VIDEO-ID

Identification de comportements et de personnes par la vidéo
Projet ANR-CSOSG 2007

Christian FEDORCZAK[1], François BREMOND[2]

[1]THALES SECURITY SYSTEMS, Direction Scientifique et Technique,
20 rue Grange Dame Rose, 78141 Vélizy-Villacoublay

[2] INRIA, 2004 route des Lucioles,

06902 Sophia Antipolis Cedex BP 93

christian.fedorczak@thalesgroup.com, Francois.Bremond@sophia.inria.fr

Résumé – Les systèmes de vidéosurveillance se développent rapidement car ils répondent à un besoin de sécurisation des biens et des lieux accueillant du public. Ces systèmes ont longtemps été rudimentaires : des caméras et des écrans de visualisation dans un poste de contrôle. Depuis plusieurs années, des systèmes de supervision sont développés pour permettre aux agents de sécurité de concentrer leur attention sur les zones d'intérêt qui ont été repérées par un déclenchement d'alarme du type contrôle d'accès ou intrusion. A quelques exceptions près, les images des caméras sont ensuite analysées directement par les agents de sécurité.
Alors que de plus en plus de caméras de vidéosurveillance vont être installées, les méthodes d'exploitations restent très « manuelles » et les informations temps réel fournies par les caméras, largement sous-exploitées.
A partir de la vidéosurveillance, le projet VIDEO-ID a pour ambition à la fois de réaliser des détections en temps réel, voire en légère avance de phase en détectant des comportements suspects avant le passage à l'acte, et d'être capable de mener des identifications formelles par reconnaissance de visage et d'iris. Grâce à cette identification, une personne pourra être suivie à travers un réseau de caméras distantes, permettant de déterminer l'historique du parcours de cette personne ainsi que sa nouvelle destination. Sans que cela soit systématique, il y a une logique de déroulement des séquences d'identifications : situations et comportements anormaux, reconnaissances de personnes (visage), identification formelle (iris).
Les acteurs de ce projet ont tous une spécialité déjà reconnue dans une des séquences d'identification mentionnées précédemment.

Abstract – Video surveillance systems are quickly developing as they provide the right answer to secure assets and sites hosting public. For long time, these systems were very basic: cameras connected to display screens in a control room. Since several years supervision systems have been developed allowing the operators to concentrate their attention on areas of interest where an access control alarm or an intrusion alarm was triggered. Except for a very few systems, the images provided by the cameras are directly analysed by the operators.
More and more video surveillance cameras are to be installed, however their operation remains very "manual" and the real time information provided by the cameras are largely underexploited.
Using video surveillance, the Video-ID project aims to detect suspicious behaviour in real time or even slightly in advance before misdemeanour and to perform a formal identification of the author by face or even iris recognition. Using these identification data, a person can be tracked over a network of distant cameras, allowing determining the history of the displacements of the person and her/his current location. Even if not systematic, the identification sequence follows a certain logic: detection of abnormal situations or behaviours, recognition of the persons (facial recognition), formal identification (iris).
All the partners of this project already have recognized skill in at least one of the identification sequences described before.

1. Contexte et Motivation

Les systèmes de vidéosurveillance se développent rapidement car ils répondent à un besoin de sécurisation des lieux qui accueillent du public. En effet, la vidéo permet la surveillance discrète d'un site, sans nécessiter un changement de comportement ni induire des contraintes pour le public. Toutefois ces systèmes ont longtemps été rudimentaires : des caméras et des écrans de visualisation dans un poste de contrôle où des opérateurs de sécurité tentent de détecter des situations à risque ou évaluent la situation lorsqu'une alerte a été déclenchée. La vigilance d'un opérateur humain étant très limitée, (on considère

qu'au delà de 20 minutes un opérateur n'est plus en mesure de détecter correctement des situations anormales), de nombreux systèmes installés se contentent d'enregistrer le signal des caméras, de manière continue ou uniquement sur alarme. En cas de besoin (événement ou plainte) ces enregistrements sont dépouillés et exploités à des fins judiciaires, sinon il sont écrasés après une période qui ne saurait légalement dépasser un mois (sauf applications particulières). Depuis plusieurs années, des systèmes de supervision ont été développés (SATHI pour Thales), qui permettent de palier à la monotonie de la tâche des agents de sécurité et de concentrer leur vigilance sur les zones d'intérêt qui ont été identifiées par le déclenchement d'alarmes détectées par le système de sécurité.

A quelques exceptions près, les images des caméras sont ensuite analysées directement par les agents de sécurité qui vont évaluer la situation et lancer les actions nécessaires (lancement et guidage d'une équipe d'intervention pour venir en aide aux victimes et pour intercepter les auteurs des désordres.

On constate que de plus en plus de caméras de vidéosurveillance vont être installées mais que les méthodes d'exploitation sont encore très «manuelles». Les informations temps réel fournies par ces caméras sont donc au mieux sous-exploitées si ce n'est purement et simplement perdues par manque de moyens humains ou automatiques fiables.

Le projet VIDEO-ID concerne l'utilisation des images de vidéo surveillance pour la détection de situations anormales, puis le suivi et l'identification de personnes qui y ont participé. Sans que cela soit obligatoire, il y a une logique de déroulement des séquences d'identification : situations et comportements anormaux, reconnaissances de personnes (visage), suivi des personnes et identification fine (iris).

Les acteurs de ce projet ont tous une spécialité déjà reconnue dans une des tâches de la séquence d'identification.

2. Aspects sociologiques

La vidéosurveillance n'est pas un phénomène nouveau. De nombreuses villes et institutions se sont dotées de tels systèmes qui connectent un réseau de caméras déportées à des salles de supervision. La nouveauté concerne aujourd'hui le statut de l'image transmise. Il ne s'agit plus en effet d'une image analogique mais d'une image reconstruite et déjà pré-analysée sur la base d'algorithmes. On parle alors d'images « intelligentes ». Les algorithmes doivent permettre une sélection en amont des informations jugées pertinentes pour l'opérateur situé dans le centre de surveillance. Ainsi envisage-t-on de détecter des comportements « anormaux » ou inhabituels ou d'identifier dans une masse ou un flux un individu particulier.

Généralement ce genre de dispositif suscite presque toujours les mêmes réactions dans l'opinion et dans la presse, celles qui renvoient au spectre de « Big Brother ». Nous avons pu montrer dans un précédent travail consacré aux technologies d'identification biométriques que ces résistances s'avèrent en fait souvent superficielles et sont vite dissipées dans la mesure où elles passent à côté de la véritable originalité de ces technologies, celle qui consiste à extraire sur un mode automatisé un individu ou une singularité d'un flux. La différence est qualitative. Or c'est bien à ce niveau que se situent les principaux enjeux sociaux, politiques et éthiques de ces dispositifs. La question centrale est moins celle de la possibilité d'un contrôle généralisé et centralisé de la population, que celle du traitement particulier réservé à certaines catégories d'individus (avec la réalisation de profils type), traitement d'autant plus insidieux qu'il opère (dans le cas de la vidéosurveillance) à l'insu des intéressés eux-mêmes. En d'autres termes « l'acceptabilité sociale » de ces technologies peut bien être élevée, voire très élevée, cela ne signifie pas que ces mêmes technologies soient sans danger ou sans effets sur la société. Il n'y a pas d'équivalence ou de commune mesure entre ces deux niveaux d'analyse. Nous savons par exemple que le renforcement de la sécurité par des dispositifs matériels plus ou moins contraignants et en tout cas extérieurs aux individus, dégrade les rapports de confiance « spontanément » tissés au sein de la société. Cette dégradation des rapports de confiance nourrit à son tour un sentiment diffus d'insécurité qui alimente la demande sociale en dispositifs sécuritaires.

Cette gestion « particularisée » ou « personnalisée » des flux opère par ailleurs sur un mode automatique qui soulève à son tour un certain nombre de questions qui ne sont ni seulement techniques ou sociales, mais techniques et sociales, autrement dit socio-techniques. Au niveau macro-social le risque consiste par exemple à faire passer pour naturelles ou objectives, des données ou des informations qui sont en fait le résultat d'un choix, d'une décision humaine et politique. Ainsi l'identification d'un comportement dit « anormal », «étrange » ou « bizarre » suppose déjà une interprétation et un jugement sur ce qu'il convient socialement de considérer comme tel. L'automatisation solidifie et « anonymise » ce jugement dans le dispositif matériel le rendant de la sorte invisible et inaccessible à la critique comme au débat public.

A l'échelle microsociale, celle des savoir-faire mis en œuvre par les opérateurs des QG de vidéosurveillance et des éventuelles résistances des usagers, la principale source d'interrogation concerne le caractère plus ou moins « proactif » de ces systèmes. Car cette fois il ne s'agit plus de saisir une infraction ou une agression au moment où elle se produit, mais de l'anticiper à partir de modèles comportementaux intégrés au dispositif. Ce que l'image est censée détecter c'est donc une infraction potentielle. En aval cela peut conduire à renforcer le pouvoir de ceux qui auront nécessairement à interpréter ces signalements pour le cas échéant déclencher une intervention, c'est-à-dire les opérateurs des centres de contrôle. Sur quelle base justifier une interpellation sur une action qui ne s'est pas encore

produite ? Dans quelle mesure les savoir-faire actuellement mis en œuvre par les opérateurs sont-ils compatibles avec la logique de ces nouveaux outils d'aide à la décision et ne court-on pas le risque de voir se multiplier les situations arbitraires ? Telles sont quelques unes des questions auxquelles cette recherche tentera d'apporter des réponses à partir d'une enquête de terrain (qualitative) dans les salles de surveillance et auprès des associations d'usagers des transports (dans le cadre d'applications destinées à la SNCF ou la RATP).

Une innovation technologique s'inscrit toujours dans un existant, un milieu particulier qu'elle modifie, mais qui la transforme aussi en retour. Comprendre un usage, c'est donc comprendre le contexte global au sein duquel celui-ci opère. C'est ce qui fait toute la différence entre l'approche ergonomique, par exemple, qui s'intéresse plus particulièrement à la relation exclusive d'un individu à un objet (rapport H/M ou IHM) et l'approche sociologique. La sociologie replace cette relation et ses fonctionnalités dans l'univers de significations où elle prend un sens pour un individu ou un groupe social. Retrouver ce sens implicite, c'est souvent ce qui permet de comprendre pourquoi une innovation est acceptée par les usagers, se diffuse, et inversement pourquoi elle suscite des résistances. L'innovation aura d'autant plus de chance d'être adoptée qu'elle s'appuiera sur les médiations sociales, symboliques et les pratiques existantes. Cela implique une marge d'indétermination ou d'incertitude au niveau de sa conception qui la rende capable de se greffer à l'existant, d'être en prise avec les pratiques mais également et de manière décisive avec le champ de significations dans lequel ces dernières s'insèrent.

3. Aspects juridiques

Etat des lieux juridique

Le premier travail de Vidéo-ID consistera à faire un état des lieux juridique autour de la vidéosurveillance, tant au regard du droit interne (législation, réglementation, jurisprudence dont jurisprudence constitutionnelle) qu'en ayant recours au droit comparé (Etats-Unis / Grande-Bretagne / Allemagne). Cet état des lieux accordera une place importante à la jurisprudence de la Cour européenne des droits de l'homme en matière de droit à la vie privée (notamment à propos du système anglais de vidéosurveillance), ainsi qu'aux textes et projets en cours dans le cadre européen.

Droits fondamentaux

Découlant de cet état des lieux, seront examinées les questions juridiques posées en matière de droits fondamentaux : jusqu'ici, la question de la vidéosurveillance a été essentiellement envisagée sous l'angle de la protection des données personnelles, alors qu'en France et dans le cadre de la directive de l'Union européenne du 24 octobre 1995, les images de vidéosurveillance sont explicitement exclues de ce domaine (position confirmée par la loi sur le terrorisme et malgré le vœu de la CNIL d'être saisie de la question). La protection des données personnelles est progressivement devenue un domaine autonome par rapport au droit à la vie privée, dont elle est issue, la vie privée restant encore la référence aux Etats-Unis par exemple, qui n'a pas (et refuse) de développer une politique juridique en matière de données personnelles. Mais si cet aspect est important, il n'est pas le seul.

Rattachée au droit à la vie privée, la question du droit à l'image, représentant une notion juridique devenue elle aussi autonome, sera également approfondi (jurisprudence abondante tant au niveau interne qu'européen) notamment sous son aspect pénal (art. 226-1 et suivants)

Mais l'une des grandes questions posées par la vidéosurveillance est celle de la liberté individuelle, essentiellement la liberté d'aller et venir, du fait du traçage/localisation des personnes : en France, existe une tendance à confondre le droit à la vie privée avec la liberté individuelle, alors qu'il s'agit de questions juridiquement distinctes. Il existe peu de réflexions sur cet aspect, alors que les réactions de rejet et de violence vis-à-vis de ce type de système renvoient aux atteintes ressenties à la liberté d'aller et venir.

Indissociable de la question des droits fondamentaux sera également traitée la question de la sécurité, qui n'est pas un droit de la personne, mais qui renvoie à la notion d'ordre public, objectif de valeur constitutionnelle.

Preuve pénale

La question de la preuve, notamment pénale sera également traitée ; il s'agit d'un des grands thèmes de réflexion dans le cadre de la coopération policière et judiciaire pénale au niveau de l'Union européenne : elle se pose d'abord du fait de la différence entre les systèmes des Etats membres (système accusatoire où la preuve est discutée ; système inquisitoire à la française, aujourd'hui contesté, où la preuve renvoie au juge d'instruction - v. les propositions de la commission d'Outreau à propos d'une modification en matière du contradictoire). Il s'agira de délimiter juridiquement le recours à la preuve par vidéosurveillance, y compris dans des domaines non pénaux (travail, etc.)

Prise en compte des évolutions de la société

Globalement, il est difficile de séparer la question juridique de certaines questions d'ordre plus sociologique, qui doivent être appréhendées juridiquement : il s'agit par exemple de la question de la suppression de la distinction espace public / espace privé, distinction qui est un des principes fondamentaux de la démocratie. Avec l'introduction des technologies de pointe dans la vie de tous les jours des individus - dans leur vie privée comme dans leur vie du travail - cette distinction est de plus en plus brouillée. D'où la question de la protection de l'anonymat dans l'espace public et de la vie privée dans l'espace privé. Est-ce toujours possible ? Quelle organisation et quelles conséquences juridiques liées au

nouveau rôle des personnes privées (entreprises, services de sécurité etc.) dans le cadre des missions de sécurité ?

Dans l'autre sens, on assiste à l'émergence de comportements nouveaux : l'irruption actuelle de la vidéosurveillance dans la vie quotidienne génère des comportements tendant à renverser le rapport classique surveillant (Etat ou entreprise) / surveillé (population ou salarié). On constate l'utilisation de plus en plus importante des caméras par les individus contre les représentants de l'ordre (policiers) et du service public (enseignants par ex.) ou les membres de la hiérarchie dans les entreprises ou entre salariés eux-mêmes. Là également se posent de nouvelles questions juridiques sur ce brouillage des rôles.

4. Objectifs scientifiques

Le suivi de personne à travers un réseau de caméras a fait l'objet de nombreuses recherches au cours des dix dernières années ; même dans le cas simple de la détection de présence d'un individu ou d'un objet à un point donné d'une vidéo, les mises en œuvre opérationnelles sont rares en raison d'un taux de fausses alarmes liées au contexte prohibitif.

Un des objectifs essentiels de l'étude est de suivre une personne à travers un réseau de caméras distantes en l'identifiant à l'aide de sa signature visuelle calculée (couleur, texture, visage, iris).

Pour cela, la détection d'une personne, sa caractérisation fiable, son identification ainsi que son suivi dans un réseau de caméras restent des problèmes ouverts auxquels le projet Vidéo-ID apportera des éléments de réponse.

Un deuxième objectif de l'étude est de cerner avec les spécialistes des sciences de l'homme, comme avec les exploitants, les domaines pour lesquels il est plausible que le cas encore plus complexe de la détection des comportements anormaux puisse à terme atteindre un taux de fiabilité acceptable, de façon à concentrer les recherches en algorithmique et en sciences de l'ingénieur sur des outils optimisés pour ces domaines.

Finalement l'étude sociétale et juridique réalisée dans le projet permettra de vérifier l'acceptabilité de systèmes intelligents de vidéosurveillance tant d'un point de vue juridique au regard de la législation actuelle que d'un point de vue usage et intégration dans un environnement recevant du public.

5. Défis et verrous scientifiques

Les défis de ce projet sont nombreux :

- Etablir les concepts de l'emploi de l'identification par vidéosurveillance déclenchée par la détection de comportements suspects ou anormaux
- Suivre une personne à l'aide de son identification à travers un réseau de caméras
- Constituer une base de données et des discriminants permettant d'identifier un comportement suspect ou anormal
- Définir les conditions d'emploi susceptibles d'apporter une amélioration de la sécurité tout en conservant les libertés fondamentales de chacun
- Améliorer les performances intrinsèques de chacune des technologies d'analyse d'image et inventer des structures de données et des architectures permettant la mise en œuvre modulaire des différents composants dans un système unique
- Définir des critères robustes de qualité qui permettront au logiciel de supervision de ne déclencher chacune des analyses que si les résultats attendus sont exploitables (en particulier s'ils ne vont pas générer un taux trop important de fausses alarmes). Ce point est essentiel pour l'acceptabilité du système par les agents de sécurité.
- En particulier, parvenir à obtenir des identifications fiables dans des situations non coopératives et dans une dynamique d'environnement acceptable (éclairement, distances, …)
- Démontrer l'enchaînement automatique de ces opérations et en mesurer les performances « sur le terrain ».

6. Etat de l'art et objectifs du projet

Détection de comportements anormaux

De nombreux travaux ont été réalisés dans la communauté vision par ordinateur afin de reconnaître des activités humaines normales et anormales [1]. De nombreux systèmes de vidéosurveillance intelligente ont été développés dans différents contextes applicatifs. Par exemple, les projets PRISMATICA [2] et ADVISOR [3] reconnaissent des scénarios anormaux (violence, vandalisme, fraude, graffiti) dans le métro, SAMSIT [4] reconnaît des scénarios similaires en embarqué dans des trains, VISOR-BASE [5] détecte des comportements de fraude dans des centres commerciaux, et AVITRACK [6] surveille les opérations de maintenance dans les aires de stationnement des avions. Bien que ces systèmes donnent de bons résultats sur les séquences vidéo et/ou audio traitées, leurs performances se dégradent rapidement dès que les conditions d'utilisation évoluent. Par exemple, il n'existe pas à l'heure actuelle de système vidéo capable de détecter automatiquement des individus dans des environnements bruités (eg. brusque changement d'illumination, forte pluies, pénombre) ou très encombrés (eg. lieux confinés, présence de foule).

L'objectif du projet Vidéo-ID consiste donc à étendre les capacités des systèmes de vidéosurveillance intelligente afin de détecter des individus avec précision dans tout type de situation et d'aller jusqu'à leur identification.

Suivi de personnes

En ce qui concerne la détection et le suivi d'individus, l'approche la plus efficace consiste à utiliser une image de référence correspondant à la scène vide [7]. Malheureusement cette approche permet uniquement de détecter les portions de la scène où le changement a lieu (correspondant à des régions mobiles). Il reste à extraire de ces régions mobiles les personnes évoluant dans la scène. Cette opération est relativement fiable lorsque l'environnement est peu bruité. Pour opérer dans des environnements bruités ou encombrés, il s'agit de compléter cette approche, par exemple, par des techniques de détection et suivi de descripteurs locaux (eg. blob de pixels avec forte texture) pouvant caractériser une personne. Ces descripteurs peuvent être ensuite mis en correspondance avec des modèles statistiques de personnes [8]. De façon similaire, des descripteurs locaux peuvent êtres utilisés pour détecter les visages de personnes faisant face à la caméra [4]. Dans le projet Vidéo-ID nous combinerons ces différentes techniques afin de pouvoir localiser précisément le visage d'une personne, même évoluant dans des environnements complexes.

Reconnaissance de visages

La reconnaissance de l'identité par le visage est un domaine de recherche très actif du fait du caractère non intrusif et sans contact, voire sans coopération.

La plupart des systèmes de reconnaissance de visage nécessite un prétraitement du signal (localisation, segmentation). Cette tâche est difficile puisqu'elle dépend fortement des conditions d'acquisition du signal (illumination, « background » complexe) et de la coopération des utilisateurs (pose). La segmentation, partie intégrante du traitement du signal, consiste à extraire d'un flux vidéo l'information pertinente et éliminer par filtrage le reste (par exemple, l'arrière-plan d'une image).

En biométrie, la reconnaissance sera différente s'il s'agit d'une identification ou d'une vérification. En mode identification, le système doit retrouver l'identité d'une personne dans une liste d'individus connus. Dans le cadre de ce projet, il s'agira de comparer un individu vu par l'une des caméras du réseau à une liste d'individu prédéfinie dans une base. Dans ce mode, le système compare l'image du visage avec les différents modèles présents dans la base de donnée et propose un classement. En mode vérification, l'utilisateur (ou une tierce personne) propose une identité au système et ce dernier doit vérifier que l'identité de l'individu est bien celle proposée.

Plusieurs approches ont été étudiées en ce qui concerne la reconnaissance de visage [9, 10, 11]. Ainsi, depuis ces dernières années les performances de tels systèmes se sont grandement améliorées. Les travaux peuvent être classés en deux catégories :

Approche discriminante : il s'agit de prendre une décision binaire (est-ce ou non le visage d'un client ?) et de considérer l'image en entier. De telles approches holistiques utilisent généralement l'image originale de visage en niveaux de gris ou sa projection dans le sous-espace de composantes principales (PCA, Fisherfaces) ou le sous-espace linéaire discriminant comme données d'entrée au classifieur discriminant du type perceptron multicouche (MLPs) [12,13], machines à support de vecteurs (SVMs) [14] ou une métrique plus classique [15,16].

Approche générative : il apparaît que les approches génératives telles que le modèle Gaussien (Gaussian Mixture Models) [17] et le modèle de Markov (Hidden Markov Models [18, 19, 20, 21]) sont plus robustes suite à une localisation automatique du visage. L'approche générative calcule la vraisemblance d'une ou d'un ensemble d'observations d'un modèle client donné et la compare avec la vraisemblance correspondante au modèle imposteur.

Finalement, la décision d'accepter ou de rejeter un individu dépend du score (distance, MLP, taux de vraisemblance) et d'un seuil donné.

Seuls quelques rares travaux s'intéressent réellement à la dynamique des visages en reconnaissance [22] ou émotion. Même si on peut signaler quelques laboratoires qui travaillent actuellement à identifier et reconnaître automatiquement quelques signes de base comme le mouvement global de la tête (haut/bas pour dire oui, ou gauche/droite pour dire non), la plupart des autres travaux utilisent et traitent les vidéos comme des séries d'images fixes (indépendantes), et donc comme une source multi images; et ce plutôt dans un contexte de reconnaissance de visages. Egalement, les techniques utilisées ne sont pas spécifiques à la vidéo mais adaptées du monde des images fixes. Autrement dit, l'axe temporel et donc dynamique est assez peu investigué.

Il sera nécessaire au cours de ce projet de préciser la notion de dynamique d'un visage, et des liens entre émotions et reconnaissance.

L'émotion comme la joie, la peur, etc. est universelle, et existe pour tous. Par contre, chacun de nous a un sourire spécifique, une façon de parler, etc. On peut donc dire que l'apparence dynamique d'un visage est le résultat de deux jeux de paramètres :

- inter classe (joie, tristesse, etc.),
- intra classe (dissymétrie du visage, manière exacte de sourire, ouverture de la bouche, langue visible ou non, etc.).

Les premiers paramètres sont plutôt liés à l'émotion tandis que les seconds à la personne. Savoir extraire les premiers pour identifier une émotion (paramètres communs à tous pour réaliser une expression donnée) ou les seconds pour identifier une personne (i.e. mimique) est indispensable pour bien analyser les expressions faciales et ensuite les interpréter.

Il est à noter que la quasi totalité des travaux publiés en analyse de visages (images fixes ou animées) opèrent habituellement sur des données de bonne qualité, ce qui ne sera pas le cas ici. Il faudra donc adapter nos algorithmes et repréciser nos objectifs en fonction des limites imposées par la qualité des données disponibles à traiter.

Reconnaissance de l'iris

La reconnaissance des personnes par l'iris est aujourd'hui considérée comme le moyen le plus fiable pour identifier des personnes en se basant sur leurs informations biométriques. L'iris est le seul organe interne (hautement protégé) du corps humain qui est visible de l'extérieur. L'iris possède une texture riche (la plus riche en biométrie), unique pour chaque personne (même l'iris droit et gauche d'une même personne sont différents), indépendante de l'ADN.

Pourtant, les systèmes existants à base d'iris requièrent de l'utilisateur une forte coopération à l'acquisition, ce qui veut dire : rester immobile, se positionner par rapport à la caméra à la distance requise pour une capture optimale de l'image de l'œil (quelques dizaines de centimètres). Ces contraintes sont vécues par certains utilisateurs comme très fortes et génèrent des rejets à l'acquisition lors de la mise en place des systèmes à grande échelle. Très peu de travaux sont publiés sur la faisabilité de la reconnaissance par l'iris que ce soit en mode non coopératif ou à distance. En réalité, seule une compagnie américaine, Sarnoff Corp. [23], propose une solution dans ce sens. Lors d'une première phase elle a développé un démonstrateur d'identification des personnes dans un environnement contrôlé mais sans contraindre l'utilisateur. Le démonstrateur s'appelle Iris On the Move (IOM). Il peut reconnaître une personne qui marche à vitesse normale à 5 ou 6 mètres de distance et doit donc résoudre des problèmes que les autres systèmes évitent généralement : le mouvement de l'œil, le port des lunettes ou des lentilles, attitude non coopérative. Le système doit aussi prendre en compte d'autres variations, comme celle de la vitesse de marche ou des tailles des individus. Enfin l'algorithme de reconnaissance opère sur des images de qualité moindre que celle obtenue sous une acquisition contrainte. Le programme de recherche IOM est encore au stade de l'expérimentation, les premiers résultats publiés montrent les difficultés de la tâche avec un taux de fausse identification de 22% sur une base de 119 personnes. Une étude plus approfondie sur les causes d'erreurs a montré que 30% des fausses identifications sont dues au fait que les personnes regardaient vers le bas ou fermaient les yeux tentant de mettre délibérément le système en échec. Le port des lunettes semble être aussi un défi important à étudier. Thales a eu l'opportunité de disposer du système durant une quinzaine de jours et a pu l'évaluer sur quelques centaines de volontaires. L'enrôlement des volontaires a été réalisé à partir un système dédié de type contrôle d'accès (enrôlement en mode coopératif et contraint avec contrôle de la qualité de la capture). L'identification 1/N a été réalisée avec le système IOM. Les résultats de cette évaluation ont démontré que cette technologie atteint des performances très correctes sans nécessiter de contraintes d'exploitation importantes (seule contrainte : passer sous un portique du type détection d'objets métalliques dans des conditions similaires à celles d'un aéroport. A ce jour le fonctionnement du démonstrateur est jugé suffisamment satisfaisant pour aller vers le développement d'un produit.

L'application visée par le programme IOM est un portique d'identification à l'entrée de zones sensibles (aéroports,...) et ne correspond pas à des applications de type vidéosurveillance où l'acquisition est faite à l'insu de la personne.

En ce qui concerne l'acquisition d'images d'iris à distance, des travaux préliminaires ont montré la faisabilité d'acquérir un iris de bonne qualité (rayon d'iris moyen : 128 pixels) à une distance de 10m à condition d'avoir une optique très performante et des projecteurs de lumière infrarouge [24]. D'autres résultat publiés par IOM ont démontré qu'une acquisition de l'iris avec un diamètre de l'ordre d'une centaine de pixels pouvait être suffisant [25].

Le contexte de la vidéosurveillance n'offre pas ces possibilités car les capteurs actuellement en usage sont non adaptés. L'évolution des systèmes de vidéosurveillance vers des solutions numériques sur IP permet le mélange d'équipements hétérogènes sur un même parc. Ainsi le format des images n'est plus limité au seul standard de télévision à 625 lignes entrelacées par image mais on trouve également des capteurs VGA progressifs ou haute résolution. En particulier les caméras « mégapixels » que l'on trouve désormais sur le marché à des prix très abordables permettent d'envisager la saisie de détails du visage ou de l'iris à distance. Un des objectifs de ce projet est d'étudier la faisabilité d'une infrastructure permettant une identification par l'iris (sur une petite base) dans un contexte vidéosurveillance.

7. Objectifs industriels

Dans le domaine de la sécurité publique, Thales occupe une position de systémier reconnue mondialement. Le développement de logiciels de supervision intégrant des fonctions « intelligentes » fait partie de sa stratégie de différentiation par rapport à la concurrence. Thales a pour vocation d'intégrer à son offre, dès qu'ils sont matures, les résultats de la recherche répondant à cet objectif. Thales a développé les systèmes de vidéosurveillance et de sécurité qui équipent de nombreux métros, aéroports, musées ou sites sensibles dans le monde entier ; Thales a également livré plus de 5000 systèmes vidéo embarqués sur véhicule en service dans le monde. Ces systèmes sont fournis clés en main à des opérateurs de transports publics ou de services de sécurité, demandeurs de solutions efficaces, fiables et de coût d'exploitation minimum.

Les retombées industrielles et économiques escomptées sont de plusieurs ordres :

- développement du marché total grâce à des avancées convaincantes pour les clients
- accroissement des parts de marché pour les sociétés ayant accès à ces technologies
- renforcement de la part interne dans l'offre globale du systémier

Ces retombées devraient avoir des effets bénéfiques sur l'emploi de plusieurs façons :

- postes techniques dans les entreprises concernées afin d'intégrer l'ensemble des résultats
- postes « business development » afin de faire la promotion des ces solutions de sécurité dont les performances auront été démontrées
- postes projet dans les entreprises concernées afin de construire les systèmes de sécurité globaux intégrant ces nouvelles fonctionnalités
- postes pérennes dans les start-up qui vont développer certaines des solutions qui auront été étudiées et dont les grandes entreprises systémières auront besoin pour enrichir leur offre.

8. Le consortium

Le consortium de Vidéo-ID comporte trois laboratoires scientifiques (INRIA Sophia, EURECOM et GET/INT), deux centres de recherches en sciences humaines (GET/INT et Paris X) et un industriel (Thales). Le comité de pilotage du projet est composé de représentants de chacun des partenaires et est également complété par des représentants des prescripteurs ou des utilisateurs finaux (Ministère de l'Intérieur, RATP, UIC et SNCF).

THALES SECURITY SYSTEMS

Thales est un leader mondial de l'électronique et des systèmes. Partout dans le monde, le groupe sert les marchés de la défense, de l'aéronautique et de la sécurité, appuyé par une offre globale de services. Le groupe optimise le développement parallèle des activités civiles et militaires au service d'un seul objectif : la sécurité des personnes, des biens et des États. Avec plus de 20 000 chercheurs de très haut niveau, Thales constitue une capacité unique en Europe pour créer et déployer des systèmes d'information critiques. Fort de 60 000 personnes dans cinquante pays, Thales a enregistré en 2005 un chiffre d'affaires de 10,3 milliards d'euros et un carnet de commandes record de plus 20 milliards d'euros.

Les activités Security Systems portent sur la conception, la fourniture et l'intégration de systèmes technologiques destinés à protéger les infrastructures critiques des entreprises, des collectivités locales et des organismes gouvernementaux du monde entier. Ces systèmes couvrent tous les aspects de la sécurité globale : sécurité des sites et des événements, centres opérationnels de sécurité et de gestion de crise, solutions d'identification sécurisée, sécurité des systèmes d'information, sécurité de l'environnement. Ces activités comprennent une large palette de prestations, qui va des services de conseil en sécurité, définition architecturale, intégration de systèmes complexes et de solutions clés en main jusqu'à la conception, installation, implémentation et maintenance de systèmes de sécurité.

Thales Security Systems assurera la coordination du projet et apportera ses compétences en matière de définition et d'intégration de systèmes de sécurité et de surveillance.

INRIA

L'INRIA, institut national de recherche en informatique et en automatique placé sous la double tutelle des ministères de la recherche et de l'industrie, a pour vocation d'entreprendre des recherches fondamentales et appliquées dans les domaines des sciences et technologies de l'information et de la communication (STIC). Jouant un rôle fédérateur au sein de la communauté scientifique de son domaine et au contact des acteurs industriels, l'INRIA est un acteur majeur dans le développement des STIC en France.

ORION (nouvellement PULSAR) est une équipe pluridisciplinaire de l'INRIA, à la frontière des domaines de la vision par ordinateur, des systèmes à base de connaissances et du génie logiciel. L'objectif d'Orion est de concevoir et de développer des techniques et des logiciels pour l'interprétation automatique d'images et la réutilisation et le pilotage automatique de programmes. Plus précisément, nos thèmes de recherche portent sur la conception des systèmes intelligents basés sur les techniques de la représentation de connaissances, de l'apprentissage et du raisonnement. En interprétation de vidéo, partant de la détection des objets mobiles (êtres humains), du suivi, et de la classification des objets, nous déterminons le comportement de personne(s) et les scenarii en nous basant sur des modèles prédéfinis. Une plate-forme générique, appelée VSIP (Video Surveillance Intelligent Platform) a été développée et appliquée avec succès dans plusieurs projets.

Six thèses de doctorat ont été achevées dans le domaine de vidéosurveillance au sein du projet ORION. L'équipe comprend deux chercheurs a plein temps, cinq ingénieurs, et sept étudiants en thèse travaillant dans ce domaine. ORION a participée et participe aux nombreux projets européens et nationaux et est reconnu comme un des leaders dans le domaine de l'interprétation vidéo. En plus, nous avons réussi 2 transferts de technologie et la création d'une start-up Keeneo.

EURECOM

L'Institut EURECOM est une école d'ingénieurs et un centre de recherche en systèmes de communication situé dans le Parc scientifique de Sophia Antipolis, dans le sud de la France, qui s'affirme aujourd'hui comme le pôle d'excellence européen dans le domaine des hautes technologies

Il a été fondé le 04 novembre 1991 par l'Ecole Polytechnique Fédérale de Lausanne (EPFL) et Télécom Paris (Ecole Nationale Supérieure des Télécommunications - ENST). C'est un GIE (Groupement d'Intérêts Economiques) qui compte aujourd'hui un certain nombre de partenaires académiques dont le Politecnico de Turin et The Helsinki University of Technology, ainsi que des industriels : Swisscom, Thales, France Telecom, SFR, Hitachi Europe, Bouygues Télécom, STMicroelectronics,

Sharp, BMW Group Research & Technology et Cisco Systems, et un Partenaire industriel associé : SAP

Son budget est de 9,2 millions d'Euros dont 3 millions sont issus des contrats de recherche. Il emploie quelques 128 personnes dont 97 sont des chercheurs, les 31 personnes restantes constituent les services informatique et administratifs

L'Institut est structuré autour de ses activités de recherche dans trois domaines :

- **Département Communications Mobiles**
  Protocoles pour Réseaux Mobiles, Radio logicielle (Plate-Forme, Traitement du Signal et Théorie de l'Information appliqués aux Communications Mobiles, Réseaux ad hoc et réseaux locaux sans fil, Réseaux Ultra Wide Band, Propagation et l'analyse de performance système.
- **Département Communications d'Entreprise**
  Protocoles et les Applications Internet (Réseau pair à pair, distribution des contenus ; Tomographie des Réseaux) ; Sécurité des Réseaux (Protocoles de sécurité, Détection d'intrusion).
  - les réseaux cellulaires de 3ème génération et au-delà
  - les réseaux locaux sans fil
  - les réseaux de communication à courte portée (capteurs)
  - les réseaux Ad Hoc
- **Département Communications Multimédia**
  Indexation Multimédia (Analyse de séquences vidéo ; Classification, recherche d'informations multimédia ; Filtrage d'informations multimédia), sur le Traitement de la Parole (Analyse acoustique, Reconnaissance du locuteur), sur l'Intelligence Emotionnelle (Interaction homme-machine) et sur l'Imagerie Multimédia (Biométrie, Tatouage, Clonage virtuel).

Le groupe image du département Communications Multimédia, dirigé par Jean-Luc Dugelay, est composé actuellement d'un professeur, d'un Ingénieur ainsi que de cinq doctorants. Il a développé ces dernières années une expertise en Biométrie, en tatouage numérique. En effet, ses travaux de recherche sur la biométrie ont débuté en 2001. Depuis, le laboratoire est un spécialiste reconnu en reconnaissance des visages 2D. Ses thèmes de recherche l'ont amené naturellement à s'intéresser fortement aux nouvelles modalités en rapport avec le visage, en particuliers dynamique (i.e. vidéo) et 3D. Eurecom est impliqué dans plusieurs projets en Biométrie au niveau Européen (NoE **BIOSECURE**) et National (Technovision **IV2**, RNRT **BIOBIMO**, **MISTRAL**). Eurecom est d'ailleurs leader du projet BIOBIMO, débuté en 2005.

Depuis sa création l'institut a participé activement à de nombreux projets de recherches nationaux et européens, dont sont actuellement en cours :

- 19 européens : Similar, ESA 4480, Divines, Biosecure, Newcom, Asia, Daidalos II, E2R II, Cascadas, Resist, Haggle, R4eGov, Kspace, Portivity, Multinet, Unite, Cruise, Chorist, Coopcom
- Et 16 nationaux : Lao Tseu, Cosinus, Technovision Argos, Technovision IV², Biobimo, Aces, Grace, Airnet, Semafor, Opus, Idromel, Mistral, Winem, Ormac, Maxsimm, Plateforme Multimodale

GET-INT

L'INT Evry apporte ses compétences reconnues en recherche sur la biométrie de l'iris et de la vérification par le visage ainsi qu'en études sociologiques.

La recherche en Biométrie au GET-INT (http://www.int-evry/biometrics) s'effectue au sein du projet Bio-Identité. Les activités déployées dans ce projet font suite à 2 projets initiateurs BIOMET et BIOLAB (concernant les usages). Le projet BIOMET : « Vérification biométrique multimodale de l'identité » (projet incitatif GET 2000-2002). Les modalités présentes sont : signature en ligne, visage, empreintes digitales, forme de la main, voix, ainsi que de leur fusion pour la vérification biométrique multimodale de l'identité).

Valorisation européenne et industrielle :

GET-INT est coordinateur (B. Dorizzi) du réseau d'excellence européen BioSecure. Les travaux issus du projet Bio-Identité sont valorisés au niveau de la France et de l'Europe au travers de collaborations industrielles (THALES, Atmel,…) et institutionnelles (Ministère de l'Intérieur…). D'autre part, le projet européen SECUREPHONE, cordonné par ATOS Origin, s'intéresse plus particulièrement aux aspects mobilité (PDA, téléphone mobile). L'ACI Bio-Mul permet aussi de se poser des questions plus fondamentales dans le cadre d'un projet de recherche amont. L'INT est aussi partenaire du projet oppidum VINSI (portage de visage et empreintes sur un terminal nomade sécurisé), du projet ANR BIOTYFUL (BIOmetrics and crypTographY for Fair aUthentication Licensing) et MyBlog3D. Le GET participe activement au sein de l'AFNOR aux travaux du groupe SC37 sur la biométrie, qui s'insère dans des efforts de normalisation internationaux conduits par l'ISO (International Standards Organisation).

Le GET/INT maîtrise un ensemble d'algorithmes biométriques, développés dans un premier temps pour des conditions d'utilisation non contrainte. Certains d'entre eux (tels que la vérification du scripteur par la signature et la vérification des personnes avec des visages 2D sont en cours d'adaptation (algorithmique et logicielle) pour un usage en mobilité. Le savoir-faire du GET touche aussi le domaine de la multi-biométrie, de l'évaluation des performances des systèmes biométriques et des usages.

Un brevet sur la vérification par signatures dynamiques et un autre sur la vérification par l'iris sont en cours de dépôt.

PARIS X-CREDOF

Le CREDOF a été créé en 2000 en vue de fédérer l'ensemble des recherches portant sur les droits de l'homme à l'Université Paris X-Nanterre. Il sert d'équipe d'accueil pour le Master « Droits de l'homme » et les doctorants. Dirigé par Danièle Lochak jusqu'en 2006, il est à présent dirigé par Véronique Champeil-Desplats.

Le CREDOF s'occupera des aspects juridiques liés à la vidéosurveillance.

Références

[1]. A. Avanzi, F. Bremond, C. Tornieri and M. Thonnat, Design and Assessment of an Intelligent Activity Monitoring Platform, in *EURASIP Journal on Applied Signal Processing, special issue in "Advances in Intelligent Vision Systems: Methods and Applications"*, 2005.

[2]. L. Khoudour, J. Hindmarsh, D. Aubert, S. Velastin and C. Heath (2001)Enhancing security management in public transport using automatic incident detection. In L. Sucharov and C. Brebbia, editors, Urban Transport VII: Proceeding of the 7th International Conference on Urban Transport and the Environment for the 21st Century, pages 619-628, Southampton, United Kinkdom. WIT Press.

[**3].** F. Cuppilard, A. Avanzi, F. Bremond and M. Thonnat (2004). Video understanding for metro surveillance. In Proceedings of the IEEE International Conference on Networking, Sensing and Control, Special Session on Intelligent Transportation Systems (IC-NSC'04), Taipei, Taiwan.

[**4].** V.T. Vu, F. Bremond, G. Davini, M. Thonnat (2006). Audio-Video Event Recognition System for Public Transport Security. ICDP 2006.

[5] J. Piater, S. Richetto and J. Crowley (2002). Event based activity analysis in live video using a generic object tracker. In J. Ferryman, editor, Proceedings of the 3rd IEEE Workshop on Performance Evaluation of Tracking and Surveillance (PETS'02), Copenhagen, Denmark.

[6]. M. Borg, D. Thirde, J. Ferryman, F. Fusier, V. Valentin, F. Brémond and M. Thonnat, "A Real-Time Scene Understanding System for Airport Apron Monitoring". The *Proceedings of 2006 IEEE International Conference on Computer Vision Systems*, New York, USA, January 5-7, 2006.

[7] A. Mittal, L. Davis (2003). M2tracker: A multi-view approach to segmenting and tracking people in a cluttered scene. International Journal of Computer Vision, 51(3): 189-203.

[8]. P. Viola, M. Jones and D. Snow (2005). Detecting pedestrians using patterns of motion and appearance. IJCV, 63(2): 153-161, 2005.

[9] R.Chellappa, C.L Wilson et C.S Barnes. Human and recognition of faces : A survey. Technical report CAR-TR-731, University of Maryland, 1994

[10] J.Zhang, Y.Yan and M.Lades. Face Recognition : Eigenfaces, Elastic Matching, and Neural Nets. In Proceedings of IEEE,volume 85, pages 1422-1435,1997

[11] K.Messer, J.Kittler, M.SAdeghi, M.Hamouz, A.Kostyn, S.Marcel, S.Bengio, F.Cardinaux, c.Sanderson, N.Poh,Y.Rodriguez, K.Kryszczuck, J.Czyz, and al. Face authentication competition on the BANCA databse. In Proceeding of the International Conference on Biometric Authentication (ICBA), Hong Kong. July 15-17 2004.

[12] S.Marcel. A symmetric transformation for lda-based face verification. In proceedings of the 6th International Conference on Autmatic Face and Gesture Recognition. IEEE Computer Society Press, 2004

[13] S.Marcel et S.Bengio. Improving face verification using skin color inforamtion. In proceedings of the 16th ICPR. IEEE Computer Society Press, 2002

[14] K.Jonsson, J.Matas, J.Kittler et Y.P.Li. Learning support vectors for face verification and recognition. In the 4th International Conference on Automatic Face and Gesture Recognition, pages 208-213,2000

[15] Y.Li, J.Kittler et J.Matas. On matching scores of LDA-based face verification. In T. Pridmore et D.Elliman, editors, Proceedings of the british Machine Vision Conference BMVC2000. British Machine Vision Association, 2000

[16] J.Kittler, R.Ghaderi, T.Windeatt et G.Matas. Face verification via ECOC. In British Machine Vision Conference (BMVC01), pages 593-602,2001

[17] F.Cardinaux, C.Sanderson et S.Marcel. comparison of MLP et GMM classifiers for face verification on XM2VTS. In Proceedings of the International Conférence on Audio- and – Video-Based Biometric Person Authentication. Springer- Verlag, 2003

[18] A.Nefian and M.Hayes. Face recognition using an embedded HMM. In Proceedings of the IEEE conference on Audio and Video-based Biometric Person Authentication (AVBPA), pages 19-24,1999

[19] A.Nefian and M.Hayes. Hidden Markov Models for Face Recognition. In International conference on acoustics Speech and Signal Proceeding (ICASSP), pages 2721-2724, 1998

[20] S.Eickeler, S.Müller et G. Rigoll. High Performance Face Recognition Using Pseudo 2D-Hidden Markov Models. In European control conference (ECC), Karlsruhe, Germany,1999

[21] F.Cardinaux, C.Sanderson et S.Bengio. face verification using adapted generative models. In the6th International Conference on Automatic Face and Gesture Recognition, Seoul, Korea, 2004, IEEE.

[22] U. Saeed, F. Matta and J.-L. Dugelay, Person Recognition based on Head and Mouth Dynamics, IEEE MMSP 2006.

[23]http://www.sarnoff.com/products_services/government_solutions/homeland_security/iris.asp

[24] C. Fancourt, L. Bogoni, Y. Guo, N. Takahashi, U. Jain, "Iris recognition at a distance", Proc. Audio and Video-Based Biometric Person Authentication Conference (AVBPA 2005), New York, 20-22 juillet 2005.

[25] J. R. Matey, O. Naroditsky, K. Hanna, R. Kolczynski, D. J. LoIacono, S. Mangru, M. Tinker, T. M. Zappia, and W. Y. Zhao, "Images for Iris Recognition in Less Constrained Environments", Proceedings of the IEEE Vol. 94,No. 11, November 2006