



# Vidéosurveillance : Trop de caméras, pas assez d'yeux ?

TECHNOLOGIE

Les recherches se multiplient pour automatiser la détection des comportements « déviants » et reconnaître les individus « à la volée ». Tour d'horizon de ces techniques sécuritaires

MARINA JULIENNE

**U**n homme déambule entre les allées d'un parking et s'attarde « anormalement » près d'une voiture. A 3 kilomètres de là, devant son mur d'écrans relié à une trentaine de caméras, un opérateur se lasse de voir passer des anonymes et s'assoupit. Mais un signal d'alarme attire son attention sur l'écran 15, où la silhouette du maraudeur est encerclée de rouge. Si l'opérateur en croit l'« intelligence » de la caméra, cet homme s'apprête à voler un véhicule. Il avertit donc la police, qui sera sur les lieux cinq minutes plus tard. Et trouvera un voleur... ou un rêveur.

Après avoir été présentées comme la panacée pour lutter contre la délinquance, les caméras de vidéosurveillance sont jugées assez inefficaces par les professionnels de la sécurité : la capacité d'attention d'un opérateur ne dépasserait pas deux heures pour huit écrans. Les caméras ont sans doute un effet dissuasif, mais n'ont pas fait chuter la délinquance (*Le Monde* du 29 octobre 2011).

Les logiciels de « détection des comportements anormaux » vont changer la donne, assurent les promoteurs de la vidéosurveillance intelligente (VSI). En 2006, l'Agence nationale de la recherche (ANR) finance un projet du laboratoire d'informatique de Lille (LIFL-CNRS) : CaNADA, acronyme de « comportements anormaux : analyse, détection, alerte... ». Objectif : « Classer les comportements d'individus afin de les interpréter en termes de menace », et « établir une typologie des signaux visuels de dangerosité ». Dès le départ, des chercheurs en droit et sciences politiques sont associés au projet, coordonné par l'Ecole des mines de Douai,

pour en discuter les questions éthiques. Leurs réflexions ont bien été publiées, mais, sur le terrain, elles ne font plus l'objet de débat...

Pour l'instant, les comportements détectés se résument à quelques actions basiques – courir, marcher, lever les bras, tomber – qui font déjà le bonheur de certains services de sécurité. A la réunion intitulée « Mise en œuvre d'une détection automatique d'anormalité », organisée en décembre 2011 par l'AN2V (Association nationale de la vidéoprotection), plusieurs sociétés présentaient des logiciels qui donnent automatiquement l'alarme si une personne chute dans une foule, si une autre déambule trop longtemps, si un petit groupe se met brusquement à courir, etc. Les villes de

coup moins en milieu urbain : en cas de crise par exemple, les protagonistes ont une fâcheuse tendance à fuir chacun dans une direction. Pas facile pour une caméra, si intelligente soit-elle, de choisir qui suivre... Le système sur lequel travaille la société Thales fonctionne d'ailleurs en mode « assisté ». Imaginez un salon grand public, truffé de caméras : sur son écran, l'opérateur sélectionne l'individu suspect, affiche sa « signature visuelle » (image entière de l'individu), ce qui permet aux différentes caméras de le traquer. Les informations de suivi sont transmises sur une tablette numérique dont dispose un second opérateur qui se trouve, lui, dans les allées du salon et peut se rapprocher de l'individu pour l'arrêter.

Pour l'instant, ce sont les utilisateurs du système qui décident que dans telle situation (un parking) le comportement normal pour un individu consiste à aller directement à sa voiture, alors que dans telle autre (un magasin) il est normal de déambuler dans les allées mais normal de rester planqué plus de cinq minutes devant un rayon. Bien sûr, il n'est pas simple de tracer la frontière entre le normal et l'anormal.

Une étude menée par le sociologue Tanguy Le Goff (Institut d'aménagement et d'urbanisme d'Ile-de-France) sur des opérateurs de vidéosurveillance a montré que les jeunes sont moins « ciblés » (suivis sur écran) pour ce qu'ils font (souvent d'ailleurs ils ne font rien, ce qui est jugé suspect) que pour leur apparence vestimentaire et physique. Les caméras donneront-elles l'alerte en cas d'apparition d'une casquette sur une personne au visage basané et, circonstance aggravante, avachie sur un banc ?

Toujours dans l'optique d'assistance aux opérateurs, certaines sociétés (Orélia par exemple) proposent d'ajouter aux caméras des capteurs audio. Un bris de ver-

re, des cris perçants, et la caméra zoomé vers la scène supposée violente. Inconvénient : la portée ne dépasse pas les 20 à 50 mètres selon le bruit de fond ambiant. A moins de truffer les réverbères de micros, l'efficacité n'est pas garantie. La société Surveillance, elle, s'est spécialisée dans la lecture automatisée des plaques d'immatriculation. Intégrée à une caméra placée sur un véhicule de police, le bloc optique infrarouge fonctionne aussi la nuit, et compare toute plaque d'immatriculation croisée dans son champ de vision au fichier des voitures volées. Il lit même les plaques d'un véhicule roulant en sens inverse à grande vitesse...

## Dates

**Février 2011** Le ministère de l'intérieur annonce un triplement, en France, du nombre de caméras sur la voie publique, qui passerait de 20 000 à 60 000.

**D'ici à la fin 2012** Le plan de vidéoprotection pour Paris prévoit d'installer 1 500 caméras dans les rues de la capitale. En regroupant les caméras de la SNCF et de la RATP, c'est près de 13 500 caméras qui composeront le dispositif de la ville. Nice, avec 744 caméras, détiendra le record de la ville la plus vidéosurveillée de France (une pour 553 habitants).

**2013** Châteauroux, avec 270 caméras, pourrait détrôner Nice. Le ratio sera alors d'une caméra pour 177 Castelroussins.

**Présentées comme la panacée pour lutter contre la délinquance, les caméras de vidéosurveillance sont jugées assez inefficaces par les professionnels de la sécurité**

Cannes, Nice, et bientôt Lyon, installent des logiciels programmés pour détecter les colis suspects, les voitures en contresens, le maraudage, les attouchements. Différentes sociétés comme Keeneo (issue de l'Inria), Kaolab, Foxstream (issue du Liris) proposent de tels produits.

Certaines caméras ont une fonction « tracking », pour suivre un individu et se passer le relais entre elles sans le perdre de vue. Ce qui fonctionne assez bien dans des lieux fermés (sur un site industriel), beau-



Dans les laboratoires de recherche, la compétition est intense. Christian Wolf, chercheur au Liris (Laboratoire d'informatique en images et systèmes d'information, CNRS-INSA de Lyon), explique : « Nous avons lancé une compétition internationale, à laquelle sont inscrits 40 laboratoires de recherches. Objectif : reconnaître des comportements complexes impliquant plusieurs individus qui agissent en même temps sur une vidéo, ou une interaction humain-objet. Nous avons réalisé des vidéos types : une personne entre et sort d'une pièce, une autre passe un petit objet à son voisin, une troisième parle au téléphone puis laisse en partant un bagage à ses pieds, etc. Dans quelques mois, nous saurons quelles équipes identifient et nomment automatiquement les scènes-clés avec le plus de précision. »

Autre défaut de la vidéosurveillance classique : en recherche de suspect à posteriori, elle mobilise des heures de l'attention d'un opérateur. « Même en défilement rapide, nos hommes regardent parfois des heures d'images pour rien ! », souligne Fabrice Sierra, chef de la circonscription de sécurité publique d'Aubenas (Ardèche). Les images des 42 caméras de Vals-les-Bains et Aubenas peuvent désormais être traitées par un nouveau logiciel (Video Synopsis, société Briefcam) qui leur mène le travail : il élimine les plages où il ne se passe rien et par ailleurs il condense et étiquette un maximum d'événements sur un temps très court (jusqu'à douze heures d'images condensées en quelques minutes).

Sur une même image sont ainsi superposées des actions qui ne se sont pas déroulées en même temps, dans lesquelles des piétons marchent donc les uns sur les autres ou des voitures roulent en même temps sur le même parcours. « Si vous souhaitez retrouver un break bleu par exemple,

vous allez regarder la synthèse des images de voitures qui se sont garées sur tel rond-point, commente Fabrice Sierra. Vous sélectionnez votre break et replongez dans la vidéo originale... où vous retrouvez l'heure à laquelle il a stationné, éventuellement vous voyez le suspect en sortir, etc. »

Laurent Mucchelli, sociologue au CNRS, spécialiste de la délinquance, doute totalement de l'efficacité de tous ces systèmes. « Ce qui m'inquiète le plus, ce ne sont pas ces techniques mais la façon dont les élus tombent dans le panneau. Il y a une croyance magique dans ces outils. Les centres de surveillance vont se heurter à une augmentation des alertes (dans les fausses alertes), et devoir augmenter encore le nombre de personnes derrière les écrans ! Cette offensive marketing va coûter très cher. »

Les professionnels de la sécurité, eux, continuent de miser sur la technique. Après la détection du comportement suspect, ils comptent bien identifier le suspect lui-même. Problème : les techniques biométriques actuelles requièrent sa coopération. Il doit poser sa main si on veut reconnaître ses empreintes digitales, ou regarder fixement une caméra pour l'identification de son iris.

L'avenir appartient donc aux biométries que l'on peut collecter à la volée, comme la forme du visage. « Les systèmes basés sur les images 2D ont montré leurs faiblesses face aux variabilités causées par le changement de la pose, des conditions d'éclairages ou encore des expressions faciales », explique Mohsen Ardabilian, maître de conférences à l'Ecole centrale de Lyon-Liris. Les caméras 3D, en revanche, permettent de reconnaître un visage même quand il n'est pas de face et s'il est illuminé différemment. Une de nos techniques consiste à repérer sur le visage des points d'intérêt ayant une information de courbure ou de texture discriminante ainsi que leur position. Ce qui nous donne des taux de reconnaissance proches de 98 %. Mais ces données s'entendent sur une base test de quelques centaines de visages, non représentatifs de la population.

Bernard Didier, directeur général adjoint de Morpho (groupe Safran), leader

mondial dans les systèmes d'identification, confirme : « En mode participatif (la personne est volontaire pour être contrôlée), nos systèmes de reconnaissance, utilisés dans les aéroports d'Australie et en Nouvelle-Zélande, sont efficaces à 98 %. En non participatif, et en environnement ouvert, nous n'atteignons pas cette performance. » Principales difficultés à résoudre : le positionnement du visage et l'éclairage. « Dans les tests effectués par le NIST [National Institute of Technology], on constate qu'une différence d'angle d'une dizaine de degrés par rapport à une photo témoin suffit à diminuer la performance de reconnaissance. Pour la reconnaissance à la volée, nous devons donc utiliser plusieurs caméras pilotées de manière coordonnée, qui reposent de façon dynamique le visage de la personne suivie et la reconnaissent sans qu'il soit besoin d'interrompre la marche », précise M. Didier.

Tout cela n'est donc pas au point, mais de nombreuses sociétés, petites (Facing It) ou plus importantes (Cognitec), mettent déjà sur le marché des produits. L'aéroport de Francfort teste EasyPass, système de contrôle aux frontières faisant appel à un

logiciel de reconnaissance des visages à la volée. Et aux Etats-Unis, la société Sarnoff commercialise Iris on the Move, capable de « traire » les iris de 30 personnes par minute, si elles marchent tranquillement devant une borne d'identification.

« Le défi réside dans l'exécution de ces tâches dans le contexte de la vidéosurveillance, plus difficile que lors des contrôles biométriques, et là, les systèmes "classiques" vont avoir de la peine », estime Rudy Guyonneau, docteur de neurosciences à Spikenet Technology. Issue du Laboratoire cerveau et cognition (CNRS-université Toulouse-III), cette start-up mise sur les réseaux de neurones. « Notre originalité et nous verrons si elle paye, est de nous inspirer des dernières avancées en neurosciences computationnelles, sur le traitement de l'information à base de décharges (spikes) de neurones, pour proposer des systèmes logiciels rapides, simples, et capables de s'adapter, même dans ces contextes-là. »

Autre programme FAST (« technologie de surveillance des attributs futurs »), les Américains souhaitent repérer à leur insu les individus ayant l'intention de commettre un délit, en développant des capteurs mesurant, à distance, le rythme cardiaque des personnes passant aux points de contrôle, les contractions des pupilles, la température du visage, les expressions faciales... Un scénario à la Minority Report, ce film de Steven Spielberg où trois mutants doués d'un don de préscience préviennent le gouvernement qu'un crime va être commis, ce qui mène à l'arrestation des meurtriers potentiels... alors qu'ils n'ont encore rien fait. Sommes-nous si loin d'un tel scénario ?

La prochaine caméra Kinect 2, de la console de jeu Xbox 360 de Microsoft, pourrait détecter les expressions du visage, mesurer notre degré de satisfaction ou d'agacement. Une fonction transposable pour des applications de vidéosurveillance ? De nombreux laboratoires y travaillent. « Mais la technologie est neutre, c'est l'usage qu'en fait qui doit être arbitré par la société », estime Bernard Didier.

Pour l'instant, la réglementation interdit totalement l'identification à la volée, et encore plus la reconnaissance des émotions. Jusqu'à ce qu'un bon argument marketing convainque nos élus qu'un individu à l'air stressé dans un parking constitue une sérieuse menace pour la sécurité... ■



DIGITAL BARRIERS FRANCE

## Assistance à la personne et marketing, des marchés d'avenir

Comme le résume Nabil Belbachir, chercheur à l'Institut autrichien des technologies à Vienne et éditeur du livre *Smart Cameras* (Springer, 2009, en anglais seulement), les systèmes intelligents pour la vidéosurveillance intelligente sont déjà très utilisés dans de nombreux secteurs : en contrôle industriel, pour détecter sur les lignes de production les défauts de fabrication ; dans les aéroports, pour compter les personnes aux guichets et réguler les files d'attente ; en indexation et archivage de vidéos.

Dans l'automobile, les constructeurs travaillent sur des outils d'aide à la conduite, par exemple des caméras qui détectent les mouvements du piéton du conducteur et sonnent l'alarme en cas d'endormissement.

Nabil Belbachir mène pour sa part des recherches dans le but d'aider au maintien à domicile des personnes âgées. L'idée est de les assister à dis-

tance, en utilisant des caméras spécialement pour donner l'alarme par exemple en cas de chute. Moins simple qu'il n'y paraît : « Il faut commencer par qualifier les chutes, depuis la chute debout, depuis la chaise, depuis le lit, etc., explique Nabil Belbachir. Puis apprendre à faire la différence entre une chute involontaire et une chute volontaire (la personne s'allonge pour faire de la gymnastique ou regarde sous son lit). »

### Analyser le regard des clients

Avec le programme Pulsar (Perception Understanding Learning System Activity Recognition), l'institut de recherche en informatique et en automatique (Inria) se focalise sur deux applications : la vidéosurveillance et le maintien des personnes âgées à domicile.

Là aussi, l'objectif est de suivre les mouvements de la personne âgée, de reconnaître ses postures et ses

activités d'intérêt par le biais de capteurs vidéo.

Mais un des marchés les plus prometteurs est sans doute celui du marketing. Chercheur au Laboratoire d'informatique fondamentale de Lille (LIFL-CNRS-université de Lille), impliqué à l'origine dans le projet CANADA, Chabane Djeraba a créé son entreprise, Anaxa-vida, et applique aux consommateurs la technologie d'analyse d'objets en mouvement. Son système mesure les réactions et la vitesse de passage d'un client dans un magasin, détaille les taux de fréquentation par rayon, la capacité d'un rayon à transformer un visiteur en client, etc.

Aux Etats-Unis, la société Prism, enthousiasmée par « les formidables informations détenues par les 40 millions de caméras de surveillance installées sur le territoire », propose de les récupérer pour les mettre sur Internet en ayant auparavant flouté

les images. Et permettre aux intermédiaires de connaître la longueur d'une file d'attente, de produire des statistiques sur le nombre de personnes qui passent devant un magasin, etc.

Etape suivante en marketing : analyser le regard des clients qui achètent sur Internet. « Grâce à la webcam, et à condition que le client accepte d'être filmé, on peut suivre précisément les mouvements de son iris », explique Chabane Djeraba, savoir quel produit il regarde et si ce regard est suivi d'un clic d'achat. Nous observons aussi s'il sourit ou fait la moue selon qu'on lui propose une tablette de chocolat à 4 ou 6 euros. De telles inventions soulèvent de nombreuses questions légales et éthiques, mais nous répondons à ces inquiétudes en « anonymisant les personnes ». Nous voilà tout à fait rassurés... ■

M.JU.

## « Un contrôle politique délégué à la machine »

**B**runo Villalba est maître de conférences à l'Institut d'études politiques de Lille, couteur, avec Jean-Jacques Laveune, de *Videosurveillance et détection automatique des comportements anormaux. Enjeux techniques et politiques* (Presses universitaires du Septentrion, 2011).

Qui est censé élaborer les normes ?

En démocratie, les normes sont perçues comme légitimes parce qu'elles ont été élaborées par le législateur et qu'elles répondent à des impératifs moraux (liberté, égalité, etc.). Dans la vidéosurveillance automatisée, la normalité renvoie à une appréciation statistique et vise à caractériser le comportement régulier.

Or, le normal est un concept dynamique, qui s'inscrit toujours dans un contexte mouvant et doit faire face à l'infinie diversité des personnes et comportements humains... Dans le cas présent, la norme serait construite non pas par la puissance publique mais par la capacité technique.

**L'automatique de la détection vous inquiète-t-elle ?**

Oui, l'automatique évacue les débats puisque le contrôle politique du comportement est délégué à la machine, et permet de passer progressivement d'une procédure d'aide à la décision (dans l'évaluation du comportement) à une décision de fait. La machine rend possible la simultanéité de l'évaluation du comportement et l'alerte.

Cette dimension technique de l'automatique crée une difficulté pour l'individu à s'opposer à un tiers : à la machine ? À son concepteur ? À son utilisateur ? Comment appréhender la pluralité et l'invisibilité de ces outils d'évaluation du comportement ?

**Qu'en est-il de la convergence des données ?**

D'une part, il devient difficile d'échapper dans l'espace public à ces systèmes de vidéosurveillance, par ailleurs nous laissons quantité de traces sur notre profil par les ordinateurs ou les smartphones, et, enfin, de nombreux fichiers de police, médicaux, etc., recensent des informations sur notre identité. Le croisement de ces différentes données est potentiellement dangereux pour les libertés individuelles.

Notamment si cette convergence s'effectue de manière automatique, en répondant à des opérations toujours plus complexes de sélections d'informations, à partir de simples critères statistiques et non plus politiques. ■

**PROPOS RECUEILLIS PAR M. JU.**