

Projet Robin: compte rendu de la réunion de lancement du 04/03/2005

La réunion de lancement du projet Technovision : Robin a eut lieu le 04/03/2005 entre 10h00 et 17h00 dans les locaux du ministère délégué à la recherche (1 rue Descarte, Paris V).

I) Liste des participants à la réunion

Nous comptons 42 personnes présentes. Les différentes entités étaient toutes représentées, seules quelques équipes de recherche n'avaient aucun membre présent à la réunion.

Etaient présents :

Les animateurs scientifiques :

- Frédéric Jurie (CNRS)
- Sylvie Philipp-Foliguet (ENSEA)

Les évaluateurs responsables de la mise en oeuvre des tests :

- Stéphane Herbin (ONERA)
- Emmanuel D'Angelo (CEP) (Matthieu Ratiéville s'est excusé de ne pouvoir être présent)

Les représentants des fournisseurs de données :

- Raphael Horak (THALES-TAS)
- Jordi Inglada (CNES)
- Alexandre Kotenkoff (MBDA)
- François Leroy (EADS)
- Jean Luc Maillart (Bertin Tech.)
- Laurent Walle (Cybernétix)
- Michel Sirieix et Michel Broekaert (SAGEM) (Jacques Lonnoy s'est excusé de ne pouvoir être présent)

Les autres experts scientifiques :

- Jean-Pierre Cocquerez (UTC)
- James Crowley (INPG)
- Roger Mohr (INPG)
- Françoise Préteux (INT)
- Gérard Giraudon (INRIA)

Henri Maître s'est excusé de ne pouvoir être présent.

Les équipes de recherche :

- Jeremy Huart (LIS-GOTA)
- Youssef Cahir (GREYC)
- Catherine Mering représentant Sebastien Gadad (Univ Paris 7)
- Lew Yan Voon (Le2i)
- Pierre Courtellemont (L3i)

- Mathieu Fauvel représentant Jocelyn Chanussot (LIS-SIN)
- Giles Keryer (ISEN)
- Christophe Rosenberger (LVR)
- Yves Caron (LIT)
- J.J. Korczak (LSIT)
- Frédéric Chausse (LASMEA)
- Mathieu Cord (ETIS)
- Gwenaëlle Toulminet (PSI)
- Philippe Montesinos (LGIP)
- Nicolas Lomenie (LSIP)
- Florent Retraint et Lionel Fillatre (ISTIT)
- Boujemaa Nozha, Valerie Gouet et Michel Crucianu (INRIA-IMEDIA)
- Bertrand Vigouroux (LISA)

Jean Ponce (Beckman Institute), Jean-François Omhover (LIP6), Maria Rifqi (LIP6), Christophe Marsala (LIP6), Marcin Detyniecki (LIP6) et Eric Novak (INRIA-LEAR) se sont excusés de ne pouvoir être présents.

Représentant du groupe de pilotage du projet Techno-Vision :

- Edouard Geoffrois

L'ingénieur expert :

- Benjamin Ninassi

Suivi du projet au ministère :

- Martine Comberousse

Invités :

- Khalid Choukri et Victoria Arranz (ELDA)

II) Déroulement de la réunion

L'ordre du jour prévu était :

Introduction à la journée (10h)

1. Objectifs de la journée
2. Présentation des participants

Evaluation d'algorithmes: Témoignages et Enseignements (10h20)

1. Françoise Prêteux (10h20)
2. James Crowley (10h30)
3. Recueil des expertises de G. Giraudon et J.P. Coquerez.

Présentation général du projet Robin (11h00)

1. Objectifs du projet
2. Chronologie du projet

Evaluation d'algorithmes: principes retenus pour le projet Robin (11h30)

1. Evaluation : objectifs et démarche. Présentation de S.Herbin
2. Présentations de E. D'Angelo / M. Ratiéville

Table ronde autour de questions clefs (avec la participation de Kalhid Choukri)
(12h15)

1. La définition des épreuves
2. Les méthodologies d'évaluation
3. L'accord de consortium
4. La formation de groupes de travail
5. Autres points...

Pause déjeuner 13H30 (sur place)

Discussions autour des bases d'images (14h30-15h)

Pour chaque base :

1. Présentation par le fournisseur : (10-15 min par base)
 - type de problème applicatif visé
 - nature et contenu des images
 - exemples d'images
 - annotations disponibles
 - taille de la base
 - date de disponibilité estimée
2. Présentation éventuelle par un ou deux laboratoires intéressés par cette base : (5-10 min par présentation)
 - type de travaux menés dans l'équipe applicables à cette base
 - adéquation de la base avec les recherches du laboratoire
 - types de bases testées
 - résultats obtenus
 - améliorations envisagées
3. Discussions autour des intérêts applicatifs/scientifiques/difficultés que les bases présentent

Conclusion (16h45)

Roger Mohr dressera les conclusions sur cette journée.

Fin de la réunion à 17h

La journée a débuté peu après 10h00. Elle a été animée par Frédéric Jurie qui, après une courte introduction présentant les objectifs de la journée, a laissé la parole aux différentes personnes présentes pour qu'elles se présentent afin que tous fassent connaissance.

Françoise Prêteux a ensuite effectué une présentation sur les questions d'évaluations et les leçons tirées des travaux des groupes MPEG, notamment sur le MPEG7. Elle a présenté, entre autre, la méthodologie utilisée dans MPEG7 et commenté le large éventail des mesures utilisées dans le cadre de MPEG7. Elle a en particulier insisté sur la nécessité d'avoir des critères d'évaluations bien définis, de préciser les besoins, les procédures, et d'avoir des jeux de données représentatifs des objectifs et des vérités terrains. Elle a également rappelé l'importance de ce type de programme et

rappelé combien il pouvait contribuer à fédérer une communauté.

Une présentation de Jim Crowley lui a fait suite, mettant l'accent sur l'importance de la performance. Il a insisté sur la nécessité d'avoir des objectifs clairs, d'ambition moyenne, qui permettent d'aboutir à des mesures claires et significatives de problèmes réels (l'évaluation doit être guidée par le besoin, le marché). Il a également soulevé l'importance de procéder à une évaluation en cycles, qui permet d'avoir un retour critique sur les procédures et les annotations et d'affiner les règles du jeu. Enfin, il a abordé le sujet de la publication des métriques utilisées et de quelques données.

Khalid Choukri a ensuite pris la parole pour nous faire part de son expérience au sein de Technolangue. Il a abordé le sujet de la capitalisation des outils de production de données annotées et la distribution (sous certaines conditions, cf <http://www.technolangue.net/>) de "packs d'évaluation" après les campagnes contenant les données et les métriques.

Par la suite, Gérard Giraudon et Jean Pierre Coquerez nous ont fait part de leur analyse critique sur le projet. Gérard Giraudon a soulevé des questions sur ce qu'on attendait de l'évaluation et commenté les points suivants :

- Evaluation d'algorithmes ou de systèmes de vision ? il s'agit pour lui d'une question clé.
- Reproductibilité, nombre de paramètres, prévision des erreurs...
- Utilisabilité (expert, étudiant, importance de l'homme dans la boucle)
- Interchangeabilité dans le cas de la prise en compte dans une chaîne de traitement en vision.

Jean Pierre Coquerez a souligné les défis:

- La diversité des bases, des objectifs, des applications...
- La généralité: ce qui est trop général va être sans intérêt, mais ce qui est trop restrictif ne trouvera pas de marché.
- Les attentes spécifiques des fournisseurs des bases, qui peuvent biaiser une application plus générale. Un soin particulier doit être apporté à leur création.

Après que Frédéric Jurie ait rappelé les objectifs du projet, les évaluateurs ont pris la parole pour faire part de leurs différentes expériences d'évaluations.

Stéphane Herbin de l'ONERA a présenté la nécessité de définir le contexte, d'avoir un formalisme homogène, de définir des métriques et des épreuves claires. Les différentes questions suivantes son intervention ont amené à conclure qu'établir un glossaire des termes employés serait une bonne chose.

La présentation d'Emmanuel d'Angelo était axée sur le projet AUT-1, un projet d'évaluation d'algorithmes de suivi de route dans un chemin peu structuré. A travers cette présentation il a tracé le bilan de ce projet en indiquant qu'il était nécessaire d'avoir une parfaite maîtrise de la base de données, de posséder des algorithmes « moulinettes » pour formater les entrées/sorties et de bien définir les métriques.

Ces présentations ont exprimé le besoin de bien connaître la finalité visée par chaque base image, afin de définir au mieux les épreuves associées.

Une table ronde a suivi afin d'avoir les opinions de chacun sur des points essentiels du projet :

- L'accord de consortium: concernant le problème de la confidentialité des bases images,

- le principe initial défini dans la proposition de projet a été retenu.
- L'ouverture du projet sur l'international : il a été décidé de retarder l'ouverture du projet à des équipes de recherche étrangères afin d'attendre d'avoir une procédure d'évaluation et des bases d'images stables, et qui fonctionne avec les participants actuels.
 - La question de la procédure d'évaluation a été abordée, en particulier afin d'éviter des biais si les participants évaluaient eux même leurs algorithmes. Il a été convenu que l'engagement moral ainsi que la non rémunération de la compétition devraient suffire à assurer le respect de la sécurité des données servant à l'évaluation des algorithmes.
 - L'intérêt des participants: outre l'accès aux bases de données images, ressources rares et nécessaires à l'avancée de la recherche en détection/catégorisation d'objets par vision par ordinateur, la question de la publication des résultats a été abordée. Il a été convenu que les résultats, si ils sont publiés, le seraient anonymement pour ce qui est du premier cycle d'évaluation. Par la suite, il faudra définir par consensus préalable aux évaluations si les résultats seront publiés de façon anonyme ou non.
 - À propos des bases images, il a été décidé de créer un groupe de travail par base. Ce groupe regroupera le fournisseur de la base, deux représentants des équipes de recherche participantes ainsi qu'un évaluateur afin de définir plus clairement la base et d'aider à la conception des épreuves.

La suite de la réunion a été consacrée aux présentations des bases images par les différents fournisseurs. Chacun d'eux (Bertin Tech. & Cybernétix, MBDA, CNES, SAGEM, THALES, EADS) a présenté brièvement comment les bases seront construites, quelles étaient les applications visées, le contenu des bases ainsi que quelques exemples d'images. Les laboratoires ont réagit après chaque présentation en exprimant leur souhaits et leurs interrogations.

À l'issus de ces présentations, Roger Mohr a dressé les conclusions de cette journée de lancement du projet Robin, qui s'est achevée à 17h00 comme prévu.

L'ordre du jour a dans l'ensemble été respecté. Il est à noter qu'un léger retard dans la matinée a entraîné un débordement de la table ronde de fin de matinée sur le début d'après midi. Pour pouvoir finir la réunion à temps, les présentations éventuelles de laboratoires sur les bases d'images (prévues après les présentations des bases) n'ont pas eut lieu.

III) Les objectifs de la journée et bilan de la réunion

Les objectifs clairement énoncés de la journée étaient :

Faire connaissance

- Connaître les fournisseurs d'images : objectif atteint.
- Connaître les évaluateurs : objectif atteint.
- Connaître les acteurs et les compétences des laboratoires : impossible à atteindre dans le temps imparti.
- Echanger sur les aspects techniques du projet : des échanges approfondis ont eu lieu, mais leur nombre a été limité par le temps.

Connaître le projet :

- Ce qu'est un tel programme d'évaluation et son intérêt : objectif atteint.
- Les objectifs du projet, c'est à dire construire une base d'images et inscrire la politique scientifique des laboratoires dans une dimension d'évaluation (définition des protocoles et des métriques, leur implémentation, annotations de la base...) : ces objectifs ont été définis dans les grandes lignes.
- Connaître les différentes épreuves et les bases associées : objectif très superficiellement atteint, les épreuves devront être précisément définies 4 mois après le lancement du projet.
- La chronologie du projet : celle-ci n'a pas été présentée faute de temps, seules quelques grandes lignes ont été évoquées (phase initiale, cycle des épreuves).

S'organiser pour spécifier les épreuves : Il a été décidé de créer des groupes par bases d'images.

IV)Récapitulatif des décisions et actions

Les décisions prises:

- Les fournisseurs d'images mettent à la disposition des laboratoires, le plus tôt possible, des exemples de leurs données, sans attendre les délais T0 + 8 ou T0 + 12.
- Autour de chaque base d'images doit se constituer un groupe de travail qui comprend :
 - le fournisseur des images qui est responsable du groupe
 - des représentants de deux groupes de recherche au moins
 - un évaluateur.
 Ces groupes de travail doivent être opérationnels fin mars au plus tard.
- Rendre les propositions d'épreuves au comité de pilotage pour fin mai au plus tard, par voie électronique.
- Une réunion du comité de pilotage aura lieu début juin pour valider ces propositions. Des navettes sont envisageables avec les groupes de travail.

Le premier livrable, c'est à dire la définition des épreuves, a pour échéance fin juin.

Frédéric Jurie devra rédiger une note destinée aux groupes de travail indiquant ce qu'il est attendu de ces groupes.