

Roboty odgadujące myśli ludzi?

9 grudnia 2023

Naukowcy z Laboratorium Informatyki i Sztucznej Inteligencji (CSAIL) MIT stworzyli system Feature Fields for Robotic Manipulation (F3RM), który pozwala robotom rozumieć i manipulować pobliskimi obiektami za pomocą dowolnych wskazówek językowych. Łącząc obrazy 2D z funkcjami modelu podstawowego w celu tworzenia scen 3D, F3RM umożliwia robotom identyfikowanie i chwytanie nieznanych obiektów, co czyni go szczególnie przydatnym w rzeczywistych środowiskach, takich jak magazyny i gospodarstwa domowe.

Jedną z kluczowych cech F3RM jest możliwość interpretowania odpowiedzi tekstowych w języku naturalnym, co pozwala robotom manipulować obiektami w oparciu o mniej konkretne żądania człowieka. Oznacza to, że nawet jeśli użytkownik poprosi robota, aby „podniósł wysoki kubek”, robot będzie w stanie znaleźć i podnieść przedmiot, który najlepiej pasuje do tego opisu. Ten poziom zdolności adaptacyjnych i generalizacji zadań jest niezbędny, aby roboty mogły skutecznie działać w rzeczywistych środowiskach.

Naukowcy z National Science Foundation-MIT CSAIL Institute for Artificial Intelligence and Fundamental Interactions, wyjaśniają, że budowanie robotów, które potrafią generalizować w prawdziwym świecie, jest niezwykle trudne.

Zastosowanie F3RM wykracza poza proste zrozumienie obiektów i manipulowanie nimi. Może być szczególnie przydatne w dużych centrach logistycznych, gdzie roboty muszą wybierać produkty z zagraconego i nieprzewidywalnego środowiska. W takich magazynach roboty często otrzymują opis przedmiotu, który muszą zidentyfikować i muszą dopasować opis tekstowy dożądanego towaru, niezależnie od różnic w opakowaniu.

Na przykład centra realizacji dużych sklepów internetowych mogą zawierać miliony produktów, z którymi roboty nigdy wcześniej się nie spotkały. Aby skutecznie działać na taką skalę, roboty muszą dogłębnie zrozumieć geometrię i semantykę różnych obiektów, nawet jeśli znajdują się one na ograniczonej przestrzeni. Zaawansowane możliwości percepcji przestrzennej i semantycznej F3RM mogą znacznie zwiększyć zdolność robota do wykrywania obiektu, umieszczania go w pojemniku i wysyłania do pakowania. Ostatecznie pomoże to pracownikom fabryki efektywniej wysyłać zamówienia klientów.

Rozwój F3RM jest znaczącym krokiem naprzód w dziedzinie robotyki. Łącząc percepcję wzrokową z przetwarzaniem języka naturalnego, roboty mogą teraz rozumieć obiekty i manipulować nimi w taki sam sposób, jak robią to ludzie. Ma to daleko idące konsekwencje dla różnych branż, w tym handlu elektronicznego, logistyki i produkcji.

Ta praca otwiera nowe możliwości współpracy i interakcji człowiek-robot. Umożliwiając robotom rozumienie zapytań w otwartym języku i odpowiadanie na nie, zbliżamy się do stworzenia naprawdę inteligentnych i elastycznych maszyn.

Źródło: [ZmianyNaZiemi.pl](https://zmiany.naziemi.pl)