

Czujniki i wykrywanie

- * Opisz różnicę między czujnikami aktywnymi i pasywnymi i podaj przykłady każdego z nich.
- * Wymień przynajmniej jedną zaletę i wadę popularnych czujników robotycznych: kamery GPS, INS, IR, RGB-D.
- * Zdefiniuj obraz, piksel i funkcję obrazu.
- * Jeśli masz mały, przeplatany obraz RGB i zakres wartości kolorów dla regionu, możesz napisać kod, aby wyodrębnić afordancje kolorów, używając 1) progu dla koloru i 2) histogramu kolorów.
- * Napisz kod wizji komputerowej, aby umożliwić robotowi nadruk i śledzenie koloru.
- * Zdefiniuj każdy z poniższych terminów w jednym lub dwóch zdaniach: czujnik zbliżeniowy, czujnik logiczny, fałszywie dodatni, fałszywie ujemny, odcień, nasycenie i widzenie komputerowe.
- * Opisz trzy rodzaje fuzji czujników behawioralnych: rozszczepienie, połączenie, moda.
- * Wymień atrybuty do projektowania zestawu czujników i zastosuj te atrybuty do określonej aplikacji.
- * Zdefiniuj obciążenie lokomocyjne i hotelowe oraz wyjaśnij, dlaczego wystarczające obciążenie hotelu ma kluczowe znaczenie dla projektowania inteligentnych robotów.

Wstęp

Przeglądając Część 7, percepcja w warstwie reaktywnej pełni dwie role: uwalnianie zachowania i wspieranie lub kierowanie działaniami zachowania. Wszystkie detekcje są specyficzne dla zachowania, przy czym zachowania mogą wykorzystywać te same czujniki, ale wykorzystują dane niezależnie od siebie. Konotacja robotów reaktywnych polega również na tym, że zachowania są najczęściej wynikiem reakcji na bodziec, polegającej na bezpośredniej percepcji, a nie wymagającej pamięci. W warstwie Deliberatywnej wykrywanie służy do rozpoznawania i rozumowania obiektów, scen i zdarzeń w celu zbudowania modelu świata. Podczas gdy warstwa reaktywna wykorzystuje, jak nazwaliby Gibson i Neisser, bezpośrednie ścieżki percepcji w mózgu, warstwa deliberatywna wykorzystuje ścieżki rozpoznawania obiektów. W warstwie interaktywnej wykrywanie służy do wykrywania i wspierania interakcji społecznych. Te interakcje społeczne mogą być dość ordynarne, na przykład unikanie inwazji na przestrzeń osobistą. Mogą również mieć wysoką rozdzielczość, w której robot może potrzebować dostrzegać i interpretować subtelne rysy twarzy i gesty. Wszechobecne wykorzystanie percepcji prowadzi do subtelnego rozróżnienia między czujnikami a wyczuwaniem w sztucznej inteligencji. Czujniki dostarczają surowe dane, podczas gdy wykrywanie jest kombinacją algorytmu(ów) i czujnika(ów), które wytwarzają percepcję lub model świata. Potrzeba percepcji prowadzi do kilku ogólnych pytań dotyczących odczuwania. Pierwsze pytanie jest najbardziej podstawowe: Jak sprawić, by robot „widział”? Poprzednia część dotycząca zachowań powinna już zasiać ziarno idei, że percepcja to nie tylko czujnik; istnieje dodatkowe przetwarzanie. Sztuczna inteligencja dotyczy tego, jak wyodrębnić lub wnioskować o percepcji lub obiekcie lub jak zinterpretować to, co dzieje się w scenie – i jak to zrobić za pomocą dowolnej kombinacji określonych czujników. Kolejne pytanie brzmi: jakie czujniki są niezbędne dla robota? Konkretna odpowiedź zależy oczywiście od misji, ale ogólnie roboty potrzebują zarówno propriocepcji, jak i eksterocepcji. Ponieważ jednak roboty powinny zawsze mieć „tylne drzwi” do teleoperacji, jak omówiono w rozdziale 5, robot zazwyczaj zawsze będzie miał kamerę światła widzialnego. Wykorzystanie tej kamery (kamer) jest polem widzenia komputerowego. Ostatnie pytanie brzmi: co to jest fuzja czujników? Poprzedni rozdział dotyczący robotyki opartej na zachowaniu wskazywał, że indywidualne zachowanie może przyjmować sygnały z więcej niż jednego czujnika. Opis rozważań w części 4 definiuje pojęcie modelu świata, który jest sumą wszystkich zmysłów. Mechanizmy

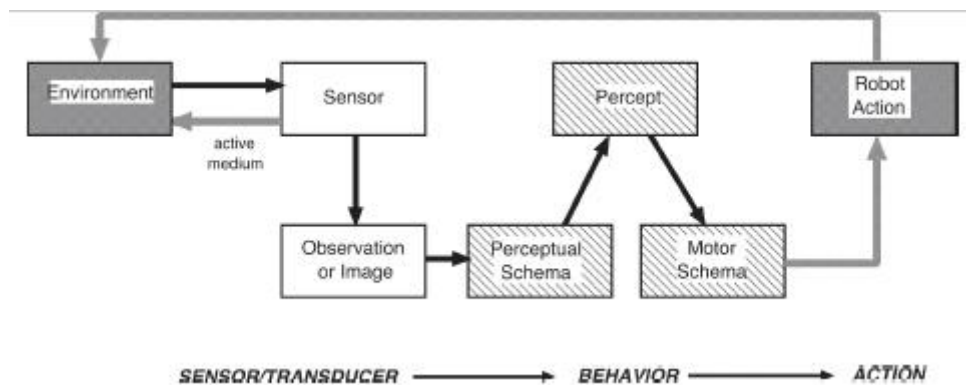
czujnikowe, które umożliwiają wielu czujnikom wytwarzanie percepcji i modeli, są ogólnie określane jako fuzja czujników. Ta część nie odpowiada na pytanie: Jak wyczuć głębię i zakres. Wykrywanie głębokości i zasięgu bez bezpośredniego kontaktu ma kluczowe znaczenie dla nawigacji i jest polem samym w sobie. W związku z tym wykrywanie zasięgu jest omówione w następnej części. W tej części omówione zostaną podstawy czujników i wykrywania z perspektywy sztucznej inteligencji. Wprowadzi trzy powszechne reprezentacje wiedzy, wykresy biegunowe, funkcje obrazu i czujniki logiczne, a także ogólny obszar sztucznej inteligencji w wizji komputerowej. Najpierw przedstawi model czujników i czujników, który zostanie wykorzystany do omówienia czujników aktywnych i pasywnych.

Model czujnika i wykrywania

CZUJNIK

TRANSDUKTOR

Niezależnie od sprzętu lub aplikacji czujnika, można pomyśleć, że czujniki i czujniki wchodzą w interakcję ze światem i robotami, jak pokazano na rysunku



Czujnik to urządzenie, które mierzy jakiś atrybut świata. Termin przetwornik jest często używany zamiennie z czujnikiem. Przetwornik to mechanizm lub element czujnika, który przekształca energię związaną z tym, co jest mierzone, w inną formę energii.⁶ Czujnik odbiera energię i przesyła sygnał do wyświetlacza lub komputera. Czujniki wykorzystują przetworniki do zmiany sygnału wejściowego (dźwięku, światła, ciśnienia, temperatury itp.) na postać analogową lub cyfrową, którą może wykorzystać robot. W przypadku zachowań reaktywnych obserwacja czujnika jest przechwytywana przez schemat percepcyjny, który wyodrębnia odpowiednią percepcję otoczenia dla zachowania. Ta percepcja jest następnie wykorzystywana przez schemat motoryczny, który prowadzi do działania. Rysunek 10.1 ilustruje wykrywanie w warstwie reaktywnej. W warstwie Deliberatywnej konfiguracja modułów schemat percepcyjny – percepcja – schemat motoryczny zostałaby zastąpiona modelem świata. Warstwa interaktywna może wykorzystywać zachowania lub wykorzystywać model świata do swoich potrzeb w zakresie wykrywania. Następnie w rozdziale omówiono odometrię, system nawigacji bezwładnościowej (INS), globalny system pozycjonowania (GPS) oraz czujniki zbliżeniowe. Wizja komputerowa i przetwarzanie obrazu są opisane w dalszej części. Pole widzenia komputerowego jest tak szerokie, że niemożliwe jest zagłębienie się w wszystkie możliwe algorytmy, dlatego część skupia się na prostych, powszechnych procedurach używanych z zachowaniami reaktywnymi w celu wzmocnienia pojęcia reaktywności. Pokazuje również, jak postrzeganie, które przyjęliśmy za pewnik, jest w rzeczywistości dość trudne. Rozdział kończy się listą atrybutów, które należy wykorzystać przy wyborze zestawu czujników i obejmuje inne koncepcje, które poprawią projektowanie percepcyjne.

Czujniki: aktywne lub pasywne

CZUJNIK PASYWNY

CZUJNIK AKTYWNY

AKTYWNE WYKRYWANIE

Czujnik jest często klasyfikowany jako pasywny lub aktywny. Czujniki pasywne opierają się na środowisku, aby zapewnić medium do obserwacji, na przykład kamera wymaga pewnej ilości światła otoczenia, aby wytworzyć użyteczny obraz. Aktywne czujniki emitują energię do otoczenia, aby zmienić energię lub ją wzmocnić. Sonar wysyła fale dźwiękowe, odbiera echo i mierzy czas podróży. Aparat rentgenowski emituje promieniowanie rentgenowskie i mierzy ilość blokową przez różne rodzaje tkanek. Chociaż aparat jest urządzeniem pasywnym, aparat z lampą błyskową jest aktywnym czujnikiem. Termin „czujnik aktywny” to nie to samo, co aktywne wykrywanie. Aktywne wykrywanie oznacza system wykorzystujący efektor do dynamicznego pozycjonowania czujnika w celu usprawnienia gromadzenia danych. Na przykład aparat z lampą błyskową jest aktywnym czujnikiem, podczas gdy aparat zamontowany na głowicy uchylno-obrotowej, z algorytmami obracania go w celu uzyskania lepszego widoku, korzysta z aktywnego wykrywania.

MODALNOŚĆ

Różne czujniki mierzą różne formy energii. To z kolei prowadzi do różnych rodzajów przetwarzania. Czujniki, które mierzą tę samą formę energii i przetwarzają ją w podobny sposób, tworzą modalność czujnika. Modalność czujnika odnosi się do surowych danych wejściowych czujnika: dźwięku, ciśnienia, temperatury, światła i tak dalej. Pod pewnymi względami modalności są podobne do pięciu zmysłów u ludzi. Modalność można dalej podzielić, na przykład widzenie można rozłożyć na światło widzialne, światło podczerwone i promieniowanie rentgenowskie.

Czujniki: rodzaje wyjścia i użytkowania

Czujniki można również podzielić na kategorie według typu wytwarzanego sygnału wyjściowego i sposobu wykorzystania wykrytych danych. Dane wyjściowe czujnika to odczyt lub obraz. Obrazy zostaną formalnie zdefiniowane w dalszej części tego rozdziału, ponieważ stanowią podstawową i popularną reprezentację. Jak opisano w rozdziale 5, wyczuwanie świata jest zwykle używane na trzy sposoby. Jednym z nich jest propriocepcja lub lokalizacja kończyn i stawów robota lub określenie, jak bardzo się poruszały. Propriocepcja jest niezbędna do sterowania platformą robota. Innym jest eksterocepcja, wykrywająca obiekty w świecie zewnętrznym i często odległość do tych obiektów. Eksterocepcja jest niezbędna, aby umożliwić robotowi poruszanie się i działanie w świecie. Trzecim jest ekspropriocepcja, wykrywająca położenie robota względem obiektów na świecie. Jest to również ważne dla manipulacji i wnioskowania o tym, dlaczego robot się nie porusza (tj. utknął). Odometria, system nawigacji inercyjnej (INS) i globalny system pozycjonowania (GPS)

ENKODERY WAŁOWE

Roboty mogą szacować swoje ruchy za pomocą czujników proprioceptywnych, takich jak kodery wału, które mierzą liczbę obrotów wykonanych przez silnik napędowy; enkoder wału działa w taki sam sposób, jak licznik kilometrów w samochodzie. Jeśli przekładnia i rozmiar kół są znane, liczba obrotów silnika może być wykorzystana do obliczenia liczby obrotów kół robota, a liczba ta może być wykorzystana do oszacowania, jak daleko robot przebył. Niestety są to tylko szacunki, ponieważ środowisko wpływa na rzeczywisty ruch robota. Robot na podłodze z płytek może ślizgać się dwa razy bardziej niż robot na suchej trawie. INS Systemy nawigacji bezwładnościowej (INS) i systemy ruchu bezwładnościowego (IMU), takie jak akcelerometry stosowane w smartfonach, również mogą dostarczać informacji o martwym naliczeniu i często działają znacznie lepiej niż odometria. Dopóki ruchy są płynne, bez nagłych wstrząsów, a próbki są pobierane często, INS może zapewnić dokładną

ocenę martwą z dokładnością do 0,1 procent przebytej odległości.⁷⁴ Niestety, twardy wstrząs lub nagły skręt może przekroczyć zakres pomiarowy akcelerometrów, wprowadzając błędy. Sojourner, łazik marsjański, miał system INS. Podczas jednej wędrówki zatrzymałby się 30 cm od skały, z której miał pobrać próbki, gdyby zastosował właśnie propriocepcję. Zamiast tego, stosując eksterocepcję, zbliżył się do 4 cm.

GLOBALNY SYSTEM POZYCJONOWANIA

KANIONY MIEJSKIE

GPS lub globalny system pozycjonowania jest powszechny w robotach pracujących na zewnątrz. Systemy GPS działają na zasadzie odbierania sygnałów z satelitów krążących wokół Ziemi. Odbiornik oblicza swoją pozycję względem co najmniej czterech satelitów GPS pod względem szerokości i długości geograficznej, wysokości i zmiany czasu. GPS nie jest czujnikiem proprioceptywnym per se, ponieważ robot musi odbierać sygnały z satelitów zewnętrznych w stosunku do robota. Nie jest to jednak również czujnik eksteroceptywny, ponieważ robot nie oblicza swojej pozycji względem otoczenia. Niezależnie od rodzaju czujnika, GPS nie jest kompletnym rozwiązaniem problemu martwego rachunku w robotach mobilnych, ponieważ GPS nie działa w pomieszczeniach w większości budynków, zwłaszcza w biurach lub fabrykach z dużą ilością żelbetonu. Struktury te zakłócają odbiór sygnałów radiowych, podobnie jak w przypadku sieci telefonii komórkowej. Podobnie GPS może nie działać na zewnątrz w dużych miastach, gdzie drapacze chmur i mosty działają jak kaniony miejskie i zakłócają odbiór.

Czujniki zbliżenia

Czujniki zbliżeniowe bezpośrednio mierzą względną odległość (zakres) między czujnikiem a obiektami w otoczeniu. Czujniki zbliżeniowe ogólnie kojarzą się z pomiarem na krótkich odległościach do obiektu, rzędu jednego metra, w porównaniu z czujnikiem zasięgu, który wykrywa na znacznie większe odległości. Czujniki zbliżeniowe mogą być aktywne lub pasywne. Aktywne czujniki zbliżeniowe umożliwiają robotowi wykrywanie obiektów przed kontaktem, natomiast pasywne czujniki zbliżeniowe wymagają kontaktu z obiektem lub powierzchnią. Czujniki podczerwieni są najpopularniejszymi aktywnymi czujnikami zbliżeniowymi, a czujniki uderzeniowe i dotykowe są najpopularniejszymi czujnikami pasywnymi.

IR

Czujniki podczerwieni (IR) emitują energię bliskiej podczerwieni i mierzą, czy jakakolwiek znacząca ilość światła podczerwonego jest zwracana. Jeśli obecna jest przeszkoda, zwraca sygnał binarny. Czujniki podczerwieni mają zasięg od cali do kilku stóp, w zależności od używanej częstotliwości światła i czułości odbiornika. Najprostsze czujniki zbliżeniowe IR mogą być zbudowane z diod LED, które emitują światło do otoczenia i mają zasięg 3-5 cali. Te emitery często zawodzą w praktyce, ponieważ emitowane światło jest często „wyprane” przez jasne oświetlenie otoczenia lub jest pochłaniane przez ciemne materiały (tj. w otoczeniu jest za dużo hałasu). W bardziej zaawansowanych czujnikach podczerwieni można wybierać lub modulować różne pasma podczerwieni, aby zmienić stosunek sygnału do szumu, co zwykle zapewnia, że obiekt w zasięgu nie pochłania światła i powoduje, że czujnik nie zauważa obecności obiektu.

DOTYKOWE

Inną popularną klasą robotycznego wykrywania jest dotyk, wykonywany za pomocą czujników uderzeniowych i czułych. Czujki lub wąsy mogą być zbudowane z wytrzymałych drutów. Czujniki uderzeń zwykle składają się z dwóch warstw tworzących wystający pierścień wokół robota. Kontakt z

przedmiotem powoduje, że dwie warstwy stykają się, tworząc sygnał elektryczny. Bardzo ważną kwestią jest umiejscowienie czujników uderzeniowych. Czujniki uderzeń w bazie Nomad 200 firmy Nomadic Technologies, Inc. miały chronić bardzo dokładny mechanizm napędu synchronicznego przed niskimi przeszkodami. Niestety, w niektórych konfiguracjach skrętu koła wystają poza osłonę, przez co czujnik wstrząsów jest całkowicie bezużyteczny w zapobieganiu uszkodzeniom. Teoretycznie czułość czujnika uderzeń lub wąsów można dostosować do różnych nacisków; niektóre roboty mogą chcieć „lekkiego” dotyku, aby wygenerować sygnał, a nie „cięższego” dotyku. W praktyce regulacja czujników uderzeń jest irytująco trudna. W „Hors d’Oeuvres, Anyone?” podczas zawodów robotów mobilnych AAI w 1997 r. ludzie otrzymywali przekąski przez robotów „kelnerów”. Ludzie mieli przekazać kelnerowi robotowi Nomad 200 w Colorado School of Mines, że skończyli jeść, kopiąc czujnik uderzeń zamontowany na spodzie robota. Czułość czujnika uderzeń była tak niska, że często wymagała wielu kopnięć, tworząc bardzo komiczną scenę z podtekstem Bruce’a Lee.

Widzenie komputerowe

Widzenie komputerowe jest głównym źródłem wykrywania eksteroceptywnego ogólnego przeznaczenia do bezpośredniej percepcji i rozpoznawania obiektów. Widzenie komputerowe może być również wykorzystywane do pomiaru odległości, metoda jest opisana w następnym rozdziale. Ta sekcja przedstawia wizję komputerową i kilka podstawowych algorytmów stosowanych bezpośrednio jako postrzeganie. Jak zauważono w Części 1, widzenie komputerowe jest odrębną dziedziną nauki od robotyki i stworzyło wiele przydatnych algorytmów do filtrowania szumów, kompensowania problemów z oświetleniem, poprawiania obrazów, znajdowania linii, dopasowywania linii do modeli, wyodrębniania kształtów i budowania reprezentacji 3D .

Definicja wizji komputerowej

OBRAZ

PIKSEL

FUNKCJA OBRAZU

Widzenie komputerowe odnosi się do przetwarzania danych z dowolnej modalności, która wykorzystuje widmo elektromagnetyczne do wytworzenia obrazu. Obraz to sposób przedstawiania danych w formie podobnej do obrazu, w którym istnieje bezpośrednia fizyczna zgodność z obrazowaną sceną. W przeciwieństwie do sonaru, który zwraca odczyt z jednego zakresu, który może odpowiadać obiektowi znajdującemu się w dowolnym miejscu stożka o kącie 30°, obraz oznacza wiele odczytów umieszczonych w dwuwymiarowej tablicy lub siatce. Każdy element tablicy mapuje się na mały obszar przestrzeni. Elementy w tablicach obrazów nazywane są pikselami, skróceniem słów „element obrazu”. Rodzaj urządzenia określa, co mierzy obraz. Jeśli używana jest kamera światła widzialnego, wartość przechowywana w każdym pikselu jest wartością światła (np. kolor). Jeśli używana jest kamera termowizyjna, przechowywaną wartością jest ciepło w tym regionie. Funkcja konwertująca sygnał na wartość piksela nazywana jest funkcją obrazu. Widzenie komputerowe obejmuje wyjście z dowolnego typu kamery, która wytwarza obrazy w tym samym spektrum elektromagnetycznym, które widzą ludzie, do bardziej egzotycznych technologii: czujników termicznych, promieni rentgenowskich, dalmierzy laserowych i syntetycznego radaru temperatury. Bardziej egzotyczne czujniki mogą być trudne dla człowieka, który próbuje zwiualizować wartość danych na obrazie. Należy pamiętać, że to wyjście reprezentacji obrazu umieszcza czujnik w kategorii widzenia komputerowego.

Odwzorowanie skali szarości i kolorów

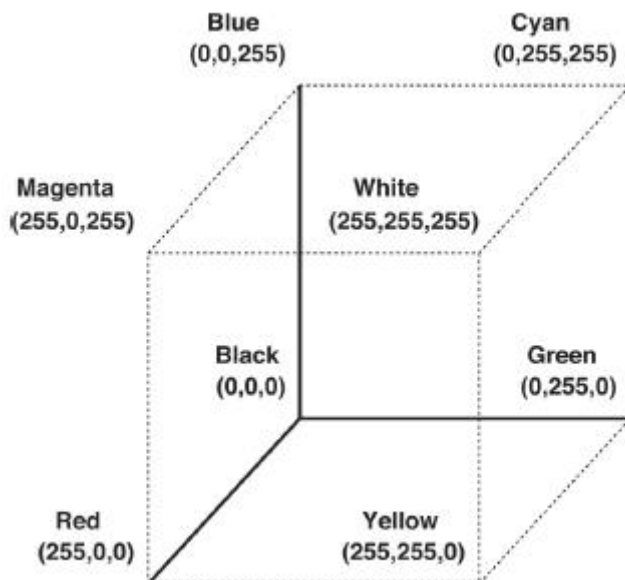
Wartość piksela jest zwykle przedstawiana w skali szarości lub jako kolor. Reprezentacje w skali szarości są standardowe, przy użyciu liczby 8-bitowej (1 bajt pamięci komputera). Prowadzi to do 256 dyskretnych wartości szarości, przy czym 0 reprezentuje czern, a 255 reprezentuje biel. (Pamiętaj, 256 wartości oznacza 0...255.)

RGB

Jednak kolor jest przedstawiany inaczej, a każda reprezentacja ma zalety i wady. Po pierwsze, istnieje wiele różnych metod wyrażania koloru. Drukarki do komputerów domowych używają metody subtraktywnej, w której cjan plus żółty dają kolor zielony. Większość urządzeń komercyjnych w Stanach Zjednoczonych korzysta ze standardu National Television System Committee (NTSC). Kolor NTSC jest wyrażony jako suma trzech pomiarów: czerwonego, zielonego i niebieskiego, w skrócie RGB.

PŁASZCZYZNY KOLORÓW

RGB jest zwykle przedstawiany jako trzy kolorowe płaszczyzny lub osie sześcianu 3D, jak pokazano na rysunku

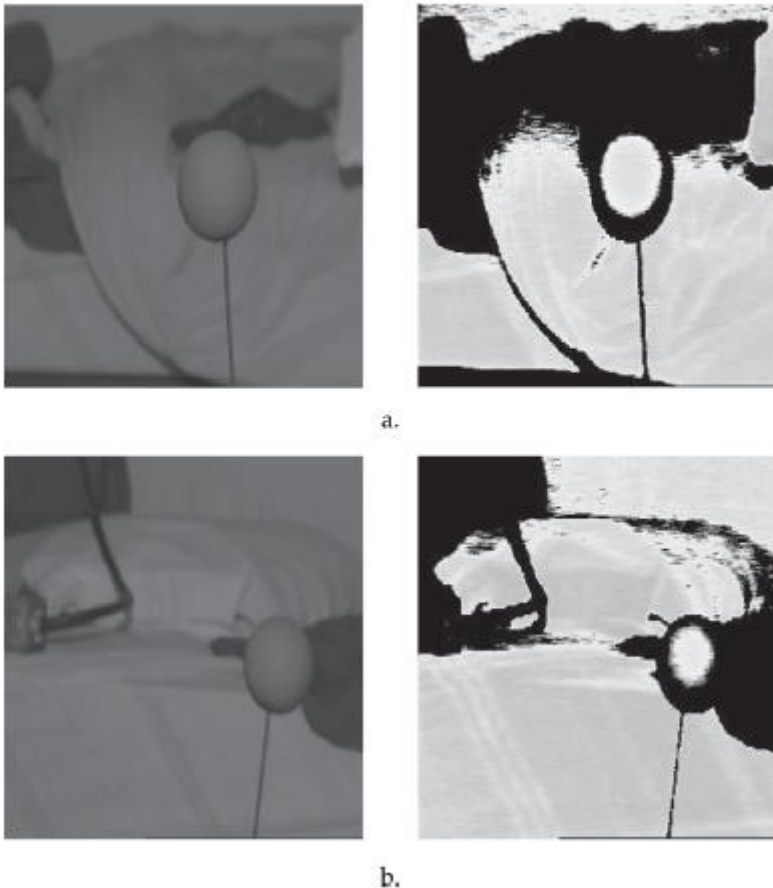


Kostka reprezentuje wszystkie możliwe kolory. Określony kolor jest reprezentowany przez krotkę trzech wartości do zsumowania: (R, G, B). Czarny to (0, 0, 0) lub 0+0+0 lub brak pomiarów na żadnej z trzech płaszczyzn kolorów. Biały to (255, 255, 255). Czyste kolory Czerwony, Zielony i Niebieski są reprezentowane odpowiednio przez (255, 0, 0), (0, 255, 0) i (0, 0, 255). Reprezentacje są takie same jak w przypadku grafiki kolorowej. Zauważ, że wymiary sześcianu na rysunku to 256 na 256 na 256, gdzie 256 to zakres liczb całkowitych, które można wyrazić za pomocą 8 bitów. Ponieważ istnieją trzy wymiary kolorów, producent może określać tę kostkę jako kolor 24-bitowy (3 x 8), aby odróżnić konkretnego framegrabbera od tych, które mapują kolor na liniową skalę szarości. 8-bitowy model kolorów jest używany do kolorowania starych czarno-białych filmów. Jest tylko 256 wartości koloru, co jest dość ograniczone, a wartości szarości są często niejednoznaczne. Wartości pikseli czerwonych ust kobiety mogą wynosić 185, podczas gdy jej ciemnoniebieska sukienka również wynosi 185. Osoba może być zmuszona ręcznie rozróżnić regiony w każdej klatce filmu, gdzie 185=czerwony i regiony, w których 185=ciemnoniebieski. Kolor 8-bitowy nie jest często używany do widzenia robota, chyba że

robot będzie działał w środowisku, w którym jedyne widoczne kolory nie będą niejednoznaczne. W przypadku robotyki zwykle wystarcza kolor 24-bitowy. W przypadku innych zastosowań wizji komputerowej, takich jak obrazowanie medyczne, prognozowanie pogody lub rekonstruacja wojskowa, rozdzielczość 8-bitowa jest często niewystarczająca. Te aplikacje mogą używać 10 bitów na każdą płaszczyznę koloru. Ponieważ 10 bitów nie mieści się na granicy bajtów, programowanie algorytmów reprezentujących i manipulujących tego rodzaju obrazami może być niewygodne. Do tych zastosowań często wykonywane są specjalne komputery.

EROZJA WIZUALNA

Reprezentacja RGB ma wady dla robotyki. Kolor w RGB jest funkcją długości fali źródła światła, sposobu, w jaki modyfikuje go powierzchnia obiektu (odbicie powierzchni) oraz czułości czujnika. Pierwszy problem polega na tym, że kolor nie jest absolutny. RGB opiera się na wrażliwości na odbite światło trójkolorowych elementów czujnikowych. Obiekt może wydawać się mieć różne wartości w różnych odległościach ze względu na intensywność odbitego światła. Rysunek



przedstawia czułość na światło odbite. Zdjęcia przedstawiają dwa zdjęcia zrobione pomarańczowego punktu orientacyjnego, który służy jako „flaga” do śledzenia małego robota. Zdjęcia zostały zrobione, gdy kamera się poruszała, zmieniając w ten sposób widok. Mimo że obrazy zostały przetworzone przez ten sam program i te same parametry w celu segmentacji tuż poza obrazem, wyniki są różne. Segmentacja RGB na rysunku a jest bardziej poprawna niż na rysunku b. Jedyne różnica polega na tym, że oflagowany robot przemieścił się, zmieniając w ten sposób kąt padania światła. To pogorszenie jakości segmentacji nazywa się erozją wizualną, ponieważ obiekt wydaje się erodować wraz ze

zmianami oświetlenia. Co więcej, aparaty cyfrowe są notorycznie niewrażliwe na czerwień. Oznacza to, że jedna z trzech płaszczyzn kolorystycznych nie jest tak pomocna w rozróżnianiu kolorów.

ODCIEŃ

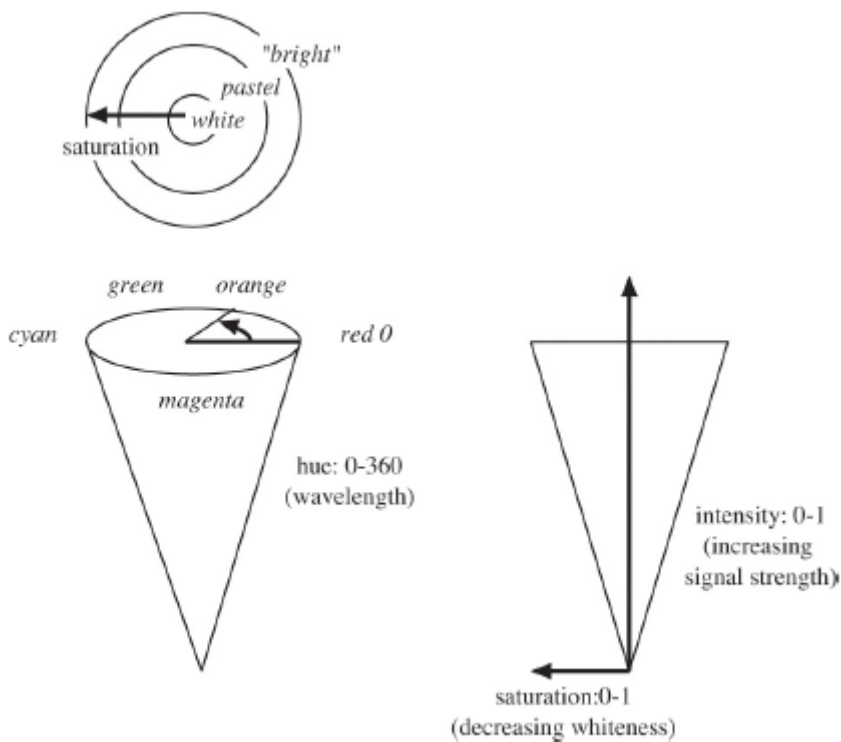
HSI

NASYCENIE

Oczywiście kamera, która jest wrażliwa na bezwzględną długość fali odbitego światła (zwaną barwą), byłaby bardziej korzystna niż konieczność obejścia ograniczeń RGB. Odcień jest dominującą długością fali i nie zmienia się wraz ze względną pozycją robota ani kształtem obiektu. Taka kamera działa na reprezentacji koloru HSIHSIHSI (odcień, nasycenie, intensywność). Nasycenie to brak bieli w kolorze; czerwony jest nasycony, różowy jest mniej nasycony. Miarą wartości lub natężenia jest ilość światła odbieranego przez czujnik, a reprezentacje HSI są często nazywane HSV. HSV to zupełnie inny schemat kolorów niż RGB.

HSV

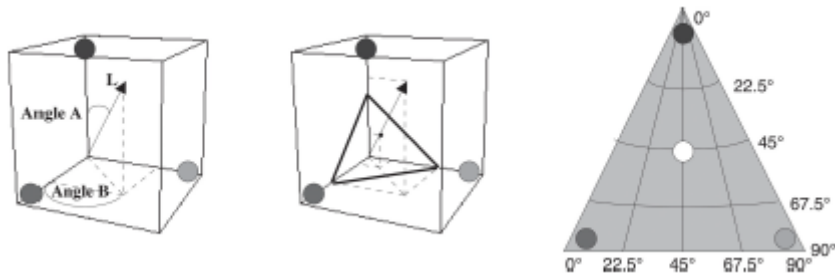
HSV jest trójwymiarową przestrzenią, ponieważ ma trzy zmienne, ale zdecydowanie nie jest reprezentacją sześcianu; jest to bardziej stożek, jak widać.



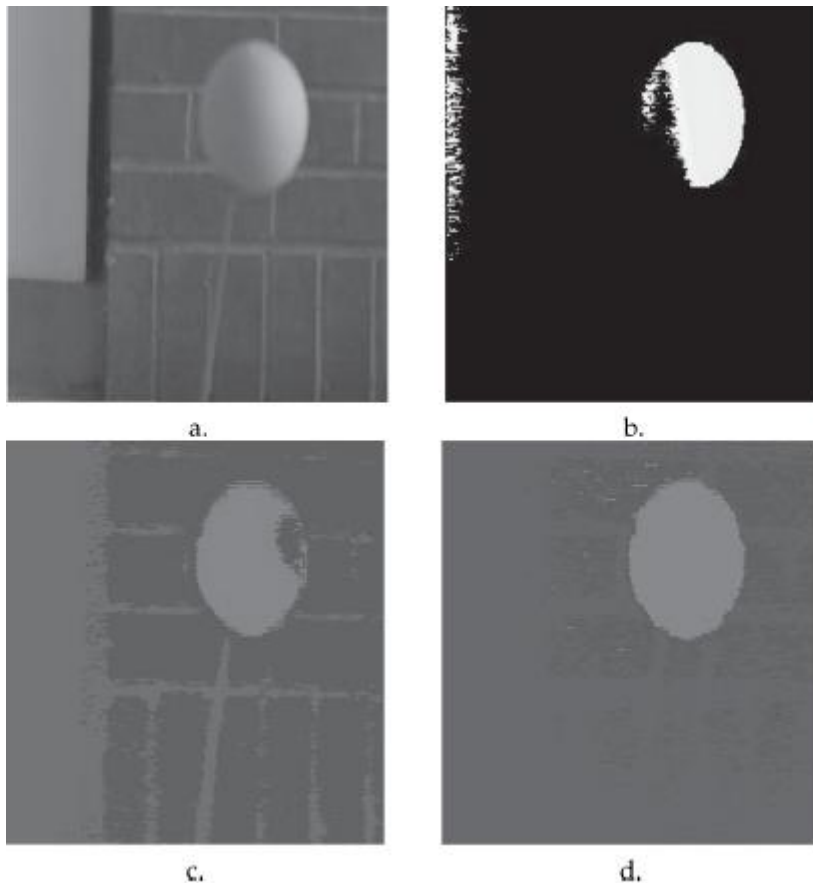
l.

Odcień lub kolor jest mierzony w stopniach od 0 do 360. Nasycenie i intensywność to liczby rzeczywiste z zakresu od 0 do 1. Są one zazwyczaj skalowane do liczb 8-bitowych. W związku z tym czerwony to zarówno 0, jak i 255, pomarańczowy to 17, zielony to 85, niebieski to 170, a magenta to 200. Przestrzeń HSV stanowi wyzwanie dla robotyków z wielu powodów. Po pierwsze, wymaga specjalnych kamer i framegrabberów do bezpośredniego pomiaru koloru w przestrzeni HSV. Ten sprzęt jest zaporowo drogi. Po drugie, istnieje konwersja programowa z przestrzeni RGB, ale nie jest ona używana, ponieważ konwersja jest kosztowna obliczeniowo i zawiera osobliwości, w których algorytm zawodzi. Te osobliwości występują w miejscach, w których trzy kolory piksela są takie same; Płaskość płaszczyzny

koloru czerwonego w kamerach CCD zwiększa prawdopodobieństwo wystąpienia osobliwości. Alternatywną przestrzenią kolorów, która jest obecnie badana w robotyce, jest transformacja współrzędnych sferycznych (SCT). Ta przestrzeń kolorów została zaprojektowana w celu przekształcenia danych RGB w przestrzeń kolorów, która dokładniej odwzorowuje reakcję ludzkiego oka. Jest stosowany w obrazowaniu biomedycznym, ale nie był powszechnie brany pod uwagę w robotyce. Przestrzeń kolorów ma kształt trójkąta, jak pokazano na rysunku



Wstępne wyniki wskazują, że jest on znacznie bardziej niewrażliwy na zmiany oświetlenia niż RGB. Rysunek przedstawia obraz i wyniki segmentacji koloru w przestrzeni RGB, HSI i SCT



Segmentacja regionów

Najbardziej wszechobecnym zastosowaniem widzenia komputerowego w robotyce reaktywnej jest identyfikacja na obrazie regionu o określonym kolorze, proces zwany segmentacją regionu. Segmentacja regionów i afordancje kolorów są podstawowymi algorytmami percepcyjnymi dla udanych zgłoszeń do wielu różnych międzynarodowych zawodów robotów, w tym RoboCup. Podstawową koncepcją jest identyfikacja wszystkich pikseli na obrazie, które są częścią regionu, a

następnie nawigacja do środka regionu (centroid). Pierwszym krokiem jest wybranie wszystkich pikseli, które mają ten sam kolor (progowanie), a następnie zgrupowanie ich i odrzucenie wszystkich pikseli, które nie wydają się znajdować w tym samym obszarze, co większość pikseli (rozrastanie się regionu).

OBRAZ BINARNY

PRÓG

Część 19 opisuje robota, którego kolorem czerwonym oznaczono czerwoną puszkę Coca-Coli do recyklingu. Idealnie, robot podczas zachowania polegającego na szukaniu puszkę postrzegałby świat jako obraz binarny (mający tylko dwie wartości) składający się z czerwonego, nie-czerwonego. Ten podział świata można osiągnąć przez progowanie obrazu i tworzenie obrazu binarnego. Przykład kodu C/C++ pokazano poniżej:

```
for (i= 0; i < numberRows; i++)
  for (j= 0; j < numberColumns; j++) {
    if ((ImageIn[i][j][RED] == redValue)
        && (ImageIn[i][j][GREEN] == greenValue)
        && (ImageIn[i][j][BLUE] == blueValue)) {
      ImageOut[i][j] = 255;
    }
    else {
      ImageOut[i][j] = 0;
    }
  }
}
```

Należy zauważyć, że wynikowy obraz z progiem w kolorze jest dwuwymiarową tablicą, ponieważ nie ma potrzeby dołączania więcej niż jednej wartości do każdego piksela. Teoretycznie obraz binarny dopuszczałby tylko wartości 0 i 1. Jednak w wielu kompilatorach nie ma szczególnych korzyści z wykonywania reprezentacji na poziomie bitów i może to skomplikować ponowne użycie kodu. Ponadto większość oprogramowania wyświetlającego jest używana do wyświetlania co najmniej 256 wartości. Ludzkie oko nie wykrywa różnicy między 1 a 0. Dlatego częściej zastępuje się 1 liczbą 255 i używa pełnego bajtu na piksel. Próg sprawdza się lepiej w teorii niż w praktyce ze względu na brak stałości koloru. Kształt obiektu oznacza, że chociaż człowiek widzi obiekt jako jednolity kolor, komputer widzi go jako zestaw podobnych kolorów. Powszechnym rozwiązaniem jest określenie zakresu wysokich i niskich wartości na każdej płaszczyźnie koloru. Kod C/C++ staje się teraz:

```
for (i= 0; i < numberRows; i++)
  for (j= 0; j < numberColumns; j++) {
    if (((ImageIn[i][j][RED] >= redValueLow)
        && (ImageIn[i][j][RED] <= redValueHigh))
        && ((ImageIn[i][j][GREEN] >= greenValueLow)
        && (ImageIn[i][j][GREEN] <= greenValueHigh))
        && ((ImageIn[i][j][BLUE] >= blueValueLow)
        && (ImageIn[i][j][BLUE] <= blueValueHigh))) {
      ImageOut[i][j] = 255;
    }
    else {
      ImageOut[i][j] = 0;
    }
  }
}
```

PIERWSZOPLANOWY

TŁO

Zmiana punktów widzenia i oświetlenia oznacza, że zakres wartości pikseli, który definiuje obiekt w bieżącej pozycji robota, prawdopodobnie zmieni się, gdy robot przejdzie do nowej pozycji. Jednym z podejść jest poszerzenie zakresu kolorów dla obiektu, aby uwzględnić zestaw wszystkich możliwych wartości pikseli dla obiektu widzianego z różnych punktów widzenia. Jeśli kolor obiektu jest unikalny dla tego środowiska, to zwiększenie zakresu kolorów jest akceptowalne. W przeciwnym razie, jeśli istnieją obiekty, których kolor jest wystarczająco zbliżony do obiektu zainteresowania, mogą one zostać pomyłone z celem. W niektórych kręgach obiekt zainteresowania nazywany jest pierwszym planem, podczas gdy wszystko inne na obrazie nazywa się tłem. Aby odnieść sukces, progowanie obrazu wymaga znacznego kontrastu między tłem a pierwszym planem. Na szczęście istnieją teraz algorytmy „camshift”, które mogą dostosować zakres kolorów regionu za pomocą statystyk. Rysunek ilustruje wynik progu dla samego koloru.



a.



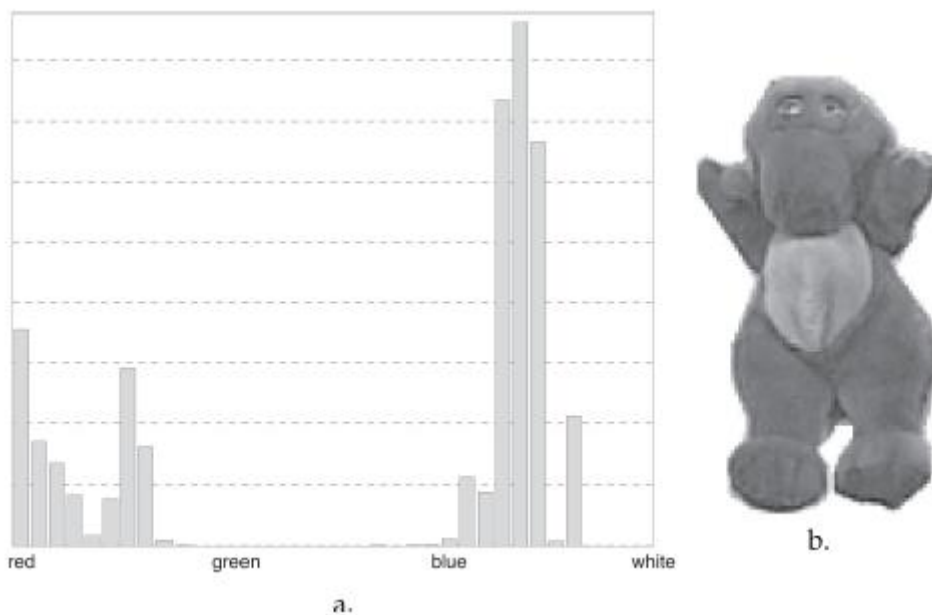
b.

Gdyby robot miał przejść do „czerwonego” na obrazie, skąd miałby wiedzieć, za którym pikselem ma podążać? Schemat percepcyjny można utworzyć dla każdego czerwonego piksela; jest to proste, ale zmarnuje wiele cykli wykonania. Schemat percepcyjny mógłby przyjąć ważoną centroidę wszystkich czerwonych pikseli. W tym przypadku środek znajdowałby się nieco blisko miejsca, w którym większość ludzi powiedziała, że była puszka, ale lokalizacja byłaby bliżej białego kubka. Albo schemat percepcyjny może próbować znaleźć największy obszar, w którym czerwone piksele sąsiadują ze sobą, a następnie obliczyć środek ciężkości tego regionu. (Region jest często określany jako „blob”, a proces ekstrakcji jest znany jako analiza blob). Kolorowe regiony mogą być również pomocne w zaśmieconym otoczeniu. Rysunek przedstawia robota mobilnego Denning symulującego przeszukiwanie zawalonego budynku. Międzynarodowa pomarańczowa kamizelka robotnika jest ważną wskazówką. Robot może sygnalizować teleoperatorowi, gdy widzi jasne kolory.



Histogram kolorów

Próg sprawdza się dobrze w przypadku obiektów składających się z jednego koloru lub jednego koloru dominującego. Inna technika, histogram kolorów, może być wykorzystana do identyfikacji regionu z kilkoma kolorami. Zasadniczo histogram kolorów jest sposobem dopasowania proporcji kolorów w regionie. Histogram to wykres słupkowy danych. Użytkownik określa zakres wartości dla każdego słupka, zwanych wiaderkami. Długość paska reprezentuje liczbę punktów danych z wartościami mieszczącymi się w zakresie dla tego segmentu. Na przykład histogram obrazu w skali szarości może mieć osiem segmentów (0-31, 32-63, 64-95, 96-127, 128-159, 160-191, 192-223, 224-251), a każdy segment zawiera liczbę pikseli obrazu, które mieszczą się w tym zakresie. Konstruowanie histogramu kolorów jest proste dla regionu w przestrzeni barw, jak pokazano na rysunku.



Histogram kolorów w RGB lub innej rozproszonej przestrzeni kolorów jest nieco trudniejszy do wizualizacji. Histogramy obrazów w skali szarości i odcienia miały tylko jedną oś dla wiader, ponieważ

te obrazy mają tylko jedną płaszczyznę, która ma znaczenie. Ale kolorowy obraz ma trzy płaszczyzny w układzie współrzędnych RGB. W rezultacie ma kubelki dla każdej płaszczyzny lub osi koloru. Zakładając, że każda płaszczyzna jest podzielona na osiem wiader, pierwszy wiader będzie liczbą pikseli mieszczącą się w zakresie (R, G, B) (0-31, 0-31, 0-31). Prawdziwą zaletą histogramu kolorów dla reaktywnych robotów jest to, że histogramy kolorów mogą być odejmowane od siebie w celu określenia, czy bieżący obraz (lub jego część), I, pasuje do wcześniej skonstruowanego histogramu, E. Histogramy są odejmowane wiadro po wiadrze (j zasobników łącznie), a różnice wskazują liczbę niezgodnych pikseli. Liczba niedopasowanych pikseli podzielona przez liczbę pikseli obrazu daje dopasowanie procentowe. Nazywa się to przecięciem histogramu:

$$intersection = \frac{\sum_{j=1}^n \min(I_j - E_j)}{\sum_{j=1}^n E_j}$$

Na przykład robot może „obudzić się” i odcisnąć znajdujący się przed nim obiekt, tworząc histogram kolorów. Następnie schemat percepcyjny wyzwalacza lub zachowania może obliczyć przecięcie histogramu kolorów bieżącego obrazu z odciskiem. Robot może użyć histogramu kolorów, aby określić, czy dany obiekt jest interesujący, czy nie. Ponieważ histogram kolorów bieżącego obrazu można dopasować do innego obrazu, technika ta wydaje się być oparta na modelu, czyli jest formą rozpoznawania. Ale systemy reaktywne nie pozwalają na rozpoznawanie rodzajów percepcji. Czy to sprzeczność? Nie, histogram kolorów jest przykładem lokalnej, specyficznej dla zachowania reprezentacji, którą można bezpośrednio wydobyć ze środowiska. Na przykład robotowi można by pokazać lalkę Barneya o wyraźnym fioletowym kolorze z zielonym brzuchem jako percepcją celu dla zachowania typu ruch do celu. Jednak robot będzie podążał za fioletowym trójkątem z zielonym obszarem, ponieważ stosunek kolorów jest taki sam. Nie ma pamięci ani wnioskowania, tylko bardziej złożony bodziec. Zauważ, że skrzyżowanie można uznać za miarę siły bodźca, co jest pomocne w robotyce reaktywnej. W jednym zestawie eksperymentów robotowi zaprezentowano plakat Sylwestra i Tweetyego. Uczył się histogramu, a następnie, po nauczaniu się obiektu (np. fiksacji na nim), zaczynał się do niego zbliżać, grając w berka, gdy osoba poruszała plakatem. Robot wykorzystał proste, atrakcyjne, oparte na potencjalnych polach zachowanie polegające na przejściu do celu, w którym schemat percepcyjny zapewniał lokalizację plakatu i procent przecięcia. Schemat silnika wykorzystywał lokalizację do obliczenia kierunku do plakatu, ale przecięcie wpłynęło na wielkość wektora wyjściowego. Gdyby osoba przesunęła plakat w ciemne miejsce lub obróciła go pod kątem, przecięcie byłoby niskie, a robot poruszałby się wolniej. Gdyby mecz był mocny, robot przyspieszyłby. Ogólnie rzecz biorąc, wytworzyło to bardzo podobne do psa zachowanie, w którym robot wydawał się grać w berka szybko (i szczęśliwie), dopóki człowiek nie uczynił tego zbyt trudnym. Następnie, jeśli człowiek przesunie plakat z powrotem do korzystniejszej pozycji, robot wznowi grę bez żadnych urazów.

Wybór czujników i wykrywanie

Projektant musi zobowiązać się do określonego zestawu czujników i algorytmów. Czyniąc to, warto znać trzy koncepcje. Pierwsza koncepcja to idea czujników logicznych lub równoważnych, w której możliwe jest wygenerowanie tego samego postrzegania z różnych czujników lub algorytmów. Drugi to fuzja czujników behawioralnych, która opisuje ogólne metody łączenia czujników w celu uzyskania pojedynczego spostrzeżenia lub wsparcia złożonego zachowania. Trzecią koncepcją są atrybuty zestawu czujników, które można wykorzystać do pomocy w projektowaniu systemu.

Czujniki logiczne

Potęzną abstrakcją czujników są czujniki logiczne, wprowadzone po raz pierwszy przez Hendersona i Shilcrata. Czujnik logiczny jest jednostką wyczuwania lub modułem, który dostarcza określoną percepcję. Składa się z przetwarzania sygnału z fizycznego czujnika i przetwarzania oprogramowania potrzebnego do wyodrębnienia percepcji; jest funkcjonalnym budulcem percepcji. Czujnik logiczny można łatwo zaimplementować jako schemat percepcyjny.

RÓWNOWAŻNOŚĆ LOGICZNA

Przeoczanym aspektem czujnika logicznego jest to, że zawiera on wszystkie dostępne alternatywne metody lub schematy uzyskiwania tego spostrzeżenia. Na przykład sposób postrzegania powszechnie używany do omijania przeszkód to wykres biegunowy danych o zasięgu. Czujnik logiczny dla erceptu może mieć nazwę range_360 i zwracać strukturę danych lub obiekt określający wykres biegunowy. Czujnik logiczny poszedłby dalej i wypisał wszystkie możliwe sposoby robota na skonstruowanie wykresu biegunowego w tej formie. Robot może być w stanie używać sonaru, lasera, wizji stereo lub tekstury. Każdy z tych modułów byłby logicznie równoważny; oznacza to, że zwracają tę samą strukturę danych percept, dzięki czemu można ich używać zamiennie. Jednak niekoniecznie musiałyby być. Dlatego czujnik logiczny zawiera funkcję selektora, która określa warunki, w których każda alternatywa jest użyteczna i dlatego powinna zostać wybrana. Potężnym praktycznym zastosowaniem czujników logicznych lub logicznej równoważności jest to, że model świata może służyć jako wirtualny czujnik. W niektórych hierarchicznych architekturach systemów do generowania globalnego modelu świata wykorzystywane jest całe wykrywanie. Architektura może mieć zachowania, ale percepcyjne dane wejściowe to schemat percepcyjny, który wydobywa percepcję z modelu świata, a nie z rzeczywistego czujnika.

REDUNDANCJA FIZYCZNA

Istnieją dwa rodzaje nadmiarowości. Fizyczna nadmiarowość oznacza, że w robocie występuje kilka przypadków fizycznie identycznych czujników.

REDUNDANCJA LOGICZNA

Nadmiarowość logiczna oznacza, że inny czujnik, wykorzystujący inną modalność wykrywania, może wytworzyć ten sam percept lub wyzwalacz. Na przykład robot mobilny Mars Sojourner był wyposażony w dwa czujniki zasięgu, parę stereowizyjną i system pasków laserowych do określania zasięgu do przeszkód. Czujniki nie są fizycznie nadmiarowe, ale dostarczają tych samych ogólnych informacji: lokalizacji przeszkód względem robota. Jednak logicznie nadmiarowe czujniki niekoniecznie są równoważne pod względem szybkości przetwarzania, dokładności i rozdzielczości. Czujnik i algorytm zakresu stereo obliczyły ten zakres znacznie wolniej niż laserowy system paskowy.

TOLERANCJA NA BŁĘDY

Redundancja fizyczna wprowadza nowe zagadnienia, które są obszarem aktywnych badań naukowych. Prawdopodobnie najbardziej intrygujące jest to, jak robot może określić, że czujnik (lub algorytm) uległ awarii i należy go wymienić. Przetwanie awarii jest określane jako odporność na awarie. W większości przypadków robota można zaprogramować tak, aby tolerował błędy, o ile potrafi rozpoznać, kiedy one wystąpią.

Fuzja czujników behawioralnych

POŁĄCZENIE CZUJNIKÓW

ZBĘDNY

UZUPEŁNIAJĄCY

KOORDYNOWANE

Fuzja czujników to szerokie pojęcie używane w odniesieniu do dowolnego procesu, który łączy informacje z wielu czujników w jedną percepcję. Motywacja do fuzji czujników wynika z trzech podstawowych kombinacji czujników: nadmiarowych (lub konkurujących), komplementarnych i skoordynowanych. Chociaż wielu badaczy traktuje fuzję czujników jako sposób konstruowania globalnego modelu świata w systemie hierarchicznym lub deliberatywnym, fuzję czujników można włączyć do zachowań poprzez podział czujników, zorientowaną na działanie fuzję czujników i modę czujników.

FAŁSZYWY POZYTYW

FAŁSZYWY NEGATYWNY

Motywacją do korzystania z wielu czujników jest to, że jeden czujnik może być zbyt nieprecyzyjny lub zbyt hałaśliwy, aby dostarczać wiarygodnych danych. Dodanie drugiego czujnika może dać kolejny „głos” na percept. Kiedy czujnik prowadzi robota do przekonania, że istnieje jakiś element, ale tak nie jest, błąd określany jest jako fałszywie pozytywny. Robot dokonał pozytywnej identyfikacji percepcji, ale była to fałszywa. Podobnie błąd polegający na tym, że robot nie trafia w percepcję, nazywany jest fałszywie negatywnym. Czujniki często generują różne wartości wyników fałszywie dodatnich i fałszywie ujemnych. To, czy robot może tolerować wyższy odsetek wyników fałszywie dodatnich lub fałszywie ujemnych, zależy od zadania. Prace nad cyfrowym przetwarzaniem sygnałów (DSP) traktują problem łączenia czujników jako problem w rejestrowaniu lub wymuszaniu podobieństw w strumieniach danych. Na przykład czujniki mogą pokrywać nieco inne obszary, a zatem problemem DSP jest wyrównanie odczytów czujnika ze sobą. Jeden czujnik może wytwarzać dane częściej niż inny. Chodzi o to, aby strumienie danych były wyrównane w czasie i przestrzeni, aby można je było zsumować.

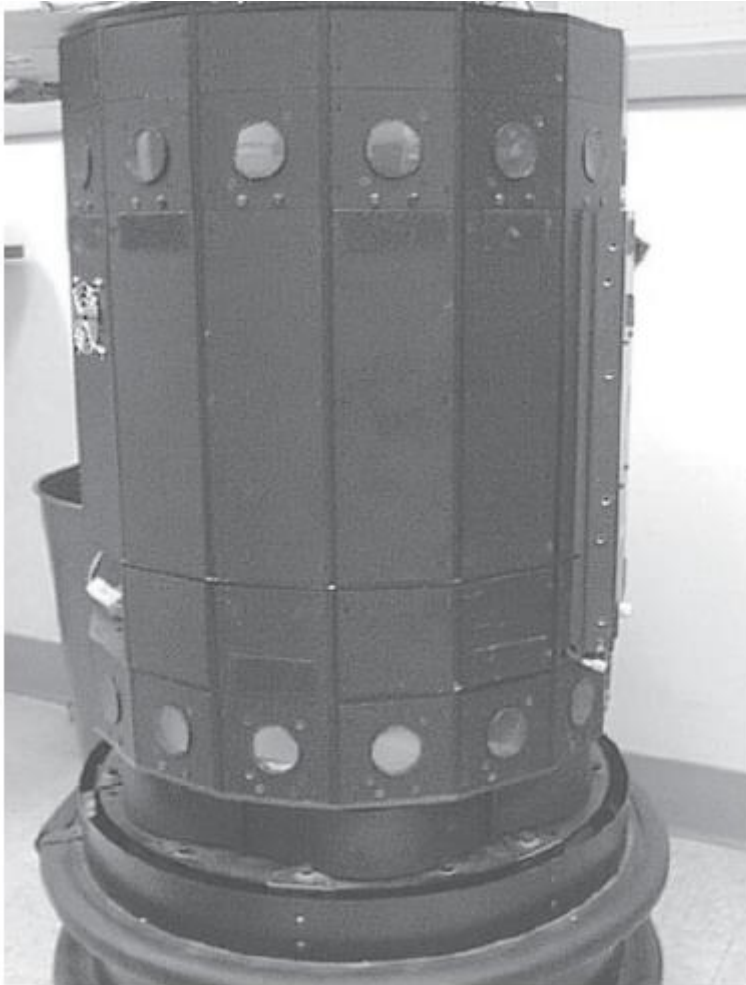
ZBĘDNY

REDUNDANCJA FIZYCZNA

LOGICZNIE ZREDUKOWANY

KONKURENCYJNE CZUJNIKI

Fuzja czujników behawioralnych ma inne podejście. Gdy oba czujniki zwracają tę samą percepcję, czujniki są uważane za nadmiarowe. Przykład redundancji fizycznej pokazano na rysunku, gdzie Nomad 200 ma dwa pierścienie sonaru.



Oprogramowanie sonaru zwraca minimalny odczyt (najkrótszy zasięg) z dwóch pierścieni, zapewniając bardziej niezawodny odczyt dla niskich obiektów, które zwykle odbijałyby wiązkę z górnego sonaru. Czujniki mogą być również logicznie nadmiarowe, gdy zwracają identyczne spostrzeżenia, ale używają różnych modalności lub algorytmów przetwarzania. Przykładem jest wyodrębnienie obrazu odległości z kamer stereo i dalmierza laserowego. Czasami czujniki nadmiarowe nazywane są czujnikami konkurującymi, ponieważ czujniki mogą być postrzegane jako konkurujące o opublikowanie „zwycięskiego” punktu widzenia.

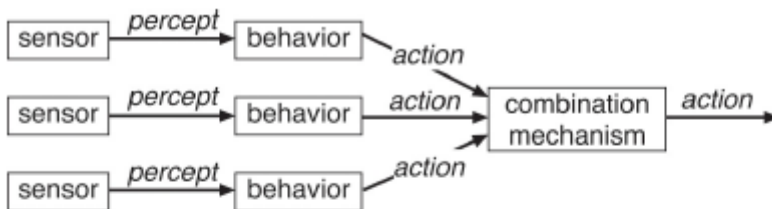
Komplementarne czujniki dostarczają rozłączne typy informacji o percepcji. W fuzji czujników behawioralnych do działań poszukiwawczo-ratowniczych w miastach robot może szukać ocalałych, łącząc obserwacje z czujnika termicznego ciepła ciała z obserwacjami z kamery wykrywającej ruch. Oba czujniki logiczne zwracają pewien aspekt „ocalałego”, ale żaden z nich nie zapewnia pełnego obrazu.

MODA CZUJNIKÓW

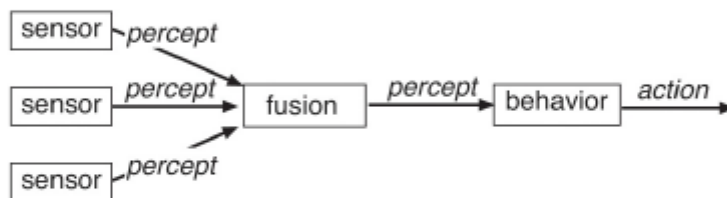
Czujniki skoordynowane wykorzystują sekwencję czujników, często do wskazywania lub zapewniania skupienia uwagi. Drapieżnik może zobaczyć ruch, powodując zatrzymanie i dokładniejsze zbadanie sceny pod kątem śladów ofiary. Tradycyjna fuzja czujników i podejścia DSP pozostawiają kategorię czujników skoordynowanych nietkniętą. Arkin wypełnił widoczną lukę, nazywając ten rodzaj mody czujnikami koordynacji, „modą” literacką z „fuzją” i zamierzając sugerować, że robot zmienia czujniki wraz ze zmieniającymi się okolicznościami, tak jak ludzie zmieniają styl ubioru wraz z porami roku.

CZUJNIK ROZCIĘCIA

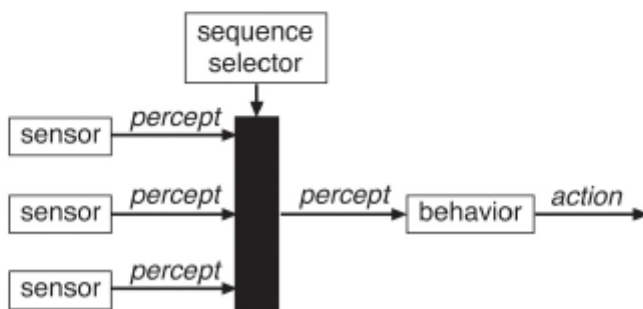
Większość prac traktuje fuzję czujników jak proces deliberatywny: taki, który wymaga globalnego modelu świata. Wczesne prace nad systemami reaktywnymi wykorzystywały roboty z zaledwie kilkoma prostymi czujnikami, sonarem lub pierścieniem sonaru do określania zasięgu i kamerą do określania koloru, tekstury lub afordancji ruchu. W rezultacie powstała filozofia projektowania polegająca na użyciu jednego czujnika na zachowanie. Zachowania mogą współdzielić strumień czujników, ale nie wiedząc o tym. Ta filozofia doprowadziła do podejścia przyjętego przez Brooksa, zgodnie z którym fuzja czujników na poziomie behawioralnym była mirażem; tak naprawdę była to kombinacja wielu przypadków tego samego zachowania, z których każda miała inne dane wejściowe czujnika. Zewnętrznemu obserwatorowi wydawałoby się, że wewnątrz robota zachodzi jakiś skomplikowany proces, ale w rzeczywistości byłaby to zwykła konkurencja z wyłaniającym się zachowaniem. Brooks nazwał to rozszczepienie czujnika po części jako początek konotacji słowa „fuzja” w fizyce jądrowej. W fuzji jądrowej energia jest wytwarzana przez zmuszenie atomów i cząstek do siebie, podczas gdy w rozszczepieniu energia jest wytwarzana przez rozdzielanie atomów i cząstek. Rysunek a przedstawia schemat rozszczepienia czujnika.



a.



b.



c.

POŁĄCZENIE CZUJNIKÓW ZORIENTOWANYCH NA DZIAŁANIE

Murphy doniósł o badaniach z psychologii poznawczej i neurofizjologii, które wykazały, że fuzja czujników behawioralnych występuje u zwierząt i dlatego powinna być częścią behawioralnego repertuaru robota. Ścieżki sensoryczne w mózgu pozostają oddzielne i mogą być kierowane do wielu zachowań w wzgórkach górnych. Dopiero gdy sygnały czujnika skierowane do części mózgu związanej z

określonym zachowaniem dotrą do tego miejsca, wydaje się, że nastąpiła jakakolwiek transformacja w nową połączoną reprezentację. Dowolny lub wszystkie z tych strumieni czujników dla danego zachowania mogą być aktywne i wpływać na wynikowe zachowanie. Rozważmy na przykład zachowanie drapieźników u kotów. Jeśli kot usłyszy hałas i zobaczy ruch, zareaguje silniej, niż gdyby otrzymał tylko jeden bodziec. Ten rodzaj fuzji czujników nazywa się łączeniem czujników zorientowanych na działanie w celu podkreślenia, że dane z czujnika są przekształcane w reprezentację specyficzną dla zachowania w celu wsparcia określonego działania, a nie konstruowania modelu świata. Rysunek b przedstawia schemat czujnika zorientowanego na działanie. Rozszczepienie czujnika i zorientowane na działanie połączenie czujników obejmuje konkurencyjne i uzupełniające się wykrywanie. Rozszczepienie czujnika jest z definicji metodą konkurencyjną, chociaż czujniki uzupełniające mogą być używane do wspierania konkretnego przypadku zachowania. Fuzja czujników zorientowana na działanie nie ogranicza się do czujników konkurujących lub uzupełniających, ponieważ zachowanie i tak powoduje transformację lokalną. Schemat budowy czujnika pokazano na rysunku c.

Projektowanie zestawu czujników

W przeszłości roboty reaktywne wykorzystywały niedrogie przetworniki IR lub ultradźwiękowe do wykrywania zasięgu. Teraz przeszli do używania kamer RGB-D lub algorytmów głębi z X. Najwcześniejsze zachowania koncentrowały się na podstawowych umiejętnościach nawigacyjnych, takich jak unikanie przeszkód i podążanie za ścianą. Spostrzeganie tych zachowań polega na znajomości odległości do zajmowanego obszaru przestrzeni. Obecnie, wraz z pojawieniem się niedrogich miniaturowych kamer i dalmierzy laserowych do zastosowań konsumenckich, korzystanie z wizji komputerowej staje się coraz bardziej powszechne. W zastosowaniach robotów reaktywnych w rolnictwie i transporcie popularna stała się również technologia GPS. W tym rozdziale podjęto próbę omówienia podstaw tych modalności wykrywania i sposobu ich wykorzystania w robotach mobilnych. Ponieważ rynek czujników szybko się zmienia, rozdział skupi się na tym, jak zaprojektować zestaw czujników do użytku przez robota, a nie na podsumowaniu szczegółów urządzenia.

WŁASNOŚĆ

WYKROCZENIE

WYWŁASZCZENIE

Sztucznie inteligentny robot musi mieć pewne wyczucie, aby można go było uznać za prawdziwego robota AI. Jeśli nie może obserwować świata i skutków swoich działań, nie może zareagować. Projektowanie zestawu czujników dla robota rozpoczyna się od oceny rodzaju informacji, które należy wydobyć z otoczenia. Informacje te można uzyskać z propriocepcji (pomiaru ruchów względem wewnętrznego układu odniesienia), eksterocepcji (pomiaru układu otoczenia i obiektów względem układu odniesienia robota) lub ekspropriocepcji (pomiar położenia robota ciała lub części w stosunku do układu środowiska). Colorado School of Mines wystawił zgłoszenie do konkursu UGV 1995; konkurs z 1994 r. omówiono w rozdziale 19. Ten wpis stanowi przykład innego rodzaju wykrywania dla robota podążającego ścieżką. W 1995 roku zachowanie podążania za ścieżką zostało rozszerzone, aby śledzić obie linie ścieżki za pomocą obiektywu szerokokątnego w aparacie. follow-path może być uważana za eksteroceptywną, ponieważ pozyskiwała informacje o środowisku. Jednak kamera robota została zamontowana na maszcie obrotowym, który miał się obracać, aby utrzymać ścieżkę w zasięgu wzroku, niezależnie od kierunku skręcania ścieżki. Dlatego robot musiał wiedzieć, gdzie kamera została obrócona względem wewnętrznego układu odniesienia robota, aby poprawnie przekształcić współrzędne obrazu położenia białych linii w kierunku sterowania. Oznaczało to, że informacje potrzebne do śledzenia ścieżki musiały zawierać zarówno elementy proprioceptywne, jak i eksteroceptywne, czyniąc percepcję nieco eksproprioceptywną. (Gdyby robot wydobywał pozycję

swojej kamery z eksterocepcji, byłoby to wyraźnie eksproprioceptywne.) Z powodu błędu programowania zachowanie podążania za ścieżką błędnie zakładało, że dane kamery eksteroceptywnej zostały przekształcone przez dane kodera proprioceptywnego z wału. panoramowanie masztu na dane eksproprioceptywne. Robot potrzebował ekspropriocepcji, aby określić, gdzie ma się dalej poruszać: skręt, aby podążać ścieżką we współrzędnych kamery, plus kompensacja bieżącego kąta obrotu kamery. Błąd programowania spowodował, że robot zachowywał się tak, jakby kamera była zawsze ustawiona w linii ze środkiem robota. Ale kamera może być lekko obrócona, aby zachować widok obu linii ścieżki przez zachowanie pancamera. Wynikające z tego polecenie nawigacyjne może polegać na skręcie, ale nie na tyle, aby zrobić różnicę, a nawet skręcić w złą stronę. Ten subtelny błąd pojawił się, gdy robot poruszał się po szpilkach do włosów, powodując, że robot konsekwentnie wychodził poza granice.

DOPASOWANIE CZUJNIKÓW

Jak widać z powyższego przykładu, roboty mogą mieć zdolność martwego liczenia, ale zawsze będą miały jakiś rodzaj czujnika eksteroceptywnego. W przeciwnym razie robota nie można uznać za reaktywnego: nie byłoby bodźca ze świata, który wywołałby reakcję. Zestaw czujników dla konkretnego robota nazywa się dopasowaniem czujników. Zgodnie z czujnikami do robotów mobilnych, aby zbudować zestaw czujników, należy wziąć pod uwagę następujące atrybuty dla każdego czujnika:

1. Pole widzenia i zasięg. Każdy czujnik eksteroceptywny ma obszar, który ma pokryć. Szerokość tego obszaru jest określona przez pole widzenia czujnika, często określane skrótem jako FOV. Pole widzenia jest zwykle wyrażane w stopniach; liczba stopni pokrytych w pionie może różnić się od liczby stopni pokrytych w poziomie. Pole widzenia jest często wykorzystywane w fotografii, gdzie różne obiektywy rejestrują obszary o różnej wielkości i kształcie. Obiektyw szerokokątny często pokrywa do 70°, podczas gdy „zwykły” obiektyw może mieć pole widzenia tylko około 27°. Długość pola nazywa się zakresem.

POLE WIDZENIA (FOV)

POZIOME FOV

PIONOWE-FOV

UPORZĄDKOWANE

Pole widzenia (FOV) można traktować w kategoriach egocentrycznych współrzędnych sferycznych, gdzie jeden kąt to FOV w poziomie, a drugi to FOV w pionie. Innym aspektem jest zasięg, czyli jak daleko czujnik może wykonać wiarygodne pomiary. We współrzędnych sferycznych pomiary te byłyby wartościami r , które określały głębokość zakresu operacyjnego. Pole widzenia i zasięg mają oczywiście kluczowe znaczenie przy dopasowywaniu czujnika do aplikacji. Jeśli robot musi być w stanie wykryć przeszkodę, gdy znajduje się w odległości ośmiu stóp, aby bezpiecznie ją ominąć, wówczas czujnik o zasięgu pięciu stóp nie będzie akceptowany.

REZOLUCJA

2. Dokładność, powtarzalność i rozdzielczość. Dokładność odnosi się do tego, jak poprawny jest odczyt czujnika. Ale jeśli odczyt dla tego samego stanu jest dokładny tylko w 20% przypadków, wtedy czujnik ma małą powtarzalność. Jeśli czujnik jest stale niedokładny w ten sam sposób (zawsze 2 lub 3 cm nisko), oprogramowanie może zastosować odchylenie (dodaj 2 cm), aby skompensować. Jeśli niedokładność jest losowa, to będzie trudno ją zamodelować, a zastosowania, w których można zastosować taki czujnik, będą ograniczone. Jeśli odczyt jest mierzony w odstępach co jeden metr, odczyt ten ma mniejszą rozdzielczość niż odczyt z czujnika, który jest mierzony w odstępach co jeden cm.

3. Responsywność w domenie docelowej. Większość czujników ma określone środowiska, w których źle funkcjonują. Środowisko musi umożliwiać wyodrębnienie interesującego sygnału z szumu i zakłóceń (np. mieć korzystny stosunek sygnału do szumu). Jak zobaczymy poniżej, sonar często nie nadaje się do nawigacji w biurowym foyer z dużą ilością szklanych okien i ścianek działowych, ponieważ szkło odbija energię dźwięku w sposób prawie niemożliwy do przewidzenia. Ważne jest, aby scharakteryzować niszę ekologiczną robota pod kątem tego, co będzie dostarczać, pochłaniać lub odychać energię.

4. Pobór mocy. Pobór mocy jest zawsze problemem dla robotów. Ponieważ większość robotów działa na bateriach, im mniej energii zużywają, tym dłużej działają. Na przykład żywotność baterii w Nomad 200, która zawiera pięć baterii, została zwiększona z czterech do sześciu godzin dzięki wyłączeniu wszystkich czujników. W większości robotów mobilnych moc jest tak ograniczona, że wielu producentów robotów przechodzi na bardziej energooszczędne układy mikroprocesorowe tylko po to, aby zmniejszyć pobór mocy. Czujniki wymagające dużej mocy są mniej pożądane niż te, które tego nie robią. Ogólnie rzecz biorąc, czujniki pasywne mają mniejsze zapotrzebowanie na moc niż czujniki aktywne, ponieważ czujniki pasywne nie emitują energii do otoczenia.

OBCIĄŻENIE HOTELU

ŁADUNEK LOKOMOCYJNY

Ilość mocy robota mobilnego wymagana do obsługi pakietu czujników (i dowolnej innej elektroniki, takiej jak mikroprocesor i łącza komunikacyjne) jest czasami nazywana obciążeniem hotelu. Pakiet czujników jest „gościem” platformy. Moc potrzebna do poruszania robotem nazywana jest obciążeniem lokomocji. Niestety, wielu producentów robotów skupia się tylko na obciążeniu lokomocji, równoważąc zapotrzebowanie na energię z chęcią zmniejszenia całkowitej masy i rozmiarów. Prowadzi to do bardzo małego obciążenia hotelu i często uniemożliwia dodanie wielu czujników do platformy.

5. Niezawodność. Czujniki często mają fizyczne ograniczenia co do tego, jak dobrze działają. Na przykład sonary Polaroid będą generować nieprawidłowy odczyt zakresu, gdy napięcie spadnie poniżej 12 V. Inne czujniki mają ograniczenia temperatury i wilgotności, które należy wziąć pod uwagę.

6. Rozmiar. Rozmiar i waga czujnika mają wpływ na ogólny projekt. Mikrołazik na zamówienie pudełka na buty nie będzie w stanie przetransportować dużego aparatu lub kamery, ale może być w stanie użyć miniaturowego aparatu typu „Quick-Cam”. Powyższa lista skoncentrowała się na rozważaniach fizycznych aspektów czujnika. Jednak czujniki dostarczają tylko obserwacji; bez programowych schematów percepcyjnych zachowania nie mogą korzystać z czujników. Dlatego oprogramowanie, które będzie przetwarzać informacje z czujnika, musi być traktowane jako część procesu wyboru czujnika.

7. Złożoność obliczeniowa. Złożoność obliczeniowa to oszacowanie liczby operacji wykonywanych przez algorytm lub program. Złożoność obliczeniowa jest często zapisywana w zapisie funkcji jako O , zwanym „porządkiem”, gdzie $O(x)$ oznacza, że liczba głównych operacji jest proporcjonalna do x . Na przykład algorytm, który przewija listę 100 liczb, może mieć 100 operacji. Jeśli ten sam algorytm otrzyma listę 1000 liczb, liczba operacji będzie wynosić 1000; liczba operacji zależy od wielkości wejścia, czyli x . x jest często samą funkcją, a n jest używane zamiast x do oznaczenia funkcji. Niższe zamówienia są lepsze. Algorytm, który jest wykonywany z $O(n)$ równo zużywającymi się operacjami, jest szybszy niż algorytm z $O(n^2)$ operacjami. (Jeśli w to wątpisz, sprawdź, czy możesz znaleźć dodatnią liczbę całkowitą n taką, że $n \leq n^2$.) Ponadto, jeśli liczba danych wejściowych do algorytmu $O(n)$ jest podwojona, złożoność jest liniowo proporcjonalna; wykonanie algorytmu zajmie dwa razy dłużej. Ale jeśli liczba

danych wejściowych do algorytmu $O(n^2)$ zostanie podwojona, wykonanie algorytmu zajmie cztery razy dłużej. Złożoność obliczeniowa stała się mniej krytyczna dla większych robotów z powodu szybkich postępów w procesorach i miniaturyzacji komponentów. Pozostaje jednak poważnym problemem dla mniejszych systemów bezzałogowych, zwłaszcza że algorytmy widzenia komputerowego mogą mieć wartość $O(m^4)$.

8. Niezawodność. Projektant powinien wziąć pod uwagę, na ile czujnik będzie niezawodny w warunkach ekologicznych i interpretacyjnych. Robot często nie ma możliwości określenia, kiedy czujnik podaje nieprawidłowe informacje. W rezultacie robot może „halucynować” (myśleć, że widzi rzeczy, których nie ma) i robić złe rzeczy. Wiele czujników wytwarza dane wyjściowe, które są trudne do zinterpretowania przez człowieka bez wieloletniego treningu; Jednym z przykładów są medyczne promienie rentgenowskie, a drugim syntetyczny radar temperatury (SAR), który tworzy wykresy polarne. Jeśli algorytm czujnika nie działa poprawnie w tych modalnościach, projektant może nie mieć wystarczających umiejętności, aby to zauważyć. Dlatego same algorytmy muszą być niezawodne.

Podsumowanie

Sukces robota zależy od przydatności jego czujników i wykrywania. Często bardziej przydatne jest myślenie o wykrywaniu w kategoriach schematów percepcyjnych lub czujników logicznych potrzebnych do wykonania zadania, niż skupianie się na cechach konkretnego przetwornika lub modalności. Prawie wszystkie roboty mobilne mają pewną formę propriocepcji, najprawdopodobniej enkodery wału lub koła używane do szacowania przebytej odległości na podstawie liczby obrotów silnika. Roboty zewnętrzne zwykle korzystają z GPS, ale dokładność i niezawodność GPS może być znacznie zmniejszona w miejskich kanionach i nie będzie istnieć w pomieszczeniach. Dlatego nawigacja w obszarach, w których nie ma GPS, pozostaje wyzwaniem. Czujniki zasięgu, omówione w kolejnej części, nie do końca rozwiązują to wyzwanie. Ze względu na niską cenę i dostępność elektroniki użytkowej oraz zwykłą potrzebę posiadania kamery na robocie do zapasowego teleoperacji, widzenie komputerowe jest coraz powszechniejsze w systemach robotycznych. Komputerowe przetwarzanie widzenia działa na obrazach, niezależnie od modalności, która je wygenerowała. Systemy współrzędnych kolorów mają tendencję do dzielenia obrazu na trzy płaszczyzny. Dwa najpopularniejsze systemy współrzędnych kolorów to RGB i HSV. HSV traktuje kolor w kategoriach bezwzględnych, ale RGB jest faworyzowany przez producentów sprzętu. Przestrzeń kolorów wykorzystywana w obrazowaniu biomedycznym, SCT, wydaje się być mniej wrażliwa na warunki oświetleniowe niż RGB i HSV pochodzące z RGB. Wiele reaktywnych robotów wykorzystuje kolor jako afordancję poprzez progowanie obrazu i identyfikowanie obszarów o odpowiednim kolorze. Metodą afordancji kolorów, która sprawdza się dobrze w przypadku obiektów o wielu kolorach, jest histogram kolorów. Projektant musi zobowiązać się do określonego zestawu czujników i algorytmów. Czyniąc to, warto znać trzy koncepcje. Pierwsza koncepcja to idea czujników logicznych lub równoważnych, w których możliwe jest wygenerowanie tego samego postrzegania z różnych czujników lub algorytmów. Drugi to fuzja czujników behawioralnych, która opisuje ogólne metody łączenia czujników w celu uzyskania pojedynczego spostrzeżenia lub wsparcia złożonego zachowania. Trzecią koncepcją jest zrozumienie atrybutów zestawu czujników, które można wykorzystać do pomocy w projektowaniu systemu. Wracając do pytań postawionych we wstępie, odpowiedź na pytanie Jak sprawić, by robot „widział”? polega na tym, że zależy to od czujników, które dostarczają dane, oraz algorytmów wykrywania, które wydobywają spostrzeżenia lub generują modele świata z tych danych. Kolejne pytanie brzmi: jakie czujniki są niezbędne dla robota? Konkretna odpowiedź zależy oczywiście od misji, ale ogólnie rzecz biorąc, roboty będą potrzebować zarówno propriocepcji, aby śledzić pozę swojego ciała fizycznego, jak i eksterocepcji, aby nawigować i rozpoznawać obiekty. Ponieważ ludzie mają tendencję do myślenia o percepcji w kategoriach oczu, teleoperator zazwyczaj ma kamerę światła

widzialnego, aby teleoperator mógł bardziej naturalnie sterować robotem. Wykorzystanie kamery (kamer) to domena pola widzenia komputera. Część dotyczyła również: Co to jest fuzja czujników? Omówiono fuzję czujników, czyli mechanizmy, które umożliwiają wielu czujnikom wytwarzanie percepcji i modeli. W przeciwieństwie do cyfrowego przetwarzania sygnałów, fuzja czujników w sztucznej inteligencji nie zakłada, że dane z czujników muszą być rejestrowane i próbkowane w tym samym czasie; zamiast tego metody sztucznej inteligencji badają zarówno zbędne, jak i uzupełniające się modalności percepcji. Pomimo różnorodności czujników robotyka wyróżnia się brakiem wyrafinowania w wykrywaniu. Może to wynikać z podziału między wizją komputerową a robotyką w pierwszych latach tej dziedziny. Zachęcamy czytelnika do dokładnego zapoznania się z literaturą dotyczącą wizji komputerowych, zamiast próbować odkrywać koło na nowo.