

Akwizycja obrazów RGB-D: czujniki

Tomasz Kornuta, Maciej Stefańczyk

Instytut Automatyki i Informatyki Stosowanej, Politechnika Warszawska

Streszczenie: Dwuczęściowy artykuł poświęcono czujnikom umożliwiającym akwizycję obrazów RGB-D. W pierwszej części uwagę skupiono na metodach akwizycji, w kolejnej części dokonano przeglądu czujników, zarówno dostępnych obecnie, jak i tych, które dopiero w najbliższym czasie pojawią się na rynku.

Słowa kluczowe: mapa głębi, chmura punktów, obraz RGB-D, czujnik RGB-D

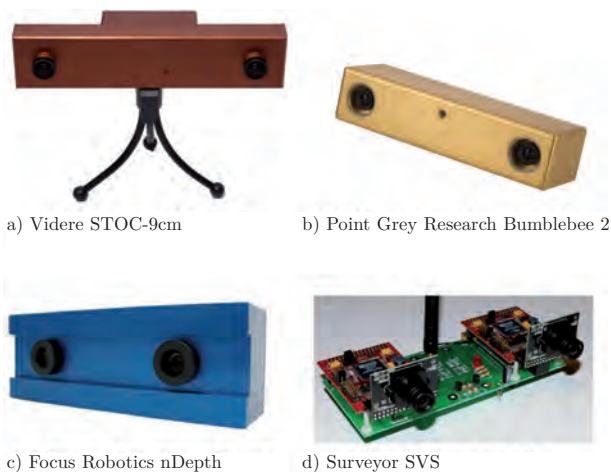
DOI: 10.14313/PAR_204/92

Rosnąca rola trójwymiarowej wizji komputerowej w robotyce związana jest z rozwojem oraz upowszechnianiem czujników umożliwiających akwizycję obrazów RGB-D. Poniższe opracowanie jest drugą częścią artykułu dotyczącego czujników tego rodzaju. W poprzedniej części omówiono podstawowe rodzaje reprezentacji głębi oraz zdefiniowano obraz RGB-D za pomocą czterech kanałów: trzy kanały kolorowe oraz kanał przechowujący odległości poszczególnych elementów sceny od kamery, zwany mapą głębi. Teoretycznie mapę głębi można zbudować korzystając z dowolnego punktowego czujnika pomiaru odległości – przez zamontowanie go na głowicy uchylno-obrotowej (przeгляд takich czujników m.in. w [17]), jednak taka akwizycja danych byłaby bardzo czasochłonna.

W niniejszym artykule skupiono uwagę na rozwiązaniach, które mogą działać poprawnie w czasie rzeczywistym, a więc mogą być montowane na poruszającym się pojeździe lub robocie. W przypadku takich zastosowań można wymienić trzy główne metody pozyskiwania głębi: stereowizję, zastosowanie światła strukturalnego oraz pomiar czasu lotu wiązki. Podstawy fizyczne tych metod omówiono dokładnie w pierwszej części artykułu [25]. W drugiej części są omówione dostępne na rynku rozwiązania sprzętowe. Czujniki sklasyfikowano wg metody ich działania oraz omówiono ich parametry. Poza samymi urządzeniami zaprezentowano także wybrane komercyjne systemy pomiarowe wykorzystujące opisywane technologie.

1. Stereowizja

Działanie stereowizji jest analogią do sposobu postrzegania otoczenia przez człowieka. Głębina obliczana jest na podstawie niezgodności obrazów pochodzących z lewej i prawej



Rys. 1. Popularne czujniki zwracające obrazy RGB-D wykorzystujące stereowizję

Fig. 1. Popular RGB-D sensors based on the stereo vision principle

kamery. Zaletą tej metody jest brak dodatkowych urządzeń (np. projektorów), natomiast wadą jest konieczność kalibracji kamer, ograniczona dokładność oraz duży koszt obliczeniowy. Z tego powodu szereg firm oferuje stereopary złożone z dwóch skalibrowanych ze sobą kamer oraz zoptymalizowane biblioteki przeprowadzające obliczenia związane z ekstrakcją mapy głębi. Dostępne są też produkty wyposażone dodatkowo w wyspecjalizowany układ, przeprowadzający odpowiednie obliczenia, tym samym odciążający procesor komputera.

Firma Videre Design [14] oferuje rodzinę stereopar STOC (rys. 1a) o różnym rozstawie kamer. Na rynku dostępne są wersje z rozstawem osi optycznych 6–30 cm, istnieje też wersja z płynną regulacją odległości. Rozstaw ten wpływa na odległość od czujnika, w której obliczana jest mapa niezgodności, dzięki czemu klient może dobrać czujnik do własnych potrzeb. Obliczenia związane z pozyskaniem mapy głębi realizowane są na zintegrowanym procesorze. STOC przesyła obraz RGB-D z rozdzielczością 640 px × 480 px z maksymalną częstotliwością 30 Hz.

Równie popularnym producentem stereopar jest firma Point Grey Research [7], oferująca m.in. czujnik Bumblebee 2 (rys. 1b). Czujnik ten jest drugą generacją popularnej stereopary Bumblebee, zwraca obrazy o rozdzielczości 640 px × 480 px z częstotliwością 48 Hz lub 1024 px × 768 px



Rys. 2. Czujniki zwracające obrazy RGB-D bazujące na technologii opracowanej przez PrimeSense

Fig. 2. The PrimeSense technology-based RGB-D sensors

z częstotliwością 20 Hz. W obu przypadkach mapa głębi wyliczana jest za pomocą oprogramowania na komputerze. Firma oferuje trzy zestawy soczewek o różnych ogniskowych, co umożliwia zmianę pola widzenia (w nawiasie podano kąt widzenia w poziomie): 2,5 mm (100°), 3,8 mm (65°), 6 mm (43°). Dodatkowo, Bumblebee 2 umożliwia zewnętrzne wyzwalanie akwizycji obrazów.

Na rys. 1c pokazano kamerę stereowizyjną nDepth firmy Focus Robotics [1]. Czujnik ten zwraca obrazy RGB-D z częstotliwością do 30 Hz z maksymalną rozdzielczością 752 px × 480 px (WVGA). Czujnik określa głębie skwantyzowaną do 92 poziomów. Razem z czujnikiem firma oferuje nDepth Vision Processor, dedykowany układ FPGA instalowany w komputerze w slotcie PCI.

Tańszą alternatywą dla powyższych trzech stereopar jest czujnik SVS (rys. 1d) (ang. *Stereo Vision System*) firmy Surveyor Corporation [11]. Układ składa się z dwóch kamer Blackfin dostarczających obrazy z rozdzielczością 1280 px × 1024 px, 640 px × 480 px, 320 px × 240 px lub 160 px × 120 px. Minusem czujnika jest brak dedykowanego układu obliczeniowego, odpowiedzialnego za rektyfikację obrazu oraz wyliczanie map niezgodności i głębi.

Konstrukcje bazujące na dwóch kamerach są najpopularniejsze, ale są inne możliwości. Istnieją rozwiązania, w których są zastosowane jednocześnie trzy kamery [18, 19] lub takie, które mogą sterować rozstawem kamer [23]. Przykładem komercyjnego rozwiązania pierwszego typu jest czujnik Bumblebee XB3 (rys. 3). Ma on trzy kamery, z których każda charakteryzuje się rozdzielczością 1,3 Mpx. Urządzenie jest dostępne z dwoma różnymi rozstawami kamer: szerszy, umożliwiający otrzymanie głębi z większą precyzją przy większych odległościach, oraz węższy, który lepiej sprawdza się dla bliższych obiektów. Czujnik ten zwraca obrazy o rozdzielczości 1280 px × 960 px z częstotliwością 15 Hz. Obecnie na rynku dostępne są dwie wersje czujnika z obiektami o ogniskowych 3,8 mm oraz 6 mm, dzięki czemu kąt widzenia w poziomie wynosi odpowiednio 66° i 43°.

Warto zwrócić uwagę na nowy czujnik PlayStation Camera firmy Sony (rys. 4), który pojawił się na rynku pod



Rys. 3. Bumblebee XB3 firmy Point Grey Research

Fig. 3. The Bumblebee XB3 sensor offered by Point Grey Research



Rys. 4. PlayStation Camera firmy Sony

Fig. 4. Sony PlayStation Camera

koniec listopada 2013 r. wraz z premierą nowej konsoli do gier PlayStation 4. Produkt ten jest następcą kamery PlayStation Eye oraz stanowi odpowiedź firmy Sony na konkurencyjne czujniki Kinect, oferowane jako kontrolery ruchu dla konsoli Xbox 360 przez firmę Microsoft. PlayStation Camera ma dwie kamery zwracające obrazy o rozdzielczości 1280 px × 800 px z częstotliwością do 60 Hz. Brak jest informacji dotyczących algorytmu do obliczania głębi, jej jakości oraz obecności dedykowanego układu obliczeniowego na samym czujniku. Wiadomo, że zakres działania czujnika wynosi od 30 cm do (teoretycznie) nieskończoności, natomiast szerokość pola widzenia w poziomie wynosi 85°. Niewątpliwą zaletą tego czujnika jest wynikająca ze skali produkcji niska cena.

2. Światło strukturalne

Największym komercyjnym sukcesem produktów wykorzystujących światło strukturalne jest niewątpliwie Kinect firmy Microsoft (rys. 2b). W ciągu pierwszych 60 dni sprzedano ponad 8 milionów egzemplarzy, a liczba obecnie wykorzystywanych na całym świecie urządzeń przekracza 20 milionów sztuk. Jest to efekt niskiej ceny, dzięki której niemal każde laboratorium badawcze mogło się w tego typu czujniki zaopatrzyć. Niską cenę czujnik Kinect zawdzięcza masowej produkcji – został wyprodukowany jako kontroler do gier, i jako taki ma zdecydowanie większy rynek zbytu niż tradycyjne, wyspecjalizowane czujniki.

Światło strukturalne zostało także użyte do wyznaczania mapy głębi przez czujniki Carmine firmy PrimeSense (rys. 2a) oraz Xtion firmy Asus (rys. 2c). W istocie stosowaną w nich technologię opracowała ta sama firma – PrimeSense. Idea polega na projekcji znanego wzorca strukturalnego za pomocą rzutnika w bliskiej podczerwieni IR (ang. *Infrared*) oraz rejestracji zniekształconej siatki za pomocą kamery działającej w tym samym paśmie. Deformacja wzorca jest stosowana do wyznaczenia obrazu głębi o rozdzielczości 300 × 200 punktów, programowo interpolowanej do rozdzielczości kolorowej kamery wizyjnej zamontowanej w czujniku (640 × 480). Mapa głębi generowana jest z częstotliwością 30 Hz, a zakres

działania zależy od kalibracji układu. Poza projektorem oraz kamerami czujniki te są wyposażone w macierz czterech mikrofonów kierunkowych oraz akcelerometr.

Czujniki Carmine mają pole widzenia $57,5^\circ \times 45^\circ$ oraz zakres akwizycji mapy głębi zależny od wersji: Carmine 1.08 ma zakres 0,8–3,5 m, natomiast zakres dla wersji 1.09 to 0,35–1,4 m. W przypadku czujników Kinect oraz Xtion zakres ten wynosi 0,6–4 m. W odróżnieniu od Carmine oraz Xtion, czujnik Kinect ma napęd umożliwiający sterowanie pochyleniem głowicy.

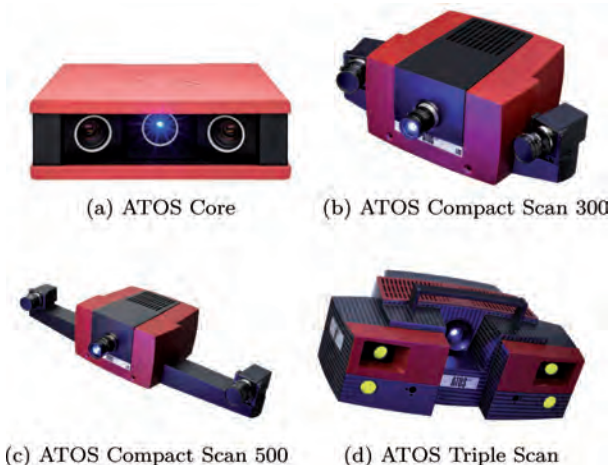


Rys. 5. Czujnik Geomagic Scan3D BLUE firmy SMARTTECH
Fig. 5. Geomagic3D Blue sensor from SMARTTECH

Na rys. 5 pokazano czujnik Geomagic Scan3D BLUE firmy SMARTTECH [9]. W odróżnieniu od technologii opracowanej przez PrimeSense, Scan3D BLUE tworzy mapę głębi na podstawie analizy odkształceń sekwencji kolejno wyświetlanych wzorców (prążków Graya). Urządzenie wykorzystuje światło widzialne – projektor wyświetla wzorce za pomocą niebieskich diod LED. Czas jednokrotnego skanowania sceny waha się w przedziale 1–5 s. Ponieważ w trakcie analizy całej sekwencji scena nie powinna się zmieniać, czujnik ten nie nadaje się do zastosowań w systemach czasu rzeczywistego, natomiast doskonale sprawdza się w aplikacjach fotogrametrycznych (np. kontrola jakości produktów). W zależności od detektora (firma oferuje czujniki o rozdzielczości 2 Mpx, 5 Mpx lub 10 Mpx) oraz od odległości kamery od analizowanej sceny, pojedynczy skan tworzy chmurę złożoną z 2–10 milionów punktów, a pole pomiaru wynosi od $150 \text{ mm} \times 200 \text{ mm}$ do $1200 \text{ mm} \times 1600 \text{ mm}$.

Na rys. 6 pokazano wybrane czujniki z rodziny ATOS firmy GOM [4]. Zasada działania bazuje na stereowizji z wykorzystaniem dodatkowego projektora oświetlającego scenę za pomocą sekwencji wzorców (warto dodać, że nie jest to pierwsze rozwiązanie łączące wizję wielokamerową ze światłem strukturalnym [16]). Podobnie jak w Scan3D BLUE, w najnowszych swoich produktach firma GOM postanowiła również wykorzystać światło niebieskie. Według danych producenta, w porównaniu do technologii opartej na świetle białym, stosowanej w starszych wersjach produktów z rodziny ATOS, wykorzystanie światła niebieskiego dało znaczną poprawę rozpoznawania powierzchni odbłaskowych (ok. 40 % więcej szczegółów).

Podstawowym produktem z tej rodziny jest ATOS Core (rys. 6a). Dostępnych jest osiem modeli ATOS Core, różniących się głównie rozstawem kamer, co wpływa na wielkość pola pomiarowego (od $45 \text{ mm} \times 30 \text{ mm}$ do $500 \text{ mm} \times 380 \text{ mm}$) oraz maksymalną odległość obiektu od czujnika (od 170 mm do 440 mm).



Rys. 6. Wybrane czujniki z rodziny ATOS firmy GOM

Fig. 6. Selected sensors from the ATOS family offered by the GOM company

W przypadku czujników ATOS Compact Scan firma oferuje dwa modele (2M oraz 5M), różniące się rozdzielczością zainstalowanych kamer ($2 \times 2 \text{ Mpx}$ lub $2 \times 5 \text{ Mpx}$). W każdym z modeli wyróżniono trzy konfiguracje (SO, 300, 500) różniące się rozstawami kamer (rys. 6b i rys. 6c). W zależności od modelu oraz konfiguracji czujnik może wyznaczyć głębię w zakresie 35–1200 mm.

ATOS Triple Scan są najbardziej zaawansowanymi produktami z rodziny ATOS. W zależności od modelu są wyposażone w kamery o rozdzielczości 5 Mpx, 8 Mpx oraz 12 Mpx, pole pomiarowe o maksymalnym rozmiarze $2000 \text{ mm} \times 1500 \text{ mm}$ oraz mogą wyznaczyć głębię w zakresie 490–2000 mm. Producent wraz z czujnikami dostarcza oprogramowanie (GOM Scan software), umożliwiające m.in. aproksymację zeskanowanej powierzchni za pomocą trójkątów, a także analizę i pomiary powierzchni.



Rys. 7. Czujnik Structure firmy Occipital

Fig. 7. The Structure sensor offered by Occipital

Czujnik Structure (rys. 7) firmy Occipital [10] generuje obraz RGB-D o rozdzielczości $640 \text{ px} \times 480 \text{ px}$ z częstotliwością 30 Hz, działa w zakresie 0,4–3,5 m, pole widzenia $58^\circ \times 45^\circ$. Jego charakterystyka jest bardzo podobna do czujników bazujących na technologii PrimeSense, natomiast różni się tym, że jest reklamowany jako pierwszy czujnik RGB-D dla urządzeń mobilnych – czujnik jest przystosowany głównie do pracy z tabletami z rodziny iPad firmy Apple. Co ciekawe, finansowanie na ten projekt firma Occipital zdobyła na portalu Kickstarter¹, zbiórka funduszy zakończyła się

¹ www.kickstarter.com/projects/occipital/structuresensor-capture-the-world-in-3d

w październiku 2013 r., a pierwsze komercyjne wersje czujnika mają być rozsyłane w marcu 2014 r.

Implikacją stosowania światła strukturalnego we wszystkich wymienionych tu czujnikach jest ich problematyczne działanie w obszarach silnie oświetlonych zewnętrznym źródłem światła, np. w świetle dziennym. Czujniki te mogą być stosowane głównie w pomieszczeniach (ang. *Indoor Environment*). Mimo tego ograniczenia oraz znacząco słabszych parametrów od np. SR3000 (w przypadku czujników działających w czasie rzeczywistym), pojawienie się technologii PrimeSense oraz czujnika Kinect wywołało bardzo duże zainteresowanie, a w konsekwencji postęp na polu trójwymiarowej percepcji komputerowej. Rozwój ten nie byłby możliwy bez odpowiednich narzędzi programistycznych, jak np. PCL (ang. *Point Cloud Library*) [24], których powstanie zbiegło się z opracowaniem czujnika Kinect.

3. Pomiar czasu lotu wiązki

Trzecim sposobem wyznaczania głębi jest pomiar czasu lotu wiązki. Jednym z najpopularniejszych czujników wykorzystywanych w robotyce mobilnej jest czujnik LMS111 firmy SICK (rys. 8a). Jego zaletami są szeroki kąt widzenia w poziomie 270° , zasięg do 30 m oraz rozdzielczość kąтова $0,25^\circ/0,5^\circ$ dla częstotliwości pomiaru równej odpowiednio 25 Hz/50 Hz. Należy podkreślić, że poza głębią danego punktu czujnik określa również intensywność światła powracającego do sensora po odbiciu od obiektu (ang. *Reflectance*) oraz odległość tzw. drugiego odbicia, dzięki czemu możliwe jest m.in. wykrycie obiektów znajdujących się za szkłem.

Podobne parametry cechują czujnik UTM-30LX firmy Hokuyo (rys. 8b). Ma zakres działania 0,1–60 m (choć producent gwarantuje poprawne odczyty dla odległości nie większych od 30 m), kąt poziomego widzenia równy 270° , rozdzielczość $0,25^\circ$ oraz częstotliwość pracy równa 40 Hz. Czujnik rotoScan ROD4 plus (rys. 8c) firmy Leutze Electronic [5] pracuje z częstotliwością 25 Hz, ma kąt poziomego widzenia równy 190° oraz zasięg do 65 m.



(a) SICK LMS111 (b) Hokuyo UTM-30LX (c) Leutze ROD4 plus

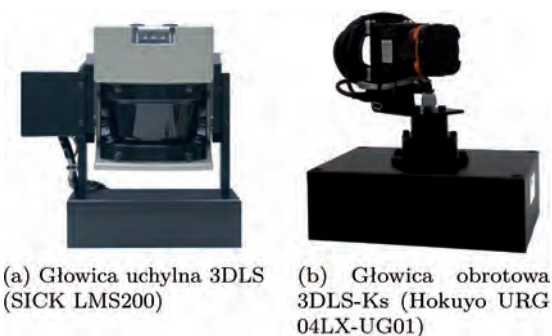
Rys. 8. Popularne lidary tworzące planarne chmury punktów
Fig. 8. Common sensors generating planar clouds of points

Oczywistym minusem wymienionych czujników jest akwizycja głębi tylko w jednej płaszczyźnie. Do otrzymania dwuwymiarowej mapy głębi niezbędne jest montowanie ich np. na obrotowych głowicach. Przykładem komercyjnie dostępnego robota usługowego, wyposażonego w skaner laserowy umieszczony na głowicy uchylnej zintegrowanej z jego korpusem, jest robot PR2 [15]. W przypadku robotów

mobilnych lidary najczęściej umieszczane są na głowicach uchylnych lub uchylno-obrotowych. Przykładami eksperymentalnych konstrukcji tego typu są głowice zamontowane na robotach Elektron [26] oraz Pioneer [21].

Obecnie na rynku dostępne są też głowice komercyjne. Jednym z producentów jest Fraunhofer IAIS [2]. Firma ta oferuje kilka typów głowic, zarówno dla lidarów SICK, jak i Hokuyo. Wraz z głowicami w ofercie znajduje się interfejs programowy do obsługi obrotnic z podłączonymi do nich czujnikami.

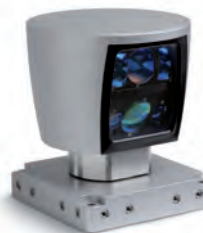
Na rys. 9a pokazano głowicę uchylną 3DLS z zamontowanym lidarem firmy SICK. Powstały układ pomiarowy ma zasięg 10 m oraz kąty widzenia w poziomie i w pionie równe odpowiednio 180° oraz 124° . Czas akwizycji całego skanu, w zależności od rozdzielczości, wynosi od 3,2 s (1°) do maksymalnie 26,64s ($0,25^\circ$). Rozwiązanie to nie nadaje się do aplikacji czasu rzeczywistego.



Rys. 9. Komercyjne głowice Fraunhofer IAIS
Fig. 9. Commercial pan-tilt units from Fraunhofer IAIS

Wady tej pozbawione są rozwiązania firmy Velodyne [13]. Czujnik HLD-64E (rys. 10) ma 64 diody laserowe, dzięki czemu oferuje pełny kąt widzenia w poziomie (360°), kąt widzenia w pionie równy $26,8^\circ$ oraz zwraca ponad milion punktów na sekundę z częstotliwością odświeżania 5–15 Hz. Zasięg pracy czujnika to 70 m z dokładnością ± 2 cm. Velodyne oferuje też mniejszy i tańszy czujnik HDL-32E, który generuje chmurę złożoną z maksymalnie 700 tys. punktów.

Z uwagi na swoje cechy, wymienione czujniki często stosowane są razem. Przykładowo, autonomiczny samochód Junior miał czujnik Velodyne zamontowany na dachu, dwa lidary firmy SICK zamontowane na bokach samochodu, dwa lidary IBEO ALASCA XT zamontowane w przednim



Rys. 10. Czujnik HDL-64E firmy Velodyne
Fig. 10. The HDL-64E sensor offered by Velodyne

zderzaku oraz pięć radarów dalekiego zasięgu BOSCH LRR2 zamontowanych na masce samochodu [22].

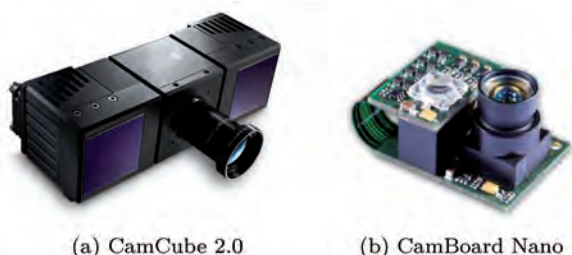
Zasadę ToF wykorzystują również czujniki rodziny SwissRanger firmy Mesa Imaging [6]. Podobnie jak produkty firmy SICK, czujniki SwissRanger, wraz ze współrzędną kartezjańską każdego punktu, generują też informację dotyczącą intensywności odbitego światła dochodzącego do sensora (obraz w odcieniach szarości). Czujnik SR3000 (rys. 11a) ma pole widzenia o zakresie kątowym równym $47,5^\circ$ na $39,6^\circ$ oraz zasięg działania do 7,5 m. Natomiast czujnik SR4000 (rys. 11b) ma sensor o rozdzielczości równej $176 \text{ px} \times 144 \text{ px}$ i jest oferowany w dwóch wersjach – ze standardowymi oraz szerokokątnymi obiektywami. W zależności od wersji pole widzenia w zakresie kątowym poziomym i pionowym wynosi odpowiednio $43,6^\circ$ i $34,6^\circ$ lub 69° i 56° . Obie wersje mogą pracować z różnymi maksymalnymi zasięgami działania: 5 m lub 10 m. Odpowiednio do zasięgu, odczyty zwracane są z częstotliwością 15 Hz lub 30 Hz.



Rys. 11. Czujniki firmy Mesa Imaging zwracające głębię z dodatkową informacją dot. intensywności (I-D)

Fig. 11. Sensors from Mesa Imaging acquiring depth images with additional intensity information (I-D)

Zasada pomiaru czasu lotu wiązki jest stosowana w czujnikach CamCube firmy PMD Technologies (rys. 12a). Mają one matrycę o rozdzielczości $204 \text{ px} \times 204 \text{ px}$, umożliwiającą obserwację pola widzenia o wymiarach $40,0^\circ \times 40,0^\circ$ oraz zasięg 7 m. Dla maksymalnej rozdzielczości czujnik pracuje z częstotliwością do 40 Hz. Analogicznie do czujników z rodziny SwissRanger, CamCube tworzy również obraz intensywności. Niestety, ten popularny czujnik został wycofany z produkcji. Firma specjalizuje się w produkcji matryc ToF PhotonICs 19k-S3 oraz ich referencyjnej konstrukcji – CamBoard Nano (rys. 12b). Sensor ten, w porównaniu



Rys. 12. Produkty firmy PMD Technologies zwracające mapę głębi z dodatkową informacją dot. intensywności (I-D)

Fig. 12. Sensors acquiring depth maps with additional intensity information (I-D) from PMD Technologies

do poprzednika, ma niższą rozdzielczość ($160 \text{ px} \times 120 \text{ px}$), zaletą natomiast jest ponad dwukrotnie wyższa częstotliwość pracy (90 Hz). Konstrukcja referencyjna cechuje się zasięgiem do 2 m (ograniczonym przez moc diody IR) oraz szerokim polem widzenia (90° w poziomie oraz 68° w pionie), dzięki czemu może on być idealnym czujnikiem do integracji z chwytakiem robota usługowego.

Kolejnym czujnikiem działającym na tej zasadzie jest SpectroScan 3D firmy Spectrolab (rys. 13). Czujnik ten ma sensor o rozdzielczości $256 \text{ px} \times 128 \text{ px}$, pole widzenia o rozmiarze 60° w poziomie i 30° w pionie oraz zasięg równy 20 m. Niestety, w porównaniu do poprzednich czujników, charakteryzuje się stosunkowo niską częstotliwością pracy (5–6 Hz).



Rys. 13. Czujnik SpectroScan 3D firmy Spectrolab

Fig. 13. SpectroScan 3D sensor offered by Spectrolab

W listopadzie 2013 r. rozpoczęto sprzedaż następnej generacji konsoli Xbox, razem z którą dystrybuowana jest nowa generacja czujnika Kinect (rys. 14a). Poza widoczną w obudowie kamerą kolorową (obraz RGB w rozdzielczości Full HD), czujnik ma również pracującą w podczerwieni kamerę ToF (rys. 14b), generującą mapę głębi wraz z obrazem intensywności w rozdzielczości $512 \text{ px} \times 424 \text{ px}$. Czujnik ten jest pierwszym na świecie zintegrowanym urządzeniem zwracającym obrazy RGB-I-D. W porównaniu z poprzednikiem, nowy Kinect może pracować z taką samą częstotliwością (30 Hz) oraz szerszym polem widzenia, co poprawi rozpoznawanie postury oraz ruchów graczy w małych pomieszczeniach.



a) Nowy czujnik Kinect w obudowie



b) Elementy nowego czujnika Kinect związane z akwizycją obrazu RGB-I-D [3]

Rys. 14. Czujnik Kinect drugiej generacji (Xbox One) firmy Microsoft

Fig. 14. Second generation of the Microsoft Kinect sensor (Xbox One)

Tab. 1. Zestawienie parametrów wybranych czujników

Tab. 1. List of parameters of selected sensors

	Rozdzielczość [px × px]	Częstotliwość pomiarowa [Hz]	Zasięg minimalny [m]	Zasięg maksymalny [m]	Dokładność dla zasięgu maks. [mm]	Odległość dla dokładności 20 mm [m]	Kąt widzenia w poziomie [°]	Kąt widzenia w pionie [°]		
Stereowizja									Rozstaw kamer [cm]	Przetwarzanie
Videre STOC	640 × 480	30	0,23 1	4 8,7	100 100	1,7 3,9	73	53	6 30	FPGA OnBoard
PointGrey Bumblebee 2	640 × 480	48	0,1 0,3	4 7	100 100	1,9 3,2	97 43	72 32	12	Programowo
PointGrey Bumblebee XB3	1280 × 960	16	0,1 0,3	7 12	100 100	3,7 6,3	97 43	72 32	12 + 24	Programowo
Focus Robotics nDepth	752 × 480	30	0,9 0,5	10 5,3	700 400	1,7 1,3	31 57	19 36	6	FPGA PCI
Surveyor SVS	1280 × 1024	15					82	61	10,75	Programowo
Sony PS 4 Camera	1280 × 800	60	0,3				85			
Oświetlenie strukturalne									Długość fali [nm]	Dodatkowe informacje
PrimeSense	640 × 480	30	0,8 0,35	3,5 1,4	20	† †	57	45	NIR 827	Obraz RGB
Microsoft Kinect (X360)	640 × 480	30	0,6	4		†	57,5	43,5	NIR 827	Obraz RGB
Occipital Structure	640 × 480 320 × 240	30 60	0,4	3,5+	3,5	†	58	45	NIR 827	Obraz RGB
ToF									Długość fali [nm]	Dodatkowe informacje
SICK LMS100/LMS111	1081 × 1 541 × 1	25 50	0,05	18	30		270	nd.	IR 905	Obraz I
Hokuyo UTM-30LX	1081 × 1	40	0,1	10 30	30 50		270	nd.	IR 905	
Leutze ROD4 plus	528 × 1	25	0	65			190	nd.	IR 905	
Fraunhofer 3DLS (LMS200)	721 × 501 181 × 125	0,0375 0,3125	0,1	10	15	10	180	124	IR 905	Obraz I
Velodyne HDL-64E	6250 × 64† 3125 × 64† 2083 × 64†	5 10 15	0,9	120			360	26,8	IR 905	
SwissRanger SR3000	176 × 144	25	0,1	7,5		3	47,5	39,6	NIR 850	Obraz I
SwissRanger SR4000	176 × 144	30 15	0,1 0,1	5 10		5 5	43,6 69	34,6 56	NIR 850	Obraz I
PMD CamCube 2,0	200 × 200	40	0,1	7,5					NIR 850	Obraz I
PMD CamBoard Nano	160 × 120	90	0,1	2		2	90	68	NIR 850	Obraz I
Spectrolab SpectroScan 3D	256 × 128	5-6	0,1	20			60	30	IR 1550	Obraz I
Microsoft Kinect (XBO)	512 × 424	30	0,8	4			70	60	IR	Obraz RGB 1920 × 1080 + obraz I

Oznaczenia: n.d. – nie dotyczy (np. w przypadku kąta widzenia w pionie czujników planarnych);

† błąd estymacji głębokości czujnika mniejszy od ±20 mm w całym zakresie pomiarowym;

‡ 16 dolnych linii czujników Velodyne ma czterokrotnie mniejszą rozdzielczość (odpowiednio 1532, 781 oraz 521);

wszystkie pola puste oznaczają brak danych.

4. Podsumowanie

Czujniki dostarczające informacje o głębi pełnią coraz istotniejszą rolę w pomiarach, automatyce i robotyce. Zakres ich zastosowań obejmuje zagadnienia od kontroli poprawności wykonania elementu, przez budowę trójwymiarowych modeli obiektów, pomiary w geodezji i kartografii, inspekcje stanu dróg, autostrad i mostów, aż po telewizory i kontrolery do konsol.

Celem dwuczęściowego artykułu był przegląd czujników do akwizycji obrazów RGB-D. W pierwszej części omówiono metody akwizycji obrazów, zwracając wybór do metod stosowanych w aplikacjach czasu rzeczywistego – do stereowizji, metody światła strukturalnego oraz pomiaru czasu lotu wiązki. Natomiast w drugiej części przedstawiono przegląd czujników bazujących na tych metodach. Parametry omówionych czujników zestawiono w tab. 1. Mimo starań, przegląd ten nie jest i nie może być pełny. Wynika to przede wszystkim z nieustannego postępu na tym polu, jak i z faktu, że niektóre z rozwiązań nie są dostępne komercyjnie, a inne są dostępne tylko w wyspecjalizowanych produktach.

Przykładem wyspecjalizowanych produktów jest np. system pomiarowy MX8 (rys. 15a) firmy Trimble, który składa się z zestawu czujników (kamera, skaner laserowy, GNSS, czujnik IMU). Wraz z komputerem integrującym dane ze wszystkich czujników MX8 przeznaczony jest do zbierania danych przestrzennych z pokładu poruszającego się pojazdu, np. podczas inspekcji stanu dróg, autostrad, mostów etc. Podobne rozwiązania oferuje firma RIEGL w postaci systemu pomiarowego VMX-250 (rys. 15b).



a) Trimble MX8 zamontowany na dachu samochodu [8]

b) Czujnik VMX-250 firmy Riegl na dachu pociągu [12]

Rys. 15. Komercyjne laserowe systemy pomiarowe 3D

Fig. 15. Commercial 3D laser measurement systems

Czujniki RGB-D stosowane są również w robotach usługowych. Na rys. 16a pokazano głowę robota PR2 firmy Willow Garage, w konstrukcji której zastosowano metodę uzyskiwania mapy głębi za pomocą stereopary oraz oświetlania sceny światłem strukturalnym. Podczas pracy projektor wyświetla (pseudo) losowe wzorce, dzięki czemu stereowizja wyznacza gęstą mapę głębi. Rozwiązanie to zostało przez autorów nazwane stereowizją z projekcją tekstury (ang. *Projective Texture Stereo*) [15, 20].

Na rys. 16b pokazano wnętrze tzw. aktywnej głowy (ang. *Active Head*) robota Velma [27], skonstruowanego



a) Aktywna stereopara robota PR2 [15]



b) Aktywna głowa robota Velma

Rys. 16. Aktywne głowy robotów wyposażone w czujniki RGB-D
Fig. 16. Robotic active heads equipped with the RGB-D sensors

w Instytucie Automatyki i Informatyki Stosowanej Politechniki Warszawskiej. Poza parą kolorowych kamer o wysokiej rozdzielczości (kamery Black Fly firmy Point Grey zwracające obrazy o rozdzielczości 1280 px × 1024 px) głowę wyposażono w dwa czujniki Kinect pierwszej generacji zamocowane w sposób, który pozwala na uzyskanie głębi dla kąta widzenia poziomego równego około 87° oraz kąta widzenia pionowego równego 57,5°. Układ wizyjny zamontowany na głowicy uchylno-obrotowej ma dodatkową jednostkę inercyjną, której zadaniem jest poprawa pozycjonowania całej głowy.

Podsumowując przegląd czujników RGB-D można stwierdzić, że przyszłość należy do czujników wykorzystujących pomiar czasu lotu wiązki. Sugeruje to nie tyle liczba różnych czujników dostępnych na rynku, co raczej fakt, że firmy pokroju Microsoft czy Samsung rozwijają własne produkty bazujące na tych technologiach. Dodatkowo, czujniki ToF mają przewagę nad stereowizją czy światłem strukturalnym (praca na zewnątrz, zakres, częstotliwość, obraz intensywności etc.), natomiast ich główną wadą jest wysoka cena. Jednak komercjalizacja kamer ToF oraz ich produkcja na masową skalę (stosowane w sterowanych gestami telewizorach typu F8000 HD TV firmy Samsung oraz w nowych kontrolerach do gier typu Kinect firmy Microsoft) spowoduje znaczne obniżenie ich ceny, a w konsekwencji popularyzację.

W 2014 r. czeka nas najprawdopodobniej rewolucja na tym polu, która być może dorówna tej z 2010 r., kiedy na rynku pojawił się czujnik Kinect pierwszej generacji.

Zaproszenie

Osoby zainteresowane tematyką poruszaną w artykule zapraszamy do udziału w sesji specjalnej *Percepcja robotów z wykorzystaniem obrazów RGB-D*, organizowanej w ramach 13. Krajowej Konferencji Robotyki, która odbędzie się w Kudowie Zdrój w dniach 2–6 lipca 2014 r. www.kkr13.pwr.wroc.pl

Podziękowania

Praca finansowana ze środków grantu Dziekana Wydziału Elektroniki i Technik Informatycznych nr 504M/1031/0044.

Bibliografia

- [www.focusrobotics.com], Focus Robotics.
- [www.3d-scanner.net], Fraunhofer IAIS.
- [www.wired.com/gadgetlab/2013/05/xboxone-development-photos], *From green light to boot-up: Behind the scenes of xbox one's development*.
- [www.gom.com/metrology-systems/3dscanner.html], Gom.
- [www.leuze.com], Leutze electronic.
- [www.mesa-imaging.ch], MESA Imaging.
- [www.ptgrey.com/products/stereo.asp], Point Grey Research, Inc.
- [www.riegl.com], RIEGL Laser Measurement Systems.
- [www.smarttech3d.com], SMARTTECH 3D.
- [http://structure.io], Structure sensor.
- [www.surveyor.com], Surveyor corporation.
- [www.trimble.com], Trimble.
- [http://velodynelidar.com/lidar/lidar.aspx], Velodyne.
- [www.videredesign.com], Videre Design.
- [www.willowgarage.com/pages/pr2/overview], *Willow Garage – PR2 robot overview*.
- Hamdan H.M., Hemayed E.E., Farag A.A., *Fast 3d object reconstruction using trinocular vision and structured light*, Proc. of SPIE, Intelligent Robots and Computer Vision XVII: Algorithms, Techniques, and Active Vision, Boston, 1998, 444–454.
- Kamiński M., *Przemysłowe bezdotykowe czujniki odległości w pomiarach wielkości geometrycznych*, „Pomiary Automatyka Robotyka”, vol. 17, nr 3, 2013, 32–40.
- Kim M.Y., Cho H., *An active trinocular vision system of sensing indoor navigation environment for mobile robots*, “Sensors and Actuators A: Physical”, vol. 125, no. 2, 2006, 192–209.
- Kitamura Y., Yachida M., *Three-dimensional data acquisition by trinocular vision*, “Advanced Robotics”, vol. 4, no. 1, 1989, 29–42.
- Konolige K., *Projected texture stereo*, International Conference on Robotics and Automation, IEEE, 2010, 148–155.
- Majek K., Pełka M., Będkowski J., Cader M., Masłowski A., *Projekt autonomicznego robota inspekcyjnego*, „Pomiary Automatyka Robotyka”, vol. 17, nr 2, 2013, 278–282.
- Montemerlo M., Becker J., Bhat S., Dahlkamp H., Dolgov D., Ettinger S., Haehnel D., Hilden T., Hoffmann G., Huhnke B., Johnston D., Klumpp S., Langer D., Levandowski A., Levinson J., Marcil J., Orenstein D., Paefgen J., Penny I., Petrovskaya A., Pflueger M., Stanek G., Stavens D., Vogt A., Thrun S., *Junior: The stanford entry in the urban challenge*, „Journal of Field Robotics”, vol. 25, no. 9, 2008, 569–597.
- Okutomi M., Kanade T., *A multiple-baseline stereo*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 15(4), 1993, 353–363.
- Rusu R.B., Cousins S., *3D is here: Point Cloud Library (PCL)*, International Conference on Robotics and Automation, Shanghai, China, 2011.
- Stefańczyk M., Kornuta T., *Akwizycja obrazów RGB-D: metody*, „Pomiary Automatyka Robotyka”, vol. 18, nr 1, 2014, 82–91. DOI: 10.14313/PAR_203/82
- Szynkiewicz W., Chojecki R., Rydzewski A., Majchrowski M., Trojanek P., *Modułowy robot mobilny elektron*, [w:] Tchoń K. red., *Postępy Robotyki: Sterowanie, percepcja i komunikacja*, vol. 1, 265–274. WKiŁ, Warszawa 2006.
- Wałęcki M., Stefańczyk M., Kornuta T., *Control system of the active head of a service robot exemplified on visual servoing*, 9th Workshop on Robot Motion and Control, 2013, 48–53. ■

Acquisition of RGB-D images: sensors

Abstract: The paper constitutes the second part of the article devoted to RGB-D sensors. In the first part the attention was focused on the methods of acquisition, whereas in the following, second part we present an overview of currently available commercial sensors.

Keywords: depth map, point cloud, RGB-D image, RGB-D sensor

Artykuł recenzowany, nadesłany 02.11.2013 r., przyjęty do druku 20.12.2013 r.

dr inż. Tomasz Kornuta

Absolwent Wydziału Elektroniki i Techniki Informatycznych Politechniki Warszawskiej. W 2003 r. uzyskał tytuł inżyniera, w 2005 r. tytuł magistra inżyniera, a w 2013 r. stopień doktora nauk technicznych. Od 2008 r. pracuje w Instytucie Automatyki i Informatyki Stosowanej, od 2009 r. pełni funkcję kierownika Laboratorium Podstaw Robotyki. Zainteresowania naukowe – metody programowania robotów, wykorzystanie informacji wizyjnej w robotyce, w szczególności aktywna wizja oraz rozpoznawanie obrazów RGB-D. Autor/współautor ponad trzydziestu publikacji dotyczących powyższych tematów. Recenzent krajowych oraz międzynarodowych konferencji robotycznych (KKR, IEEE mmAR, IEEE ICAR, IFAC SYROCO) oraz czasopism (Sensor Review, International Journal of Advanced Robotics). Członek IEEE RAS.

e-mail: tkornuta@ia.pw.edu.pl



mgr inż. Maciej Stefańczyk

Absolwent Wydziału Elektroniki i Techniki Informatycznych Politechniki Warszawskiej. W 2010 r. uzyskał tytuł inżyniera, w 2011 r. tytuł magistra inżyniera, obydwa z wyróżnieniem. W 2011 r. rozpoczął pracę nad doktoratem na temat zastosowania aktywnej wizji wraz z systemami opartymi na bazie wiedzy w systemie sterowania robotów. Główne zainteresowania naukowe – zastosowanie informacji wizyjnej w robotyce oraz w systemach rozrywki komputerowej.

e-mail: stefanczyk.maciek@gmail.com

