

Savoir au delà de voir: vision artificielle et raisonnement logique

Roberto Marroquin, Julien Dubois, Christophe Nicolle

Univ. Bourgogne Franche-Comté, Laboratoire Le2i-FRE 2005, Dijon, France
{roberto-enrique.marroquin-cortez, julien.dubois, cnicolle}@u-bourgogne.fr

Il existe de nombreux travaux sur l'analyse numérique d'une image ou vidéo. Ces travaux concernent soit l'amélioration du signal pour améliorer la qualité de l'image, soit la reconnaissance d'éléments contenue dans l'image.

Au delà des travaux d'extraction d'information à partir du signal numérique construit par les caméras, nous nous sommes intéressés, à la déduction de nouvelles connaissances par agrégation d'informations provenant de sources hétérogènes multiples. Notre ambition est d'extraire des connaissances en couplant les informations issues des algorithmes habituels de traitement d'images avec des connaissances contextuels et des savoir-faire liés à l'usage d'un bâtiment. Cette interopérabilité est réalisée par l'intermédiaire d'une agrégation d'ontologies.

L'ontologie utilise les axiomes définis dans le langage OWL-2. Elle fournit un vocabulaire pour intégrer, re-organiser et analyser sémantiquement des sources de données hétérogènes. Notre ontologie est constituée d'un ensemble de termes dérivés de plusieurs ontologies :

- L'ontologie `DUL`, fournit un ensemble de concepts utilisés pour permettre l'interopérabilité entre différentes ontologies.
- L'ontologie `event`, traite de la notion d'événement et les propriétés associées telles que la localisation le temps, les agents, les facteurs et les produits.
- L'ontologie `ifcowl`, est une représentation sémantique du schéma IFC (standard pour la représentation des données du bâtiment).
- L'ontologie `person` fournit la classe pour décrire une personne physique.
- L'ontologie `ssn` permet la description des capteurs, des observations, des traitements de détection, les capacités de mesures et tout autre concept relatif.
- L'ontologie du temps `time` fournit les concepts pour décrire les propriétés temporelles des ressources.

Toutes les sources de données hétérogènes sont intégrées dans l'ontologie du système `WiseNET` (Marroquin et al., 2016) en utilisant des techniques issues du `linked data` : les URIs (Uniform Resource Identifiers) et RDF (Resource Description Framework).

Le traitement des connaissances issues des images est organisé en deux étapes. La première étape concerne la partie extraction et la seconde concerne la partie gestion (voir figure 1).

L'extraction des connaissances consiste à extraire une partie des données d'un ensemble, ensuite filtrer les données pour ne conserver que les données pertinentes et finalement, à enrichir les données avec de la connaissance. Dans notre cas, les caméras intelligentes ajoutent de la connaissance aux parties d'images sélectionnées à l'aide d'algorithmes de traitement d'images. Après avoir réalisé l'extraction des connaissances à partir de ce que voit le réseau

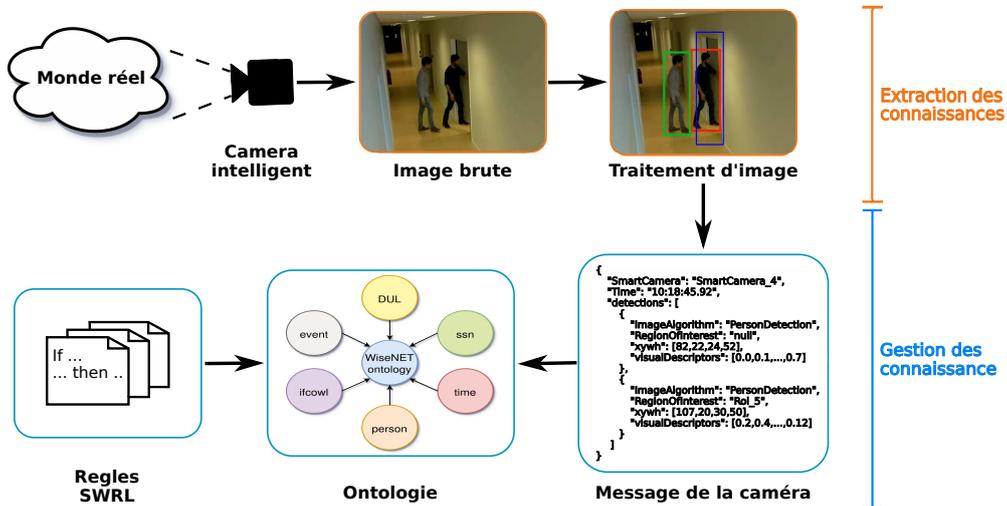


FIG. 1 – Traitement d'extraction et de gestion des connaissances. Ce traitement commence par l'extraction de données issues du monde réel, représentées sous la forme d'une vidéo. Ensuite, ces données sont converties en connaissances et traduites sous la forme d'un vocabulaire contrôlé défini par l'ontologie WiseNET. Enfin, une étape de peuplement permet d'insérer ces connaissances dans l'ontologie pour réaliser le raisonnement.

de caméras intelligentes, plusieurs traitements sont réalisés pour combiner cette connaissance avec d'autres, pour réaliser ensuite le raisonnement.

La gestion de la connaissance consiste à convertir les données extraites en un ensemble de termes structurés selon le vocabulaire de l'ontologie. Ce résultat, est envoyé au serveur pour peupler l'ontologie. A partir de l'ontologie peuplée, nous avons réalisé des inférences ou raisonnements. Ces raisonnements permettent de proposer des services aux utilisateurs ou aux gestionnaires du bâtiment, telles que compter le nombre de personnes qui sont dans le bâtiment, monitorer automatiquement le déplacement de ces personnes (heatmap), savoir quelle est l'espace le plus utilisé dans la journée, ou quelle porte est la plus utilisée (dans le but de planifier les opérations de maintenance liées à cet usage). Tout ces service sont fournis en respectant la vie privée de gens (pas de transmission vidéo).

La connaissance réunie peut aussi permettre de résoudre des limitations intrinsèques à la vision par ordinateur, telles que des détections manquées, erronés ou des occultations. Ceci est réalisé par la construction de règles métiers sous la forme de règles SWRL. L'ensemble complet des notre approche est accessible à l'adresse <http://wisenet.checksem.fr>.

Références

Marroquin, R., J. Dubois, et C. Nicolle (2016). WiseNET-smart camera network interacting with a semantic model : PhD forum. In *Proceedings of the 10th International Conference on Distributed Smart Camera*, pp. 224–225. ACM.