
RAPPORT DE CONJONCTURE

DU COMITÉ NATIONAL DE LA RECHERCHE SCIENTIFIQUE

ÉDITION 2014

Extrait



CNRS ÉDITIONS

15, rue Malebranche – 75005 Paris

SECTION 07

SCIENCES DE L'INFORMATION : TRAITEMENTS, SYSTÈMES INTÉGRÉS MATÉRIEL-LOGICIEL, ROBOTS, COMMANDES, IMAGES, CONTENUS, INTERACTIONS, SIGNAUX ET LANGUES

Extrait de la déclaration adoptée par le Comité national de la recherche scientifique réuni en session plénière extraordinaire le 11 juin 2014

La recherche est indispensable au développement des connaissances, au dynamisme économique ainsi qu'à l'entretien de l'esprit critique et démocratique. La pérennité des emplois scientifiques est indispensable à la liberté et la fécondité de la recherche. Le Comité national de la recherche scientifique rassemble tous les personnels de la recherche publique (chercheurs, enseignants-chercheurs, ingénieurs et techniciens). Ses membres, réunis en session plénière extraordinaire, demandent de toute urgence un plan pluriannuel ambitieux pour l'emploi scientifique. Ils affirment que la réduction continue de l'emploi scientifique est le résultat de choix politiques et non une conséquence de la conjoncture économique.

L'emploi scientifique est l'investissement d'avenir par excellence

Conserver en l'état le budget de l'enseignement supérieur et de la recherche revient à prolonger son déclin. Stabiliser les effectifs ne suffirait pas non plus à redynamiser la recherche : il faut envoyer un signe fort aux jeunes qui intègrent aujourd'hui l'enseignement supérieur en leur donnant les moyens et l'envie de faire de la recherche. On ne peut pas sacrifier les milliers de jeunes sans statut qui font la recherche d'aujourd'hui. Il faut de toute urgence résorber la précarité. Cela suppose la création, sur plusieurs années, de plusieurs milliers de postes supplémentaires dans le service public ainsi qu'une vraie politique d'incitation à l'emploi des docteurs dans le secteur privé, notamment industriel.

Composition de la section

Michèle BASSEVILLE (présidente de section); Sophie ACHARD (secrétaire scientifique); Pierre-Olivier AMBLARD; Philippe BLACHE; Emmanuel BOUTILLON; Annick CHOISIER; Jamal DAAFOUZ; Christine FERNANDEZ-MALOIGNE; Christophe FONTE; Philippe FRAISSE; Christian GENTIL; Jérôme IDIER; François LE CHEVALIER; Éric LECOLINET; Philippe LOUBATON; Paolo MASON; Gilles MOURROT; Mathias PAULIN; Hyewon SEO; Thierry SIMEON; Jocelyne TROCCAZ.

Résumé

Ce rapport est structuré en huit thèmes : automatique ; traitement du signal ; traitement des images et vision numérique ; informatique graphique, IHM et réalité virtuelle ; traitement automatique des langues et de la parole ; robotique ; systèmes sur puce ; interactions avec les sciences du vivant. Chacun de ces thèmes est présenté en cinq rubriques : les grandes avancées actuelles et les thèmes émergents ; les interactions avec d'autres disciplines ; la place de la France dans le contexte de la recherche mondiale ; les points forts et points faibles de la recherche en France ; les retombées sociales, économiques (valorisation) et culturelles. Nous concluons en soulignant les lignes de force fondamentales transverses à la section et quelques caractéristiques liées au contexte.

Introduction

Pour présenter l'état des lieux et les sujets en émergence dans le cadre du périmètre thématique de la nouvelle section 07, l'une des deux sections de sciences de l'information, nous avons choisi de structurer le rapport selon la liste des huit thèmes que nous utilisons dans nos travaux : automatique ; traitement du signal ; traitement des images et vision numérique ; informatique graphique, IHM et réalité virtuelle ; traitement automatique des langues et de la parole ; robotique ; systèmes sur puce ; interactions avec les sciences du vivant. Chacun de ces thèmes est présenté en cinq rubriques : les grandes avancées actuelles et les thèmes émergents ; les interactions avec d'autres disciplines ; la place de la France dans le contexte de la recherche mondiale ; les points forts et points faibles de la recherche en France ; les retombées sociales, économiques (valorisation), et culturelles. Nous concluons en soulignant les lignes de force fondamentales transverses à la section et quelques forces et faiblesses liées au contexte.

I. Automatique

A. Les grandes avancées actuelles et les thèmes émergents

L'automatique joue un rôle majeur pour l'amélioration et l'évolution des produits et des systèmes. Elle est également considérée comme fondamentale pour les visions et ambitions du futur dans les domaines émergents tels que les infrastructures critiques, la transition énergétique, l'écologie industrielle, la biomédecine ou les neurosciences. Les grandes avancées à l'échelle mondiale, répertoriées dans le rapport⁽¹⁾ «*The Impact of Control Technology*», 2nd edition, T. Samad and A.M. Annaswamy (eds.), IEEE Control Systems Society, 2014, available at www.ieeecss.org, visent à répondre aux défis industriels actuels et futurs, qui impliquent des processus dont les modèles, de nature continue, à événements discrets ou hybride, sont de complexité croissante. La dynamique des systèmes est décrite par des modèles mathématiques, par exemple issus des lois de la physique. Ces représentations de systèmes, ou systèmes de systèmes ou systèmes en réseau, peuvent être des équations différentielles ordinaires, des équations aux dérivées partielles, des équations algébro-différentielles, non linéaires, hybrides, à événements discrets, stochastiques, ou une combinaison hétérogène de ces différents modèles. Elles peuvent tenir compte de contraintes sur le système lui-même, de l'environnement, de perturbations et d'incertitudes sur les modèles, de retards de transmission de l'information ou de pertes de données. Les objectifs de ces systèmes ont également une expression mathématique, tels ceux de suivi de trajectoire, de contrôle optimal, de performance, de robustesse, de prédiction de comportements, de diagnostic et pronostic de défauts, de surveillance et de supervision. Les solutions de l'automatique, qui s'appuient sur

des concepts mathématiques, ont aussi leurs concepts et développements propres, comme les notions de rétroaction et d'observateurs, d'identification de paramètres, et elles s'inspirent du contexte réel de l'application. Ce va-et-vient constant entre outils fondamentaux et exigences réelles fait de l'automatique une discipline fondamentale à caractère fortement pluridisciplinaire.

Le rapport⁽²⁾ «*Contribution de l'Automatique aux défis, à l'HORIZON 2020, des Sciences et Technologies de l'Information et de la Communication et de leurs Interactions*», écrit par la communauté française avec le soutien de nombreux industriels européens, définit plusieurs thèmes de développements fondamentaux émergents qui posent de nombreux défis à la communauté scientifique. La complexité des problèmes est liée à la prise en compte des aspects contrôle et optimisation, des composantes matérielles et logicielles, des interactions avec l'humain, des environnements plus ou moins coopératifs, dans le cadre d'applications variées concernant aussi bien les micro- et nano-systèmes que les systèmes industriels de grande dimension, les systèmes vivants que les systèmes mécatroniques. Répondre à ces défis nécessite de résoudre des problèmes fondamentaux importants et de développer de nouvelles méthodes et nouveaux algorithmes. Les approches centralisées classiques sont inopérantes face à l'organisation en réseau omniprésente. Les capteurs et les actionneurs distribués engendrent des retards, des incertitudes et une forte dégradation des signaux de mesure voire une perte de paquets d'information. Les sous-systèmes qui composent le processus global sont souvent hétérogènes et dotés d'une grande autonomie. La maîtrise de ces processus pose des problèmes théoriques importants allant de la modélisation à la synthèse de lois de commande avec garantie de stabilité, performance et fiabilité, et prise en compte des aspects complexes liés à l'autonomie, à la grande dimension, à la nature multiéchelles, et aux interactions. La modélisation implique de traiter de grands ensembles de données, en grandes dimensions, et fait appel à de nouvelles méthodes non paramétriques. La garantie de stabilité et de performance

de systèmes dans lesquels la communication ou le couplage entre sous-systèmes est limité, va bien au-delà des recherches classiques qui intègrent des retards ou des pertes de paquets dans les communications. La conception de systèmes de commande performants nécessite le développement de méthodes de contrôle et d'optimisation distribuées. La surveillance, la reconfiguration, la sûreté et la fiabilité sont des aspects d'une importance primordiale dans le développement des architectures de contrôle efficaces et sûres, par exemple pour les grands réseaux de l'automatisation industrielle. Les nouvelles technologies de capteurs permettant l'accès à un état distribué, le développement de réseaux sans fil permettant l'enregistrement synchronisé de mesures distribuées, et les avancées en calcul scientifique (e.g. modèles multiphysiques) renouvellent en profondeur les problématiques de la surveillance, telles celle de l'intégrité des infrastructures civiles.

B. Interactions avec d'autres disciplines

De nombreuses recherches multidisciplinaires sont menées avec d'autres domaines scientifiques, comme l'informatique, les mathématiques, la physique, les nanosciences, la biologie, les sciences humaines et sociales. Citons par exemple : avec les mathématiques, les recherches sur la commande et l'observation des systèmes à paramètres distribués, la commande optimale des systèmes non linéaires, l'optimisation ; avec la physique, les résultats fructueux obtenus récemment sur les systèmes quantiques, en particulier sur le contrôle des photons de l'électrodynamique quantique en cavité en collaboration avec l'équipe de Serge Haroche (Prix Nobel) ; avec la recherche biomédicale, le contrôle de la maladie de Parkinson, le contrôle de signaux neuronaux, la modélisation et le contrôle du VIH, la détection de tumeurs cancéreuses.

C. Place de la France dans le contexte mondial

De nombreux chercheurs des établissements universitaires et de recherche (EPST) sont des leaders internationaux, très impliqués dans des comités éditoriaux de revues prestigieuses et dans l'organisation des conférences internationales, notamment sous l'égide de l'International Federation of Automatic Control (IFAC⁽³⁾), l'IEEE Control Systems Society (CSS⁽⁴⁾), et l'European Control Association (EUCA⁽⁵⁾). Les chercheurs français occupent des responsabilités importantes au sein de l'IFAC et de ses comités. Grâce à cette dynamique, la France assure la présidence de l'IFAC pour la période d'août 2014 à juillet 2017 et l'organisation de son 20^e congrès mondial en 2017 à Toulouse. Dans le cadre des activités internationales du GDR MACS, l'école annuelle organisée par l'European Embedded Control Institute (EECI⁽⁶⁾) propose des cours dispensés par les meilleurs chercheurs internationaux sur les systèmes complexes, non linéaires, hybrides, embarqués, ou en réseau.

D. Points forts/points faibles de la recherche en France

L'automatique française jouit d'une visibilité internationale et d'une attractivité certaines tant au niveau doctoral et postdoctoral qu'au niveau des chercheurs confirmés effectuant des séjours courts, postulant à des postes permanents voire les occupant. La position de l'automatique à l'interface de plusieurs disciplines, ses approches génériques, le soutien institutionnel (GDR MACS), et les partenariats industriels forts et pérennes, en particulier dans les domaines aéronautique, spatial et automobile, sont également des points forts. Par contre, si l'automatique est omniprésente en réalité, elle n'est pas perçue par les non spécialistes et apparaît peu en tant que telle dans les appels d'offres, au niveau tant national qu'europpéen. Elle souffre

aussi d'un manque de visibilité dans les formations : peu de diplômés affichent l'automatique comme spécialité principale, contrairement à ceux de Suède ou d'Italie.

E. Retombées sociales, économiques (valorisation), culturelles

L'automatique est une discipline fondamentale à très fort impact dans le monde socio-économique. Les interactions se traduisent par de nombreux contrats de collaboration, des projets ANR ou Européens communs, des thèses CIFRE, etc. De nombreux docteurs en automatique sont recrutés dans l'industrie à l'issue de leur thèse, ce qui accroît l'impact des recherches fondamentales et facilite leur transfert vers les applications.

II. Traitement du signal

A. Les grandes avancées actuelles et les thèmes émergents

Les signaux numériques sont de plus en plus nombreux et de nature variée. Ainsi, la structure en réseaux de certaines données (réseaux de capteurs, de neurones, de gènes, d'ordinateurs, de transport ou même de contacts sociaux, etc.) est à l'origine du développement récent de divers modèles sur graphes, et amène à revisiter dans ce contexte des problématiques telles que l'analyse multi-échelle, l'inférence statistique, l'apprentissage ou l'optimisation. Des travaux du même type sont également en cours pour être en mesure de représenter et traiter des signaux sur des

variétés. L'un des grands défis actuels est le traitement de très grands volumes de données qui nécessite de repenser les techniques d'estimation, de détection, de classification ou d'optimisation classiques. En particulier, les algorithmes standards d'échantillonnage stochastique de type Monte Carlo sont mis en difficulté par les grandes dimensions, et des solutions ont été trouvées par des techniques de proposition, de mutation ou de sélection. Ces avancées ont permis l'émergence des modèles Bayésiens non paramétriques, pour la classification non supervisée ou la séparation de sources par exemple. Le traitement efficace de grandes masses de données nécessite de concevoir de nouvelles méthodes d'apprentissage (de et sur dictionnaires, par renforcement, profond, etc.) qui visent à s'affranchir des modèles physiques. Les grandes dimensions conduisent également au développement de techniques basées sur la théorie des grandes matrices aléatoires. L'utilisation de la parcimonie est par ailleurs très populaire dans les problèmes de grandes tailles, et aboutit le plus souvent à la résolution de problèmes d'optimisation difficiles (non lisse, non convexe, combinatoire) qui font l'objet de nombreux travaux. Les algorithmes développés sont également étudiés dans des versions incrémentales (ou adaptatives) pour les mettre en œuvre dans le contexte des grandes masses de données. L'ensemble des problèmes inverses, omniprésents dans la conception de systèmes d'acquisition de signaux et d'images, bénéficie de ces avancées méthodologiques et algorithmiques (cf. section III). La variété des capteurs à disposition rend fondamental le développement de méthodes de fusion de données multimodales qui reste un enjeu majeur particulièrement pour des applications en imagerie satellitaire ou biomédicale.

Les nouveaux usages de l'informatique et des communications (nomadisme et informatique en réseau) requièrent l'ubiquité des ressources (calcul, données) afin d'en assurer l'exploitation sur des infrastructures hétérogènes. Ils nécessitent de coupler le traitement des mesures et la prise de décision, pour traiter complètement le problème de l'estimation et de la détection conjointes, en concevant des

traitements multiples en parallèle (multi-processing) avec décision de type fusion ; ce qui entraînera aussi des conséquences sur l'adéquation avec l'architecture de traitement (architectures de traitement dites adaptatives et/ou « knowledge-based »). Ces nouveaux usages nécessiteront aussi le développement de techniques de traitement distribué de l'information prenant en compte la nature distribuée des données échangées au sein d'un réseau de télécommunications, qui influence très fortement les nouvelles orientations de la théorie de l'information et du codage. La complexité croissante des opérations réalisées au niveau des nœuds (possibilité de faire des calculs par opposition aux opérations classiques de relaiage ou routage) et leur nature distribuée induisent de nouvelles interactions avec la théorie du contrôle, l'optimisation distribuée, la théorie de la complexité et l'apprentissage automatique. Beaucoup de problèmes de stockage distribué de sources et de codage réseau, avec ou sans contraintes de sécurité, peuvent bénéficier grandement de ces interactions.

Un enjeu majeur est le développement d'une architecture de réseau intelligente et flexible capable de comprendre les besoins des utilisateurs et de se reconfigurer dynamiquement en prenant en compte les contraintes d'énergie, de miniaturisation et de grande masse d'appareils hétérogènes. Parmi les thèmes émergents figurent les réseaux green, les réseaux auto-organisant, les réseaux denses (techniques de MIMO en réseau ou virtuelles), la virtualisation (stockage et du calcul distribué dans le réseau). La prise en compte de l'hétérogénéité dans le contexte multi-utilisateurs ou multicapteurs est aussi essentielle dans la conception même des méthodologies de codage, qui doivent intégrer des fonctionnalités d'indexation et de protection (insertion de données cachées, chiffrement, authentification, intégrité) – cf. section III.

Pour faire face au volume de données et à l'augmentation de la complexité des algorithmes, des architectures parallèles sont proposées. Un défi est d'exprimer le parallélisme lors de la conception des méthodes de traite-

ment afin d'implanter efficacement les algorithmes. Pour cela un travail commun à l'interface algorithme-architecture est nécessaire. La modélisation par flots de données est une voie prometteuse car elle favorise l'expression du parallélisme entre les traitements.

B. Interactions avec d'autres disciplines

Le traitement du signal est en interaction avec de nombreuses disciplines telles que les sciences physiques et sciences de la vie (interaction ondes-matière, physique des instruments de mesures, méthodes bio-inspirées), les mathématiques appliquées (optimisation, calcul variationnel et EDP, probabilités et statistiques, analyse harmonique, théorie des jeux, etc.), et l'informatique (algorithmique, parallélisme, réseaux sans fils, reconfigurables et dynamiques), où nous assistons à une forte convergence avec les télécommunications à travers la volonté de virtualiser le stockage et le calcul (« cloud storage and computing »). Au niveau des applications, le traitement du signal analyse et traite les données issues de nombreux autres domaines scientifiques. Au-delà des domaines traditionnels (biomédical, spatial, sécurité, astrophysique, environnement), des domaines tels que les sciences sociales offrent de nouveaux types de données sur lesquels un investissement de la communauté est nécessaire.

C. Place de la France dans le contexte mondial

La France a une compétence reconnue mondialement dans le domaine des mathématiques appliquées, la théorie de l'information et les aspects méthodologiques du traitement du signal et de l'image. La France est présente dans la majorité des conférences du domaine, dans les comités d'organisation et de pro-

gramme, les comités techniques de l'IEEE Signal Processing Society (SPS), et les comités éditoriaux des grands journaux.

D. Points forts/points faibles de la recherche en France

Les points forts de la recherche sont les compétences sur les aspects théoriques, reconnues par les experts mondiaux et qui s'appuient sur des formations universitaires solides. Le soutien institutionnel (GDR ISIS) doit également être souligné. Un des points faibles est une certaine difficulté d'exploitation et de transfert vers l'industrie des résultats issus de la recherche. En outre, la France manque de centres interdisciplinaires rassemblant chercheurs en traitement du signal et de l'image, médecins, biologistes, biophysiciens, électroniciens, et associés à des grands centres de calcul haute performance, à l'instar du projet Cardioïd du LLNL. Les forces existent mais sont dispersées, malgré des efforts de création de structures communes. Par ailleurs, le nombre de chercheurs travaillant sur l'implémentation d'algorithmes reste faible, par rapport à l'Allemagne par exemple. Enfin le secteur des télécommunications est en position de sous-investissement industriel dans la compétition mondiale, en particulier asiatique.

E. Retombées sociales, économiques (valorisation), culturelles

En dehors des retombées dans les domaines de la santé, la sécurité, l'environnement, les innovations dans le domaine des systèmes embarqués et du traitement du signal et des images sont importantes pour un grand nombre de secteurs (automobile, télécommunications, robotique, cinéma, loisirs, etc.). Des enjeux économiques cruciaux sont liés à de

nouveaux standards, tels la nouvelle norme de compression HEVC adaptée à la vidéo Ultra HD, ou le standard de télécommunication 5G qui aura un impact sur les autres domaines de recherche (3D, Internet des objets, etc.) par les débits offerts tout en tenant compte de l'hétérogénéité des appareils.

III. Traitement des images et vision numérique

La distinction entre le traitement du signal et le traitement des images est actuellement moins marquée qu'auparavant. Cette convergence est liée au développement d'outils mathématiques récents fonctionnant aussi bien sur les signaux que sur les images (techniques de restauration basées sur la parcimonie par exemple), ainsi qu'à l'accroissement des capacités de calcul qui permet de les utiliser dans les deux contextes. Ainsi, un certain nombre de points qui vont être abordés dans cette section ont déjà été évoqués en section II. Il subsiste bien entendu des différences importantes entre les problématiques signal et images et vision, notamment la prise en compte de divers aspects géométriques dans le domaine des images et de la vision.

A. Les grandes avancées actuelles et les thèmes émergents

Le contexte des « Big Data » conditionne, encore une fois, beaucoup les avancées actuelles. Nous les synthétiserons par trois avancées fondamentales sur le plan méthodologique et théorique : le développement des approches d'apprentissage supervisé, semi-supervisé, transductif, « profond », pour répondre aux problèmes de reconnaissance,

de recherche, et d'interprétation ; la prise en compte de très grandes masses de données et de données de très grandes dimensions avec la multiplication des capteurs, la disponibilité de grandes collections de données hétérogènes ; et la fusion dans le processus de reconnaissance (noyaux multiples, sélection des caractéristiques). Certaines de ces avancées s'appuient sur la construction de dictionnaires et de représentations creuses des données. Par ailleurs, les images couleur, spectrales, multi- ou hyperspectrales nécessitent des approches vectorielles tant dans la représentation que dans le traitement et l'analyse, favorisant des recherches interdisciplinaires avec les équipes de mathématiques. Ces spécificités des données, à la fois géométriques et multivaluées, ont été le point de départ de nouvelles familles de méthodes comme les modèles de représentation par graphes et par algèbres vectorielles. De plus, des outils de traitements, comme les opérateurs de morphologie mathématique, ont été généralisés aux images multi-composantes en s'appuyant sur de nouvelles relations d'ordre, qui sont encore à l'étude. On notera également que les problèmes inverses en imagerie posent encore de nombreuses difficultés, liées à la dimensionnalité des données et à leur caractère souvent non régulier. De nombreux travaux (algorithme Proximal par exemple) offrent néanmoins maintenant des solutions dans différents cadres applicatifs comme l'inpainting, la restauration et l'échantillonnage comprimé. Outre l'intérêt des problèmes inverses, on peut souligner celui des représentations parcimonieuses, en particulier pour la co-conception de systèmes d'imagerie intelligents, alliant acquisition et traitements (avec les aspects théoriques, algorithmiques, méthodologiques, et hardware), en s'appuyant sur la super-résolution, l'échantillonnage comprimé, l'acquisition 3D ou 4D (voir les nombreux systèmes d'imageries médicales et biologiques, les caméras plénoptiques, panoramiques, etc.). Les réflexions sur les capteurs de demain posent des questions ouvertes notamment sur la notion de « bon modèle » pour une image.

Sur le plan de l'analyse, on ne peut que répéter l'importance de l'apprentissage statistique par analyse de masses d'images acquises

dans des conditions non contrôlées pour répondre à des problèmes de reconnaissance d'objets, et aussi d'évaluation de la qualité des images sans référence. Cette prise en compte de la qualité est devenue un leitmotiv des constructeurs comme des utilisateurs, au même titre que les enjeux de sécurité, prérequis indispensable à l'exploitation des masses de données, y compris en conditions de mobilité. Dans ce contexte, des recherches originales continuent d'émerger sur la définition de nouveaux protocoles d'identification de la source d'acquisition d'une image ; de nouvelles techniques de cryptographie et de tatouage de grandes masses de données visuelles 2D, 3D, 4D, à haute résolution ; de schémas robustes d'authentification et de vérification de l'intégrité des contenus. Les protocoles de sécurité, développés avec les informaticiens et les mathématiciens, doivent également être pensés avec les nouvelles approches de compression d'images et de vidéos haute définition, multivues et multi-profondeurs, 3D voxelliques, surfaciques, 4D, holographiques, etc. Pour tous ces types de données en émergence, des approches de compression en rupture telles que le codage bio-inspiré suscitent l'intérêt. D'une façon générale, les modèles bio-inspirés sont amenés à prendre une place importante dans le développement des nouveaux outils de traitements d'images en dimension N. En ce qui concerne le 3D, les techniques de reconstruction de scènes statiques à partir d'images ou de photographies sont passées au stade industriel (e.g. Autodesk et Acute3D en France). L'introduction ou le développement de nouvelles caméras non-standard apporte à la fois de nouvelles applications et de nouveaux problèmes scientifiques, avec notamment le fort développement de la photographie computationnelle. Les efforts de recherche en vision 3D se focalisent donc soit sur la reconstruction 3D+t de scènes non rigides, soit sur l'utilisation de ces nouvelles modalités d'imagerie optique. Le principal verrou de l'utilisation massive de ces technologies repose sur l'appréhension du système visuel humain qui nécessite, là encore, une approche pluridisciplinaire.

B. Interactions avec d'autres disciplines

Le traitement des images et la vision numérique sont, comme le traitement du signal, en interaction avec un grand nombre d'autres disciplines : sciences physiques et sciences de la vie, mathématiques, informatique, neurosciences computationnelles (modèles du système visuel humain, notamment de saillance visuelle de plus en plus étudiée), informatique graphique (rendu basé image, réalité augmentée, cf. section IV), automatique, robotique. Notons que le fort développement des applications où robots et humains partagent le même espace exige de nouvelles capacités de perception, notamment avec des capteurs visuels (reconnaissance de gestes et d'actions, navigation visuelle, etc.). On peut insister également sur l'importance croissante du traitement des images (et des signaux en général) pour les secteurs applicatifs du biomédical, de la santé, des problèmes liés au vieillissement, au handicap et au bien-être, contribuant au développement de la télémédecine et de la e-santé. L'imagerie joue également un rôle primordial pour l'observation de la Terre que ce soit pour des applications spatiales, environnementales, écologiques ou encore de sécurité des biens, des lieux et des personnes. L'habitat, les transports, que l'on veut désormais durables, sont aussi impactés par des applications de vision numérique. Par ailleurs, la priorité donnée aux objets connectés rend indispensable l'existence de liens forts avec les chercheurs des sciences humaines et sociales pour étudier les verrous essentiels au couplage entre l'homme et les dispositifs technologiques. Les plates-formes de réalité virtuelle, champ exploratoire en pleine expansion, sont des outils de simulation et de démonstration très propices à ces interactions (cf. section IV).

C. Place de la France dans le contexte mondial

La France est présente dans les comités d'organisation ou de programme des meilleures conférences du domaine, dans les comités techniques IEEE, et les comités éditoriaux des grands journaux internationaux. Elle est aussi représentée dans des campagnes d'évaluation internationales (telles que TRECVID) et participe activement aux standards (ISO, compression image – JPEG, vidéo – MPEG ou 3D – MPEG4). On notera enfin qu'en analyse d'images biomédicales, et notamment cérébrales, la France est très visible au travers de grosses structures. Elle est également pionnière en indexation multimodale pour la santé.

D. Points forts/points faibles de la recherche en France

Les points forts et points faibles sont ceux cités en section II. On peut ajouter d'une part un soutien institutionnel marqué (GDR ISIS et MIA), et d'autre part le retard de la France dans le domaine de la photographie computationnelle qui requiert, là encore, de la transdisciplinarité comme au sein du MIT Media Lab.

E. Retombées sociales, économiques (valorisation), culturelles

Le monde du numérique est présenté comme la « troisième révolution industrielle » et les dernières études sectorielles montrent que la dématérialisation et la valorisation des ressources numériques présentent des perspectives de croissance très élevées sur tous les secteurs. Ceci est particulièrement vrai pour les images, dans des secteurs variés : vision et robotique, automobile, éducation, culture, etc.

IV. Informatique graphique, IHM et réalité virtuelle

A. Les grandes avancées actuelles et les thèmes émergents

Informatique graphique. Les objets 3D, ou plus généralement les mondes virtuels, sont l'enjeu de recherches qui visent à capturer ou créer les formes géométriques, leurs apparences, leurs mouvements, pour offrir des moyens d'interagir avec ces mondes afin de fabriquer une réalité virtuelle. Soutenue par les grands domaines d'application que sont la CAO, la production de contenu pour le divertissement et la simulation, la recherche en informatique graphique et géométrique vise à concevoir de nouvelles structures de données discrètes, géométriques, topologiques, combinatoires ou évolutionnaires permettant de prendre en compte la dimension dynamique des environnements 3D. De fortes préoccupations algorithmiques dans les outils développés (complexité, NP-complétude, décidabilité, prédicats géométriques exacts), montrent l'ancrage informatique de la communauté. De même, la forte coloration mathématique des modèles développés (géométrie, analyse, algèbre, probabilités) positionne les travaux aux interfaces des mathématiques. Ainsi, la communauté d'informatique graphique s'intéresse aux fondements mathématiques de la géométrie différentielle ou algébrique, aux fondements de la topologie combinatoire ou discrète, aux propriétés arithmétiques, aux théories de l'approximation et de l'optimisation. Abordés principalement dans le cadre 3D et 3D+t, ces différents concepts sont étudiés et développés avec l'objectif de la généralisation en grande dimension.

Les avancées de la recherche se déclinent autour de « la 3D pour tous » avec la démocratisation des dispositifs d'affichage 3D et l'apparition d'outils d'impression 3D. Les capacités de

capture de données réelles et de simulation suivant une progression exponentielle, le problème du traitement des gigantesques flots de données 3D qui en résultent devient une priorité de recherche. Pour accompagner cette évolution rapide, la communauté doit attaquer un certain nombre de verrous scientifiques et technologiques : capture rapide et automatique d'environnements 3D réels, dans toute leur complexité ; passage à l'échelle pour modéliser, animer et visualiser en temps réel de gros volumes de données 3D, parfois incertaines et en flux ; mise en place de méthodes de création intuitive et d'édition rapide de contenus 3D, permettant le prototypage et l'impression de nouveaux objets réels ; extension des méthodes intuitives d'édition d'images aux images ou vidéos de synthèse en surmontant le clivage traditionnel édition 2D/édition 3D ; intégration des méthodes d'apprentissage statistique en modélisation géométrique, rendu et animation pour utiliser des données réelles comme contrôle de la génération de nouveaux contenus ; couplage modélisation 3D/rendu et analyse/visualisation pour la fouille visuelle et la visualisation collaborative.

Interaction homme-machine et réalité virtuelle. L'omniprésence des ordinateurs sous toutes leurs formes (PCs, tablettes, smartphones, petits objets communicants, etc.) et la popularisation des capteurs nécessitent le développement de nouvelles modalités d'interaction exploitant par exemple les possibilités offertes par l'interaction gestuelle ou vocale ou la prise en compte des émotions, des intentions ou du contexte social. Ces interactions peuvent éventuellement être implicites, en inférant des situations d'interaction à partir de capteurs biophysiques ou répartis dans un système ambiant. Les interfaces cerveau-ordinateur ouvrent également de nouvelles potentialités en donnant accès à des informations jusque-là inaccessibles. La multimodalité, qui exploite les couplages multisensoriels et moteurs de l'utilisateur, et l'hybridation réel-virtuel, qui réalise une intégration fine d'éléments de ces deux mondes, amènent des ruptures dans nos modes d'interaction. Leurs applications touchent l'interaction en situation de mobilité, et l'interaction « en grand » avec le développement

de salles immersives et interactives. La généralisation des dispositifs personnels connectés renouvelle les problématiques de l'activité et de l'interaction collaboratives colocalisées ou à distance (réseaux sociaux, partage et édition collaborative, collaboration dans les environnements virtuels, crowdsourcing, etc.).

Face à l'accroissement exponentiel des masses de données, la visualisation de l'information vise à faciliter l'exploration, la comparaison et la compréhension de données abstraites ne présentant pas de structures géométriques naturelles. À la frontière entre visualisation interactive et fouille de données, la visualisation analytique couple ces domaines pour mieux filtrer les données, faire apparaître des relations ou des motifs caractéristiques, et ainsi produire du sens à partir de données hétérogènes.

La difficulté de modéliser le comportement humain dans un contexte complexe rend la validation des systèmes homme-machine particulièrement délicate. Ceci constitue un enjeu majeur du fait de la généralisation de ces systèmes dans des environnements critiques. Des approches comme la validation de modèles de simulation de comportement ou l'étude de modes de répartition ou délégation de fonction visent à conjuguer des systèmes automatisés et des interactions humaines. Le partage ou la coréalisation de tâche sont des questions critiques nécessitant de nouveaux modèles et méthodes pour fiabiliser et évaluer les dispositifs de commande. La conception centrée sur l'utilisateur vise une meilleure adaptation des systèmes aux besoins humains ainsi qu'une plus grande intégration des dimensions esthétiques et affectives de l'interaction. L'étude expérimentale et la définition de mesures appropriées sont nécessaires pour valider l'utilisabilité et l'acceptabilité des systèmes, ainsi que leur plausibilité visuelle dans le cas de la réalité virtuelle.

B. Interactions avec d'autres disciplines

Outre l'informatique, la communauté est en relation avec un grand nombre de disciplines : mathématiques appliquées (optimisation et interpolation, calcul variationnel et équations aux dérivées partielles, probabilités et statistiques, analyse harmonique, algèbre et combinatoire), sciences physiques et sciences de la vie (interaction ondes-matière, mécanique et dynamique du solide et des corps déformables, méthodes bio-inspirées), robotique, et sciences humaines (psychologie perceptive, ergonomie, analyse des émotions, sociologie, design, etc.) afin d'intégrer les usages dans la validation des modèles théoriques et pratiques.

C. Place de la France dans le contexte mondial

La recherche française est de très haut niveau avec une excellente visibilité internationale. Les membres de ces communautés publient régulièrement dans les meilleurs conférences et journaux et sont fortement impliqués dans les comités de pilotage des associations savantes, les comités éditoriaux des grands journaux et les comités d'organisation et de programme des grandes conférences internationales, dont plusieurs ont eu lieu en France ces dernières années. On peut enfin remarquer la présence en France du leader mondial pour la CAO (Dassault Systèmes) et de nombreux studios de jeux vidéo, de production de films et d'effets spéciaux, domaine dans lequel la créativité scientifique et artistique de la France est mondialement reconnue.

D. Points forts/points faibles de la recherche en France

Outre le soutien institutionnel via les GDR (IG-RV, MAGIS), ces thématiques sont soutenues par l'ANR, la DGCIS et les pôles de compétitivité, et participent activement à plusieurs Labex. Les étudiants français de profil math/info constituent un excellent vivier pour ces disciplines. Les points à améliorer concernent d'une part les partenariats industriels qui pourraient être renforcés par une plus grande implication du CNRS dans la valorisation des travaux menés dans les UMR, et d'autre part le soutien institutionnel. De plus, contrairement à l'Amérique et l'Europe du Nord, la France ne dispose pas de centres multidisciplinaires dédiés à ces thématiques. Enfin, les excellents résultats scientifiques en termes d'impact national et international ne doivent pas masquer le manque notoire de chercheurs CNRS dans ces communautés.

E. Retombées sociales, économiques (valorisation), culturelles

Tous ces domaines ont connu une croissance extrêmement rapide ces 30 dernières années et leurs retombées industrielles et sociétales sont considérables. L'informatique interactive a été l'un des moteurs du développement des sciences et technologies de l'information dans tous les domaines de la société en permettant de démocratiser l'usage de l'informatique (ordinateurs, smartphones, tablettes) grâce aux interfaces utilisateurs. Les applications industrielles sont extrêmement variées : CAO, animation 3D, jeux vidéo, simulateurs d'apprentissage, applications pédagogiques, culturelles ou artistiques, supervision et contrôle, interaction mobile, visualisation financière, etc. Les domaines d'application concernés sont tout aussi nombreux : aéronautique, médecine et biologie, énergie, bâtiment,

science des matériaux, réseaux sociaux, Web et multimédia. L'informatique étant aujourd'hui omniprésente dans les sociétés modernes, être en capacité de comprendre, d'analyser, de modéliser et d'interagir avec les images, formes spatiales ou données numériques ambiantes à la cadence où elles sont produites, est un véritable enjeu sociétal.

V. Traitement automatique des langues et de la parole

Le traitement automatique des langues et de la parole (TALP) vise à développer des outils, méthodes et applications mettant en jeu le langage sous ses différentes formes (modalités orales, écrites, gestuelles), en prenant en compte la variété des langues, des locuteurs, des usages. Le TALP a par ailleurs également conduit à de nouvelles interactions, notamment vers le traitement de la musique.

A. Les grandes avancées actuelles et les thèmes émergents

Le TALP s'est structuré autour de problématiques et d'applications variées : traduction automatique, transcription automatique de parole, analyse de contenu, recherche d'information, traitements audionumériques, traitement automatique des langues des signes en sont des exemples. Parmi les avancées actuelles, notons le rapprochement des traitements de l'écrit et de l'oral. Ces travaux s'appuient sur le développement de ressources (lexiques, corpus, grammaires, etc.) et d'outils (transcription, segmentation, alignement, etc.) exploitant notamment les indices extra-linguistiques. Plusieurs nouvelles applications sont ainsi ouvertes, comme la traduction de parole, la recherche d'information multimédia, etc. De même, la

question du multilinguisme constitue un enjeu décisif, avec des applications variées vers le résumé automatique ou la recherche d'information à partir de documents multilingues. Cette thématique est en lien étroit avec la réutilisation des ressources, ainsi que des perspectives pour le traitement des langues peu dotées (ne disposant pratiquement pas de ressources ou d'outils pour leur traitement automatique). De même, la prise en compte du langage en contexte naturel rend nécessaire le traitement de la multimodalité, intégrant les aspects verbaux et non verbaux de l'interaction (gestes, attitudes, etc.). Les applications sont nombreuses : traitement des données multimédias, agents conversationnels, etc. Proche des aspects non verbaux, le traitement de la musique repose sur des besoins spécifiques en termes de précision et de qualité dans les domaines du signal, du langage et de l'information (audiovisuel 3D, analyse des contenus, classification, etc.).

La dimension humaine occupe une position centrale dans ce domaine. Un axe de travail particulier concerne la prise en compte de la personnalité du locuteur dans les interactions situées (compétences et performances linguistiques, sociales, interactionnelles et émotionnelles). De même, les enjeux de la robotique cognitive, développementale, sociale, conduisent au développement de modèles de mécanismes interactifs (traitement sensoriel et perceptif, modèles biomécaniques, modèles de catégorisation, modèles de reconnaissance conjointe parole-geste, etc.). Les approches bio-inspirées sont largement utilisées dans ce domaine. Un nombre croissant de travaux aborde par ailleurs la modélisation des aspects physiologiques de la production et la perception du langage ainsi que ses bases cérébrales, y compris en contexte pathologique. Du point de vue méthodologique, la modélisation des réseaux neuronaux, des mécanismes d'oscillation collective, de codage multiplexe, de codage prédictif, de liage neuronal, sont ainsi des enjeux importants. De son côté, le domaine du « Big Data » pose de nouvelles questions, allant de l'infrastructure (entrepôts de données) à l'accès et la visualisation. Il s'agit de problèmes méthodologiques (parallélisation, traitement à grande

échelle, etc.) qui se posent de façon spécifique en articulation avec l'émergence en apprentissage automatique de nouveaux paradigmes (réseaux neuro-mimétiques profonds, découverte de connaissances).

B. Interactions avec d'autres disciplines

Le TALP est un domaine fortement interdisciplinaire, alliant informatique et linguistique, psycholinguistique, médecine, physique, neurosciences ou arts. Le TALP connaît une évolution vers la prise en compte de données naturelles, éventuellement non standard (parole, tweets, SMS) et dans un contexte situé. Il convient pour cela de prendre en compte la variabilité des données, leur contexte social, physiologique, communicationnel. Pour cela, les données sont non seulement plus volumineuses, mais également plus riches, comportant des informations de haut niveau, reposant sur une expertise humaine provenant de l'ensemble des disciplines mentionnées.

C. Place de la France dans le contexte mondial

La France est historiquement très bien positionnée dans la communauté internationale du TALP. En traitement de la parole, nous occupons une position centrale grâce à notre participation à des organismes comme l'ISCA, ainsi que notre rôle actif dans la structuration de la recherche au niveau européen (laboratoires internationaux, création d'agences de diffusion de ressources, organisation de campagnes d'évaluation, etc.). Le traitement automatique de l'écrit est également fortement ancré au niveau international avec des responsabilités dans les associations internationales généralistes (ICCL, ACL, EACL) ainsi que dans

des domaines de spécialité comme l'information biomédicale, la recherche d'information, la traduction automatique, etc.

D. Points forts/points faibles de la recherche en France

Au-delà des contributions méthodologiques autour des techniques d'apprentissage automatique ainsi que leurs applications (statistique textuelle, recherche d'information, traduction automatique, etc.), plusieurs points forts de la communauté TALP proviennent de l'intégration de différents niveaux d'information. La France est ainsi pionnière sur la question de l'intégration écrit/oral, avec un effort de développement de ressources communes. Nous sommes très bien positionnés vis-à-vis des enjeux actuels : traitement de données naturelles, boucle compréhension/production, multimodalité, variabilité, données physiologiques, etc. Par ailleurs, nos équipes sont très actives sur la question de l'interface entre le TALP et les autres disciplines : linguistique formelle et descriptive, neurosciences en particulier. Plusieurs projets d'envergure nationale et internationale abordent ces questions.

Une direction à renforcer porte sur la sémantique pour laquelle il n'existe toujours pas de modèle véritablement utilisable, permettant de passer d'une approche de niveau lexical à l'utilisation de représentations structurées. Il devient de plus en plus nécessaire de compléter les paradigmes d'apprentissage par des explications sur les informations produites, en rapprochant les techniques à base de règles avec les méthodes d'apprentissage. Cette évolution repose notamment sur la capacité d'appliquer l'apprentissage à des processus ou structures plus complexes, à partir de sources d'information multiples. L'évolution vers des systèmes complexes rendant plus difficile la tâche d'évaluation et l'établissement d'une référence, il devient nécessaire de mettre en œuvre de nouveaux paradigmes, reposant par exemple sur des mécanismes de crowd-sourcing. Il convient également de prendre en compte la notion d'acceptabilité des utilisateurs.

E. Retombées sociales, économiques (valorisation), culturelles

Les applications du TALP sont nombreuses : traduction automatique, recherche d'information, systèmes conversationnels, traitement de la langue des signes, transcription de parole et de la musique, synthèse de parole, de musique et de gestes, aide aux handicapés, localisation de logiciels, etc. Les industriels sont ainsi devenus des acteurs importants du TALP, et sont en relation étroite avec nos laboratoires. Leur contribution réside en particulier dans la robustesse et le passage à l'échelle. Le TALP est ainsi désormais largement utilisé dans le cadre de l'analyse des réseaux sociaux (recommandations, dynamique des communautés, fouille d'opinion), en synergie avec le développement durable, la e-citoyenneté, ou la sécurité publique (cybercriminalité) et civile (gestion de catastrophe ou crise sanitaire). De nouvelles applications apparaissent dans des domaines variés : dialogue homme-machine, gestion de l'information, aide au diagnostic, suppléances sensorielles, réalités virtuelles et augmentées, aide à l'apprentissage, nouvelles interfaces, nouveaux instruments, etc.

VI. Robotique

A. Les grandes avancées actuelles et les thèmes émergents

L'activité de recherche de la robotique se structure autour du mouvement et de l'autonomie des robots. On peut citer la modélisation, la conception de mécanismes, le contrôle, la perception du mouvement, la planification, la

navigation autonome de véhicules (aériens, terrestres, marins ou sous-marins), l'apprentissage du mouvement et l'interaction du robot avec son environnement. Les récentes avancées technologiques permettent de créer des plates-formes robotisées ayant des performances et des capacités de mouvement suffisantes pour devenir des outils performants dans les secteurs tels que production, transport, logistique, technologies de la santé et des services à la personne, environnement et exploration. Parmi les défis scientifiques figure l'étude des propriétés géométriques, cinématiques et dynamiques des nouveaux mécanismes à structures parallèles, séries ou arborescentes, conduisant à de nouvelles approches multidisciplinaires de conception multi-objectifs. L'amélioration des performances des robots passe par la recherche de nouvelles architectures cinématiques pour les mécanismes à composants déformables, les structures hyper-redondantes, les robots compliants, les micro- et nano-robots.

Un autre grand défi concerne la commande avec asservissement visuel et vision active, en plein renouvellement grâce aux nouvelles modalités de perception : effort, toucher (haptique), mesures inertielles, afin d'atteindre les performances recherchées dans de nombreuses applications allant du médical à l'exploration en passant par le domaine industriel. Un autre problème de commande difficile concerne la compréhension du mouvement anthropomorphe et le développement de méthodes globales et optimales de commande des robots humanoïdes en situation de mouvement libre ou contraint. La modélisation d'interactions complexes entre le robot et son environnement, brique constitutive des systèmes de commande, doit encore progresser pour les interactions instruments/organes, les contacts rigides multiples et les tissus mous. Il en est de même pour l'intégration de ces modèles dans les algorithmes de contrôle.

L'autonomie des véhicules, autre défi important, a vu son développement s'accélérer récemment grâce aux avancées en localisation et cartographie simultanée (SLAM) et à l'apparition de voitures autonomes (Google Car), de

drones, de véhicules sous-marins télé-opérés ou semi-autonomes. Néanmoins, des efforts importants sont à réaliser pour la conception de nouvelles méthodes de localisation, planification, re-planification temps réel, coordination de flottille de véhicules hétérogènes, supervision et communication multi-robots. Les architectures de contrôle, importantes pour l'autonomie, contribuent plus généralement au développement des systèmes embarqués. Le défi principal est la modélisation des fonctionnalités et du comportement dynamique de tous les éléments (capteurs, actionneurs, noyau de contrôle, etc.) pour concevoir et implanter une architecture et un système robotique par composition fine et explicite de ces modèles.

La manipulation d'objets fait également partie des défis à relever dans les prochaines années. En effet, de nombreuses applications nécessitent de la manipulation d'objets. Cependant, le succès de manipulateurs peut dépendre fortement des caractéristiques de l'objet manipulé, en particulier ses dimensions : les effets volumiques et inertiels sont prépondérants à l'échelle macroscopique, les effets surfaciques prédominent pour des dimensions submillimétriques. Le principal défi aux échelles microscopique et nanoscopique est de porter autonomie et complexité fonctionnelle aux mêmes niveaux qu'à l'échelle macroscopique, ouvrant ainsi un nouveau champ d'applications biomédicales où la dextérité des micro-outils intracorporels est un enjeu majeur. À plus long terme, la conception de nano-robots autonomes peut ouvrir un nouveau champ de recherche sur l'exploration à l'échelle nanoscopique, pour le diagnostic médical par exemple.

B. Interactions avec d'autres disciplines

La robotique entretient historiquement de fortes interactions avec la mécanique, l'informatique, l'automatique, le traitement du signal et des images et les mathématiques appliquées. On peut également souligner une interaction

avec les neurosciences pour identifier les processus neurobiologiques permettant aux animaux et à l'homme de construire une représentation unifiée de l'environnement à partir des données multisensorielles afin de planifier et exécuter des actions de manière autonome et adaptative. Pour l'aide à la personne, la robotique d'assistance et la rééducation motrice, des interactions existent aussi avec la biomécanique et les neurosciences, pour la compréhension de l'apprentissage moteur et les manières de le favoriser par une interaction physique. L'interaction homme-robot joue un rôle croissant qui dépasse largement ces domaines et adresse des problématiques nouvelles : interprétation, fusion et prédiction de signaux de communication (verbale et non verbale, intention, émotion), synthèse de ces signaux pour les robots, interfaces homme-robot (interfaces cerveau-machine ou haptiques). La coopération et la coordination entre l'humain et le robot pour les autres domaines d'application (robots en milieu industriel, robots portables de type exosquelettes, véhicules autonomes pour le transport des personnes) nécessitent de prendre aussi en compte les questions de confort et de sécurité. Ces travaux impliquent des coopérations avec des spécialistes des sciences cognitives, à l'instar de ceux de la réalité virtuelle et augmentée.

C. Place de la France dans le contexte mondial

La France fait partie des grandes nations de recherche en robotique, elle possède une compétence et une reconnaissance internationale dans l'ensemble des grands domaines de cette discipline. Ses chercheurs sont présents dans les comités éditoriaux des grands journaux internationaux, les comités de programmes et d'organisation des grandes conférences, les comités techniques et le comité exécutif de l'IEEE Robotics and Automation Society (RAS⁽⁷⁾). La France est également présente dans les consortiums des grands projets et défis robotique européens (*H2020, Factory of the Future*).

D. Points forts/points faibles de la recherche en France

La France est une destination attractive pour de nombreux chercheurs internationaux, visiteurs, doctorants ou post-doctorants. La robotique française possède des points forts dans les grands domaines de la robotique, et bénéficie d'un soutien institutionnel (GDR Robotique). Les interactions avec la médecine et les secteurs industriels (aéronautique, automobile, PME, startups) sont fortes bien qu'il existe trop peu de relais entre la recherche académique et le monde des entreprises. Un grand défi est à relever pour l'aide au transfert et à l'innovation de la robotique vers les entreprises, et le soutien des laboratoires pour accompagner ces transferts dans cette période charnière de développement de la robotique dans l'industrie.

E. Retombées sociales, économiques (valorisation), culturelles

La robotique a un impact économique et social important, actuellement en pleine croissance. Une grande mutation dans le secteur industriel est en cours, avec l'objectif pour 2020 d'avoir des robots ayant un certain degré d'autonomie, possédant les capacités de travailler à proximité de personnes en toute sécurité dans des modes de coopération ou de collaboration, et ce avec des efforts d'installation et de programmation minimum. D'autres attentes sont affichées dans le domaine médical avec des systèmes chirurgicaux dédiés au contact du patient : robots porte-aiguilles, robots guides pour la chirurgie orthopédique, robots pour la chirurgie transluminale ou à trocart unique, robots autonomes de type capsules ingérables. Enfin, les véhicules autonomes routiers ouvrent de nombreuses perspectives pour améliorer la sécurité routière et la mobilité des personnes.

VII. Systèmes sur puce

A. Les grandes avancées actuelles et les thèmes émergents

L'avènement de l'internet des objets et de l'informatique en nuage s'appuie sur des systèmes embarqués à ultra-faible consommation pour instrumenter et interagir avec l'environnement physique en temps réel et de manière autonome, ainsi que sur des centres de calcul haute performance pour manipuler et stocker, de manière sécurisée, les exa-octets de données recueillies. Ces systèmes ont également besoin d'une infrastructure de communication capable de véhiculer des centaines de Mbps par terminal mobile ou des centaines de Gbps par nœud de calcul. Les systèmes sur puce au cœur de ces maillons de la révolution numérique comptent aujourd'hui plusieurs dizaines de milliards de transistors élémentaires de dimensions autour de 20 nm. Ils sont organisés en architectures comptant quelques dizaines (voire centaines) de cœurs, et sont capables de se reconfigurer pour s'adapter aux besoins applicatifs. D'une complexité inouïe, ils nécessitent des méthodes de conception, de vérification et de test extrêmement puissantes pour garantir un fonctionnement efficace.

La recherche en architectures de systèmes intégrés, en interface avec la recherche en informatique embarquée, permet la conception des objets connectés et intelligents tirant parti des nouvelles technologies. L'architecture idéale serait ainsi : adaptative (offrant le meilleur compromis de calculs par unité d'énergie); complète (contenant les capteurs, les calculateurs et la connexion aux réseaux); sûre (résistante aux attaques visant la confidentialité et l'intégrité des informations qu'ils traitent); traçable (garantissant l'origine des composants contre les fraudes et les vols; on parle de confiance

matérielle pour la sûreté et la traçabilité); résiliente et hautement disponible (fonctionnant malgré les fautes, le vieillissement et les pannes); virtualisée (proposant aux applications des cellules de traitement génériques); et interopérable (capable de communiquer avec plusieurs réseaux suivant plusieurs protocoles). Les architectes doivent proposer des solutions pour gérer: le vieillissement des circuits; les variations des procédés de fabrication; l'intégration de technologies hétérogènes (analogiques, mécaniques, etc.); l'émergence de la troisième dimension des circuits; la montée du parallélisme des systèmes; et enfin l'énergie. Ils doivent sans cesse repenser leurs outils de conception et de validation. Les futures solutions architecturales seront nécessairement matérielles et logicielles, et intégreront des systèmes d'exploitation embarqués.

La conception des systèmes cyber-physiques nécessite la prise en compte de l'interaction entre le monde électronique, des systèmes issus de l'ingénierie hors numérique, et leur environnement physique. Cette interaction passe par des interfaces qui rendent les systèmes électroniques technologiquement hétérogènes. Cette hétérogénéité amène le besoin de langages de modélisation et de simulation adaptés à des fonctions numériques et non numériques. D'autre part, la conception des circuits analogiques/mixtes et RF exige de nouvelles idées au niveau architectural ou au niveau circuit pour s'adapter aux évolutions des besoins et des technologies disponibles. La gestion de l'énergie (consommation et production locale) devient aussi une problématique de recherche nouvelle, en particulier pour les réseaux de capteurs.

Les dispositifs d'e-santé bénéficient d'avancées technologiques permettant d'en faire des systèmes autonomes et adaptatifs, sous des contraintes spécifiques aux applications biomédicales (innocuité des dispositifs, sécurité des données, faible coût). Elles trouvent leur prolongement naturel dans des interactions multidisciplinaires avec la médecine, la biologie, la biochimie et les matériaux.

Au niveau des technologies et modèles de calcul, plusieurs thèmes en rupture émergent

et doivent être pris en compte: exploitation des nouvelles technologies (mémoires non-volatiles MRAM ou RRAM, photonique silicium, etc.), intégration de méthodes d'apprentissage neuro-inspirées robustes, impact sur les systèmes et les applications, notamment en reconnaissance de formes, fouille de données, etc.

B. Interactions avec d'autres disciplines

Depuis ses origines, le domaine des systèmes intégrés est transdisciplinaire entre informatique et électronique. L'évolution des technologies de fabrication conduit à de nouvelles interactions avec la physique: modélisations multiphysiques (thermique, mécanique pour les MEMS, photonique, etc.), émergence de nouveaux procédés physiques (spinttronique, électronique moléculaire, etc.). Par ailleurs, l'explosion de la complexité amène de nouvelles interactions avec l'informatique: compilation, langages, logiciels embarqués, systèmes d'exploitation, vérification, optimisation, etc. Enfin, les interactions avec le traitement du signal sont renforcées par l'émergence de nouveaux algorithmes et de nouvelles applications. L'évolution vers les systèmes embarqués augmente aussi les interactions avec les mathématiques (cryptographie, tolérance aux dispersions, etc.); la biochimie (biomatériaux), les neurosciences (traitement de l'information neuro-inspiré) et la médecine (neurologie, cardiologie, orthopédie, urologie, etc.).

C. Place de la France dans le contexte mondial

La recherche française est reconnue dans de nombreux domaines des thématiques des systèmes intégrés matériels-logiciels, en particulier: architecture (manycore, réseau sur puce (NoC), synthèse de haut niveau); sys-

tèmes cyber-physiques (systèmes hétérogènes, langages et modélisation, capteurs, MEMS, systèmes AMS&RF spécifiques, gestion et conversion d'énergie); confiance matérielle (attaques physiques et contre-mesures, cryptographie appliquée); test. Dans le domaine des objets connectés, elle a une expertise en systèmes à basse consommation d'énergie et dans les dispositifs médicaux intelligents. Il en est de même dans le domaine du calcul émergent, et pour l'intégration de fonctions neuro-mimétiques.

D. Points forts/points faibles de la recherche en France

En ce qui concerne les points forts, trois traits ressortent. Tout d'abord, cette recherche montre une grande richesse, avec des forces pluridisciplinaires complémentaires menant des recherches aux frontières et aux interfaces avec les autres disciplines. Ensuite, cette communauté est dynamique, expérimentée et pérenne, et elle bénéficie d'un soutien institutionnel (GDR SoC-SiP). Enfin, la recherche française dans ce domaine est menée de longue date en collaboration avec un tissu industriel national qui couvre toute la chaîne : technologie, composants, caractérisation, circuits, systèmes, outils de CAO, équipements, sociétés de service. L'essentiel des effectifs (90 %) est constitué d'enseignants-chercheurs, cette recherche est ainsi très liée à l'enseignement, en liaison avec les besoins industriels (GIP CNFM). Toutefois, la transdisciplinarité de cette recherche est compliquée par un découpage académique interne complexe : trois sections du CNU (27, 61, 63), deux sections du Comité National (7, 8) et deux instituts du CNRS (INS2I, INSIS). Une piste d'amélioration serait de dépasser le clivage entre logiciel et matériel, particulièrement marqué en France. Cette évolution serait en accord avec la vision internationale du domaine, clairement illustrée par le regroupement de ENIAC, ARTEMIS et EPoSS dans ECSEL au niveau européen.

E. Retombées sociales, économiques (valorisation), culturelles

Les retombées économiques sont très importantes pour les entreprises des secteurs électronique (STMicroelectronics, PME) et équipementier, et pour les startups en création dans le secteur des objets connectés (domotique, e-santé, tissus intelligents, jeux, etc.). Il faut aussi noter le renouveau d'une industrie autour des architectures manycores en France (KALRAY, CORTUS, etc.) et en Europe (ARM, etc.). Enfin, la maîtrise de la conception et de la fabrication des composants, des systèmes intégrés et des systèmes embarqués critiques joue un rôle stratégique pour les enjeux de souveraineté nationale.

VIII. Sciences de l'information et sciences du vivant

A. Les grandes avancées actuelles et les thèmes émergents

Les neurosciences sont, pour le traitement du signal et des images, à la fois un vaste champ d'application des avancées récentes en optimisation et en inférence et une source importante de nouveaux développements. Par exemple, l'étude de la connectivité cérébrale renouvelle l'intérêt et suscite de nouveaux travaux sur l'inférence de structures de graphes pour des signaux multivariés. Les problèmes inverses sont prégnants dans les questions de séparation, extraction et localisation de sources à partir de mesures EEG (crânien, intra-

crânien), EMG, IRMf, MEA (multielectrode arrays), d'amélioration d'images, de reconstruction, etc. Enfin, des problèmes spécifiques à de nouvelles méthodologies expérimentales apparaissent : la prise en compte des mesures multimodales en neurosciences devient un enjeu croissant, les comparaisons interindividuelles posent des questions de recalage difficiles, notamment en traitement d'images et de formes.

D'autres domaines liés à la médecine et à la biologie posent des questions aux sciences de l'information et profitent des avancées récentes. De gros progrès ont été faits dans le monitoring du fœtus grâce aux techniques de séparation et de localisation de sources. Des méthodes récentes d'analyse du génôme génèrent des données structurées qui trouvent un cadre naturel de modélisation dans la théorie des signaux sur graphes. Réciproquement, les sciences de l'information s'inspirent des systèmes biologiques : approches évolutives d'optimisation, robotique bio-inspirée, etc.

À l'interface avec les sciences du vivant, la bioinformatique et la modélisation des systèmes biologiques représentent des domaines applicatifs importants de nombreuses thématiques de la section : automatique, traitement du signal et de l'image, algorithmique, robotique. Par exemple, l'imagerie de processus biologiques nécessite le développement d'algorithmes d'analyse d'images performants, capables de traiter de manière robuste des quantités considérables d'images volumiques bruitées. L'étude des systèmes dynamiques de l'automatique non linéaire fournit des outils théoriques pour la modélisation, l'analyse et la simulation du fonctionnement dynamique à l'échelle globale des systèmes complexes et multiéchelles de la biologie systémique : macromolécules, cellules, organes, populations organisées en réseaux d'interactions, réseaux de régulation biologique, etc. La bioinformatique structurale, qui s'intéresse au lien structure-fonction de macromolécules biologiques, représente un enjeu pour la biologie du futur et un défi algorithmique en raison de la complexité des objets moléculaires en jeu. Elle constitue un champ d'application privilégié

de techniques de géométrie computationnelle pour la modélisation, la visualisation et le calcul d'interfaces d'interaction entre macromolécules, et aussi de techniques issues de l'algorithmique robotique pour la simulation du mouvement et de la flexibilité macromoléculaire, avec comme enjeu à terme l'aide au développement de nouvelles méthodes de CAO moléculaire pour les bio-nano-technologies.

La robotique médico-chirurgicale a connu une très forte évolution en trois décennies d'existence tant sur le plan de l'architecture (de robots polyvalents inspirés de la robotique industrielle à des robots plus spécialisés mieux adaptés aux contraintes cliniques), de leur mode de contrôle (de l'automatisation de sous-tâches à la téléopération ou à la co-manipulation laissant l'opérateur au cœur de l'action), de leur taille (depuis des systèmes encombrants et fixés au sol ou au plafond jusqu'à des dispositifs posés sur le patient ou même suffisamment petits pour être intracorporels) et enfin des applications (des structures osseuses aux tissus mous et aux approches mini-invasives de type NOTES). Les défis de cette robotique d'assistance pour l'opérateur clinicien sont à poursuivre tout d'abord en continuant de mettre au premier plan l'information médicale au cœur des dispositifs robotisés ; le robot médico-chirurgical est en effet avant tout un robot guidé par l'imagerie multimodale. Des progrès sont également attendus en direction de systèmes de type MEMS ou de systèmes cyber-physiques afin de faire des outils robotisés plutôt que des robots porteurs-outils. Mais ce n'est là qu'un des aspects du potentiel robotique dans le domaine du soin : l'assistant pour la rééducation du patient, le membre ou l'organe artificiels et enfin le robot assistant de vie pour la personne handicapée ou vieillissante sont autant d'opportunités de mettre les laboratoires français et leurs partenaires au premier plan de ce domaine fortement interdisciplinaire. La collaboration des scientifiques, des cliniciens et des industriels est certainement une clé des progrès passés et à venir.

C. Place de la France dans le contexte mondial

En robotique médicale, les laboratoires français ont joué un rôle clé dans les évolutions récentes et ont proposé des alternatives à l'état de l'art qui se sont imposées lors de ces différentes évolutions. Des entreprises ont d'ailleurs fortement valorisé ces travaux aux plans national et international. En traitement du signal et des images, de nombreuses équipes françaises collaborent avec médecins et biologistes pour appliquer les développements récents.

D. Points forts/points faibles de la recherche en France

L'interaction entre les sciences de l'information et les sciences du vivant, bien que soutenue au niveau institutionnel (e.g. GDR STIC-Santé), souffre de l'insuffisance de centres dédiés. Notons toutefois que des efforts sont faits dans cette direction, avec la présence d'équipes spécialisées dans quelques instituts (ICM, GrenobleIN, TimoneIN, par exemple) ou laboratoires. Ces thématiques souffrent de leur absence dans les formations supérieures. Les recherches à cette interface, très interdisciplinaires, nécessitent un fort investissement des chercheurs et ingénieurs dans plusieurs domaines, et leur développement bénéficierait amplement de formations croisées.

E. Retombées sociales, économiques (valorisation), culturelles

Les travaux à l'interface des sciences de l'information et des sciences du vivant sont fondamentaux pour les progrès en médecine, pour la compréhension du cerveau et des

maladies neurodégénératives, en obstétrique (monitoring du fœtus), en imageries (IRM, échographie, etc.), en chirurgie, etc.

Conclusion

Les thématiques qui composent la section 07 participent pleinement à la discipline que constituent les sciences de l'information. Elles relèvent pour la plupart d'une très longue et forte tradition en France, tant dans le monde académique que dans le monde industriel. Elles donnent lieu à des recherches couvrant tout le spectre depuis la recherche fondamentale jusqu'à la recherche au plus près des applications réelles, en passant par de nombreuses et fortes interactions avec de très nombreuses autres disciplines, de toutes les autres sciences. La mise en évidence dans la section VIII de l'interaction entre les sciences de l'information et les sciences du vivant, se retrouve également dans tous les thèmes de la section. Par ailleurs, les travaux de la section sont souvent caractérisés par de fortes, voire très fortes interactions avec les mathématiques du continu, du discret, de l'aléatoire ; les sciences de l'ingénierie et des systèmes ; la physique ; les sciences de l'univers ; les sciences de l'environnement ; les sciences humaines et sociales. Dans leur grande variété, les thèmes de la section partagent des lignes de force fondamentales, autour des modèles stochastiques et méthodes statistiques, de l'optimisation, et des modèles et méthodes distribués.

L'attractivité internationale de la France pour des chercheurs de haut niveau est forte, grâce notamment aux concours ouverts par le CNRS et aussi par Inria. Une meilleure coordination entre les organismes, et avec les établissements, a clarifié stratégies et politiques scientifiques. La mise en place de très gros laboratoires (plus de 500 personnes) et une complexité évidente dans certaines configurations locales constituent néanmoins des écueils potentiels importants.

La communauté est structurée en GDR : GDR717 MACS, GDR720 ISIS, GDR2340 MAGIS, GDR2647 STIC-Santé, GDR2995 SoC-SiP, GDR3000 IG-RV, GDR3072 Robotique, GDR3045 Vision. La plupart d'entre eux ont contribué à ce rapport, et doivent en être

remerciés. La section 07 bénéficie d'une forte structuration nationale et d'une très bonne visibilité internationale. Le potentiel de recherche élevé permettra d'encore renforcer ce point, notamment par une participation accrue aux projets européens.

Notes

(1) <http://ieeecss.org/general/IoCT2-report>

(2) http://www.univ-valenciennes.fr/gdr-macs/sites/fr.gdr-macs/files/document_id_19834.pdf

(3) <http://www.ifac-control.org/>

(4) <http://www.ieeecss.org/>

(5) <http://www.euca-control.org/>

(6) <http://www.eeci-institute.eu/>

(7) <http://www.ieee-ras.org/>

Comité national de la recherche scientifique. « Section 07- Sciences de l'information : traitements, systèmes intégrés matériel-logiciel, robots, commandes, images, contenus, interactions, signaux et langues ». *Rapport de conjoncture 2014*, [édition PDF en ligne]. ISBN : 978-2-271-08746-1.
Disponible sur : <http://rapports-du-comite-national.cnrs.fr/>