



CHALLENGING LIGHT  
FOR OUR PARTNERS

WHITE PAPER

# ANALYSE VIDÉO

octobre 2011

# ANALYSE VIDÉO

En raison de son pouvoir discriminatif, l'analyse vidéo est une technique très prometteuse pour améliorer le niveau de performance des systèmes avancés de vidéo surveillance. Toutefois, la détection et le pistage vidéo d'objets en temps réel dans un environnement naturel extérieur constituent une tâche difficile qui pose, de façon générale, le problème du pistage de cibles multiples. Les systèmes de pistage de cibles multiples se prêtent à un grand nombre d'applications, notamment dans le domaine du contrôle de la circulation aérienne. La mise en œuvre d'un système fonctionnel de contrôle au sol de la circulation aérienne capable de fournir des pistes de bonne qualité à partir d'enregistrements vidéo exige de régler les difficultés posées par un arrière-plan dynamique et des changements d'éclairage. Ces difficultés incluent l'impact visuel du balancement des arbres dans le vent, de même que la position et l'intensité changeantes des ombres projetées selon la position du soleil et des nuages dans le ciel. Le pistage en continu d'une zone extérieure suppose de s'adapter au passage de conditions diurnes à des conditions nocturnes, et vice versa, et de réduire au maximum l'impact négatif du brouillard, de la pluie ou de la neige sur la qualité visuelle des images.

Par conséquent, le plus grand défi à relever en pistage vidéo de cibles multiples en zone extérieure consiste à traiter des données brutes bruitées (aussi appelées images) pour obtenir en temps réel des observations consistantes image par image. La fiabilité de cette première étape aura un effet important sur le rendement du système, notamment sur ses taux de détections et de fausses alarmes. Or, pour illustrer simplement notre propos, nous nous limiterons au cas d'une caméra unique et fixe fournissant des images à une cadence de prise de vue constante. L'accès à un réseau complet de caméras permet au système de profiter d'un plus grand nombre de perspectives sans que la capacité habituelle de détection des objets par vidéo en soit affectée.

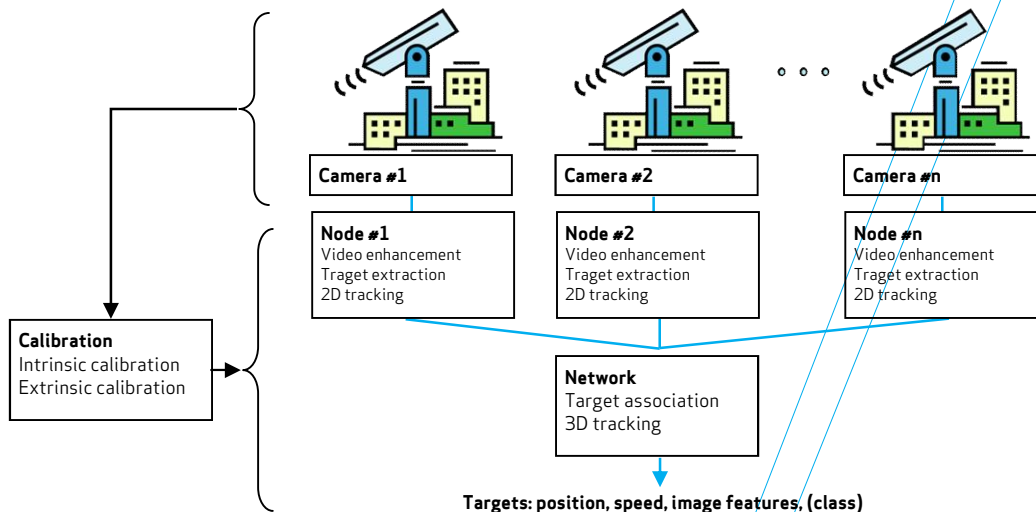


Figure 1: Architecture du système

tante. L'accès à un réseau complet de caméras permet au système de profiter d'un plus grand nombre de perspectives sans que la capacité habituelle de détection des objets par vidéo en soit affectée.

Une fois les images convenablement rehaussées, on peut s'attaquer au problème de la détection d'objets proprement dit. Il existe différentes techniques de détection d'objets et, la conception de procédés pour les améliorer et les rendre plus efficaces constitue d'ailleurs un champ de recherche toujours très actif. Dans le cas des caméras fixes, les techniques de détection d'objets les plus précises et efficaces procèdent par soustraction du modèle d'arrière-plan. Ce procédé exige donc une représentation de l'arrière-plan mise à jour automatiquement. Ainsi, en comparant, pixel par pixel, les images entrantes avec le modèle de l'arrière-plan mis à jour, on tient pour acquis que les objets détectés font partie de l'information de l'avant-plan.

La segmentation en temps réel de l'information d'avant et d'arrière-plan exige la mise en place d'un modèle très efficace capable de s'adapter à des conditions environnementales changeantes, et cela, à la cadence de la vidéo. Cette technique doit aussi tenir compte de la présence d'objets à l'avant-plan durant la phase d'initialisation du modèle d'arrière-plan. À cet égard, les techniques les plus rapides et les plus précises s'inscrivent dans la catégorie des modèles non paramétriques. Ceux-ci utilisent un processus guidé par les données pour produire et actualiser une représentation en pixels et plurimodale du modèle d'arrière-plan. Il faut également tenir compte de considérations pratiques sur le plan de la conception pour maximiser le rendement du procédé et, dans le cas des images en couleurs, choisir judicieusement un code de transformation de couleurs, lequel peut réduire de manière significative les délais de traitement.

Reste ensuite à déterminer si le pixel appartient à l'arrière ou à l'avant-plan en évaluant si la valeur du pixel en question relève ou non d'un des modèles d'arrière-plan établis. Dans les cas où sa valeur représente une innovation par rapport à l'information du modèle d'arrière-plan, on établit ce pixel comme appartenant à l'avant-plan. Par la suite, il s'agit de soumettre les éléments d'avant-plan d'une image à une série d'étapes de filtrage de manière à former des groupes de pixels apparentés, ou nuées. L'objectif de l'opération est d'éliminer les éléments d'avant-plan obtenus par des bruits parasites ou qui n'ont pas pu être analysés en tant qu'objets significatifs dans le contexte de l'application par suite de certaines contraintes sur le plan des attributs, par exemple, en raison d'une localisation ou d'une taille invalide.

À cette étape du processus, chaque nuée correspond à une observation. Il s'agit ensuite d'associer les nuées à des séries de caractéristiques géométriques (ex. la position dans l'image), d'attributs morphologiques (ex. l'étendue en pixels, l'intensité moyenne ou la couleur) ou, plus généralement, d'attributs statistiques ou de descripteurs. On peut aussi effectuer une classification des observations à cette étape. Le choix des attributs d'observation joue, ici, un rôle important dans l'efficacité générale du système et l'on doit faire ce choix en visant un compromis entre la complexité algorithmique et la richesse descriptive de l'observation.

COMM-110035

À cet égard, la position dans le monde réel est un attribut d'observation important. Dans le contexte d'une zone de surveillance relativement planaire, comme une aire de manœuvre ou une piste d'aéroport, il est possible de déterminer la position d'un objet au sol en calculant l'intersection du champ de vision projeté de la caméra avec le plan correspondant à la hauteur de cet objet à partir du sol. On appelle ce calcul une estimation monoscopique de la position. Par ce procédé, les observations d'objets au sol sont associées à une estimation de position dans le monde réel au moment même de la détection.

Cependant, l'exploitation des images visant à transformer les estimations de position dans le monde réel en des résultats hautement précis exige de procéder d'abord à des étalonnages intrinsèques et extrinsèques de la caméra. L'étalonnage intrinsèque a pour but d'établir les paramètres de fonctionnement de la caméra. Il est effectué hors ligne, avant l'installation de la caméra. L'étalonnage terminé, les algorithmes de traitement des images peuvent prendre en compte des départs dans le monde réel à partir de cette caméra idéale, en repérant par exemple les désalignements optiques qui envoient des signaux d'imagerie et les aberrations optiques qui causent des déformations d'images. Enfin, la capacité de transformer les coordonnées pixels d'un objet en une position dans le monde réel suppose que l'emplacement et la position de la caméra par rapport au système de référence du monde réel ont été bien établis. Cet étalonnage extrinsèque ne peut s'effectuer qu'après avoir déterminé une série de points de repère dans le monde réel apparaissant dans le champ de vision de la caméra.

À partir de là, on peut alimenter le système de pistage par des observations vidéo de bonne qualité et procéder aux opérations de pistage proprement dites. Il reste un problème, celui de l'attribution des observations à des pistes d'enregistrement vidéo. Or, la description d'un système de pistage de cibles multiples général et pratique dépasse de loin la portée de cette discussion. Le pistage de cibles multiples a suscité beaucoup de recherches et fait l'objet d'une activité de développement intense, notamment au sujet des applications de poursuite radar. À la différence de la poursuite radar, toutefois, le pistage vidéo est plus riche parce qu'il prend en compte un ensemble de caractéristiques morphologiques décrivant l'étendue des objets à repérer.

Il est possible d'améliorer le pistage vidéo d'objets en exploitant le potentiel de fusion des caractéristiques géométriques et morphologiques des observations. Cet ensemble enrichi de caractéristiques permet à son tour d'interpréter les situations d'occlusion partielle ou complète qui se présentent dans le flux d'images lorsque des pistes se rencontrent et forment momentanément un « groupe ». Une fois les pistes séparées, ce sont les caractéristiques morphologiques des observations qui permettent de résoudre les ambiguïtés concernant l'identité de ces pistes. À la lumière de ces considérations, nous croyons que si nous réussissons à résoudre les difficultés que pose un système de pistage en zone extérieure, nous aurons accès aux capacités d'imagerie inhérentes à un tel système et pourrons mettre au point des systèmes efficaces permettant la détection et le pistage vidéo de cibles multiples.

## CONTACT INO

[www.ino.ca](http://www.ino.ca) / [info@ino.ca](mailto:info@ino.ca)

### INO (HEADQUARTERS)

2740, Einstein Street  
Quebec City, Quebec  
G1P 4S4 Canada  
418 657.7006 / 1 866 657.7406

### ONTARIO

175 Longwood Rd. S  
Suite 316 A  
Hamilton, ON  
L8P 0A1 Canada  
905 529.7016

### WESTERN CANADA

403 875.1521