Un système de vision pour machine apprend de lui-même à reconnaître les objets

19 octobre 2018, 02:00

Peter Paulissen
Jan Kempeneers

Des robots qui seraient capables à l'avenir de s'adapter par eux-mêmes à leur environnement trouveraient de multiples applications dans les entrepôts et ateliers de production.

A l'heure actuelle, les systèmes de vision permettent à une machine de faire un mouvement déterminé, par exemple, saisir un objet à la sortie d'une ligne de production pour le déposer dans un bac. Pour faire davantage qu'une opération aussi élémentaire, les robots devraient être équipés d'un système de vision discriminant. C'est l'objectif du système 'DON' (Dense Object Nets), en cours de développement au CSAIL (Computer Science and Artificial Intelligence Laboratory - MIT).

Tâche spécifique / Tâche générique

Les deux méthodes les plus utilisées actuellement pour permettre aux robots de saisir des objets sont basées sur l'apprentissage d'une tâche spécifique ou sur la création d'un algorithme générique de préhension. Ces techniques présentent toutefois des limites : les méthodes spécifiques à la tâche sont difficilement applicables telles quelles pour d'autres tâches, quant à l'opération générique de préhension, elle n'est pas suffisamment 'large' pour inclure les nuances des tâches spécifiques, par exemple, déposer les objets à des endroits bien précis.

Aujourd'hui, le maximum que l'on puisse attendre d'un système de vision numérique est la faculté de détection d'un objet. 'Pixel labeling' est le niveau suivant. Ces deux techniques sont mises en oeuvre dans les voitures autonomes. Toutefois, dès lors qu'une interaction est requise entre la machine et son environnement, par exemple, saisir un objet déterminé sur une courroie de transport, la détection des pixels n'est pas suffisante.

Feuilles de routes visuelles

Le système DON est une nouvelle méthode de vision pour machine, qui génère des feuilles de route visuelles, ce qui revient à collecter des points de données visuels ordonnancés sous forme de coordonnées. Les sets individuels de données sont assemblés pour former des sets de coordonnées qui seront présentées dans un format 3D, selon un procédé comparable à celui utilisé par un appareil photonumérique pour créer une vue panoramique avec plusieurs photos. Le système de vision est ainsi capable de mieux appréhender la forme d'un objet quelconque et la manière dont il se déplace dans le contexte de son environnement.

Le système DON perçoit l’objet avec tous les détails de son aspect extérieur, c’est le type d’informations requises pour rendre possible l’exécution de tâches précises. Le système permettra à un robot, par exemple, d’examiner une tasse de café sous tous les angles, de s’orienter comme nécessaire pour manipuler cet objet, et finalement de déterminer par lui-même que le fond de la tasse doit toujours être dirigé vers le bas, pour éviter l’épanchement du contenu. En d’autres termes, il s’agira d’un système qui, de toute façon, permettra à un robot de saisir un objet spécifique dans une pile d’objets partageant une même apparence.

Toutefois, l’identification des parties spécifiques d’un objet n’indique pas nécessairement quels sont les angles d’approche en vue de la préhension de l’objet. Les algorithmes existants ne permettent pas d’indiquer au robot que la tasse doit être saisie par l’anse.

Auto-apprentissage

Le système DON est basé sur un capteur RVB en combinaison avec une caméra 3D, et il est capable d’apprendre par lui-même, sans aucune intervention humaine, même s’il s’agit de nouvelles tâches, ou de tâches concernant des objets que le système n’a jamais vus. Autrement dit, il n’est pas nécessaire d’archiver au préalable des centaines ou des milliers d’images. Par exemple, si le système doit être capable d’identifier des chaussures de couleur marron, il suffit de placer le robot de préhension dans une pièce où se trouve une chaussure de cette couleur. Le système prendra des photos de référence, sous tous les angles, pour générer des points de coordonnées qui serviront de base à son auto-apprentissage, une procédure qui nécessite moins d’une heure.

Dans un atelier de production, la fiabilité de fonctionnement des robots est tributaire des systèmes d’amenée et positionnement correct des composants à assembler. Le système DON sera en mesure de ‘comprendre’ dans quel sens les objets doivent être orientés, ce qui permettra de les manipuler d’une manière adéquate.



(Source : MIT)

Essais menés avec succès

Dans une série d’essais de préhension (sur un ours en peluche), un bras robotique équipé du système DON s’est avéré capable de saisir le jouet par l’oreille droite, à partir de configurations différentes. Le système a prouvé ainsi sa capacité, en présence d’objets symétriques, de faire la distinction entre le côté gauche et le côté droit.

Dans une série d’essais de préhension dans une boîte contenant plusieurs sortes de casquettes de baseball, le système s’est révélé capable de sélectionner un modèle déterminé, malgré la forte ressemblance entre les diverses casquettes et malgré que les photos n’avaient pas été montrées pendant la séance d’auto-apprentissage réalisée la veille.

La technologie est encore dans la phase des premiers balbutiements, mais les chercheurs espèrent qu'ils seront un jour en mesure d'équiper les robots d'un système amélioré de vision et coordination permettant la réalisation de tâches spécifiques nécessitant une compréhension approfondie du contexte d'utilisation, par exemple, dans les ateliers de production, mais également dans les logements d'habitation et bureaux.
Regardez le clip vidéo :

Sources

- http://news.mit.edu
- https://www.engadget.com
- https://arxiv.org

Auteurs



Peter Paulissen



Jan Kempeneers