

ЛОКАЛІЗАЦІЯ ТА КЛАСИФІКАЦІЯ ОБ'ЄКТІВ НА ЗОБРАЖЕННІ**LOCATION AND CLASSIFICATION OF IMAGE OBJECTS**

Процес розпізнавання предмету у неструктурованих сценах – складна сфера поточних досліджень комп'ютерного зору. Робототехніка є одним із важливих напрямків, де здатність чітко і швидко ідентифікувати об'єкти, що цікавлять, має вирішальне значення для універсальних роботів, що виконують завдання в неструктурованому повсякденному середовищі, такій як будинок, офіс або склад.

Традиційна система розпізнавання об'єктів завжди опрацьовує значне число різних об'єктів. Один із способів упоратися з цим – запровадити ієрархію, виконуючи розпізнавання на рівні категорії, а не на рівні примірника. Завдання узагальнення від кількох екземплярів до цілої категорії об'єктів залишається складним, і існують численні тести та завдання, такі як Caltech 256 та PASCAL VOC, щоб сприяти прогресу у цій галузі. З іншого боку, для конкретного робота в конкретному середовищі кількість унікальних об'єктів відносно невелика (можливо, близько сотень). Це дозволяє розглядати завдання як проблему розпізнавання екземплярів, збираючи великий обсяг навчальних даних кожного об'єкта. Крім того, робот може використовувати дані, одержані від різних сенсорних пристроїв, таких як камери та датчики глибини.

Локалізація об'єктів у робототехніці також є додатковими проблемами, яких немає в тестах розпізнавання об'єктів на рівні категорій. Середовищ реального світу дуже захарашені, містять багато перекриттів і часто містять більшу кількість різних об'єктів в одній сцені. Роботам часто доводиться взаємодіяти з одними об'єктами у своєму середовищі, оминаючи інші. Це означає, що роботизована система сприйняття повинна точно локалізувати об'єкти після виявлення. Крім того, щоб робот міг швидко реагувати на зміни у навколишньому середовищі, роботизована система сприйняття має працювати у режимі реального часу. В даний час існує ряд різних підходів до розпізнавання та локалізації предметів у складних неструктурованих сценах. Першим є використання сенсорів для збирання кольорових та глибинних зображень. Прикладами даних сенсорів є різні лідари, або більш прості давачі типу Kinect. Під час навчання розпізнаванню та локалізації об'єктів на основі інформації про глибину вдається точніше виділити контури кожного об'єкта. На основі контурів, кольору та глибини будується повна хмара точок, з якої виділяються тривимірні сіткові моделі об'єктів. Потім з отриманих моделей витягуються і реєструються локальні особливості об'єктів.

Розпізнавання об'єктів за локальними ознаками на зображеннях відбувається у кілька етапів. Спершу знаходяться відповідності. Цей етап відноситься до точного зіставлення характеристик зображення з характеристиками, що належать певному предмету. Далі йде процес оцінки пози об'єкта, яка буде геометрично узгодженою зі знайденими відповідностями. Після цього відповідно до масштабу предмета у базі даних і зображення знаходиться точна відстань до об'єкта. Важливо, що вибір камери та об'єктива може сильно вплинути на точність оцінки пози. Тим не менш, камери стають якіснішими та дешевшими, а їх габарити зменшуються. Це призводить до того, що стають легко доступними паралельне використання декількох камер для отримання виду сцени з різних ракурсів та високою роздільною здатністю. Цей підхід ефективний, тому що він забезпечує оцінку глибини, стійкість до оклюзій прямої видимості та зростання ефективного кола огляду.