacement d'un membre permanent et d'un doctorant par réunion d'animation, et ce en plus des déplacements des orateurs. Cette opportunité offerte aux laboratoires de recherche illustre la volonté de rassembler la communauté et de permettre un même accès aux réunions à tous les doctorants et collègues, quel que soit leur laboratoire ou équipe d'accueil. Ainsi les équipes éloignées géographiquement et disposant de peu de moyens financiers ne sont pas pénalisées.

Ce système a pu être mis en place et perdure grâce aux conventions signées avec les membres du Club des partenaires du GdR ISIS (voir paragraphe 2.4 ci-dessous). C'est pourquoi il est important de définir précisément quels sont les laboratoires membres de droit et les laboratoires partenaires qui cotisent. Le GdR étant une structure CNRS, soutenue financièrement par le CNRS, tous les laboratoires ayant pour tutelle le CNRS sont des membres de droit. Afin de conserver le caractère fédérateur des activités du GdR ISIS, nous inscrivons aussi comme laboratoires membres de droit, les équipes d'accueil des universités, même si le GdR ISIS ne bénéficie plus du soutien financier du MESR. Les laboratoires privés, les grands organismes et les industriels peuvent s'inscrire après la signature d'une convention s'accompagnant d'une cotisation.

2.4 Club des partenaires

Président :

- Daniel Duclos
SAGEM DS

Les relations fortes avec le monde industriel se traduisent par l'existence depuis de nombreuses années d'un Club des Partenaires industriels associé au GdR ISIS, regroupant de grands groupes et de plus petites entreprises tels que Thales, Safran, EDF, CNES, MBDA, Orange Labs, Huawei, Mitsubishi, ISEP.

Les échanges sont fructueux dans les deux sens, les industriels y trouvant des journées d'animation scientifique leur permettant de faire de la veille technologique et d'avoir une bonne connaissance du tissu académique français, et inversement de présenter leurs

6. <http://gdr-isis.fr/>

problématiques, nourrissant et orientant ainsi les sujets de la recherche académique. Il s'agit d'un lieu privilégié d'échange assez informel entre les chercheurs de la communauté française et les membres industriels du Club des Partenaires. Du fait de l'évolution des applications du traitement du signal, de l'image et de la vision, de nouveaux partenaires industriels sont devenus membres dans le mandat en cours tels que Huawei et MBDA, et de nouvelles conventions sont en cours de négociation. Par ailleurs, afin de faire perdurer le système d'aide au financement des déplacements pour assister aux réunions d'animation organisées par le GdR ISIS, les laboratoires privés et les grands organismes autres que le CNRS sont également membre du Club des Partenaires et à ce titre cotisants.

Sur la période à venir, afin de renforcer plus encore les interfaces entre la recherche académique et la R&D privée, le Club continuera à soutenir les actions du GdR ISIS et à être force de propositions en termes de thématiques scientifiques et techniques. Il aura aussi à cœur de soutenir toute initiative qui visera à promouvoir les métiers de la recherche publique ou privée et susciter des vocations dans le périmètre du GdR ISIS.

La liste des membres du club des partenaires est disponible sur le site web du GdR ⁷.

2.5 Réseau des doctorants

Responsables :

- Audrey Giremus
MCF61, IMS (UMR CNRS 5218), Bordeaux
- Barbara Nicolas
CR CNRS 07, Creatis (UMR CNRS 5220), Lyon

Le GdR ISIS considère que les doctorants constituent une base importante de la communauté de recherche en signal, image et vision. Il souhaite continuer à soutenir des opérations à destination des doctorants et se doit de leur porter une attention d'autant plus spécifique qu'aucune structure nationale thématique ne leur est dédiée spécifiquement. En plus de veiller à ce que les doctorants puissent se mêler aux chercheurs confirmés dans la programmation de journées thématiques, leur offrant ainsi une tribune où présenter leurs meilleurs travaux, le GdR ISIS propose de développer les opérations du réseau de doctorants dans les directions qui suivent.

Le GdR ISIS soutiendra financièrement, dans la continuité de son contrat précédent, des missions inter laboratoires de doctorants, en France ou à l'étranger. Cette opération a pour objectif d'apporter une aide financière aux doctorants pour faire de courts séjours dans des laboratoires différents du leur, afin de travailler avec d'autres chercheurs sur un projet de recherche. Ces demandes d'aide aux missions seront traitées, acceptées et financées au fil des requêtes, dans la mesure des budgets disponibles et dès lors que le projet de recherche montrera l'intérêt de la collaboration, après accord du directeur de thèse et du chercheur concerné dans le laboratoire d'accueil. Au-delà de l'aide financière, l'aspect incitatif à la mobilité est important. C'est pourquoi nous continuerons ce soutien.

En outre, une journée doctorale GdR ISIS biennal est organisée. Elle permet aux étudiants en thèse de s'informer sur les carrières de la recherche, dans le public et le privé, au travers d'exposés de représentants de comités institutionnels (CNU, CoCNRS, etc.), de chercheurs venant d'EPICs (CNES, Onera, etc.), et enfin de chercheurs de centres de R&D industriels. En marge des présentations, une large partie de la journée est

7. <http://gdr-isis.fr/organisation/partenaires/annuaire-des-partenaires.html>

dédiée à des interactions directes entre doctorants et chercheurs dans le cadre d'un forum (stands/posters). Ce forum permet en particulier aux doctorants de rencontrer les partenaires industriels du GdR ISIS et de créer potentiellement des opportunités d'embauche.

En complément de cette journée, il est envisagé une journée, dont le périmètre reste à définir, autour du thème de l'encadrement doctoral en signal et image, notamment destinée à nos jeunes collègues fraîchement recrutés. Une telle proposition viendrait compléter thématiquement les nouveaux dispositifs autour de l'encadrement doctoral proposés par les universités et les EPST (accompagnement de l'Autorisation à diriger une thèse (ADT)).

Enfin, une plus grande attention sera portée à la représentation des doctorants au sein du GdR ISIS, notamment au niveau du site internet. Des mails annuels rappelleront aux directeurs de thèse l'importance d'inscrire leurs doctorants au GdR. Ces derniers seront invités à communiquer des informations sur leur thèse (sujet, mots-clés) sur leur profil GdR, et à mettre un lien vers leur mémoire à l'issue de leur doctorat. Des outils informatiques seront mis en place afin de s'assurer de l'actualisation régulière de ces informations.

2.6 Actions inter-GdRs, ouverture internationale

Responsable :

- Nelly Pustelnik
CR CNRS 07, Laboratoire de Physique (UMR CNRS 5672), ENS Lyon

Inter-GdR – Les thématiques du GdR ISIS sont à l'interface de plusieurs autres domaines scientifiques comme les mathématiques, la physique et l'informatique, avec des domaines d'application très variés. Il est ainsi naturellement en interaction régulière avec de nombreux autres GdRs avec lesquels il se coordonne pour l'organisation d'activités communes. A titre d'exemples, des rencontres annuelles avec le GDR IG-RV ont été organisées entre 2013 et 2015 sur la thématique « De l'acquisition à la compression des objets 3D », tout comme avec les GdRs Robotique, SoC-SIP et Biocomp intitulées « Journées NEUROSTIC ». Le GdR ISIS entend poursuivre sa politique d'ouverture au cours du prochain contrat, comme en atteste la création du Thème T (« Apprentissage », cf. Section 3.5) qui fait déjà l'objet d'échanges fructueux avec les GdRs IA et Robotique.

Comme auparavant, le GdR ISIS entretiendra les liens existants avec des GdRs tels que Ondes, Phenix (Physique) et MIV (Imagerie du Vivant), offrant l'occasion de confronter certaines méthodes à des scénarios réels. Il nouera également des contacts plus ponctuels avec des GdRs pour l'organisation de journées communes sur des sujets suscitant un intérêt marqué de la part de la communauté sur une période donnée. On peut citer notamment, de façon non exhaustive, les liens déjà existant avec les GdRs à coloration mathématique comme MIA, MOA, Théorie des Jeux ou Analyse Multifractale pour l'organisation de journées à grande audience sur des sujets théoriques.

A titre d'exemples, des GdRs avec lesquels des journées pourront être envisagées sont énumérés ci-dessous :

- GdR Analyse multifractale
- GdR Biocomp « Implémentations matérielles du calcul naturel »
- GdR IA « Aspects formels et algorithmiques de l'Intelligence Artificielle »
- GdR IG-RV « Info. géométrique et graphique, réalité virtuelle et visualisation »

- GdR IM « Informatique mathématique »
- GdR MACS « Modélisation, analyse et conduite des systèmes dynamiques »
- GdR MaDICS « Masses de données, informations et connaissances en Sciences »
- GdR MIA « Mathématiques de l'imagerie et de ses applications »
- GdR MIV « Microscopie et imagerie du vivant »
- GdR MOA « Mathématiques de l'optimisation et applications »
- GdR Ondes
- GdR Phenix « Phénomènes hors d'équilibre »
- GdR Robotique
- GdR SEEDS « Systèmes d'énergie électrique dans leur Dimension Sociétale »
- GdR SOC-SIP « System-on-chip, system-in-package »
- GdR Théorie des Jeux : Modélisation mathématique et applications
- GdR Vision
- Communauté STIC-Santé

International – Suite à la demande de communautés étrangères, un format d'inscription spécial à destination des laboratoires étrangers (sans financement de déplacement) a été élaboré. Il leur permettra d'apparaître comme partenaires du GdR ISIS et de recevoir toutes les informations par sa liste de diffusion. Nous travaillerons à étendre cette initiative, marginale pour le moment. Par ailleurs, des rapprochements plus spécifiques seront envisagés avec certains partenaires étrangers. Ainsi, une action en cours d'initiation concerne le Royaume-Uni où une forte activité en traitement du signal et des images existe. Des journées thématiques communes pourraient être organisées en collaboration avec les Profs S. McLaughlin et M. Peyrera (Edinburgh) ainsi que les Profs C.-B. Schönlieb et J. Boulanger (Cambridge). Cette initiative, envisagée à court terme, pourrait être étendue à d'autres pays voisins. Enfin, le GdR ISIS envisagera le développement de ses relations internationales par des réseaux existants et labellisés par le CNRS tels que les LIA et UMI.

3 Présentation par thème

3.1 Thème A : Méthodes et modèles en traitement du signal et de l'image

Directeurs scientifiques adjoints :

- Rémy Boyer
MCF61, L2S (UMR CNRS 8506), CentraleSupélec, Gif-sur-Yvette
- François Desbouvries
PU, SAMOVAR (UMR 5157), Telecom SudParis, Evry
- Nicolas Le Bihan
CR CNRS 07, GIPSA-lab (UMR CNRS 5216), Grenoble

3.1.1 Présentation générale du Thème A

Les domaines de recherche couverts par le Thème A portent sur les questions théoriques et méthodologiques en traitement du signal et de l'image. Dans la structuration 2014-2018 du thème, ces questions étaient abordées au travers de quatre axes définissant le périmètre scientifique du thème :

- modélisation, représentation, identification ;
- détection, classification, décision, estimation ;
- reconstruction, restauration, débruitage ;
- traitements adaptatifs, en ligne, distribués.

L'animation scientifique autour de ces axes permettent, au travers notamment des journées organisées, d'identifier les verrous méthodologiques rencontrés dans un large spectre applicatif. En particulier, depuis la précédente assemblée générale, il n'est plus fait de distinction entre les questions méthodologiques rencontrées en traitement du signal et celles rencontrées en traitement des images dit de bas niveau. A titre d'exemple, l'optimisation, au cœur des problématiques de reconstruction d'image, a toute sa place au sein des prérogatives du Thème A puisque nécessitant des contributions méthodologiques importantes, en interaction avec la communauté de mathématique appliquée.

Plus généralement, de par la place centrale du Thème A, ces interactions avec des communautés scientifiques connexes seront indispensables. Irriguant des domaines aussi variés que ceux de la physique ou de l'analyse de données, le thème sera parallèlement le point d'entrée des avancées récentes en mathématiques, mathématiques appliquées et algorithmique. Il continuera donc d'entretenir des relations étroites avec les autres thèmes du GdR ISIS, mais aussi avec les GdR Ondes, MADICS, MIA, MOA, IA, IM et les sociétés savantes, comme la Société Savante Francophone d'Apprentissage Machine.

Suite à une large consultation effectuée auprès d'acteurs reconnus de la communauté, et afin de rendre compte des évolutions méthodologiques récentes de la discipline, il est proposé une nouvelle structuration du Thème A mettant en exergue les grands enjeux scientifiques dont doit se saisir la communauté lors des prochaines années. Cette structuration autour de quatre nouveaux axes est présentée dans le paragraphe suivant.

3.1.2 Perspectives du Thème A

Conformément à la stratégie générale du GdR ISIS exposée plus haut, une partie importante de l'activité du Thème A consiste en l'organisation et l'animation de journées scientifiques. Pour la sélection des problématiques abordées lors de ces réunions, le Thème A s'attache d'une part à être présent sur les grands sujets du moment, en accordant notamment une attention particulière aux intérêts exprimés par le Club des Partenaires, d'autre part à fournir une ouverture vers des thématiques émergentes dans des réunions réunissant des spécialistes des domaines concernés. Il est tout d'abord important de souligner que l'animation de journées thématiques du GdR est ouverte à tous. Aussi le Thème A, comme les autres thèmes, sollicite-t-il les contributions venant de l'ensemble des membres du GdR et couvrant l'ensemble des champs disciplinaires évoqués dans ce document. Chacun est invité à proposer des réunions, faire remonter des thématiques porteuses, et organiser des événements conjoints avec d'autres partenaires, par exemple à l'occasion de visites de spécialistes étrangers.

Afin de dégager de nouvelles perspectives, en vue de l'organisation de journées sur des sujets déjà importants mais aussi prometteurs, les responsables du Thème A ont sollicité la communauté à l'automne 2017, en amont de l'assemblée générale, pour faire émerger les orientations les plus pertinentes pour le Thème A. Sont particulièrement remerciés les acteurs de la communauté qui ont contribué par écrit à ces perspectives : Roland Badeau, Rémi Bardenet, Jean-François Bercher, Pascal Bianchi, Arnaud Breloy, Rodrigo Cabral Farias, Caroline Chaux, Emilie Chouzenoux, Pierre Comon, Valentin Emiya, Rémi Flamary, Gérard Favier, Guillaume Ginolhac, Rémi Gribonval, José Henrique de Morais Goulart, Sylvain Le Corff, Olivier Lezoray, Philippe Loubaton, Nicolas Papadakis, Frédéric Pascal, Jean-Christophe Pesquet, Gabriel Peyré, Gaël Richard, Salem Said, François Septier, Nicolas Tremblay, Konstantin Usevich, Steeve Zozor.

Une synthèse de ces retours a permis de dégager quatre axes scientifiques majeurs qui structurent la prospective scientifique du Thème A.

1. Données massives et grande dimension

(a) Simulation et optimisation pour les problèmes de grandes tailles

Les méthodes de **simulation Monte Carlo** sont des outils numériques fondamentaux pour l'estimation de modèles à données latentes et la résolution de problèmes inverses. Cependant les problématiques d'apprentissage modernes supposent le traitement de données massives, pouvant être elles-mêmes de très grande dimension.

Dans ce double contexte de données massives et/ou de grande dimension, d'importants défis restent encore à relever. En effet, peu de travaux ont été dédiés à l'étude des méthodes de Monte Carlo séquentielles ou des méthodes MCMC lorsque la dimension du problème augmente ; les défis tiennent à la taille des supports à explorer, à la possible multi-modalité des lois, et à une structure de dépendance possiblement complexe entre les données. Il est nécessaire de développer de nouveaux algorithmes capables de traiter efficacement, rapidement et de façon robuste les problèmes impliquant une grande masse de données ou portant sur des données de grande dimension.

Les progrès récents dans le domaine des capteurs d'images engendrent des masses de données toujours plus grandes, ce qui met en difficulté les algorithmes traditionnels d'**optimisation**. Dans ce contexte, il devient préférable de concevoir des algorithmes d'optimisation capables de prendre en compte la

structure du problème ; de gérer des jeux de données incomplets, distribués et variables dans le temps ; et/ou d'utiliser des infrastructures de calcul et de stockage massivement parallèles. En particulier, un défi actuel consiste à étendre les stratégies parallèles alternées par blocs dans un contexte plus général (distribué, non convexe ou non lisse).

Une autre façon de traiter les problèmes de grande dimension consiste à utiliser une stratégie de minimisation au fil de l'eau. Cependant de nombreuses méthodes du premier ordre (tels que les algorithmes de descente de gradient stochastique) pâtissent d'une vitesse de convergence lente. Il convient donc de rechercher des stratégies d'accélération applicables à une large classe d'algorithmes.

- (b) Le scénario commun à l'ensemble des travaux sur l'**optimisation distribuée** consiste en l'existence d'entités indépendantes (les *agents*), dotées de capacités de calcul et capables de communiquer entre elles au sein d'un graphe dont les arêtes correspondent aux paires d'agents capables d'échanger des messages. Une telle architecture distribuée apparaît naturellement dans les réseaux de capteurs, mais aussi en apprentissage statistique, où la volumétrie des données impose de résoudre des problèmes d'optimisation en tirant parti de plusieurs entités de calcul œuvrant simultanément.

Au cours la dernière décennie, de nombreux algorithmes ont été proposés afin de permettre à des agents d'atteindre un consensus sur la solution d'un problème d'optimisation, dont la fonction de coût ne serait que partiellement connue de chaque agent. Les méthodes ont bénéficié de fortes interactions avec le domaine de l'optimisation (algorithmes primaux, primaux-duaux, etc.) et du traitement adaptatif du signal (filtrage, approximation stochastique, etc.). L'optimisation distribuée entre aujourd'hui dans une phase de maturation où il convient de démontrer l'applicabilité des méthodes à des contextes applicatifs précis : cluster computing, localisation dans les réseaux de capteurs, clustering d'utilisateurs dans les réseaux mesh, contrôle ou planification de flotte, etc. D'un point de vue théorique, il s'agit de pousser plus avant l'extension des méthodes nouvelles d'optimisation, déterministes ou stochastiques, dans un contexte distribué. Enfin, un enjeu important consiste à établir des rapprochements théoriques avec la théorie des jeux.

- (c) **Matrices aléatoires**

Compte tenu du développement considérable des dispositifs d'acquisition et des réseaux de capteurs, il est aujourd'hui courant d'être confronté à des signaux de grandes dimensions dans des contextes applicatifs aussi divers que les communications numériques (systèmes multi-antennes massifs, OFDM), le traitement des signaux radar (STAP, MIMO), le bio-médical (génomique, détection de biomarqueurs), la surveillance de l'environnement ou des procédés industriels, l'analyse des séries temporelles financières, etc. En pratique, le nombre d'observations disponibles des signaux pouvant être relativement limité (durée et/ou stationnarité courtes), on est alors confronté à des signaux dont la dimension M est du même ordre de grandeur que le nombre d'échantillons N (avec M, N potentiellement grands). Dans ce régime où $M \asymp N$, nombre de méthodes statistiques classiques ne se comportent plus comme dans le cas standard où $N \gg M$. Il s'agit du cas des méthodes mettant en jeu des fonctionnelles de la matrice de covariance empirique des observations, en particulier

ses valeurs propres et vecteurs propres. Les avancées théoriques vers la caractérisation des vecteurs propres (et non plus seulement des valeurs propres), ainsi que de certains modèles nouveaux (tels que les modèles dits *spiked*), ont permis au spectre applicatif de s'étendre aux méthodes de détection et d'estimation en traitement du signal (par exemple les méthodes de sous-espaces de type MUSIC). Il est à noter que peu d'études ont été réalisées jusqu'à présent pour le modèle dit factoriel dynamique, où le signal observé est une version bruitée de la sortie d'un système linéaire K entrées / M sorties ($K < M$) excité par un signal utile de dimension K non observable. Les problèmes d'inférence statistique liés à ce modèle très populaire ont bien sûr été très abondamment étudiés depuis une cinquantaine d'années (estimation de paramètres de modèles ARMA vectoriels, de modèles d'état et des filtres de Kalman associés, estimation non paramétrique des densités spectrales, tests d'indépendance mutuelle entre les M composantes de l'observation, test de blancheur temporelle de l'observation, tests de détection d'un signal utile, etc.) dans le contexte des petites dimensions. Cependant, le cas où M et N tendent vers l'infini n'a été que très peu abordé. Pour aborder ces problématiques, il va s'avérer nécessaire d'étudier finement, dans le régime des grandes dimensions, le comportement de diverses fonctionnelles de statistiques telles que les matrices de covariance spatio-temporelles empiriques, les matrices d'auto-covariance empiriques entre passé et futur, les estimateurs non paramétriques usuels de densité spectrale, et d'en déduire des techniques d'inférence statistique adaptées aux modèles linéaires factoriels de grandes dimensions.

2. Signal et géométrie

(a) Géométrie de l'information

Depuis quelques années, de nouveaux problèmes ont émergé dans de nombreuses applications liées au traitement du signal et des images où les techniques standards basées sur des résultats d'algèbre linéaire matriciel ne sont plus adaptées ou transposables facilement.

En géométrie de l'information, l'objet d'intérêt (signal, image, vidéo, etc.) est vu comme un point sur un espace métrique (variétés, groupes, etc.) et les problèmes classiques (estimation, filtrage, classification, etc.) doivent être reformulés sur ces espaces qui, en général, possèdent une courbure non nulle. Par exemple, les outils de probabilité et statistique nécessaires à la résolution des problèmes de traitement du signal et des images sont alors définis de manière intrinsèque sur ces espaces courbes. La notion de barycentre ou de mesure s'appuie alors sur la métrique de l'espace courbe, ce qui nécessite, du fait de la nature non-linéaire des problèmes engendrés (absence de structure d'espace vectoriel pour ces espaces) de repenser les algorithmes pour la résolution des problèmes d'intérêt en signal et image. La communauté signal et image s'intéresse depuis quelques années à proposer de nouveaux algorithmes pour le traitement géométrique de l'information (optimisation, estimation, etc.). Ce thème de recherche est dans une phase de maturation avec des applications bien identifiées dans de nombreux domaines (radar, traitement d'images IRM, etc.) et les contributions de la communauté ISIS y sont grandissantes.

(b) Inférence sur variétés

L'inférence statistique a été généralisée à beaucoup de types d'espaces linéaires,

mais la transposition des outils d'estimation au cas des espaces non-linéaires est plus compliquée et n'a été abordée que récemment en mathématiques.

Dans de nombreux domaines d'application, la matrice de covariance est un objet central à estimer. Dans un contexte de données massives, les observations de taille très importante rendent l'estimation de ces matrices compliquée statistiquement et algorithmiquement. Par exemple, l'écart au modèle gaussien ainsi que le faible nombre d'échantillons soulèvent de nouveaux problèmes. Les lois elliptiques symétriques complexes (CES) permettent de prendre en compte les hétérogénéités des données mais de nombreux problèmes persistent pour caractériser les performances des traitements associés à ces estimateurs.

Les méthodes d'optimisation sur variétés peuvent alors permettre de réduire le support d'apprentissage, devenu non convexe par l'inclusion d'information de structure sur la matrice (par exemple Toeplitz, rang faible, Kronecker, etc.).

La multiplication des domaines d'application de ce type de problèmes d'estimation liés aux matrices de covariance empiriques et leur estimation nécessite l'exploration de nouvelles pistes méthodologiques et algorithmiques (convergence, algorithmes rapides).

(c) Traitement du signal sur graphe

Le traitement du signal sur graphe est un axe de recherche récent qui connaît un essor important. Si les applications se multiplient, il reste encore plusieurs verrous identifiés qui monopolisent l'attention de chercheurs de notre communauté. En particulier, il est admis qu'il existe encore un manque d'homogénéité/universalité pour certaines définitions (transformée de Fourier, translation). Plusieurs définitions existent par exemple pour les ondelettes sur graphes, du fait des multiples approches pour définir la translation sur un graphe. Ce manque d'universalité dans les définitions actuelles engendre l'absence d'algorithmes rapides pour le calcul sur graphes, soulevant des problèmes importants pour les grandes masses de données. Parmi les autres grandes questions sur lesquelles doit se pencher la communauté signal, trois points semblent particulièrement bloquants : le passage à l'échelle, les aspects liés à la dynamique de graphes et les graphes hétérogènes.

Les efforts méthodologiques de la communauté devraient permettre ensuite de démocratiser le traitement du signal sur graphes, et de confirmer les pistes d'applications potentielles prometteuses telles que, parmi d'autres, l'apprentissage et les neurosciences.

3. Modèles structurés et inversion

(a) Décompositions canoniques structurées

Les décompositions structurées visent à représenter des données potentiellement de grande dimension sous la forme d'une combinaison d'éléments simples. Un exemple typique est la factorisation en matrices non-négatives (NMF). Ces approches sont aujourd'hui très utilisées dans des applications de compression, fouille et reconstruction de données, de recommandation, de vision par ordinateur, etc. Certains aspects méritent encore d'être approfondis, comme la sélection automatique du rang de factorisation, la prise en compte de la phase dans la NMF, les factorisations couplées entre jeux de données hétérogènes, ou les décompositions adaptatives/en ligne de données non stationnaires. Les aspects algorithmiques présentent aussi un enjeu important :

réduction du temps de calcul pour le traitement de données massives (algorithmes rapides, méthodes stochastiques), robustesse à l'initialisation, calcul parallèle/distribué.

(b) Modèles et décompositions tensoriels

A l'ère des big data, les outils de traitement du signal et tout particulièrement les outils tensoriels devraient jouer un rôle de tout premier plan, dans de nombreuses applications, pour résoudre des problèmes de compression multimodale, de fouille de données (data mining), d'estimation de données manquantes, de représentation et d'analyse de données de très gros volume. De nouveaux modèles tels que les trains de tenseurs, les réseaux de tenseurs, ou encore les tenseurs sur des graphes, qui ont été développés récemment, ouvrent des perspectives de recherche très intéressantes. L'unicité, l'identifiabilité, la visualisation graphique, ainsi que le développement d'algorithmes d'optimisation permettant d'estimer tant la structure que les paramètres de ces modèles, en prenant en compte des contraintes de temps de calcul et de place mémoire, constituent des champs de recherche importants. Des traitements décentralisés et en temps réel devraient déboucher sur de nouvelles architectures multi cœurs pour l'implémentation de ces algorithmes. Enfin, le développement de benchmarks disponibles au sein du GdR ISIS serait d'une grande utilité pour la validation, l'évaluation et la comparaison de nouvelles techniques tensorielles de modélisation ainsi que des algorithmes d'optimisation associés, pour les big data, dans les domaines d'application relevant par exemple du biomédical, de l'environnement ou encore des télécommunications.

(c) Représentations parcimonieuses et acquisition compressée

La notion de problème inverse demeure centrale pour aborder les problèmes prototypiques de traitement du signal et de l'image (notamment le cœur de métier en échantillonnage, débruitage, déconvolution, séparation de sources, etc.). Des avancées importantes ont été réalisées au cours de la décennie grâce à la régularisation parcimonieuse et aux techniques d'optimisation non-convexe et/ou non-lisses. Grâce à l'éclosion du traitement du signal sur graphe, des extensions remarquables des outils classiques du traitement du signal ont aussi été obtenues pour de nouvelles classes de données dont la structure irrégulière est capturée par un graphe. Les approches basées sur la notion de parcimonie au sens traditionnel ont atteint une maturité méthodologique certaine, et celles exploitant la notion de dictionnaire et l'apprentissage ont démontré un fort potentiel pour améliorer la qualité de reconstruction. L'enjeu majeur aujourd'hui est d'en assurer tout le potentiel de déploiement applicatif, en abordant en particulier le verrou de la complexité algorithmique.

4. Signal, image et apprentissage

L'apprentissage automatique revêt des enjeux scientifiques pluridisciplinaires, mobilisant des concepts issus des mathématiques appliquées (probabilités, statistique, optimisation) et de l'informatique (algorithmie notamment). Les applications de l'apprentissage automatique sont nombreuses et concernent naturellement le traitement du signal et le traitement d'image. La classification, la régression, le partitionnement, l'identification de sous-espaces sont, par exemple, autant de tâches d'apprentissage qui se retrouvent depuis longtemps au cœur de nombreuses problématiques auxquelles est confrontée la communauté du GdR ISIS. Plus récemment, sous

la dénomination anglophone de *Machine Learning*, l'apprentissage automatique connaît un regain d'intérêt important car il est l'une des disciplines indissociables des recherches en Intelligence Artificielle (IA), au même titre que l'automatique, la robotique ou la logique. Les avancées récentes en IA, qui se confondent parfois avec celles réalisées en apprentissage automatique, ont reçu un écho médiatique important, notamment dans la presse à grand public à la suite des défis relevés par les systèmes d'apprentissage profond. Les réseaux de neurones profonds ont en effet révolutionné l'apprentissage supervisé, en particulier pour la classification de données structurées, en vision par ordinateur notamment. Ces techniques sont également en train de devenir l'état de l'art pour résoudre certains problèmes inverses en imagerie.

Pour structurer la communauté scientifique française travaillant sur l'IA au sens large, un pré-GdR IA a récemment été créé, avec pour objectif l'étude des aspects formels et algorithmiques de l'intelligence artificielle. Il n'en est pas moins crucial que le GdR ISIS doit continuer à se positionner comme le lieu d'animation incontournable en apprentissage, capitalisant sur plusieurs décennies de connaissances, savoir-faire et contributions. L'action spécifique « apprentissage automatique » menée au sein du Thème A au cours du précédent plan quinquennal a connu un succès indiscutable, avec de nombreuses réunions d'animation qui ont attiré un public toujours plus assidu. Afin de renforcer ce positionnement d'acteur incontournable sur la discipline, le GdR ISIS propose de faire émerger un thème transverse "Apprentissage" structurant les activités relevant du domaine et menées dans les différents thèmes historiques. En soutien à ce nouveau thème T décrit à la Section 3.5, le Thème A propose également de dédier un axe consacré aux questions fondamentales, c'est-à-dire méthodologiques et théoriques, de l'apprentissage. Plus spécifiquement, sans pour autant négliger les grands verrous scientifiques traditionnellement étudiés, trois enjeux majeurs seront en priorité considérés.

(a) **Robustesse et confiance en la décision**

Une problématique récurrente dans la conception de systèmes d'apprentissage automatique réside dans leur robustesse au bruit ou, plus précisément, aux données aberrantes. En particulier, dans un contexte d'apprentissage supervisé, la présence de données mal labellisées dans l'ensemble d'apprentissage peut influencer significativement les performances de décision. Parallèlement, la prise de décision concernant une donnée aberrante, dans le sens où elle ne correspond que très peu aux données rencontrées dans la phase d'entraînement, est généralement peu fiable. Une difficulté réside également dans le constat qu'il est difficile de proposer des méthodes d'apprentissage capables d'accompagner ces prises de décision par un ou des indicateurs de confiance performants concernant ces décisions. En particulier, cette question de l'incertitude de prédiction représente un obstacle à la compréhension formelle des réseaux profonds dont le Thème T s'est également saisi (cf. Section 3.5.2). Il conviendra donc de répondre à ces exigences de robustesse et de mesure de confiance.

(b) **A priori, physique et topologie**

Un des atouts majeurs des méthodes d'apprentissage automatique, comme notamment les réseaux de neurones profonds qui atteignent aujourd'hui des performances inégalées, est leur généricité et leur versatilité. Elles peuvent donc être mises en œuvre pour réaliser un grand nombre de tâches variées, ayant parfois peu en commun. Cette généricité s'accompagne d'une tendance à négliger les processus d'acquisition ou la nature même des données à traiter, qui sont

pourtant au sein de la communauté du GdR ISIS généralement intrinsèquement liées à un médium définissant le concept même de signal ou d'image. L'exploitation de connaissances a priori, qui peuvent notamment résulter de la topologie ou de processus physiques sous-jacents, représente une opportunité certaine pour repousser encore un peu plus loin les performances des systèmes d'apprentissage. Réciproquement, la compréhension fine des mécanismes de décision mis en jeu par les systèmes d'apprentissage les plus performants pourrait permettre d'extraire une connaissance complémentaire des signaux et des images, jusqu'alors inaccessible par les modèles paramétriques ou non-paramétriques traditionnellement utilisés.

(c) **Apprentissage de représentations**

En lien avec l'extraction de connaissance physique suggérée au point précédent, une question importante en traitement du signal et traitement d'image est le choix d'une représentation optimale en vue de la réalisation d'une tâche dédiée (inférence, classification, compression par exemple). Grâce aux récentes avancées dans le domaine, les méthodes d'apprentissage automatique offrent des possibilités nouvelles d'identifier des représentations structurées ou non-structurées, probabilistes ou paramétriques, dans des contextes supervisé ou non-supervisé. Ces questions pourront notamment être abordées au sein du Thème A en soutien de l'action « Apprentissage des représentations – *Deep learning* » proposée par le Thème T (cf. Section 3.5.3).

3.1.3 Actions proposées

De façon plus ciblée, afin de mettre des coups de projecteur sur des points saillants de cette prospective scientifique, des actions spécifiques seront menées autour de trois thématiques. Par ces actions, le Thème A s'engage à assurer l'organisation régulière de journées dédiées à ces thématiques.

1. Action *Transport optimal*

Animateur(s) : Rémi Flamary (MCF, Laboratoire Lagrange, Nice).

Description : Le transport optimal a été formulé par Gaspard Monge au 18e siècle et a connu de grandes avancées théoriques, notamment médiatisées par les travaux de Cédric Villani. Il s'agit d'optimiser le coût de transport depuis un ensemble de sources vers des consommateurs. Ce problème très ancien a connu plusieurs révolutions successives. Très récemment, le transport optimal et les distances associées connues sous le nom d'EMD (pour Earth Mover's Distance) dans la communauté de la vision par ordinateur, ont connu un essor certain avec les méthodes de régularisation, notamment entropique, du plan de transport, permettant de calculer rapidement des plans de transports « lissés » pour des données de grandes dimensions. Dès lors, ils tendent à devenir des outils standards pour le traitement de données et on peut maintenant anticiper le fait que ces techniques joueront prochainement un rôle important dans la résolution de problèmes inverses difficiles en imagerie (petite dimension) mais aussi en Machine Learning (grande dimension). Plusieurs verrous subsistent néanmoins, dont la communauté du GdR ISIS peut se saisir :

- Le calcul rapide et peu coûteux en mémoire du transport régularisé est néanmoins restreint à certains types de coût de transport. Ainsi, les méthodes d'optimisation modernes basées sur la régularisation ne permettent pas toujours un passage à l'échelle et des algorithmes accélérés et/ou stochastiques sont encore rares.

- Un aspect essentiel du transport, le coût, nécessite une connaissance approfondie des données pour être choisie de manière appropriée. Le choix de cette fonction coût pour des signaux temporels et/ou sur graphe reste une question ouverte.
- L'extension du transport optimal au problème dit *non balancé* est également un enjeu majeur pour l'application de ces outils à des problèmes rencontrés par exemple en traitement d'image, tels que le recalage ou l'interpolation.
- L'utilisation du transport optimal comme « fonction de perte » est enfin un sujet particulièrement actif en apprentissage profond. Les récents premiers travaux publiés sur le sujet (notamment par Google) laissent présager un déploiement futur de la théorie du transport vers les architectures d'intelligence artificielle.
- Enfin, malgré des succès pratiques, il reste encore à prouver des résultats théoriques sur la qualité de l'estimation à base de distance de Wasserstein.

Le transport optimal ouvre ainsi de nouveaux axes de recherche dont la communauté française du traitement de l'information, du signal et des images doit se saisir, et il apparaît nécessaire au GdR ISIS de soutenir cette thématique prometteuse.

2. Action *Simulation pour les problèmes de grande taille*

Animateur(s) : François Septier (MCF, laboratoire CRISAL, Villeneuve d'Ascq) et Sylvain Le Corff (CR CNRS, LMO, Orsay).

Description : L'objectif de cette Action est de développer de nouvelles méthodologies permettant de résoudre les problèmes posés par l'analyse de grands jeux de données et de données de très grande dimension. L'organisation de journées thématiques dédiées aux méthodes de Monte Carlo, aux algorithmes stochastiques ainsi qu'à l'optimisation numérique favorisera les interactions et les collaborations entre chercheurs afin d'améliorer la maîtrise des coûts nécessaires à l'étude de systèmes dynamiques complexes. L'un des enjeux majeurs est la mise en place d'algorithmes permettant une exploration efficace de l'espace d'état dont la vitesse de convergence peut être contrôlée en fonction de la dimension du problème considéré.

L'une des difficultés associées aux très grands jeux de données provient des coûts de calcul prohibitifs dus à l'évaluation de la vraisemblance des observations ou à la mise en œuvre d'algorithmes du type Markov Chain Monte Carlo (MCMC). Très récemment, des approches « diviser pour conquérir » ont été proposées afin de représenter la loi a posteriori par une combinaison de lois plus simples pour lesquelles des méthodes de simulation sont disponibles, puis de combiner les résultats associés à chacune de ces lois plus simples. Bien que l'efficacité de ces approches distribuées ait été illustrée sous certaines hypothèses, leur utilisation dans un cadre général reste un problème largement ouvert, la façon optimale de recombinaison des sous-problèmes introduits posant des difficultés tant théoriques que pratiques. D'autre part, les approches variationnelles, populaires dans les applications de grande dimension en Machine Learning, permettent de simuler de façon approchée des lois a posteriori de façon numériquement très efficace. Elles n'ont toutefois été que très peu utilisées en simulation Monte Carlo, typiquement dans les méthodes séquentielles, afin de calibrer les lois instrumentales cruciales pour l'efficacité de ces méthodes (problème pour lequel il n'existe à ce jour que très peu de solutions). D'un point de vue théorique, la convergence des approximations variationnelles en termes de concentration autour des lois a posteriori a été peu étudiée bien qu'elle permettrait d'apporter des garanties pour l'utilisation de ces méthodes.

Le second thème mis en avant dans cette action spécifique concerne les espaces d'états de très grande dimension. L'efficacité des méthodes de simulation et d'optimisation

est très sensible aux paramètres d'implémentation dans ce cadre et il semble judicieux de prendre en compte les informations apportées par la géométrie de la loi cible. De récents travaux sur les méthodes de scaling donnent les paramètres optimaux pour l'implémentation d'algorithmes Metropolis-Hastings à marche aléatoire sous des hypothèses raisonnables. Cependant, les algorithmes de type Langevin n'ont pas été traités bien qu'ils soient bien plus performants en grande dimension car ils prennent en compte le gradient de la log-vraisemblance pour guider les déplacements de la chaîne de Markov. Par ailleurs, de nouvelles méthodes permettant le couplage d'algorithmes de simulation de type MCMC avec des outils d'optimisation proposés dans la littérature fournissent des résultats prometteurs : MCMC proximaux, couplage déterministe pour transformer le mécanisme d'acceptation-rejet de type Metropolis (marche aléatoire, Langevin) en loi de proposition non gaussienne permettant ainsi une meilleure exploration de la loi cible. Assurer la flexibilité de ces approches pour traiter des problèmes d'apprentissage réels en traitement du signal est d'un intérêt crucial, au même titre que l'établissement de la convergence de ces méthodes. Une autre piste très prometteuse est l'étude des algorithmes MCMC irréversibles dont les performances numériques sont frappantes en terme d'oubli et de variance asymptotique. Enfin, les stratégies récentes consistant à exploiter des schémas Monte-Carlo pour approcher certaines quantités requises dans un autre algorithme Monte-Carlo (particle MCMC, nested sequential MC, etc.) constituent également une solution intéressante aux problèmes à grande dimension. Bien que ces méthodes soient devenues très populaires en statistiques et en apprentissage au cours des cinq dernières années, leur potentiel reste largement sous-exploité en traitement du signal et de l'image.

3. Action *Modélisation et optimisation à l'interface signal/apprentissage*

Animateur(s) : Valentin Emiya (MCF, LIF, Marseille), Caroline Chaux (CR CNRS, I2M, Marseille), Konstantin Usevich (CR CNRS, CRAN, Nancy).

Description : Le traitement du signal et l'apprentissage ont en commun un grand nombre de fondamentaux, dont des modèles et méthodes d'optimisation. Les objectifs scientifiques de cette action sont de dégager ces objets d'intérêt pour les deux disciplines, de faire ressortir leurs points de convergence, au-delà des différences de formalisme ; de s'intéresser aux contextes et histoires respectifs ; de confronter l'originalité des développements dans chaque discipline. Cet espace de rencontre entre les deux communautés signal/apprentissage favorisera en particulier l'émergence de travaux à l'interface des deux disciplines et permettra à la communauté du traitement du signal de se positionner scientifiquement par rapport aux avancées en apprentissage automatique. L'action, portée également par le Thème transverse « Apprentissage » (Cf. Section 3.5.3), se concrétisera par l'organisation de journées scientifiques sur des thèmes d'interface tels que l'optimisation, les modélisations statistiques, la factorisation de matrices et tenseurs, le traitement de séquences spatio-temporelles. Ces journées porteront par exemple sur les sujets suivants :

- *optimisation en signal et en apprentissage*. Le traitement du signal et l'apprentissage se rejoignent sur de nombreux problèmes et méthodes d'optimisation : optimisation convexe, optimisation non-convexe, modèles parcimonieux, optimisation en grande dimension, factorisation de tenseurs.
- *traitement du signal statistique vs. apprentissage statistique*. Les deux disciplines se sont construites sur un ancrage statistique fort ; il s'agira de confronter les fondements des approches et modèles statistiques respectifs.

- *factorisation de matrices et de tenseurs en signal et en apprentissage.* Les approximations et décompositions de rang faible des matrices et tenseurs sont utilisées en TS depuis longtemps (méthodes de sous-espace, NMF, statistiques d'ordre supérieurs/ACI) dans plusieurs domaines applicatifs (par exemple, spectroscopie, imagerie hyperspectrale ou identification de systèmes non linéaires). Elles sont aussi présentes dans plusieurs domaines de l'apprentissage : réduction de dimension, systèmes de recommandation, clustering, modèle thématique régression, « factorization machines ». Il est courant d'utiliser les mêmes contraintes de non-négativité, orthogonalité, parcimonie, structure, décompositions couplées, données manquantes.
- *traitement/apprentissage sur séquences spatio-temporelles.* Les avancées marquantes en apprentissage profond se sont en grande partie faites sur des séquences spatio-temporelles : vidéos, image, audio, données sur graphes dynamiques, etc. On se retrouve ainsi aujourd'hui avec des approches concurrentes venant de disciplines distinctes (apprentissage, traitement du signal, séries temporelles) sur les mêmes données et applications. Cette action permettra à ces communautés de se rencontrer autour de ce sujet d'intérêt commun, et de mettre en évidence ce qu'une discipline peut apporter à l'autre.
- *apprentissage dans le plan temps-fréquence.* Les outils classiques de l'apprentissage ne s'appliquent pas facilement dans le plan temps-fréquence, où les données sont dépendantes et à valeurs complexes, et où les matrices ne se prêtent pas aux mêmes factorisations. Un objectif sera de faire ressortir les enjeux et verrous de l'apprentissage dans le plan temps-fréquence et de mettre en valeur des approches appropriées.

Perspectives d'actions

Au cours du prochain programme quinquennal, les animateurs du Thème A resteront à l'écoute de la communauté du GdR ISIS afin de donner un éclairage particulier à toute question scientifique qui le mériterait. A titre d'exemple, des réflexions sont d'ores et déjà engagées pour la mise en place future d'une action sur les processus ponctuels. En particulier, les processus ponctuels déterminantaux (DPP) occupent une position particulière à l'intersection des mathématiques, de la physique et de l'informatique. Au sein du périmètre scientifique du Thème A, les DPP apparaissent dans au moins trois domaines d'application :

- modèles statistiques répulsifs,
- méthode d'intégration numérique,
- caractérisation de zéros de transformées temps-fréquence.

Aujourd'hui, le traitement du signal représente donc un champ d'expérimentation naturel pour l'objet transdisciplinaire que sont les DPP, nés en physique statistique, et passés ensuite par les matrices aléatoires, l'apprentissage automatique et les statistiques spatiales. Ils connaissent actuellement un engouement important qu'il conviendra de soutenir à brève échéance.

3.2 Thème B : Image et vision

Directeurs scientifiques adjoints :

- Cédric Demonceaux
Chargé des journées thématiques « Vision 3D, Géométrie »
PU61, Le2i, VIBOT (ERL CNRS 6000), Université Bourgogne Franche-Comté
- François Goudail
Chargé des journées thématiques « Co-conception et télédétection »
PU63, Laboratoire Charles Fabry (UMR CNRS 8501), Institut d'Optique Graduate School
- Edoardo Provenzi
Chargé des journées thématiques « Interaction avec l'humain : qualité, perception »
PU26, IMB (UMR CNRS 5251), Université de Bordeaux
- Su Ruan
Chargée des journées thématiques « Analyse de données »
PU61, LITIS, Normastic (FR CNRS 3638), Université de Rouen Normandie

3.2.1 Présentation générale du Thème B

Le Thème B fédère l'ensemble des acteurs sur les thématiques "Image et Vision". Il est actuellement organisé autour de quatre axes, dont chacun traite d'un aspect de la chaîne de l'image, en allant du capteur jusqu'à l'extraction d'information de haut niveau.

- Techniques de formation et de traitement des images en aval immédiat du capteur, en forte interaction avec la physique des scènes observées et les propriétés des systèmes d'imagerie
- Analyse de scène par vision par ordinateur, en particulier les nouveaux modèles mathématiques, notamment les géométries, permettant la structuration des contenus 3D et dynamiques
- Systèmes de vision intégrant des aspects perceptifs et cognitifs
- Analyse et traitement des masses de données images et vidéos, aussi bien sur le plan de la pertinence de la description de leur contenu que sur les méthodes de reconnaissance et de recherche à base de nouvelles approches en apprentissage

Ce découpage était jusqu'alors bien identifié et correspondait à la structuration de la communauté en Image et Vision. Il avait amené à la création de quatre axes fédérant les disciplines et les compétences lors du précédent contrat quinquennal :

- Axe 1 : Physique, capteurs, traitements (commun avec le GdR Ondes)
- Axe 2 : Vision, géométrie 3D, mouvement
- Axe 3 : Systèmes de vision, perception et connaissance
- Axe 4 : Recherche d'information et masses de données image et vidéo

L'ensemble de ces thématiques a connu ces dernières années des mutations très importantes en intégrant des techniques aussi variées que le calcul intensif et distribué, le traitement de masses de données ou l'apprentissage. D'autre part, l'évolution des architectures (notamment GPU) et de la puissance de calcul a permis de développer des méthodes dans lesquelles l'interactivité prend une place de plus en plus importante. Enfin, des domaines tels que l'imagerie du vivant, la télédétection, l'art et le patrimoine, la vision pour la robotique ou pour l'industrie du multimédia sont des applications directes de notre thème qui exploitent aujourd'hui tout le spectre du capteur à l'interprétation des données visuelles, recouvrant ainsi tous les axes autour desquels le Thème B était structuré. Pour ces raisons, et après de larges discussions lors de l'assemblée générale, il nous a semblé opportun de faire évoluer la structure du thème en passant de la structuration en axes thématiques à une structuration en action, qui permettra de favoriser l'animation de journées dépendant intrinsèquement de plusieurs axes et ainsi de gagner en réactivité. Cette nouvelle structure permettra de favoriser encore davantage l'interaction avec d'autres GdR (Ondes, MIV, IG, MIA, Vision, IM, Robotique, IA, etc.), de proposer des journées avec le Thème transverse « Apprentissage » (Cf. Section 3.5.3) et le Club des Partenaires, qui est fortement investi dans ce thème riche en enjeux socio-économiques.

3.2.2 Perspectives du Thème B

Comme évoqué précédemment, les activités du Thème B « Image et Vision » consistent en l'organisation et l'animation de journées scientifiques sur des thèmes allant du capteur aux interprétations des données visuelles sans distinction d'axes. L'organisation de journées dépend des diverses problématiques actuelles et tient compte des intérêts que le Club des Partenaires peut manifester. Notons que l'animation de journées est ouverte à tout membre du GdR souhaitant qu'un sujet porteur de notre thème soit traité.

Pour faire émerger de nouvelles perspectives, et ainsi faire apparaître de nouvelles actions sur des sujets prometteurs de l'Image et de la Vision, nous avons fait appel à notre communauté. A ce titre, nous tenons à remercier plus particulièrement les collègues ayant contribué à l'écriture de ces perspectives : Adrien Bartoli, Patrick Bouthemy, Pierre Gurdjos, Vincent Lepetit, Chen Liming, Pascal Vasseur, Philippe Réfrégier, Frédéric Champagnat, Elise Koeniguer.

Suite à ces retours, nous avons retenu cinq thématiques :

1. **Capteurs d'images innovants** : Nous voyons apparaître sur le commerce de plus en plus de capteurs dédiés aux applications. A titre d'exemple, depuis maintenant une dizaine d'années, il est possible d'acquérir en une seule prise de vue la colorimétrie et la profondeur de la scène avec des capteurs de vision 3D (kinect, caméra temps de vol, etc.). Les champs de vue s'élargissent puisqu'il est aujourd'hui possible d'avoir une vue à 360° de l'environnement à partir d'une seule prise de vue à l'aide de capteurs sphériques (Gear 360°, GoPro, Giroptic, Ricoh, etc.). Des caméras événementielles (*event-based*), qui ne perçoivent que les changements d'illumination dans la scène, sont désormais disponibles. Citons encore les caméras plénoptiques qui permettent d'acquérir des images de différentes focales à partir d'une seule caméra et ainsi de remonter à la 3D de la scène. Tous ces capteurs ont leurs propres particularités et amènent de nouveaux verrous scientifiques qu'il faut traiter tels que la reconnaissance d'objets sur des images 360°, l'utilisation conjointe des informations 2D et 3D des imageurs 3D, de la géométrie multi-vues pour des capteurs plénoptiques et *event-based*. Il est également important de choisir, pour une

application donnée, les informations pertinentes afin d'obtenir des systèmes à la fois efficaces, rapides et économes.

2. **Vision 3D** : Dans le cadre de scènes fixes ou rigides, les méthodes géométriques pour reconstruire le 3D de la scène ou estimer la pose des caméras (on parle de Structure From Motion, SfM) ont été largement étudiées. Par contre, le problème reste encore ouvert lorsqu'il s'agit de traiter des scènes déformables (Non Rigid Structure from Motion) ou fortement dynamiques. De même, lorsque l'on dispose de capteurs de plusieurs modalités (colorimétrie, temps de vol, polarimétrie, multispectral), les approches multi-vues sont encore difficiles à mettre en œuvre étant donné l'hétérogénéité des données. De nombreux problèmes restent donc à résoudre dans les prochaines années afin de répondre à de forts besoins applicatifs. On peut citer notamment le domaine de l'imagerie médicale, où les objets sont déformables et faiblement texturés.
3. **Colorimétrie** : Les performances d'un algorithme de traitement d'image en couleurs dépendent souvent d'une manière remarquable du choix de l'espace de couleur dans lequel on met en place l'algorithme. Une meilleure compréhension de la géométrie des espaces de couleur est nécessaire pour simplifier ce choix et pour permettre d'adapter les modèles à des situations plus générales. Les outils de la géométrie différentielle, de la topologie algébrique et de l'analyse harmonique sont particulièrement adaptés pour ce type d'étude, bien que leur utilisation en traitement des images en couleurs ne soit pas courant. Concernant la vision des couleurs par un observateur humain, il est établi que l'espace de vision des couleurs n'est pas euclidien. Certains auteurs ont proposé des modèles riemanniens (Helmholtz, Schrödinger, Stiles), mais la tendance actuelle est d'utiliser des espaces à courbure négative pour lesquels le prototype est la géométrie hyperbolique. L'étude de ces espaces et de leurs propriétés pour la vision des couleurs pourra permettre une reformulation rigoureuse de la géométrie des espaces couleur et d'apporter des réponses aux nombreuses objections que suscitent les modèles empiriques utilisés actuellement. Ceci permettra de définir des algorithmes de traitement d'image couleur qui soient intrinsèquement adaptés à la vision humaine.
4. **Vision par ordinateur pour les interfaces Humain/Machine** : La vision par ordinateur a un grand rôle potentiel à jouer pour rendre les futures interfaces humain-machine plus naturelles. Par exemple, les technologies de réalité augmentée et virtuelle se développent très rapidement actuellement, et les compagnies développant des casques de visualisation pour ces applications présentent des avancées majeures. Néanmoins, des progrès sont encore nécessaires pour obtenir la précision et la robustesse nécessaires. De façon plus générale, la vision sera sans doute d'une grande aide pour interagir avec l'utilisateur grâce au développement de la compréhension des gestes, de la posture du corps, ainsi que des expressions faciales. En particulier, la reconnaissance des expressions faciales n'en est encore qu'à ses balbutiements, les méthodes existantes étant limitées à la reconnaissance d'expressions clairement exagérées alors que souvent seuls de micro-indices indiquent l'émotion d'une personne. La prédiction et l'anticipation des actions futures de l'utilisateur sont aussi importantes et encore peu explorées. Pour des systèmes dangereux comme les machines-outils, cette anticipation est pourtant cruciale. La reconnaissance vocale devrait aussi être couplée avec l'analyse des images pour une meilleure compréhension des intentions de l'utilisateur par la machine. Les techniques d'ap-

prentissage sont évidemment un outil important pour développer ces méthodes pour les interfaces humain-machine.

5. **Les techniques d'apprentissage pour l'image et la vision** : A l'ère numérique, la dimension de grands ensembles de données, parfois hétérogènes, s'est imposée rapidement dans tous les domaines. Des méthodes d'analyse, de reconnaissance, d'indexation et de fouille des données images et vidéo évoluent pour répondre à ces nouveaux besoins en tenant compte, d'une part des notions classiques, mais aussi de ces nouvelles échelles.

Dans ce contexte, les techniques d'apprentissage sont devenues un outil incontournable. Les approches supervisées, où des exemples sont annotés par un expert humain, sont les plus couronnées de succès, mais ce sont aussi celles qui demandent le plus d'investissement en temps et en argent. Le problème est exacerbé en vision par ordinateur, où les annotations peuvent être complexes, comme des régions de l'image 2D ou des mesures 3D. La dimension temporelle apportée par la vidéo reste un problème ouvert pour les techniques d'apprentissage où, pour les problèmes inverses comme le flot optique, les approches end-to-end ne surpassent pas encore les approches plus classiques ou mixtes (benchmark MPI-Sintel). Les méthodes géométriques peuvent également être utiles dans un contexte d'apprentissage en couplant conjointement les mesures géométriques 3D aux données sémantiques. Enfin, des méthodes alternatives, spécifiques à la vision par ordinateur, commencent à émerger pour diminuer la quantité de données à annoter. Une première approche possible est l'auto-apprentissage (*self-learning*) dont le principe est d'exploiter la cohérence temporelle entre images d'une séquence vidéo. Par exemple, la segmentation d'un objet dans des images consécutives doit être cohérente avec le flot optique. Une autre approche est le *domain transfer*, où différentes techniques commencent à être développées, pour pouvoir utiliser des images de synthèse pour l'entraînement tout en offrant des propriétés de généralisation aux images réelles. Le *transfert learning*, les Generative Adversarial Networks sont des méthodes possibles pour le *domain transfer*.

3.2.3 Actions proposées

Au vu des problèmes énumérés ci-dessus, nous avons extrait huit actions qui répondent aux enjeux de la communauté. Ces actions feront l'objet d'au moins une journée d'animation par an. Tout comme les autres thèmes du GdR ISIS, le Thème B se laisse bien évidemment la possibilité d'organiser des journées en dehors de ces actions selon les demandes de la communauté. Les trois dernières actions proposées ci-après relevant des techniques d'apprentissage pour la vision par ordinateur, elles seront transverses au Thème B et au Thème T « Apprentissage » (Cf. Section 3.5.3).

Les actions retenues jusqu'à la prochaine assemblée générale à mi-parcours pour l'activité du Thème B sont les suivantes :

1. Action Co-conception

Animateur(s) : Matthieu Boffety (MCF, Laboratoire Charles Fabry, Institut d'Optique Graduate School), Julien Fade (MCF, Institut FOTON, Université Rennes 1), Corinne Fournier (MCF, Laboratoire Hubert Curien, Université de St Etienne), Pauline Trouvé (IR, ONERA).

Description : La conception de systèmes d'acquisition d'images connaît un renouveau grâce aux approches où le dispositif d'imagerie est étroitement associé aux algorithmes

de traitement des données. On parle de conception conjointe ou co-conception. Cette démarche est particulièrement utile aujourd'hui où les capteurs peuvent délivrer beaucoup d'informations : spatiales, 3D, temporelles, spectrales, polarimétriques, etc. Il est très important de définir des approches qui permettent d'acquérir l'information utile pour une application donnée de la manière la plus efficace, simple et économe possible. Elles trouvent des applications dans des domaines à fort potentiel d'innovation tels que l'imagerie biomédicale, le multimédia, la télédétection ou le contrôle des processus industriels.

Pour atteindre cet objectif, on doit prendre en compte la nature physique de la scène observée, de l'instrument qui acquiert les données et des algorithmes de traitement qui extraient l'information pertinente. Ce domaine est donc par essence multidisciplinaire, et l'objectif de cette action est de favoriser l'interaction entre les utilisateurs des informations (physiciens, biologistes et médecins, spécialistes de l'environnement, etc.), les concepteurs de systèmes d'imagerie et les traiteurs de signaux/images, dans le but d'améliorer l'efficacité globale des systèmes d'imagerie. Cet axe se situant à l'interface entre la physique, les mathématiques appliquées et le traitement d'image, la plupart des journées seront organisées en collaboration avec des GdRs tels que Ondes, MIV (Microscopie et imagerie du vivant), ou MIA.

2. Action Nouveaux enjeux en télédétection

Animateur(s) : Abdourrahmane Atto (MCF, LISTIC, Université Savoie Mont Blanc), Loïc Denis (MCF, Laboratoire Hubert Curien, Telecom Saint-Etienne), Stéphane May (CNES).

Description : Les nouvelles générations de satellites d'observation de la Terre sont déployées en constellations pour augmenter de manière très significative les résolutions spatiales, spectrales et temporelles des images numériques des sites observés. Il en est de même en imagerie astrophysique : les dernières générations de systèmes de télédétection sont déployées par exemple en constellations de télescopes imageurs pour améliorer l'observabilité de rayonnements rares dans le temps. D'autre part, la généralisation de l'usage d'imageurs embarqués sur des drones révolutionne l'imagerie aéroportée au sens où le coût faible (en comparaison avec l'utilisation d'avions ou d'hélicoptères dans les précédentes décennies) de tels systèmes rend possible des revisites régulières de sites sensibles et/ou peu accessibles.

La grande quantité des données acquises et les nouvelles caractéristiques de ces images constituent de nouveaux défis pour leur traitement et l'extraction de l'information dans ces images de télédétection. Ces défis imposent à la communauté de se coordonner et de partager ses compétences sur de nouvelles approches en analyse des images de télédétection. Les nouveaux défis soulevés sont relatifs, entre autres, à la multi-modalité, au traitement de séries temporelles d'images de plus en plus denses, à l'apprentissage statistique de régularités ou d'évolutions spatio/spectro-temporelles caractérisant différents phénomènes dynamiques.

3. Action Vision guidée par les capteurs émergents

Animateur(s) : Pascal Vasseur (PU, LITIS, Université de Rouen Normandie), Guillaume Caron (MCF, MIS, Université de Picardie Jules Verne), Franck Ruffier (CR CNRS, ISM, Aix Marseille Université).

Description : De nombreuses entreprises s'impliquent de plus en plus dans la fabrication et la commercialisation de nouveaux types de caméras comme les caméras à temps de vol, les caméras plénoptiques, les caméras omnidirectionnelles, les caméras

événementielles et autres capteurs visuels. Ces nouvelles caméras et les nouveaux usages qu'elles laissent entrevoir sont le fruit de progrès d'intégrations techniques et technologiques récents. Ces caméras permettent d'obtenir des informations de natures différentes comme la profondeur, le mouvement, la couleur ou encore de longueurs d'ondes différentes (IR, FIR, SWIR).

L'usage de ces capteurs fait émerger de nombreuses questions sur les outils et les méthodes associées. En premier lieu, se pose le problème de la modélisation et de l'étalonnage. Ce thème a été profondément traité pour les caméras classiques, voire également les caméras omnidirectionnelles, avec des résultats significatifs et des outils performants. Néanmoins, il existe un intérêt important pour l'amélioration de ces modélisations géométriques et/ou photométriques, ainsi que pour le développement d'outils de mise en œuvre rapides, précis et robustes. De même, la nature hétérogène des informations mesurées nécessite le développement de principes spécifiques pour l'étalonnage extrinsèque des caméras comme ce fut le cas pour les caméras et les Lidars par exemple.

Du point de vue du traitement de l'information, de nombreuses pistes de recherche sont également à explorer pour l'extraction de primitives, la caractérisation de ces primitives, la mise en correspondance. Ainsi, il est pertinent d'étudier si les traitements classiques sont applicables à ces modalités et s'ils sont optimaux dans tous les cas. Une question intéressante concerne les traitements précoces qu'il est possible voire souhaitable d'effectuer au plus proche du capteur. La collaboration entre capteurs interroge sur l'usage de l'information d'une modalité pour optimiser le traitement d'une autre. L'étude des gains et des pertes en termes de robustesse, précision, incertitude, temps de calcul avec ces modalités par rapport aux méthodes de vision classiques ainsi les conditions optimales d'utilisation sont d'un intérêt fondamental.

Sur ces questions, plusieurs journées seront organisées en lien avec le GdR Robotique.

4. Action Interaction avec l'humain : qualité et perception

Animateur(s) : Qualité : Patrick Le Callet (PU, LS2N, Polytech'Nantes), Frédéric Dufaux (DR, L2S, Gif-sur-Yvette), Christophe Charrier (MCF, GREYC, Université de Caen Normandie), Hakim Saadane (MCF, XLIM-SIC, Université de Nantes), Giuseppe Valenzise (CR, L2S, Gif-sur-Yvette), Aladine Chetouani (MCF, PRISME, Polytech Orléans) ; *Perception* : Jesus Angulo (CR, CMM, MINES ParisTech), Michel Berthier (PU, MIA, Université de La Rochelle), Philippe Carré (PU, XLIM, Université de Poitiers), Laurent Condat (CR, GIPSA-lab, Grenoble), Julie Delon (PU, MAP5, Université Paris-Descartes), Edoardo Provenzi (PU, IMB, Université Bordeaux) .

Description : L'objectif de cette action est de mettre en relation les chercheurs travaillant sur la modélisation des propriétés de la vision humaine et ceux intéressés par les applications de ces propriétés dans le domaine du traitement des images et de la vision par ordinateur. L'avantage de cette confrontation est double : les applications bénéficient souvent des connaissances sur le fonctionnement du système visuel humain ; inversement, des simulations et des expériences avec des images numériques sont souvent indispensables pour inférer des propriétés visuelles. Un exemple récent de ce type d'interaction est donné par le workshop EUVIP (European Workshop on Visual Information Processing) qui s'est tenu à Marseille en octobre 2016, dans lequel plusieurs chercheurs en neurosciences, informatique, psychologie cognitive et mathématiques appliquées au traitement d'images ont présenté leurs travaux et entamé des collaborations. Une liste non-exhaustive de champs de recherche intéressés par cette action est donnée par :

- l'imagerie à gamme dynamique étendue (HDR - High Dynamic Range)
- l'étude de la qualité des images et des vidéos
- les statistiques des images naturelles
- les métriques et la géométrie des espaces des couleurs
- le rehaussement du contraste
- la restauration d'images et des vidéos
- la colorimétrie
- la métrologie des attributs perceptuels des surfaces.

Cette thématique demande évidemment des compétences pluridisciplinaires, ce qui implique des interactions avec d'autres GdRs, notamment les GdRs Vision et MIA. Des contacts pour l'animation de journées communes ont déjà été formalisés avec Jean-François Aujol (PU, IMB, Université de Bordeaux) et Nicolas Papadakis (CR, IMB, Université de Bordeaux) du GdR MIA, et avec Pascal Mamassian (DR CNRS, LSP, ENS Ulm), Laurent Madelain (PU, SCALab, Université Lille), Simon Thorpe (DR, CerCo, Toulouse), Anna Montagnini (CR, INT, Marseille) et Laurent Perrinet (CR, INT, Marseille) du GdR Vision. Notamment, Laurent Madelain a répondu avec enthousiasme à la proposition d'une rencontre avant le forum annuel du GdR Vision, à Paris, en automne 2018, organisé par Thérèse Collins et Céline Paeye.

L'évaluation de la qualité des images et des vidéos joue un rôle central dans cette action. Elle se concentre sur l'intégration de l'observateur dans la chaîne de traitement et d'évaluation de la qualité des images et des vidéos. Les domaines d'application visés par cette action sont : la réalité virtuelle ou augmentée, l'imagerie à grande gamme dynamique (High Dynamic Range Imaging ou HDRI) et autres formats immersifs, les dispositifs d'affichage d'images 3D ou lightfield, les processus de protection des données et des personnes (biométrie, vidéo-protection), l'imagerie bio-médicale (tomographies), l'évaluation de la qualité avec ou sans référence. Les problématiques et enjeux concernent les métriques, les protocoles d'évaluation, les techniques d'évaluation automatique par apprentissage, les mesures de saillance, la qualité d'expérience et l'intégration de l'observateur dans la chaîne d'évaluation de la qualité. La conférence annuelle Qomex (Qualité de l'expérience visuelle), qui gagne sans cesse en importance, est entièrement dédiée à la recherche sur la qualité.

5. Action Analyse de données massives en imagerie du vivant

Animateur(s) : Carole Lartizien, (CR, CREATIS), Caroline Petitjean (MCF, LITIS, Normastic, Université de Rouen), Su Ruan (PU, LITIS, Normastic, Université de Rouen).

Description : Les sciences de l'information tiennent aujourd'hui une place primordiale dans l'analyse des données produites en Sciences du Vivant. Il s'agit d'une part d'organiser les masses de données hétérogènes, multi-échelles et multimodales produites dans les hôpitaux et les centres de recherche, et d'autre part d'extraire l'information pertinente pour mieux soigner ou comprendre les mécanismes physiopathologiques.

Les travaux de recherche dans ce domaine doivent être réalisés dans un contexte pluridisciplinaire rassemblant des acteurs des sciences des données, de l'ingénierie et du vivant, afin de concevoir et développer des méthodes adaptées aux spécificités des données biomédicales. L'objectif de cette action est de favoriser les interactions entre les différentes communautés scientifiques autour de thèmes de recherche émergents et fédérateurs en science de l'information.

Cette action s'articulera autour de plusieurs journées scientifiques associant des chercheurs seniors pour des exposés tutoriels et des jeunes chercheurs pour la présentation de travaux plus focalisés afin de favoriser la qualité des échanges et la création de dynamiques communes. Certaines journées, organisées autour de thèmes ciblés des sciences des données particulièrement intéressants pour l'analyse des données biomédicales, viseront à favoriser le transfert des connaissances. D'autres journées seront centrées autour de grands thèmes de l'analyse des données biomédicales. Elles viseront à identifier les verrous scientifiques ou technologiques dans ce domaine, et à dresser un état de l'art des méthodes des sciences de l'information les plus performantes pour les résoudre ainsi que des axes de recherche à développer, notamment en extraction d'information, en modélisation d'image, fusion de données, fouille de données, et apprentissage.

6. Action Vision 3D et apprentissage

Animateur(s) : Adrien Bartoli (PU, Institut Pascal, Université Clermont Auvergne), Pierre Gurdjos (IR, IRIT, ENSEEIHT), Vincent Lepetit (PU, LABRI, Université de Bordeaux).

Description : Cette action s'intéressera à la reconstruction 3D d'un objet ou d'une scène rigide ou déformable (SfM, NRSfM - MVS), au recalage d'images multimodales (RGB, RGB-D, depth), au calcul de pose d'un objet ou d'une scène rigide ou déformable (pose, SfT) à partir de modèles texturés ou non-texturés. Alors que ces problèmes peuvent sembler résolus, ils restent néanmoins ouverts par certains aspects relatifs aux limites des méthodes géométriques actuelles.

A titre d'exemples, on peut citer :

- La vision 3D à partir d'images endoscopiques pour la réalité augmentée. Il s'agit d'une application clef de la vision 3D. Les images impliquées nécessitent de faire des progrès fondamentaux en vision 3D : structures déformables, surfaces peu texturées ou texture répétitive, présence de fumée et de fluides et changements de topologie.
- Le recalage rigide d'images et d'un modèle 3D non texturé obtenu par exemple par un scanner 3D, permettant la reconstruction de scènes d'extérieur à grande échelle. Les capteurs laser vont être accessibles à des prix ultra compétitifs et seront dans un futur proche utilisés pour raccourcir le cycle d'acquisition de scènes 3D. L'intégration des données 3D (souvent non texturées) qu'ils fourniront pose de nouveaux problèmes dans le cadre du SfM et du MVS (Multi-View Stereo).

Notons qu'aujourd'hui, les méthodes d'apprentissage profond peuvent également être développées en vision 3D et font l'objet de nombreux travaux. Les algorithmes sont capables de prédire d'importantes quantités de cartes de profondeur et de normales à partir d'une seule image. Il est à présent intéressant d'utiliser ces données pour généraliser les approches sur des données temporelles. De la même manière, l'étude du concept du self-learning, où l'apprentissage est effectué grâce à une tâche secondaire différente de la tâche primaire de reconstruction 3D, peut être fort utile dans le cas où les données d'apprentissage ne sont pas disponibles. Enfin, les algorithmes de SLAM récents se concentrent sur l'association d'un label sémantique sur chaque objet tout en reconstruisant la scène de manière automatique. Ce sujet de recherche récent et prometteur alliant reconstruction 3D et reconnaissance d'objet sera également étudié ici. Ainsi, dans cette action, il s'agit d'étudier les techniques géométriques de vision 3D en tenant compte de la spécificité des données (multi-modales, dynamiques, déformables)

et d'utiliser, dans une certaine mesure, les méthodes d'apprentissage en vision 3D de manière conjointe à la géométrie.

7. Action Visage, Geste et comportement

Animateur(s) : Catherine Achard (MCF, ISIR, Université Pierre et Marie Curie), Christophe Ducottet (PU, Laboratoire Hubert Curien, Université de Saint-Etienne), Olivier Alata (PU, Laboratoire Hubert Curien, Université de Saint-Etienne).

Description : Cette action concerne l'extraction du mouvement humain (visage, mains, corps, gestes) à partir de séquences vidéo, et son analyse à plus haut niveau (tâches, comportement), notamment pour des applications interactives ou de surveillance. Une attention particulière sera portée aux données hétérogènes (couleur, thermique, 3D, etc.) et/ou acquises dans des situations non contraintes.

Elle s'intéresse notamment à la prédiction de posture 2D ou 3D de la main ou du corps avec notamment des approches utilisant l'apprentissage profond, à la modélisation statistique de gestes, à la modélisation du visage (forme, apparence, émotion), et à la reconnaissance d'activités et de comportements. Des journées thématiques, co-organisées éventuellement avec le GdR Robotique, seront organisées sur ces sujets.

8. Action Apprentissage profond pour l'indexation multimédia

Animateur(s) : Jenny Benois-Pineau (PU, LABRI, Université de Bordeaux), Alexandre Benoit (MC, LISTIC Polytech'Savoie), Hervé Le Borgne (CR, Vision and Content Engineering Lab CEA Saclay), Hervé Bredin (CR, LIMSI, Orsay), Liming Chen (PU, LIRIS, Centrale Lyon).

Description : La structuration des contenus multimédia, notamment ceux impliquant une modalité visuelle, suscite un intérêt de la communauté scientifique mais aussi des acteurs industriels et institutionnels pour leurs besoins propres. Il émerge un besoin d'organiser des objets non seulement plus complexes ou de manière plus fine, mais aussi plus composite par le nombre de modalités mises en jeu. Le traitement et l'indexation de chaque modalité ont fait des progrès importants. Ainsi la recherche et la conception des descripteurs du contenu adaptés a cédé la place au renforcement des méthodes d'apprentissage, spécifiquement avec la popularisation des réseaux d'apprentissage profond. La combinaison des techniques propres à chaque domaine reste cependant un défi.

Le cœur de l'action portera sur la fusion et la combinaison des traitements de diverses modalités avec la modalité visuelle. Elle s'intéressera aux liens en termes de traitement et d'indexation par apprentissage statistique et plus particulièrement par apprentissage profond. Les défis propres à cette problématique concernent notamment l'hétérogénéité des approches en fonction des modalités, les différences de performances obtenues selon les médias. Dans le contexte de l'apprentissage profond, les modes de fusion optimale restent à étudier afin d'obtenir des scores suffisamment élevés pour l'utilisation des outils dans des problèmes réels avec une quantité de données d'apprentissage souvent faible et des annotations bruitées.

3.3 Thème C : Adéquation Algorithme Architecture pour le TdSI

Directeurs scientifiques adjoints :

- Christophe Jégo
PU61, IMS (UMR CNRS 5218), INP de Bordeaux, Université de Bordeaux
- Jean-François Nezan
PU61, IETR (UMR CNRS 6164), INSA de Rennes

3.3.1 Présentation générale du Thème C

Le Thème C joue un rôle transversal vis-à-vis des autres thèmes du GdR ISIS, dont sont issus les algorithmes à implanter, mais également vis-à-vis d'autres GdRs qui traitent des supports d'exécution tels que SOC2. Ce rôle est primordial puisqu'il consiste à rapprocher l'application de son support d'exécution. Ce positionnement permet d'ouvrir de nouvelles explorations architecturales, aussi bien au niveau algorithmique qui bien souvent est restreint par des capacités de calcul limitées, qu'au niveau architectural où des propriétés applicatives sont exploitables. Le cœur d'activité des membres du Thème C est l'**Adéquation-Algorithme-Architecture** (AAA) qui consiste à étudier simultanément les aspects algorithmiques et architecturaux en prenant en compte leurs interactions dans le sens algorithme vers architecture, et vice-versa. Ce positionnement original n'existe pas dans les GdRs connexes avec une telle diversité de domaines applicatifs. Cela constitue donc la force du Thème C au sein du paysage scientifique français.

Le positionnement du Thème C au sein du GdR ISIS permet à ses membres une forte interaction avec les membres des autres thèmes (A, B, D et T). Cette interaction vise à identifier, pour les concepteurs d'architectures ou les programmeurs de processeurs, les nouvelles applications et les défis associés. Elle permet également aux algorithmiciens de suivre l'évolution des technologies et des nouveaux paradigmes architecturaux. Cette veille technologique est indispensable afin de lever des verrous algorithmiques qui seraient essentiellement liés à une trop faible puissance de calcul.

Le Thème C a également pour objectif de proposer des méthodologies d'évaluation de la qualité des implémentations développées, aussi bien au niveau des performances (consommation, débit, latence, etc.) que des contraintes de conception (exécution temps réel, disponibilité des ressources, etc.). Ces méthodologies doivent permettre de mesurer le plus finement possible l'adéquation de l'architecture à l'algorithme et réciproquement.

3.3.2 Perspectives du Thème C

L'évolution des architectures d'exécution s'organise autour de deux grandes catégories de systèmes. Une première, bien connue, est constituée de systèmes possédant une grande puissance de calcul sans prise en compte d'autres contraintes. Une seconde qui se renforce jour après jour est la constitution d'une informatique dématérialisée où l'organe central n'est plus l'ordinateur traditionnel mais un ensemble de systèmes nomades coopérants entre eux pour réaliser les traitements liés à une application. Dans ce second cas, les contraintes d'exécution sont multiples et ne se résument pas à la seule considération du nombre d'opérations effectuées par seconde.

Les algorithmes de traitement du signal, ceux des communications numériques et les méthodes de traitement vidéo doivent être associés à l'une ou l'autre de ces catégories selon l'application et son domaine d'utilisation. Dans ce contexte, les nouvelles architectures de type GPU (Graphics Processing Unit) ou MPSoC (*MultiProcessor System on Chip*)

sont naturellement de bonnes candidates. L'adéquation entre ces architectures massivement parallèles et des algorithmes souvent séquentiels est loin d'être triviale. Il est dans ce cadre nécessaire, d'une part d'étudier la synergie existante entre l'algorithme à exécuter et le support d'exécution, et d'autre part d'étudier les méthodes permettant une extraction efficace du parallélisme présent dans ces algorithmes.

L'utilisation des architectures parallèles (FPGA, GPU, MPSoC) pour l'implémentation des algorithmes de traitement du signal et des images est en pleine mutation. De nouvelles méthodes de programmation parallèles sont proposées (OpenMP, OpenCL, OpenVX, etc.) par les fournisseurs de composants. D'un point de vue méthodologique, les développeurs se heurtent à l'évaluation de ces méthodes et sont dans l'impossibilité de pouvoir les tester toutes et sur toutes les plateformes disponibles sur le marché. D'autre part, ces méthodes de programmation ne sont pas supportées sur toutes les cibles et les codes ne sont pas portables d'une cible matérielle à une autre. Le Thème C doit donc permettre aux développeurs de partager leurs expériences sur ces nouvelles méthodes de programmation parallèles. La définition de modèles de programmation « hardware independant » reste un objectif d'actualité. L'exploration architecturale à partir de ces modèles doit permettre de faciliter le choix d'une cible matérielle pour une application donnée. De plus, la génération de code optimisé pour une cible devra permettre de passer des modèles de programmation « hardware independant » vers les nouvelles méthodes de programmation parallèles disponibles.

Au cours du prochain contrat, le Thème C organisera des journées permettant d'approfondir ces sujets dans la lignée des actions thématiques des dernières années. Pour traiter ces problématiques scientifiques, le Thème C sera organisé autour de trois actions thématiques :

1. Architectures et applications
2. Méthodes et outils pour le TdSI
3. Capteurs et caméras

3.3.3 Actions proposées

1. Action *Architectures et applications*

Animateur(s) : Fan Yang (PU, LE2I, Université de Bourgogne), Daniel Ménard (PU, IETR, Université de Rennes), Matthieu Gautier (MCF, IRISA, Lannion).

Description : Cette action a pour vocation de proposer des journées communes avec les autres thèmes du GdR ISIS afin de coupler les domaines algorithmes et architectures. L'objectif est d'explorer, d'une part les architectures dédiées, et d'autre part les processeurs multi-cœurs et many-cœurs pour des algorithmes de traitement du signal et des images. Les architectures dédiées à une seule application n'utilisent pas de jeu d'instructions. La spécificité qui les caractérise leur confère des performances maximales (surface, consommation, débit, latence) et un niveau de flexibilité très faible. Les processeurs généralistes présentent au contraire un degré de flexibilité maximale mais un débit moindre. Le temps de développement d'une application sur processeur avec jeu d'instructions est réduit car il correspond à l'écriture d'un programme dans un langage de plus ou moins haut niveau qui est compilé avant d'être exécuté sur l'architecture. Dans tous les cas, l'exploration se doit d'être menée en prenant en compte conjointement les contraintes algorithmiques et architecturales.

2. Action Nouveaux modèles de programmation et méthodologies pour le TdSI

Animateur(s) : Virginie Fresse (MCF, Laboratoire Hubert Curien, Université de Saint-Etienne), Loïc Lagadec (PR, LabSTICC, Brest), Olivier Muller (MCF, TIMA, Grenoble).

Description : La problématique des algorithmes, de l'arithmétique et des architectures robustes sera traitée dans cette action afin de couvrir les besoins croissants de sûreté de fonctionnement et de fiabilité des systèmes électroniques. Le développement de méthodes et d'outils pour le choix et l'utilisation de nouvelles plateformes matérielles doit prendre en compte les nouveaux modèles de programmation parallèles disponibles. Ces nouveaux modèles doivent être évalués et comparés pour être en mesure d'en proposer de nouveaux plus performants.

3. Action Capteurs et caméras

Animateur(s) : Gilles Sicard (CEA LETI, Grenoble), François Berry (PU, Institut Pascal, Université Clermont Auvergne), Matthieu Thevenin (CEA LIST, Saclay).

Description : Cette action s'intéresse aux capteurs 3D, qui permettent les traitements intégrés au sein du capteur, et aux architectures massivement parallèles afin d'aboutir à des solutions hautes performances à partir des informations provenant des pixels. Le problème de la faible consommation est également central au sein de cette action et sera abordé régulièrement. Les notions de *smart sensors*, *smart caméra*, *visual sensor network* (traitement distribué) seront également traitées avec la volonté d'intégrer des algorithmes intelligents de traitement du signal (décision, classification, etc.) au plus près du capteur.

3.4 Thème D - Télécommunications : compression, protection, transmission

Directeurs scientifiques adjoints :

- Chaker Larabi
Chargé de la thématique « Compression et protection »
MCF61, XLIM (UMR CNRS 7252), Univ. Poitiers
- William Puech
Chargé de la thématique « Compression et protection »
PU27, LIRMM (UMR CNRS 5506), Univ. Montpellier
- Ghaya Rekaya-Ben Othman
Chargée de la thématique « Information et communication »
PU, Communications and Electronics Department, Telecom ParisTech
- Maxime Guillaud
Chargé de la thématique « Information et communication »
Huawei Technologies, Paris

3.4.1 Présentation générale du Thème D

Le Thème D est un espace de rencontre pour la communauté travaillant sur les couches basses des télécommunications (théories de l'information et de la communication, traitement du signal, codage des données, réseaux radio), sur les problèmes de compression de source (image, vidéo, 3D) et sur la protection des données visuelles par insertion de données cachées (tatouage, stéganographie, *data hiding*), par détection de manipulation ou par chiffrement spécifique à ce type de données. Cette communauté représente actuellement plus de 200 chercheurs répartis dans une cinquantaine d'équipes de tailles variées.

Les thématiques de recherche choisies privilégient la transversalité et les collaborations avec d'autres GdRs, en particulier les GdRs ASR (au travers des problèmes de communications en réseau), IG (Informatique Graphique pour la compression géométrique et la 3D), IM (Informatique Mathématique), MACS (pour les problèmes de contrôle appliqués aux réseaux), Ondes (pour les recherches sur les canaux de propagation, ou la conception numérique-analogique), SEEDS (systèmes de distribution électrique), SOC-SIP (systèmes matériels-logiciels intégrés), et Sécurité Informatique (chiffrement, insertion de données, authentification, criminalistique).

Les domaines abordés en codage de source et en théorie des communications mènent à des travaux de recherche amont, mais aussi à des recherches plus appliquées, tournées vers une collaboration étroite avec les partenaires industriels, souvent impliqués dans les comités de normalisation (JPEG, MPEG, JVET, etc.).

Le projet scientifique du Thème D est organisé selon deux axes. Pour chacun des axes, nous avons différencié d'une part les grandes thématiques actuelles de notre communauté, et d'autre part celles qui sont émergentes et sur lesquelles la recherche française entend s'investir dans les prochaines années. Cette répartition n'a aucune prétention d'exhaustivité, et d'autres thématiques intéressantes peuvent bien entendu faire l'objet d'actions et de journées du Thème D.

Axe 1 : Compression et protection

Cet axe regroupe les chercheurs qui s'intéressent au codage de source. En plus des problématiques de base que sont la réduction de redondance, la quantification et le

codage, la communauté concernée est active dans les domaines de l'évaluation de la qualité des images, l'insertion de données cachées et la protection dans des études normatives et dans la prise en compte du canal de transmission. Plusieurs thématiques de recherche sont abordées dans cet axe dont : la compression/codage d'images, de vidéos, d'objets 3D et 3D+t (volumiques et surfaciques), le multi-échelle, le codage multi-sources, le codage de sources distribué, la protection basée contenu, l'insertion de données cachées, la compression « intelligente » (basée contenu, ROI, etc.), le bio-inspiré, les optimisations perceptuelles à différents points de la chaîne de codage, les nouvelles modalités d'acquisition ou de représentation des données visuelles (ultra HD, 360°, plénoptique, lightfield) et les approches conjointes.

Axe 2 : Information et communication : de la théorie à l'ingénierie

Cet axe regroupe les thématiques en relation avec la théorie de l'information, le traitement du signal, le codage de canal, l'estimation, la synchronisation, la détection, le calcul et l'optimisation distribués, ainsi que les problèmes énergétiques dans les réseaux de communications. Les réseaux de cinquième génération (5G), en cours de normalisation, et en particulier les nouvelles applications que ces réseaux vont supporter (réalité virtuelle, contrôle à distance, véhicules autonomes, etc.) impliquent une évolution drastique en termes de fiabilité, de latence, de tolérance aux pannes, et d'efficacité énergétique. Des outils théoriques, de nouvelles méthodes et algorithmes répondant à ces nouvelles contraintes devront voir le jour. Rajoutons à cela l'augmentation du nombre d'appareils connectés, qui amènent à considérer des réseaux de grande taille. Ceci pose la question du passage à l'échelle des approches théoriques actuelles, avec des algorithmes de couche physique classiquement développés pour un ou quelques utilisateurs, et qui devront bientôt supporter des millions de terminaux par kilomètre carré.

3.4.2 Perspectives du Thème D

Axe 1 : Compression et protection

De plus en plus de données numériques visuelles (images, vidéos et objets 3D) sont transmises, archivées et visualisées. Même si les bandes passantes des réseaux continuent d'augmenter, il reste obligatoire de passer par une étape de compression. D'une part, les résolutions et les fréquences des contenus ne cessent d'augmenter. D'autre part, les exigences des utilisateurs en bout de chaîne sont de plus en plus importantes par leur souhait d'accéder à des données en très haute qualité et en temps réel. De même, concernant la protection, de plus en plus de données sont extraites de réseaux sécurisés protégés par des DRM (*Digital Right Management*). De ce fait, il est nécessaire de protéger ces données à la source par chiffrement spécifique ou insertion de données cachées. Le Thème D souhaite accompagner des ruptures à venir dans ces domaines. Des thèmes futurs sont à envisager autour de la compression et la protection de grandes masses de données visuelles 2D, 3D, 4D à haute résolution pour des applications interactives, variées et ambitieuses.

Les actions « Compression et qualité » et « Compression et perception » des précédents contrats donnent ici le jour à une nouvelle action intitulée « Nouvelles tendances en compression des signaux visuels ». Cette action favorisera les échanges avec le Thème B, le Thème C et l'action « Sécurité et données multimédia ». La perspective principale de cette action est de s'intéresser plus particulièrement aux problématiques des approches de compression en rupture, du multi-vues, de l'imagerie 360, de l'holographie numérique, des approches bio-inspirées, du JPEG Pleno, du JPEG XR/XT/XL/**, du Post-HEVC. Des applications spécifiques seront mises en avant telles que la vidéo-surveillance, la réalité

virtuelle/augmentée/mixte, le médical et le multimédia. Au-delà des applications grand public, cette action s'intéressera à des applications de niche, qui nécessitent souvent une recherche amont plus importante et ont un impact sur les autres schémas de compression.

Concernant les perspectives de l'action « Protection des données visuelles » qui donne le jour à l'action « Sécurité et données multimédia », de nombreuses pistes sont encore à explorer. En partenariat avec le Thème B, la protection basée contenu, permettant d'aborder différemment la robustesse aux attaques désynchronisantes, semble très prometteuse, en particulier par l'utilisation des nouvelles méthodes d'extraction de points d'intérêts et des composantes couleurs. Le chiffrement sélectif, les fonctions de hachage perceptuel et les traitements des données visuelles dans le domaine chiffré (en partenariat avec le GdR IM) sont un autre exemple de perspectives à poursuivre. En association avec l'action « Nouvelles tendances en compression des signaux visuels » et le Thème B, de nouvelles métriques pour la protection des données visuelles doivent être développées en s'appuyant sur le système visuel humain, que ce soit pour la très haute qualité comme pour la très basse qualité (métriques de confidentialité). Les domaines autour de la criminalistique, la biométrie et l'authentification de données visuelles, très étudiés actuellement, seront également à explorer. Enfin, un effort important sera porté sur les données visuelles de types vidéo, objets ou scènes 3D.

Au cours des prochaines années, le thème D se concentrera sur les actions suivantes :

- Sécurité et données multimédia
- Nouvelles tendances en compression des signaux visuels

L'objectif de ces actions sera de fédérer des chercheurs autour d'une thématique spécifique et de créer ainsi une synergie. Pour ce faire, un travail important sera mené afin de re-cartographier le paysage français sur la compression et la protection de données. L'idée est de recenser toutes les équipes académiques ou industrielles travaillant dans les thématiques proches ou connexes à l'Axe 1 du Thème D. Ce travail minutieux permettra d'identifier les thématiques nouvelles et les nouvelles équipes travaillant sur ces sujets, mais également de sensibiliser sur l'importance du GdR ISIS en tant que forum d'échange scientifique. Des journées sur les approches en rupture pourront voir le jour, notamment avec l'implication des partenaires industriels. Les actions susmentionnées n'ont pas de vocation à être pérennes puisqu'elles doivent s'adapter au contexte de la recherche académique et industrielle. Une évaluation à mi-parcours (2 ans) conditionnera leur poursuite ou non.

Conscients que d'autres pistes doivent être explorées, nous souhaitons aussi favoriser l'organisation de journées thématiques en dehors du fonctionnement des actions. Nous pensons notamment à des journées sur les sujets suivants :

- Architectures pour le bio-inspiré *en partenariat avec le Thème C*
- Méthodes et outils pour le signal, l'image et la 3D *en partenariat avec le Thème C*
- Indexation – compression – protection *en partenariat avec le Thème B*
- Chaînes de distribution *en partenariat avec l'Axe 2 du Thème D*
- Application au domaine médical
- Etc.

Axe 2 : Information et communication : de la théorie à l'ingénierie

Cet axe riche et dynamique se nourrit des applications émergentes dans le domaine des télécommunications et du développement des domaines connexes. Au cours des dernières années, on a pu assister à l'essor de nouvelles thématiques telles que le codage en réseau (*network coding*), le Massive MIMO, la sécurité pour la couche physique, le stockage distribué, etc. Ces évolutions s'accompagnent du renouvellement et de l'utilisation de méthodes et outils mathématiques. A titre d'exemple d'appropriation positive de résultats théoriques, pour le développement d'applications pratiques, on peut citer l'utilisation des matrices aléatoires dans l'étude des systèmes Massive MIMO au cours de la dernière décennie. Ce dynamisme est également porté par une recherche en forte interaction avec le monde industriel, notamment les membres du Club des Partenaires du GdR ISIS.

Actuellement, sans prétendre à l'exhaustivité, les thématiques émergentes structurant les perspectives de l'Axe 2 concernent :

1. Apprentissage pour la couche physique (en lien avec le Thème T)

L'apprentissage s'avère aujourd'hui être un outil puissant à même de jouer un rôle dans le domaine des communications. En particulier, la couche physique est en mesure d'en bénéficier pour résoudre des problèmes d'optimisation, aussi bien à un niveau local qu'à un niveau global. Plus précisément, l'apprentissage peut être utilisé dans l'optimisation de paramètres d'un bloc de la chaîne de transmission tel que le décodage. Il peut aussi être mis en œuvre pour reconsidérer toute la chaîne de transmission, au travers d'une optimisation globale de bout en bout, incluant tous les blocs (codage, modulation, décodage, etc.), comme l'ont montré les développements théoriques récents autour des auto-encodeurs.

2. Transfert et récupération d'énergie

Le transfert d'énergie sans fil est un domaine mature connaissant déjà des applications commerciales, par exemple pour le rechargement d'un mobile. Toutefois, le transfert simultané d'énergie et d'information sur des bandes de fréquences a priori réservées aux communications reste un sujet de recherche ouvert. Les applications envisagées sont l'alimentation à distance d'appareils à faible consommation tels que des capteurs, qui devront pouvoir fonctionner pendant une à plusieurs décennies sans accès physique ni connexion filaire à un réseau de distribution d'énergie. Dans ce domaine, de nombreux problèmes restent à résoudre : formation de voie avec des contraintes simultanées de débit et d'énergie, conception des composants radiofréquences capables de traiter simultanément l'information tout en collectant de l'énergie, algorithmes de gestion dynamique des réserves d'énergie.

3. Lien entre cryptographie et couche physique

Les problèmes de confidentialité et de sécurité sont critiques lors d'échanges d'information au travers de tout système de communications. En complément des systèmes de cryptographie classiques, des schémas de codage sécurisés pour la couche physique ont été introduits récemment. Ils permettent ainsi de renforcer la sécurité globale des systèmes. Jusqu'à présent, le lien entre la sécurité en couche physique et la cryptographie n'a pas été mis en avant. Il est toutefois à remarquer qu'un outil mathématique unit déjà ces deux systèmes de sécurité, à savoir « les réseaux euclidiens » (*lattice*). Le codage utilisant les lattices est déjà mis en œuvre dans plusieurs applications comme les systèmes MIMO, le codage réseau, les schémas de cryptogra-

phie LWE et GGH. Il paraît donc opportun de mettre plus en avant le lien entre ces deux domaines.

4. Théorie des graphes dans les réseaux de télécommunications (en lien avec le Thème A)

Les réseaux de télécommunications se prêtent naturellement à une représentation sous forme de graphe, dans lesquels les sommets correspondent aux composants physiques (terminaux mobiles, antennes relais, nœud de routage et de calcul) ou logiques (services virtualisés). Les arêtes capturent les caractéristiques physiques et/ou logiques de l'ensemble du réseau (terminal en contact avec une antenne-relais, existence d'un lien de communication entre deux nœuds du réseau, présence d'interférence entre deux terminaux, etc.). Dans ce domaine, l'essor actuel de la théorie des graphes (fournissant des outils d'inférence de la topologie, d'échantillonnage, de filtrage, de transformée de Fourier définie sur le graphe — voir le thème A) permet d'envisager des possibilités nouvelles en termes de contrôle des performances et de la santé des réseaux de télécommunications, et d'optimisation de leur fonctionnement (allocation de ressources spectrales, association optimale entre mobiles et antennes relais, planification de déploiement, etc.).

5. Architectures Cloud RAN et Fog RAN, virtualisation et aspects distribués

L'augmentation de la puissance de calcul, et l'utilisation de composants matériels flexibles et réutilisables, ouvrent des perspectives nouvelles dans le domaine de l'architecture des réseaux de télécommunications. Dans ce contexte, la virtualisation des fonctions telles que le traitement du signal en bande de base, le routage, et certaines fonctions relatives aux applications en ligne (transcodage audio et vidéo, voire génération d'environnements virtuels dans les applications de réalité augmentée) permet une grande flexibilité : l'exécution des algorithmes qui fournissent ces fonctionnalités virtuelles peut être déplacée dynamiquement d'un centre de calcul à l'autre dans le réseau en fonction de leur charge ; elle peut être implémentée de manière centralisée (le *cloud RAN*, énergétiquement efficace), ou bien traitée au plus près de l'utilisateur (*edge computing*) si la latence est une contrainte forte pour l'application en question. Le terme *Fog RAN* désigne quant à lui la gestion dynamique des ressources de communication, de stockage et de calcul nécessaire constituant le réseau de télécommunications et ses services, depuis les couches basses jusqu'aux applications offertes à l'utilisateur. Il s'agit donc d'une thématique extrêmement riche en problématiques relatives à l'optimisation et au calcul distribués.

6. Latence et « Age of Information »

Les nouveaux champs d'application attendus dans la 5G mettent l'accent sur les aspects de fiabilité et de latence des communications. Cette dernière en particulier est critique, notamment pour les applications de réalité virtuelle avec retour haptique. Cette contrainte de latence est relativement nouvelle dans un domaine où une grande partie des technologies actuelles sont bâties sur la théorie développée par Shannon, dont une des contributions majeures était l'idée de transmettre de larges volumes de données en utilisant des codes asymptotiquement longs, incompatibles avec une faible latence. Le développement d'algorithmes pour le traitement du signal en bande de base incorporant des contraintes de latence adaptatives au type de trafic, sera par conséquent un défi majeur du développement de la 5G. Ce sujet est à rapprocher de la notion récemment popularisée d'« Age of Information » (âge

des données), qui consiste à tenir compte du fait que la valeur de certaines données expire dans le temps ; en particulier, certains paramètres de la couche physique des systèmes de communications ont une valeur ne dépassant pas l'ordre de la dizaine de millisecondes. Dans ce contexte, il peut être préférable d'obtenir rapidement un résultat approximatif, plutôt qu'un résultat exact qui serait disponible trop tard ; de la même manière, un traitement local peut être préférable à une fusion avec des données distantes, qui fournirait un résultat plus précis mais nécessiterait un coût en communication supplémentaire. Les outils analytiques et algorithmes incorporant de manière rigoureuse le compromis entre latence et précision de calcul restent à développer.

7. Codage et formes d'ondes en communications optiques

Les communications optiques ont connu une révolution ces dernières années grâce au développement de transducteurs optoélectroniques cohérents. Cette révolution était nécessaire pour suivre l'énorme augmentation des débits demandés dans les réseaux de base. Pour continuer de répondre à cette demande de débit sans cesse croissante, les communications sur fibres optiques ont besoin de fournir de nouvelles solutions afin d'éviter la congestion dans l'ensemble du réseau. Les principales solutions seraient d'une part l'utilisation de plus de canaux parallèles à travers des fibres multi-mode et multi-core. Pour cela de nouveaux schémas de codage et de décodage spécifiques seront nécessaires ; d'autre part, en atténuant voire supprimant les effets non linéaires notamment par l'utilisation de la transformée de Fourier non linéaire avec des récepteurs à retour de décision.

8. Décision distribuée dans les réseaux de communications

Les systèmes de communications de demain seront composés d'un grand nombre d'objets communicants entre eux. De nouvelles problématiques relatives au codage, stockage, calcul, sécurité, routage, allocation de ressources, optimisation dans les réseaux distribués seront posées. Cela nous amène naturellement à revoir la manière dont les décisions doivent être prises dans de tels réseaux. La décentralisation des traitements est un défi pour l'optimisation des réseaux de communications de demain afin de proposer des solutions, aussi bien localement que globalement pour le réseau distribué, qui offriraient de bons compromis d'efficacité, flexibilité, rapidité, performance et complexité.

3.4.3 Actions proposées

Axe 1 : Compression et protection

De façon plus ciblée, afin de mettre en lumière des points saillants des perspectives scientifiques qui viennent d'être dressées, des actions spécifiques seront menées dans l'Axe 1. Ces actions nous engagent à assurer l'organisation régulière de journées dédiées à ces thématiques. Elles n'excluent aucunement l'organisation d'autres journées, à la demande de la communauté.

1. Action *Nouvelles tendances en compression des signaux visuels*

Animateur(s) : M. Antonini (DR CNRS, I3S, Sophia Antipolis), C. Larabi (MCF, XLIM, Poitiers).

Description : Cette action s'intéressera aux approches de codage bio-inspirées, aux nouveaux schémas de compression, à l'optimisation qualitative et la prise en compte de la

qualité d'expérience, la stéréovision, le 3D, la réalité virtuelle et augmentée, les images omnidirectionnelles. Les mots-clés sont les suivants :

- Approches de compression en rupture
- Multi-vues, imagerie 360°, holographie numérique, etc.
- Approches bio-inspirées, JPEG Pleno, JPEG XR/XT/XL/**, Post-HEVC, etc.
- Applications : video-surveillance, médical, multimédia, etc.

Au minimum, deux journées sont prévues pour 2018-2019 :

- **Imagerie plénoptique : état des lieux et enjeux** (début 2018)
Chaker Larabi, Marc Antonini
- **Codage et implémentation post HEVC** (commune avec le Thème C, fin 2018)
Daniel Menard et Wassim Hamidouche (IETR, Rennes)

2. Action Sécurité et données multimédia

Animateur(s) : C. Fontaine (CR CNRS, Lab-STICC, Brest), W. Puech (PU, LIRMM, Montpellier).

Description : Cette action s'inscrit dans la continuité de l'action « Protection multimédia ». Elle a pour objectif d'animer la communauté autour de la sécurité des images, des vidéos et des objets 3D. Des journées sont envisagées sur la criminalistique, la biométrie, la stéganalyse, l'authentification et la protection en partenariat avec le Thème B, la protection basée contenu (robustesse aux attaques désynchronisantes, utilisation de points d'intérêts, composantes couleurs, psycho-visuel, en partenariat avec l'action « Compression et perception »), ainsi que l'évaluation de la qualité pour le tatouage (HR) et le brouillage (LR) avec le développement de métriques de confidentialité. Les mots-clés sont les suivants :

- Insertion de données cachées, stéganographie et stéganalyse
- Multimedia forensics (détection de manipulations, identification de capteurs)
- Traçage de traître et robust hash (fingerprintings actif et passif)
- Biométrie et sécurité des traits biométriques
- Vidéo surveillance
- Traitement des images dans le domaine chiffré
- Apprentissage sous contrainte d'adversaire
- Connections avec la sécurité de la couche physique

Au minimum, deux journées sont prévues pour 2018-2019 :

- **Comment concilier Big Data, identification des personnes, traçabilité des contenus et respect de la vie privée**
Caroline Fontaine (Lab-STICC, Brest), Teddy Furon (IRISA, Rennes)
- **Biométrie** (fin 2018)
Cathel Zitzmann (LM2S, UTT), Christophe Charrier (GREYC, Caen)

Axe 2 : Information et communication : de la théorie à l'ingénierie

Après concertation avec les participants à cet axe lors de l'Assemblée Générale, il a été décidé de ne pas structurer cet axe par des actions. Une animation autour des thématiques identifiées dans la section précédente sera favorisée. Cette animation consistera en l'organisation de journées thématiques intégrant si possible des tables rondes afin de favoriser les échanges et les discussions.

Les journées programmées à court terme (courant 2018) sont :

- **Strategic information transmissions – Interplay between network information theory and game theory**
Mael Le Treust et Tristan Tomala
- **New waveform for 5G**
Jean-Claude Belfiore et Ingmar Land
- **Coding theory and its applications to data storage, communication and security**
Iryna Andriyanova et Charly Pouillat
- **Plateformes d'évaluation des communications radio en réseau**
Hicham Khalife, Jean-Marie Gorce, Leonardo Cardoso

A moyen terme (2019), les journées que nous envisageons d'organiser porteront sur :

- Apprentissage pour la couche physique (en collaboration avec le Thème T)
- Transfert et récupération d'énergie
- Lien entre cryptographie et couche physique

Enfin, à plus long terme, sont envisagées des animations autour des thèmes suivants :

- Théorie des graphes dans les réseaux de télécommunications (avec le Thème A)
- Architectures Cloud RAN et Fog RAN, virtualisation et aspects distribués
- Latence et « Age of Information »
- Codage et formes d'ondes en communications optiques
- Décision distribuée dans les réseaux de communications

Il convient de souligner que cet axe suscite historiquement un très fort intérêt de la part des partenaires industriels du GdR ISIS, les sociétés actives dans les communications étant fortement représentées au sein du Club des Partenaires. Afin de nourrir ce lien privilégié, l'organisation d'animations telles qu'une journée « Problèmes industriels en Télécoms » sont envisagées, lors desquelles les industriels pourront présenter à la communauté académique les verrous technologiques actuels pour lesquels des avancées théoriques sont nécessaires.

3.5 Thème T - Apprentissage pour l'analyse du signal et des images

Directeurs scientifiques adjoints :

- Christian Wolf (MCF HDR, LIRIS, CITI, INRIA, Lyon)
- Nicolas Thome (PU, CNAM, Paris)

3.5.1 Présentation générale du Thème T

L'apprentissage automatique est une force motrice majeure dans les récentes évolutions des Sciences de l'Information. Les avancées effectuées, par exemple en vision par ordinateur, ont eu un impact à la fois sur le monde académique et sur le monde industriel. Les méthodes d'apprentissage profond se sont établies comme référence pour un grand nombre de problèmes en remportant les plus prestigieuses compétitions scientifiques. A ce succès académique et industriel s'ajoute une visibilité croissante pour le grand public, témoignée par les nombreux articles parus dans la presse. Les conférences de vulgarisation se multiplient et, depuis peu, l'ensemble des médias se sont saisis du sujet.

L'apprentissage automatique connaît une forte croissance interne (doublement de la taille des congrès en quelques années) et externe, à la fois dans des disciplines recourant naturellement à des méthodes d'apprentissage (signal, image, vision, texte, intelligence artificielle, etc.), que dans des disciplines contribuant au développement d'outils nécessaires à sa mise en œuvre et à sa compréhension (optimisation, statistique, etc.). Le traitement du signal et de l'image a profondément muté grâce à ce succès, fort de l'avènement du deep learning et, plus généralement, des développements de l'Intelligence Artificielle. Cette réussite repose à la fois sur des avancées en termes de modélisation de l'information et de représentation des connaissances, sur des méthodes d'apprentissage statistiques exploitant des données massives annotées, et enfin sur des calculateurs puissants parfaitement adaptés aux modes de calcul associés (GPU, accélérateurs dédiés). Les méthodes de modélisation recourant à des architectures neuronales profondes, en particulier les réseaux de neurones convolutifs et les réseaux de neurones récurrents, sont aujourd'hui omniprésentes. De nombreux champs s'ouvrent dans la discipline du traitement du signal et des images afin de revisiter certaines tâches, en vision par exemple pour la segmentation sémantique, la détection d'objets ou la reconnaissance d'actions, en faisant collaborer ces modèles avec d'autres comme les modèles graphiques ou plus généralement les méthodes par raisonnement. Il est intéressant de noter que l'émergence de l'apprentissage profond comme thématique phare de l'Intelligence Artificielle est généralement associée aux progrès en rupture obtenus en reconnaissance de la parole (les avancées en 2011 par Microsoft) et en classification d'images (la compétition ILSVRC/ImageNet gagnée en 2012 par l'Université de Toronto). Enfin, il convient de mentionner que les trois exemples remarquables cités dans le livre blanc sur l'Intelligence Artificielle publié en 2016 par l'INRIA relèvent du traitement d'images.

La discipline du traitement du signal et des images connaît une évolution de paradigme qui s'est progressivement installée au cours de la dernière décennie. Compte tenu du nombre sans cesse croissant de chantiers ouverts qui lui sont dédiés, le GdR ISIS a souhaité en tenir compte et accompagner les évolutions de son métier en consacrant une place entière à l'apprentissage par son nouveau Thème T. Celui-ci est par nature **transverse** au sein de l'organisation du GdR, compte tenu des multiples interactions possibles en apprentissage automatique avec les Thèmes A, B, C et D. L'intérêt porté par le GdR ISIS sur l'apprentissage automatique est à deux niveaux :

- promouvoir l'utilisation des dernières techniques d'apprentissage, en particulier les réseaux de neurones profonds, dans des tâches de traitement de signaux qui, jusqu'ici, y avait peu recours ;
- développer et analyser de nouvelles classes de méthodes pour l'apprentissage.

En tant que structure transverse, le Thème T aura des interactions fortes avec les quatre autres thèmes du GdR ISIS :

- Le Thème A porte sur les méthodes et modèles en traitement du signal et des images. Un lien naturel existe essentiellement avec les travaux méthodologiques du domaine : apprentissage et optimisation, généralisation, apprentissage faiblement supervisé, adaptation et transfert de connaissances, modélisation des structures, explicabilité des résultats, etc.
- Le Thème B, « Image et Vision », est sans nul doute le thème le plus en prise avec l'apprentissage automatique. A l'heure actuelle, l'intégralité des traitements en vision par ordinateur est réévaluée suivant ce prisme, même si l'apprentissage est souvent couplé avec des méthodes issues de la géométrie et du traitement d'image.
- Le Thème C, « Adéquation Algorithme Architecture », se trouve au cœur des efforts consacrés à la maîtrise de l'apprentissage automatique pour les grandes masses de données à l'aide d'architectures massivement parallèles (GPU, CPU *many/hyper cores*) et d'architectures dédiées (Google TPU, IBM True North, NVIDIA Volta, Intel Nervana, etc.).
- Le Thème D, « Télécommunications », a sans doute été historiquement moins touché par l'apprentissage profond que la vision par ordinateur. Il existe néanmoins des travaux dans le domaine, et une accélération nous semble probable.

Une propriété très intéressante de l'apprentissage réside dans le caractère générique des modèles dans les divers champs applicatifs. Ainsi, un grand nombre de modèles sont utilisés à la fois pour la vision par ordinateur, pour le traitement automatique des langues, pour la reconnaissance de la parole, pour la robotique, etc. En conséquence, il nous a paru important de garder des liens forts avec les communautés respectives et d'organiser des journées communes.

Interactions avec d'autres GDRs — De par son sujet, le thème T a des interactions fortes avec le **GDR IA** (« **Aspects Formels et Algorithmiques de l'Intelligence Artificielle** »). L'apprentissage automatique étant à l'intersection des deux GDRs, une partie des activités sera co-organisée par les deux structures, impliquant plusieurs actions du Thème T. Les discussions sont en cours.

Nous proposons également une action commune avec le **GDR Robotique** sur l'apprentissage (Cf. Section 3.5.3, Action n° 2.).

3.5.2 Perspectives du Thème T

Le succès récent de l'apprentissage statistique et du deep learning en particulier est essentiellement dû à des aspects opérationnels, principalement liés à l'augmentation massive du volume de données annotées et de la puissance de calcul (e.g., GPU). Au-delà de ce constat, de nombreuses problématiques restent ouvertes pour en étendre et généraliser la mise en œuvre à d'autres types de problèmes et données.

Afin de dégager de nouvelles perspectives pour le Thème T, nous avons sollicité la communauté à l'automne 2017 afin de faire émerger des orientations pertinentes. Sont particulièrement remerciés les acteurs de la communauté qui ont contribué par écrit à ces perspectives : Francis Bach, Matthieu Cord, Marc Sebban, Amaury Habrard, Rémi Bardenet et Stéphane Canu.

Une synthèse de ces retours a permis de dégager les cinq axes scientifiques majeurs qui structurent la prospective scientifique du Thème T :

1. **Formulation d'apprentissage et architecture** : le succès spectaculaire de l'apprentissage profond a été particulièrement marqué pour des tâches d'apprentissage supervisé et de classification. Dans ce contexte, la capacité du deep learning a bénéficié d'une formulation d'apprentissage (discrimination) clairement définie. Pour d'autres problématiques classiques d'apprentissage tel que le contexte non supervisé, la formulation d'un critère d'optimisation clair pour l'apprentissage de représentations reste encore largement ouverte. Des solutions prometteuses récentes d'« auto-supervision » exploitent la nature des données ou du problème pour convertir un problème d'apprentissage non-supervisé en apprentissage supervisé. On peut citer par exemple word2vec pour l'apprentissage de représentations textuelles ou les très populaires modèles génératifs adversaires (Generative Adversarial Networks, GAN). Le coût des annotations est également un problème qui motive la mise en place de formulations d'apprentissage alternatives. Ainsi, développer des modèles permettant de pouvoir bénéficier des données massivement disponibles mais moins massivement annotées est un enjeu actuel majeur. Développer des solutions d'apprentissage semi-supervisé ou faiblement supervisé est une piste pour cela, mais la formulation de ces problèmes est aujourd'hui également largement ouverte. L'apprentissage avec labels bruités est également une question centrale dans le contexte actuel du big data.

En ce qui concerne l'apprentissage profond, l'architecture des réseaux convolutifs profonds pose des questions cruciales pour répondre à une problématique particulière : détection et reconnaissance à large échelle, segmentation sémantique et segmentation d'instances, détection et estimation de la posture 3D, etc. Le choix d'une architecture est également étroitement lié à des questions d'optimisation, de sur-apprentissage et d'empreinte mémoire des algorithmes qui conditionnent leur bon comportement et déploiement.

Enfin, la mise en place de solutions d'apprentissage permettant de prédire des sorties « structurées », c'est-à-dire pour lesquelles les variables de sortie sont corrélées, est une question importante pour dépasser les succès récents obtenus pour des problèmes simples de catégorisation. L'utilisation de modèles graphiques pour la résolution des problèmes centraux en vision, comme la segmentation d'images ou l'estimation de pose, a été intensivement étudiée. L'apprentissage structuré pose des questions spécifiques dans le contexte du deep learning moderne, notamment au niveau de la capacité d'inclure l'étape de prédiction (inférence) lors de l'entraînement des modèles afin de pouvoir bénéficier d'une optimisation globale du réseau ("end-to-end learning").

2. **Passage à l'échelle et transfert** : Un enjeu important en Machine Learning concerne la capacité à traiter des volumes d'informations diverses à grande échelle, et pour cela, à mettre en place des algorithmes d'optimisation efficaces. Si les méthodes de descente de gradient stochastique ont été la clé du déploiement des algorithmes

d'apprentissage pour les réseaux de neurones profonds, l'accélération de la convergence reste un enjeu crucial, et de nombreuses variantes ont été proposées ces dernières années. L'apprentissage distribué est également une solution pour le passage à l'échelle des algorithmes qui s'inscrit naturellement dans le contexte du big data et des objets connectés. Les systèmes potentiellement décentralisés sont une piste intéressante par exemple pour préserver les aspects privés.

Un autre point important pour l'apprentissage sur données plus au moins massives est la possibilité d'exporter les connaissances apprises d'une tâche à l'autre (« transfer learning »). Pour la classification d'images, la base ImageNet a été un élément majeur permettant l'extraction de représentations apprises par des réseaux convolutifs (« Deep features ») dont la capacité de transfert s'est avérée spectaculaire pour un très grand nombre de tâches de reconnaissance visuelle. Pour aller plus loin que cette approche de transfert naïve, de nombreux travaux restent à mener pour modéliser la corrélation entre la base cible et source, ou pour chercher explicitement à mettre en place des algorithmes d'adaptation de domaine (par exemple en utilisant des méthodes de transport optimal). Dans ce cadre, la formulation d'apprentissage inclut souvent différents objectifs ou tâches (apprentissage multi-tâches), ce qui nécessite une réflexion particulière pour le développement des algorithmes.

3. **Compréhension formelle des réseaux profonds** : en dépit de leur performances prédictives spectaculaires, la compréhension des architectures profondes est encore fragile et limitée, notamment aux niveaux suivants :
 - (a) Optimisation non convexe : frein historique à l'utilisation des réseaux de neurones profonds, la nature non convexe de la fonction objectif d'apprentissage ne semble pas être un problème majeur pour bon nombre de problèmes concrets. Des travaux récents ont apporté certains éléments de réponse à cette observation empirique, à travers l'étude des minima locaux et des points selles de la fonction objectif. D'une manière générale, la compréhension plus fine de la structure de la fonction d'apprentissage constitue assurément une source d'inspiration importante pour la mise en place de formulations d'apprentissage ou d'architectures bénéficiant de garanties de convergence fortes.
 - (b) Incertitude de prédiction : les réseaux de neurones profonds ne bénéficient pas naturellement d'une mesure de confiance de prédiction. Ceci est une limitation rédhibitoire pour pénétrer certains champs applicatifs critiques comme la conduite autonome, la médecine ou le nucléaire. Le recours aux réseaux de neurones bayésiens, bien que solution naturelle pour modéliser l'incertitude, se trouve cependant rapidement limité par un coût calculatoire prohibitif. Proposer des solutions permettant à la fois de modéliser l'incertitude d'un réseau profond et d'assurer un entraînement raisonnable sur données massives est un enjeu actuel très important. En ce sens, des travaux préliminaires établissent des liens entre la technique de dropout (régularisation moderne de référence) et les méthodes d'inférence variationnelles approchées.
 - (c) Stabilité : la stabilité à différentes déformations est une propriété importante dans système d'apprentissage et liée à la capacité de « *manifold untangling* » des réseaux profonds. Certaines architectures profondes particulières inspirées de la littérature du traitement du signal et des ondelettes (scattering) ont montré leur capacité de stabilité et/ou l'invariance par rapport à certaines transformations particulières (difféomorphismes). À l'inverse, les exemples adversaires illustrent l'incapacité des réseaux à assurer une stabilité par rapport

à d'autres familles de déformations. Une analyse de la stabilité vis-à-vis de déformations plus générales et apprises permettra certainement de mieux comprendre le fonctionnement des réseaux de neurones profonds.

- (d) Théorie de la généralisation : étonnement, les réseaux de neurones profonds semblent robustes au sur-apprentissage. Les outils classiques de l'apprentissage statistique comme la PAC semblent insuffisants pour expliquer ces très bonnes performances prédictives. Le développement d'un cadre théorique pour les méthodes d'apprentissage de représentations semble crucial, bien qu'encore à ses balbutiements.

4. Traitement de données multi-modales : le succès de l'apprentissage automatique a fait émerger de nouvelles applications et thématiques traitant de données variées, hétérogènes et multi-modales :

- (a) Vision et langage, comprenant des sujets tels que l'indexation multi-modale et le VQA (*Visual Question Answering*). Au cœur de cette thématique se trouve la fusion de texte et de signaux visuels, nécessitant des compétences de domaines différents. Les verrous scientifiques vont de l'alignement de ces sources d'information, la prédiction d'une source à partir d'une autre jusqu'au raisonnement.
- (b) Données audio-visuelles : il s'agit sans doute du cas le plus classique intervenant dans de nombreuses applications telles que les interactions humain-machine, la robotique et l'indexation multi-média. Au cœur des défis sur ces questions, se trouvent la fusion des modalités et l'exploitation de leurs complémentarités.
- (c) Géométrie et sémantique : la géométrie est traditionnellement utilisée pour la reconstruction 3D et pour la génération de cartes métriques dans le cadre d'applications telles que la robotique, les véhicules autonomes, la mobilité, etc. L'utilisation d'étiquettes sémantiques a émergé comme une alternative ou comme une source complémentaire à la géométrie. Souvent prédites par des réseaux de neurones profonds, les étiquettes sémantiques permettent de raisonner sur la présence d'objets spécifiques et sur la classe d'objets.

5. Apprentissage et connaissances : les problèmes de raisonnement ont longtemps été traités principalement par des méthodes symboliques basées sur un ensemble de règles, déterminées par des experts ou apprises à partir de données. Le raisonnement sur des données complexes (images, sons, grands corpus de textes) a démontré les limites de ces types d'approches. L'apprentissage statistique semble être une alternative puissante, capable de gérer le bruit et les incertitudes inhérents à ces données.

Les interactions avec le domaine TSI se trouvent autour du raisonnement à partir de signaux de bas niveau (images et sons), où l'interprétation des signaux (le fossé sémantique) est un enjeu important.

3.5.3 Actions proposées

1. Action Apprentissage des représentations — « Deep learning »

Animateur(s) : Christian Wolf (MCF HDR, LIRIS, Lyon), Nicolas Thome (PR, CNAM, Paris).

Description : L'objectif de l'apprentissage de représentations (surtout connu par son anglicisme Deep Learning) est l'apprentissage automatique, à partir de données diverses, de modèles hiérarchiques et basés sur des niveaux d'abstractions. Contrairement aux modèles classiques, les caractéristiques/représentations sont apprises à partir

de données et non pas conçues manuellement à partir de connaissances de métiers. L'apprentissage à partir de grandes masses de données étiquetées est actuellement la norme. Diminuer la dépendance des données est un enjeu essentiel. Entre autres, cette action s'intéresse aux verrous scientifiques suivants :

- Les divers formalismes de l'apprentissage sont au cœur de cette action : apprentissage supervisé, non supervisé, semi-supervisé, faiblement supervisé. L'enjeu principal consiste à diminuer la dépendance des masses de données annotées, l'annotation par des experts humains étant considérée comme difficile et chronophage. Dans les situations où les performances obtenues dépassent les performances humaines, l'annotation par les humains est impossible. Si pour certaines applications, les annotations peuvent être obtenues à partir de sources alternatives, comme cela est le cas pour la lecture labiale ou pour l'estimation de l'âge d'une personne⁸, pour d'autres applications cela s'avère difficile.

Le défi phare consiste à apprendre, de manière entièrement non supervisée, des représentations uniques et riches, permettant de répondre à des tâches multiples et diverses de la vie quotidienne (reconnaissance visuelle, prédictions, etc.). Similaire à l'apprentissage humain, effectué en grande majorité de manière non supervisée, il s'agit de découvrir les régularités qui gouvernent notre monde physique pour apprendre des abstractions utiles pour le raisonnement.

- L'apprentissage contemporain est étroitement lié aux données massives, ce qui soulève la question de leur acquisition. La création de données synthétiques, souvent par simulation, permet de répondre à ce problème dans certains cas. Dans ce contexte, la notion de transfert de connaissances, supervisé ou non-supervisé, joue un rôle important (transferts entre distributions de données sources et cibles, entre modèles, entre tâches, entre applications, etc.).
- Cette action considère également les modèles permettant d'apprendre des sorties structurées et les modèles génératifs permettant de générer du contenu (GAN, VAE, Pixel CNN et Pixel RNN, etc.). Le contenu prédit peut servir à des applications diverses telles que le graphisme (cinéma, jeux vidéos), la modification automatique et créative du contenu, la prédiction automatique du « futur » dans une vidéo, etc. L'espace de sortie étant de très grande dimension, les verrous scientifiques sont multiples : manque de stabilité des problèmes d'optimisation sous-jacents (estimation de l'équilibre de Nash des GAN), démélange des représentations/estimation et identification des variables latentes pertinentes pour l'application.
- L'action s'intéresse aux incertitudes des systèmes d'apprentissage et à la conception de modèles tractables permettant d'obtenir des mesures de confiance sur les prédictions données, par exemple les formalismes bayésiens. Les enjeux sont forts dans certains contextes applicatifs tels que l'imagerie médicale et la conduite autonome. Nous étudierons également l'explicabilité des modèles issus de l'apprentissage.

Cette action interagira fortement avec les autres actions du Thème T et des autres thèmes.

8. Pour les deux applications, l'apprentissage automatique a récemment permis de dépasser les performances humaines.

2. Action Vision, robotique et apprentissage – Co-animation ISIS/Robotique

Animateur(s) : David Filiat (PU, ENSTA ParisTech), Christian Wolf (MCF HDR, LIRIS, CITI, INSA-Lyon).

Description : La vision par ordinateur et la robotique partagent un passé commun qui remonte à la naissance de ces deux domaines. En effet, parmi l'ensemble des capteurs dont peut disposer un robot moderne, la vision peut répondre aux exigences liées à un grand nombre de problèmes applicatifs, par exemple en navigation, en manipulation ou dans les cas où la présence d'humains est un facteur important. Les tendances actuelles montrent que les liens entre les deux domaines se resserrent pour plusieurs raisons : (i) d'un point de vue méthodologique, l'émergence de l'apprentissage profond a permis de faire des liens nouveaux entre la perception et le contrôle (*Deep Reinforcement Learning* par exemple), et (ii) d'un point de vue applicatif, la maturation de certaines thématiques telles que les véhicules autonomes et les UAV, gourmandes en technologies de perception, renforce les liens avec la vision.

Cette action sera co-animée par les GdRs Robotique et ISIS. Elle s'intéressera aux aspects liant la robotique, l'image et le signal, et l'apprentissage automatique. Un accent particulier est mis sur les boucles perception/action, notamment sur la modélisation par MDP/POMDP et l'apprentissage par renforcement (profond). Elle couvre un champ divers d'applications autour de la robotique humanoïde, mobile, industrielle, des véhicules autonomes ou des drones. Parmi les thèmes abordés, nous proposons par exemple :

- Apprentissage de stratégies de navigation (en lien avec la perception)
- Apprentissage conjoint de tâches de navigation et de reconnaissance visuelle
- Interactions humains-robots
- Navigation sociale de robots mobiles par perception
- Traitement de signaux sociaux
- UAV et robots terrestres mobiles : perception, coordination
- Perception et apprentissage pour la cobotique
- Manipulation et saisie : contrôle par perception et apprentissage

3. Action transverse (A+T) : Modélisation et optimisation à l'interface signal/apprentissage

Animateur(s) : Valentin Emiya (MCF, LIF, Marseille), Caroline Chauv (CR CNRS, I2M, Marseille), Konstantin Usevich (CR CNRS, CRAN, Nancy).

Description : Cette action, commune avec le Thème A, consiste à encourager les rapprochements et discussions entre les communautés de l'apprentissage automatique et du traitement du signal. Ces deux disciplines partagent en effet nombre de fondamentaux, dont des modèles et méthodes d'optimisation. Une description détaillée de l'action est donnée à la Section 3.1.3, Action n° 3.. Le Thème T structurera en particulier l'organisation des journées autour des thématiques suivantes :

- *Optimisation en signal et en apprentissage*. Le traitement du signal et l'apprentissage se rejoignent sur de nombreux problèmes et méthodes d'optimisation : optimisation convexe, optimisation non convexe, modèles parcimonieux, optimisation en grande dimension, factorisation de tenseurs.

- *Compréhension des réseaux profond.* Comme mentionné dans les perspectives du Thème T, la compréhension théorique des mécanismes d'apprentissage mis en œuvre dans les réseaux profonds est aujourd'hui limitée. Afin de guider cette analyse formelle, des outils à la convergence des méthodes de traitement du signal et d'apprentissage statistique seront explorés. On peut mentionner par exemple l'analyse harmonique pour les notions de stabilité (scattering), la convergence entre les modèles bayésiens et les techniques de régularisation structurelle pour modéliser l'incertitude de décision, ou encore l'utilisation de méthodes de la théorie de l'information pour analyser les performances de généralisation des modèles.
- *Convolution sur grilles non régulières.* L'opérateur de convolution est un élément essentiel dans le succès spectaculaire obtenu par les réseaux profonds pour la reconnaissance de données de type signaux pauvres en sémantique comme les images ou les sons. Nous étudierons la manière de définir et d'adapter ces opérateurs de convolution pour traiter des grilles définies par des graphes d'adjacence non réguliers (e.g., réseaux de capteurs, données 3D).

4. Action transverse B+T : Vision 3D, géométrie et apprentissage

Animateur(s) : Vincent Lepetit (PU, LABRI, Bordeaux), Adrien Bartoli (PU, Institut Pascal, Clermont-Ferrand).

Description : Cette action Vision 3D et Apprentissage s'intéressera à la reconstruction 3D d'un objet ou d'une scène rigide ou déformable (SfM, NRSfM - MVS), au recalage d'images multimodales (RGB, RGB-D, depth), au calcul de pose d'un objet ou d'une scène rigide ou déformable (pose, SfT) à partir de modèles texturés ou non-texturés.

Complémentaires aux approches classiques basées sur la géométrie (voir la description de cette action transverse dans le Thème B : cf. Section 3.2.3, Action n° 6.), aujourd'hui les méthodes d'apprentissage profond peuvent également être développées en vision 3D et font l'objet de nombreux travaux. A l'heure actuelle, les algorithmes sont capables de prédire d'importantes quantités de cartes de profondeur et de normales à partir d'une seule image. Il est aujourd'hui intéressant d'utiliser ces données pour généraliser les approches sur des données temporelles. De la même manière, l'étude du concept du self-learning, où l'apprentissage est effectué grâce à une tâche secondaire différente de la tâche primaire de reconstruction 3D, peut être fort utile dans le cas où les données d'apprentissage ne sont pas disponibles. Enfin, les algorithmes de SLAM récents se concentrent sur l'association d'un label sémantique sur chaque objet tout en reconstruisant la scène de manière automatique. Ce sujet de recherche récent et prometteur alliant reconstruction 3D et reconnaissance d'objet sera également étudié ici. Ainsi, dans cette action, il s'agit d'étudier les techniques géométriques de vision 3D en tenant compte de la spécificité des données (multi-modales, dynamiques, déformables) et d'utiliser, dans une certaine mesure, les méthodes d'apprentissage en vision 3D de manière conjointe à la géométrie.

Une description plus détaillée de cette action est donnée dans la section consacrée au Thème B., c.f. Section 3.2.3, Action n° 6.

5. Action transverse B+T : Visages, gestes, activités et comportements

Animateur(s) : Catherine Achard (MCF, ISIR, Paris), Olivier Alata (PU, LHC, St. Etienne), Christophe Ducottet (PR, LHC, St. Etienne).

Description : Cette action concerne l'extraction du mouvement humain (visage, mains, corps, gestes) à partir de séquences vidéo, et son analyse à plus haut niveau (tâches,

comportement), notamment pour des applications interactives ou de surveillance. Une attention particulière sera portée aux données hétérogènes (couleur, thermique, 3D, etc.) et/ou acquises dans des situations non contraintes.

Elle s'intéressera notamment à la prédiction de posture 2D ou 3D de la main ou du corps avec des approches utilisant l'apprentissage profond, à la modélisation statistique de gestes, à la modélisation du visage (forme, apparence, émotion), et à la reconnaissance d'activités et de comportements. Des journées thématiques seront éventuellement co-organisées avec le GdR Robotique sur ces sujets.